Habit Formation and Political Persuasion: A Behavioral and Statistical Approach

by

Amir Tohidi Kalorazi

B.S., Sharif University of Technology, 2017

Submitted to the Institute for Data, Systems, and Society
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN
SOCIAL AND ENGINEERING SYSTEMS AND STATISTICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2023

Authored by: Amir Tohidi Kalorazi
Institute for Data, Systems, and Society
September 1, 2023

Certified by: Ali Jadbabaie
JR East Professor of Engineering
Thesis Supervisor

Certified by: Dean Eckles
Mitsubishi Career Development Professor of Marketing
Thesis Supervisor

Accepted by: Fotini Christia
Ford International Professor of the Social Sciences
Chair, Department Committee on Graduate Students

Habit Formation and Political Persuasion: A Behavioral and Statistical Approach

by

Amir Tohidi Kalorazi

Submitted to the Institute for Data, Systems, and Society
on September 1, 2023, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY IN
SOCIAL AND ENGINEERING SYSTEMS AND STATISTICS

Abstract

This thesis explores the complex dynamics of human behavior across diverse contexts, integrating perspectives from behavioral science and statistical analysis. The central focus of this study revolves around the analysis of repetitive behavior in various scenarios including shopping, social media use, and news sharing.

The initial study investigates the influence of habits on the in-store shopping experience. By leveraging store closures as a disruptive event, we examine how these closures prompt individuals to alter their purchasing patterns. We propose that such disruptions encourage people to engage in more deliberate decision-making processes, leading them to explore alternatives that they might have previously overlooked due to established habits. Employing a difference-in-differences framework, we estimate the causal impact of habits on brand loyalty. Our findings reveal a significant role of habits, with households exhibiting stronger habits experiencing a temporary disruption in their shopping routines following store closures. Over time, these households appear to develop new habits in different stores, resulting in lasting changes in preferred brands. This suggests that the formation of shopping habits can lead to suboptimal consumer behavior. These insights have practical implications for businesses, including pricing strategies, advertising approaches, and product placement within stores.

The second study introduces an innovative methodology for quantifying habitual behavior in the context of social media usage. Interactions with social media platforms often yield psychological rewards, fostering the development of habitual behaviors driven by cue-response associations. By leveraging entropy as an implicit measure of behavioral regularity, this study aims to uncover the intricate relationship between habit formation and digital routines. Through empirical analyses, we establish the validity of the entropy metric, demonstrating its effectiveness in capturing distinct behavioral patterns beyond mere frequency. Our results highlight the nuanced connection between entropy and future app engagement, indicating a positive association for lower entropy values and a significant decline for excessively irregular patterns. These findings contribute to theoretical understanding of habitual behavior and offer practical insights for managing digital habits. Ultimately, this work advances our comprehension of how habits manifest in the digital realm and provides a robust tool for predicting long-term user behavior.

The third study delves into the intricate interplay between individuals' beliefs and their ability to anticipate the persuasive impact of climate change news articles. The central aim is to determine whether climate change deniers or believers possess varying capacities to predict the persuasive consequences of articles emphasizing climate change severity. Through a series of surveys, we gather predictions about the impact of such articles on climate change deniers. Surprisingly, findings reveal discordant predictions: deniers anticipate a backfire effect among peers, climate believers foresee negligible effects. We rigorously test these predictions with a randomized survey experiment involving deniers, uncovering an unexpected positive opinion shift towards climate change after article exposure. Notably, this effect does not translate into discernible changes in stated or revealed support for climate change actions. In the context of the pressing climate challenge, our study offers insights to inform targeted communication and interventions that foster consensus and meaningful action.

Thesis Supervisor: Ali Jadbabaie
Title: JR East Professor of Engineering

Thesis Supervisor: Dean Eckles
Title: Mitsubishi Career Development Professor of Marketing

Acknowledgments

My work and the creation of this thesis owe their existence to the unwavering support of numerous friends and colleagues. I wish to extend my heartfelt gratitude to my esteemed supervisors, Prof. Ali Jadbabaie and Prof. Dean Eckles. Throughout my journey in pursuit of my PhD, I was fortunate to experience a remarkable degree of independence. This privilege can be attributed to Ali's consistent encouragement, which motivated me to nurture my unique ideas and research identity. Embarking on the social and engineering systems PhD program without a background in empirical social science, I gradually developed a burgeoning fascination for empirical work as I engaged in more classes and seminars. This evolution wouldn't have been possible without Ali's steadfast guidance and the provision of resources that were unparalleled elsewhere. Ali's diverse knowledge spanning multiple disciplines and his extensive network of collaborators enriched my understanding of various fields, while also connecting me with individuals from diverse research backgrounds. His insights equipped me with the aptitude to formulate pertinent societal inquiries and to approach them with scientific rigor, an invaluable lesson that I am grateful to have imbibed.

I extend my deepest gratitude to Dean for his mentorship and unwavering support throughout this journey. Dean consistently provided ingenious solutions, particularly during moments of project despair, when a path forward seemed elusive. He played a pivotal role across all phases of our research, from shaping hypotheses to refining estimation procedures. Dean's role as a distinguished social scientist shines through his broad spectrum of interests and profound expertise across diverse domains. He's not only an adept statistician but also scrupulously upholds the validity and rigor of statistical methods in research. His capacity to swiftly comprehend research methodologies, coupled with his unwavering dedication to accurate inferences, is awe-inspiring. I've always admired his meticulous attention to detail, all while maintaining a panoramic view of the research landscape—an attribute I aspire to emulate. It's a genuine honor to have collaborated with Dean, and I eagerly anticipate the prospect of our continued partnership, wherein I aim to perpetually learn from his wealth of insights.

I want to extend my heartfelt appreciation to the other members of my thesis committee, Prof. Adam Berisky and Prof. Drazen Prelec. Their consistent availability and willingness to engage in discussions about research questions have been invaluable. And their insightful feedback and knowledge of essential relevant references have greatly enriched my work.

My collaborators have been an unceasing wellspring of inspiration, and I would like to express my deepest gratitude for their unwavering support and substantial contributions to my work. A special acknowledgment is owed to my collaborators at IDSS: Manon Revel and Mohsen Bahrami. I am equally thankful to my esteemed friends both at IDSS and across MIT, with whom I engaged in numerous insightful conversations. These include Eaman Jahani, Paolo Bertolotti, Marie Charpignon, Yan Jin, Laon Yao, Yuan Yuan, Madhav Kumar, and Jeremy Yang. A Special thanks goes to Christos Nicolaides for fruitful discussions about habits that laid the foundation for the ideas developed in chapter 3 of this thesis. My appreciation also extends to collaborators beyond MIT, particularly my mentor during my internship at the World Bank, Samuel Fraiberger. It was during this internship that the fourth chapter of this thesis took root. The collaborative experience with Sam, alongside his team comprising Stefano and Anca Balietti, and Manuel Tonneau, has been both enriching and enjoyable. I eagerly anticipate the continuation of our collaborative journey.

I am profoundly grateful for the unwavering love and support extended to me by my cherished friends both within the vibrant MIT community and the broader Boston area. Our shared experiences have formed the bedrock of unforgettable memories and cherished moments. To my dear friends Sajjad Mohammadi, Alireza Fallah, Mehrdad Khani, Zahra Hejrati, Maryam Aliakbarpour, Tina Torkaman, Ali Fahimniya, Soroush Araei, Mohammad Khorshidian, Arash Nasr, Pouya Hamedanian, Tara Broushaki, Setayesh Radkani, Hengameh Bagherian, Alek Bedroya, Ali Vakilian, Sepideh Mahabadi, Ardavan Saeedi, Asma Ghandeharioun, Majid Sabbagh, Davoud Ebrahimi, Mahshid Pourmand, Tahereh Marvdashti, Mohammad Alizadeh, Ameneh Shamekhi, Mohammad and Vahid Rashidian, Zohreh Farsi, Mohsen Goudarzi, Narges Afshordi, Hamid Ebrahimi, Leila Keyvani, Latifeh Avani, Nickolas Boylston, and Payam Mohseni, your camaraderie has illuminated even the most challenging of times, and I am indebted to the enriching conversations and laughter we've shared. Your presence has been an anchor, reminding me of the beauty in forging connections while navigating the academic landscape.

A heartfelt and special thank you goes to Babak Fotouhi and Naghmeh Momeni, along with their adorable son Elia, for the unwavering love and support they've showered upon me. Their friendship transcends ordinary bounds, resembling that of a brother and sister, always present and especially so during moments of need and challenges. Their guidance and valuable advice, grounded in their profound wisdom, have been instrumental in navigating the intricacies of life's journey. Babak and Naghmeh, your presence has been a beacon of light, and your genuine care has left an indelible mark on my heart.

A special and heartfelt expression of gratitude is reserved for Dr. Rouzbeh Touserkani, whose exceptional podcasts and speeches have left an indelible impact on me. His podcasts have served as a wellspring of knowledge and wisdom, imparting invaluable insights across a spectrum of fields. Particularly notable is his ability to foster critical and unconventional thinking, which has been both enlightening and transformative. Through his podcasts, I came face-to-face with the boundaries of my own understanding and discovered a multitude of subjects previously unbeknownst to me. Beyond theoretical knowledge, Dr. Touserkani's words have had a profound influence on my lifestyle and have fundamentally altered my perspective of the world. My debt to him is immeasurable, and his contributions will forever resonate in my life's journey.

Lastly, and with the utmost significance, I extend my deepest and most profound gratitude to my family. Their unwavering support, boundless encouragement, and unending belief in me have been the bedrock of my journey. Through every challenge and triumph, their love has been a constant source of strength. To my parents,

brother, and extended family, your sacrifices and unwavering faith have shaped me into the person I am today. Your presence in my life is a gift beyond measure, and I am endlessly thankful for the foundation you've provided, both in my personal and academic pursuits. Your enduring love is my greatest blessing, and I cherish it beyond words.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In an ever-evolving world shaped by complex economic, social, and political dynamics, understanding human behavior remains a central challenge. The intricacies of human decision-making processes hold the key to comprehending and predicting behaviors across various domains. In recent decades, there has been a notable increase in the integration of psychological theories into other social science fields, including economics and political science. Specifically within the domain of economics, this integration has given birth to the field of behavioral economics, which has transformed our understanding of economic decision-making processes (Rabin 1998, Thaler 2016). Behavioral economics diverges from neoclassical economics, which assumes that individuals have well-defined preferences and make rational, self-interested decisions based on those preferences. The groundbreaking contributions of influential psychologists and economists, exemplified by seminal works like Simon (1955), Kahneman and Tversky (1979), and Thaler and Sunstein (2009), have shed light on the importance of psychological factors, cognitive biases, and heuristics in shaping economic behavior. By integrating psychological principles into economic analysis, behavioral economics provides a more comprehensive framework for explaining and predicting individual and collective economic choices (Chetty 2015, Mullainathan and Thaler 2000). Moreover, it paves the way for practical applications in policy design and decision-making processes (Madrian 2014).

In tandem, the field of political science has increasingly recognized the significance of drawing insights from psychological concepts to enhance its depth of inquiry (Huddy *et al.* 2013). This interdisciplinary approach, frequently termed behavioral political science or political psychology, is aimed at comprehending the psychological determinants that modulate political behavior, attitudes, and decision-making processes (Sears 1987). This approach's origins can be attributed to seminal works such as *Human Nature in Politics* by Wallas (1921), highlighting the longstanding interaction between these disciplines. Prominent research in political psychology has illuminated a wide range of topics, from voter behavior and political ideology to the formation of public opinion. Notably, Philip Converse's landmark work on "The Nature of Belief Systems in Mass Publics" (Converse 1964) elucidated the role of cognitive consistency and political sophistication in shaping individuals' political attitudes. Furthermore, research by Shanto Iyengar and Donald R. Kinder in their book "News That Matters" (Iyengar *et al.* 1987) revealed how media framing and selective exposure can substantially sway public opinion and political behavior. With the digital age's advent, research has further expanded into understanding political behaviors in online spaces. For example, Van Bavel and Pereira (2018) have delved into the dynamics of online discussions on social media concerning politically charged topics. Their study underscores how the utilization of moral and emotional language in these discussions contributes to escalating political polarization. This body of research serves to showcase the vast implications of incorporating psychological perspectives into political science, thereby enriching our understanding of contemporary political phenomena.

The integration of multiple social science fields has been matched by a marked rise in data availability. This 'data revolution', bolstered by advancements in computational capacities, has transformed our capacity to comprehend and dissect complex economic, social, and psychological phenomena (Einav and Levin 2014, Lazer *et al.* 2009). The availability of large-scale behavioral datasets, coupled with advancements in analytical tools, has significantly expanded our ability to study and predict human behavior, cognition, and mental processes in real-world contexts. For instance, Kosinski *et al.* (2013) demonstrated the potential of digital records of behavior, such as Facebook Likes, to accurately predict a broad range of sensitive personal attributes including sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, parental separation, age, and gender. These results underscore the predictive power of digital behavioral data, opening up new research avenues in understanding complex human behaviors and traits within our increasingly digitized world.

Historically, psychological research relied heavily on limited sample sizes and controlled laboratory experiments, thus constraining the generalizability of results (Henrich *et al.* 2010). Now researchers increasingly explore expansive quantities of real-world behavioral data sourced from social media, online platforms, and digital devices (Kosinski *et al.* 2016, Salganik 2019). This migration towards studying behavior within naturalistic contexts furnishes a more ecologically valid understanding of human behavior, encapsulating the intricacies and subtleties that emerge in real-world situations. Furthermore, access to real-time and longitudinal data has substantially enhanced our ability to track dynamic processes and temporal patterns. Researchers can now monitor behaviors and psychological states as they unfold over time, facilitating the examination of variations and alterations in response to contextual factors or interventions (Allcott *et al.* 2022, Amir and Levav 2008, Braghieri *et al.* 2022). This temporal perspective affords invaluable insights into the dynamics of human behavior, thereby enriching our understanding of the evolution and adaptation of individuals and societies. Finally, the use of online platforms and internet-based recruitment strategies has simplified the process of recruiting large and diverse samples (Birnbaum 2004). This expansion of participant pools enables a more inclusive understanding of human behavior across various contexts and populations, thereby enhancing the generalizability and applicability of research findings.

The advent of increased data availability and enhanced computational capabilities has been accompanied by the advancement of rigorous statistical methods and the utilization of cutting-edge machine learning techniques. The use of statistical methods to explore and analyze psychological phenomena has a deep-rooted history in the field of psychology. Tracing back to the 19th century, the early use of probability-based modeling and inference in psychological investigations was initiated by Gustav Fechner's seminal work on psychophysics (Fechner 1860, Stigler 1992). The early 20th century saw further advancements with the introduction of factor analysis by Charles Spearman as a means to investigate the structure of intelligence (Spearman 1904). This innovative approach set the groundwork for modern psychometrics, significantly enhancing methods for psychological measurement and assessment.

In recent years, the advancement of rigorous statistical methods and machine learning techniques has further transformed psychological research. Researchers now have the ability to extract meaningful insights from complex and unstructured data (Yim and Ramdeen 2015). Machine learning algorithms, in particular, have proven to be valuable in predicting behavior and uncovering subtle patterns that may not be apparent through traditional analyses (Yarkoni and Westfall 2017). A noteworthy example of this is the work by Buyalskaya *et al.* (2023), where a machine learning methodology was introduced to identify context variables associated with habitual behavior and infer the pace at which habits form. This fusion of psychological research and advanced statistical methods has opened up new pathways for in-depth, comprehensive explorations of human behavior.

This thesis capitalizes on recent advancements at the intersection of rigorous statistical methods, the use of large-scale datasets, and behavioral science to delve into a comprehensive understanding of human behavior across diverse economic, social, and political contexts. By embracing interdisciplinary approaches, our aim is to unravel intricate patterns, establish causal relationships, and illuminate the underlying drivers of human behavior. Our primary focus revolves around the examination of settings characterized by repetitive behavior, as a substantial portion of our daily lives is shaped by actions and behaviors that are recurrently performed over time. These repetitive actions hold significant implications, influencing various facets of our personal and social behavior through diverse psychological mechanisms. Extensive research has been conducted across different domains to explore the nature and effects of repeated behavior. For instance, investigations have delved into the habitual nature of actions such as hand washing (Hussam *et al.* 2017), exercising (Aarts *et al.* 1997), social media use (Allcott *et al.* 2022), news reading and sharing (Ceylan *et al.* 2023, Yadamsuren and Erdelez 2011), and shopping (Sheth 2020). Understanding these behaviors is crucial for obtaining insights into the functioning of individuals, institutions, and societies, thereby enabling informed decision-making processes that contribute to positive outcomes.

The following chapters focus on investigating three common and practical scenarios that we encounter in our daily lives: shopping, social media use, and news sharing. The first two chapters are rooted in the extensive literature on habit formation in psychology, which provides a foundation for understanding the underlying mechanisms of repetitive behavior. The final chapter explores the intersection of the theory of mind and political persuasion literature, aiming to uncover the psychological processes involved in political persuasion and attitude change. The subsequent sections of the introduction chapter delve into the motivation and significance of these projects, elucidating the underlying psychological theories that serve as the bedrock of our research. By drawing upon these theories, we strive to provide a strong theoretical framework and establish a solid conceptual foundation for our studies.

1 Habits

One of the most influential psychological mechanisms associated with repetitive behavior is the construct of habit. It is important to differentiate habits from habitual behavior, as habits refer to the underlying process that influences behavior (Rebar *et al.* 2020). Since William James (1890) argued that "habit covers a very large part of life", psychologists have posited a large role for habits, with Wood *et al.* (2002) concluding that more than a third of people's daily decisions could be considered to be habitual. Habits are of great importance because they allow individuals to perform tasks effortlessly, without requiring much mental effort or conscious deliberation. This cognitive efficiency enables the allocation of mental resources to other activities such as problem-solving and decision-making (Macrae *et al.* 1994). Moreover, habits can contribute to stress reduction and improved decision-making by providing a reliable set of behaviors to rely upon, thereby reducing the need for constant evaluation of options (Welle and Graf 2011). Additionally, habits have long-lasting effects on behavior due to their resistance to change. Once established, habits can persist over time and significantly influence individuals' actions (Wood *et al.* 2014). Even small positive changes in our thoughts and behaviors, when maintained over the long term, can result in substantial benefits for populations as a whole. Conversely, seemingly harmless negative routines in the short term can have detrimental long-term consequences (Wood *et al.* 2005). Recognizing the importance of fostering long-term change, researchers have dedicated efforts to studying habits in different contexts.

The study of habits extends to areas such as health and diet, where understanding habitual behaviors is useful for promoting healthier lifestyles (Gilbert and Khokhar 2008, Schwartz *et al.* 2011). Habits have also been explored in the context of hygiene practices, such as sanitization and hand washing, particularly relevant in the domain of public health (Hussam *et al.* 2017). Additionally, the role of habits in exercising and physical activity has been investigated, as researchers seek to understand how to promote consistent engagement in these beneficial behaviors (Aarts *et al.* 1997, Buyalskaya *et al.* 2023, Gardner and Lally 2013). Furthermore, habits have been studied in the context of news consumption and sharing, as individuals develop routines and automatic behaviors around their information-seeking and dissemination practices (Ceylan *et al.* 2023, Yadamsuren and Erdelez 2011). The examination of habits extends to consumer behavior, where understanding the habitual nature of choices and purchasing patterns can inform interventions aimed at promoting sustainable and responsible consumption (Sheth 2020, Verplanken and Wood 2006). Lastly, the influence of habits in social media usage has been a topic of interest, as individuals develop ingrained patterns of behavior in their digital interactions and content consumption (Allcott *et al.* 2022). However, despite extensive research on habits in psychology, neuroscience, and cognitive science, there is no universally agreed-upon definition of habits.

Indeed, older accounts of habit formation emphasized the role of automaticity and stimulus-response associations (Berridge 2021). William James, in his influential work "Principles of Psychology," described how habits eliminate the need for conscious will and deliberation by establishing automatic chains of successive nervous events (James 1890). According to James, as habits develop, each event in a chain of actions becomes linked to its appropriate successor, bypassing alternative options and conscious decision-making. During the first half of the 20th century, behaviorists further expanded on the notion of habits and addiction using the framework of automatic stimulus-response associations (Berridge 2021). From this perspective, the strengthening of associations between stimuli and responses occurs through the "law of effect", whereby actions that lead to satisfaction or reward are more likely to be repeated (Thorndike 1898). This reinforcement-based learning process is often referred to as habitual learning (Dickinson 1985). As a behavior becomes habitual, the repetition of the behavior in a specific context strengthens the association between the context and the behavior, resulting in automatic or habitual performance and reducing the need for deliberation before making a choice (Wood and Rünger 2016). The same approach to defining habits is also adopted in some of the contemporary research on the topic (Vandaele and Ahmed 2021).

While these accounts underscore the significance of automaticity and stimulus-response associations in habit formation, contemporary perspectives have expanded beyond a strict stimulus-response framework and embrace a more comprehensive understanding of the psychological mechanisms that underlie habits. Some psychologists have adopted a different approach to defining habits, primarily contrasting them with goal-directed behavior (Dickinson and Balleine 1994, Verplanken and Aarts 1999). Goal-directed behavior involves individuals actively considering the value of their current goals, evaluating environmental conditions, and weighing different contingencies. In contrast, habits are characterized by their insensitivity to changes that diminish the value of the habitual action. Experimental paradigms have been developed to assess habits by examining the persistence of behavior even in the absence of a reward or when the reward is no longer desirable (Adams and Dickinson 1981, Beshears *et al.*

2021, de Wit *et al.* 2012, Dickinson 1985). This insensitivity to changes in reward availability or desirability distinguishes habitual behavior from goal-directed actions.

Habit formation can arise from the repetition of goal-directed behavior in similar contexts, leading to a shift in behavioral control from goal-dependence to context-dependence (Danner *et al.* 2008). However, it is important to note a limitation of this paradigm, as it assumes that behavior is strictly either goal-directed or habitual, implying that low sensitivity to reward devaluation indicates a strong habitual response (Watson and de Wit 2018). It is crucial to recognize that behavior independent of goals does not necessarily depend solely on the context (Foerde 2018). To address this limitation, an alternative approach involves integrating reward devaluation and context change in a paradigm. Accordingly, a behavior can be classified as habitual if it demonstrates insensitivity to reward devaluation in a familiar context while remaining responsive to rewards in a novel environment (Neal *et al.* 2011, Thrailkill and Bouton 2015). This framework offers a more nuanced understanding of habit formation by considering both the contextual and reward aspects of behavior, capturing the complexities inherent in habitual behaviors.

One central issue in the ongoing debates surrounding habit definition is the challenge of measurement. Theory and measurement have a reciprocal relationship, where existing theories inspire measurement methods with strong construct validity (Haynes *et al.* 1995), and experimental findings further refine theories. However, the habit literature faces a significant challenge in the absence of a definitive measure for "true habit" (Rebar *et al.* 2018). Various measurement approaches have been proposed to capture the essential constructs inherent in the definitions of habitual behavior. Initially, early habit research relied on past behavioral frequency as a proxy for habit (Bagozzi 1981, Ronis *et al.* 1989). However, relying solely on past behavior lacks explanatory power, as behavioral frequency alone cannot differentiate between habit and non-habitual behaviors (Gardner *et al.* 2012). For instance, a physician might routinely prescribe the same medication daily to individuals afflicted with similar symptoms. However, merely adhering to this pattern of regularity does not necessarily classify this behavior as habitual.

Recognizing that habits are formed through the repetition of specific actions within a stable context, Wood *et al.* (2005) introduced the concept of a frequency-in-context measure. This measure represents the product of behavior frequency and context stability. However, it falls short in capturing the automaticity observed in many habitual behaviors. To address this limitation, researchers have proposed self-reported habit measures such as the Self-Reported Habit Index (SRHI) (Verplanken and Orbell 2003) and the Self-Reported Behavioral Automaticity Index (SRBAI) (Gardner *et al.* 2012). These measures rely on participants' self-assessment of their own automaticity. However, an important limitation of these measures is their dependence on individuals' subjective evaluation of their own automaticity, which may be influenced by factors such as memory recall and introspective accuracy (Hagger *et al.* 2015). To overcome the limitations of self-reported habit measures, an emerging direction in habit measurement research is the utilization of implicit measures. Implicit measures, widely employed in cognitive and social psychology, provide indirect assessments that do not rely on participants' subjective evaluations (Gawronski and De Houwer 2014). These measures are less susceptible to response biases and are not reliant on introspection, making them more objective and reliable compared to self-report measures (Greenwald *et al.* 2002).

1.1 Overview of Chapter 2

The second chapter of this thesis builds upon these established theoretical frameworks to investigate the phenomenon of habit in the context of grocery store shopping, contributing to the fields of psychology, consumer behavior, quantitative marketing, and economics. A key issue in quantitative marketing and economics is the presence of strong autocorrelation or "inertia" in consumption patterns across different product categories (Cunningham 1956, Dubé *et al.* 2010, Guadagni and Little 1983, Keane 1997, Seetharaman *et al.* 1999). The underlying causes of this inertia have been attributed to various factors (Liu-Thompkins and Tam 2013), including psychological switching costs (Dubé *et al.* 2010, Farrell and Klemperer 2007, Klemperer 1987), although the precise mechanisms generating these costs have not been fully articulated. Drawing from psychological theory and empirical studies, it is suggested that in-store shopping behavior tends to be habitual (Machín *et al.* 2020). However, no empirical study has explicitly demonstrated and measured the role of habits in this particular context. This motivates our research to explore how sudden disruptions to the shopping context, such as store closures, can influence consumers' decision-making processes. We propose that these disruptions prompt consumers to engage in more deliberate decision-making, potentially leading them to select products that align with their current preferences.

To examine the impact of shopping habits on in-store decisions, we leverage store closures as an exogenous shock that disrupts households' shopping behavior. The core idea is that each store closure presents an opportunity for households to explore new store environments where familiar contextual cues are absent, prompting a more considered decision-making process. This exploration may lead consumers to consider alternative options that are typically overlooked in their regular store. Our research design, which utilizes a significant context change to gauge the strength of habits, aligns with the habit discontinuity hypothesis (Verplanken *et al.* 2008, Verplanken and Wood 2006). According to this hypothesis, when a habit is blocked or suspended due to a change in context, individuals may actively seek information or advice and become more open to alternative choices. We employ Nielsen retail scanner and consumer panel data to identify store closures and detect changes in households' purchase decisions following the closure of a local store. State-of-the-art methods in econometrics and causal inference are applied to rigorously estimate the role of habits in shaping consumer brand choices. Our findings demonstrate that households with higher proxy measures for antecedent habits (i.e., a higher frequency of visiting the closing store) experience a temporary disruption in their shopping habits immediately following the closure. Subsequently, they appear to develop new habits over time in the newly visited stores. This finding remains robust even after accounting for the unavailability of their preferred brands. Furthermore, the temporary disruption in shopping habits leads to lasting changes in households' favorite brands, suggesting that formation of shopping habits could lead to sub-optimal behavior. These findings have significant managerial implications for firms in various areas such as pricing, advertising strategies, and the allocation and location of goods inside stores.

1.2 Overview of Chapter 3

The third chapter of this thesis extends the existing literature on habit measurement by introducing entropy as an implicit measure of behavioral regularity. Entropy, a well-established measure of uncertainty and randomness in probability distributions, is applied to the distribution of behavior over a 24-hour clock time period to quantify the level of regularity. Clock time is chosen as a relevant contextual factor, closely tied to daily routines and activities such as sleeping, eating, working, and commuting. Additionally, it serves as a proxy for other contextual cues, such as location. By calculating the entropy of the behavior distribution, we capture the extent of regularity in behavior patterns throughout the day. A low entropy value indicates that individuals tend to repeat certain actions at fixed times, reflecting a high level of regularity or habituality. On the other hand, a high entropy value suggests greater randomness or variability in the timing of behaviors, indicating a lower degree of habit formation. In essence, entropy serves as a metric for quantifying the stability of the contextual factors surrounding behaviors, which can facilitate the formation of habits. Thus, entropy estimates habit indirectly by assessing the likelihood that habit has formed under conducive conditions.

Furthermore, we delve into various estimation methods for entropy and illustrate the utility of this metric in the context of interactions with social media applications. Interacting with social media platforms often entails receiving psychological rewards, such as social validation, entertainment, or gaining new information. The repetitive experience of these rewards in response to specific cues can foster the formation of cue-response associations that drive habitual behavior (Anderson and Wood 2021). Given the potential for social media usage to become habitual, this setting holds particular relevance to our research. Considering that learned habits are relatively unaffected by changes in goal structures, we propose the hypothesis that entropy, serving as a proxy for habit, can effectively predict long-term user behavior. Specifically, we aim to examine its predictive power in relation to the frequency of app use and the amount of time spent on the app. To empirically validate this hypothesis, we employ multiple machine learning methods that carefully control for other predictors of user behavior. Our analysis leverages a comprehensive and extensive dataset on mobile usage, which captures timestamped entries each time a user opens an app. This dataset offers valuable insights into the specific apps utilized by the participants and is particularly well-suited for studying social media habits due to its accuracy and comprehensive behavioral data. The findings from our analyses carry significant implications for both the academic understanding of habitual behavior and the practical application of this knowledge in the design of digital technologies and interventions.

2 Theory of Mind

Repetitive behaviors not only shape our actions and choices through the formation of habits but also have a profound impact on our perception of the world and our interactions with others. In the digital era we inhabit, the act of sharing information and news has become a pervasive and repetitive behavior that permeates our daily lives. The proliferation of social media platforms, online news outlets, and instant messaging services has fueled a constant cycle of information dissemination among individuals (Bakshy *et al.* 2012). This repetitive behavior holds immense significance, as it directly influences how we perceive, interpret, and respond to the world around us (Conover *et al.* 2011). Understanding the dynamics of information sharing is important for unraveling human communication patterns, cognitive processes, and the intricate mechanisms underlying belief and attitude formation. Furthermore, the study of this phenomenon and its consequences, such as political polarization, offers valuable insights into the complex interplay of social, psychological, and cultural factors that shape our interactions, decision-making processes, and ultimately, the fabric of our society as a whole. Investigating information sharing behaviors in the digital landscape therefore represents a vital avenue for comprehending the underlying dynamics of human behavior and its societal implications.

In the process of sharing news, we implicitly rely on a mental model of how others react to certain information. Understanding the cognitive processes and mindset of others is a fundamental aspect of human social cognition. This cognitive ability, commonly referred to as theory of mind (ToM) or mentalizing, enables individuals to attribute mental states such as beliefs, desires, and intentions to oneself and others (Happé *et al.* 2017). The term "theory of mind" was initially coined by Premack and Woodruff (1978) when studying the ability of chimpanzees to distinguish between an agent's intentions and its overt behavior. Since then, ToM has been extensively investigated in various fields including psychology, cognitive science, and neuroscience (Frith and Frith 2006, Schurz *et al.* 2014). The significance of ToM lies in its role in acknowledging that others possess their own thoughts, perspectives, and emotions, which may differ from one's own. By understanding the mental states of others, individuals can predict and interpret their behavior, engage in perspective-taking, and adapt their own behavior accordingly (Astington and Jenkins 1995). While most studies on theory of mind focus on infants and toddlers to understand its developmental trajectory (Tomasello 2018), it has also been investigated in adults. For instance, Clutterbuck *et al.* (2023) explored the correlations between various socio-demographic factors, political beliefs, and ToM in adults. Their findings revealed that participant gender was the most significant predictor of ToM, while political beliefs did not exhibit a significant association with ToM.

Theory of mind provides a valuable framework for understanding how individuals share, interpret, and respond to information on social media platforms. A well-developed theory of mind allows individuals to empathize with others and engage in effective communication by considering their perspectives and mental states (Baron-Cohen 1997). Neuroscientific research supports the idea that the value of information sharing is influenced by communicators' thoughts about the mental states of the receivers. For instance, (Baek *et al.* 2017) conducted a study examining the neural mechanisms underlying the decision to share health articles with others. They found that individuals exhibited greater activity in the brain's mentalizing system when deciding whether to share health articles with others compared to when making other types of decisions, such as reading the articles themselves or evaluating their content. Importantly, the study also revealed a positive association between neural activity in these mentalizing-related regions and individuals' self-reported intentions to share health news information. These findings suggest that communicators adjust their sharing strategies based on their anticipation of receivers' mental states, aiming to enhance the intended impact of the shared content. This adaptive behavior, known as audience tuning, has been empirically demonstrated in numerous studies. For instance, research conducted by Barasch and Berger (2014) revealed that participants consistently adjusted their information-sharing behavior based on audience characteristics, such as the number of people receiving their messages.

Given the significance of theory of mind in comprehending others' mental states and predicting behavior, it is reasonable to explore its implications for political persuasion. Understanding the dynamics of persuasion is crucial for comprehending users' behavior on social media and the phenomenon of polarization. Psychological studies have demonstrated that adopting the perspective of others can enhance the effectiveness of communication (Traxler and Gernsbacher 1993), while neural research has shown that successful persuaders engage brain regions associated with understanding others' minds more effectively than unsuccessful persuaders (Falk *et al.* 2013). For instance, in the domain of sales, where understanding customers' mindsets is crucial, more successful salespeople exhibit stronger mentalizing abilities and greater neural activation in regions associated with mentalizing (Dietvorst *et al.* 2009). A comprehensive review article by Falk and Scholz (2018) provides further insights into the connections between theory of mind and persuasion, shedding light on the underlying mechanisms and implications of this relationship. This body of research underscores the importance of theory of mind in political persuasion and its potential to shape communication strategies and outcomes in the digital era.

2.1 Overview of Chapter 4

The fourth chapter of this thesis advances our understanding of the interplay between individuals' viewpoints on a controversial political topic and their capacity to anticipate and understand the mental processes of others. Inspired by the theoretical foundations of theory of mind and the dynamics of persuasion, this chapter takes a novel approach to operationalize and investigate this phenomenon. Specifically, it examines how individuals' own perspectives on contentious issues influence their ability to accurately predict the opinion shifts of others following exposure to relevant news articles. It is hypothesized that personal opinions may hinder one's understanding of others' viewpoints, particularly when there are fundamental disagreements on key issues. The context chosen for this exploration is climate change, a highly polarized topic with significant implications for society and the environment (Dunlap *et al.* 2016). The aim of this chapter is to determine whether climate change deniers or believers exhibit different levels of accuracy in predicting the persuasive impact of news articles that emphasize the significance of climate change issues. The hypothesis suggests that individuals who share similar beliefs (i.e., climate change deniers) may demonstrate greater accuracy in predicting these persuasive effects. This can be attributed to their empathetic understanding of their peers' perspectives and their ability to simulate potential reactions to new information. Conversely, it is expected that individuals who hold opposing beliefs (such as climate change believers) may face challenges in comprehending others' reactions, potentially hindering their predictive accuracy.

To accomplish this, we conducted a survey to carefully select a set of news articles that predominantly highlight the significance of climate change issues and the need for action. Subsequently, a follow-up survey was administered to a diverse pool of respondents representing different climate change stances, including believers, neutrals, and deniers. In this second survey, participants were asked to indicate how they expect a hypothetical reader with a positive, indifferent, or negative attitude towards climate change would change their opinion after reading the corresponding news article. This question aimed to capture individuals' expectations of others' opinion shift. Based on the survey results, individuals who believe in climate change do not anticipate persuading climate change deniers through the selected news articles. However, deniers themselves hold a different perspective. They predict a backfire effect, wherein fellow deniers are expected to become even more opposed to climate change policies after reading articles that emphasize the severity of climate change. To test these conflicting predictions, a randomized experiment was conducted. A group of climate change deniers was invited to read the selected articles and assess their persuasiveness. Their opinions were measured before and after engaging with the material, allowing for an evaluation of the extent to which their opinions shifted. Contrary to everyone's expectations, deniers show a positive opinion shift in favor of climate change after reading a newspaper article on climate change. The effect coincides with no impacts on either stated or revealed support for actions that fight climate change. No backfire effects are documented. The findings contribute to a deeper understanding of the interplay between personal beliefs, theory of mind, and the dynamics of persuasion in the context of climate change discourse.

Chapter 2

Habits in Consumer Purchases: Evidence From Store Closures

1 Preface

This chapter examines the impact of habits on consumer purchase behavior in the grocery shopping context. Previous research has shown that consumer choices often exhibit inertia, but it remains unclear how much of this inertia can be attributed to habits. Drawing from theories in psychology, consumer behavior, and quantitative marketing, we investigate the formation, persistence, and disruption of habits. By leveraging panel data on households' purchases across various product categories and using store closures as a natural experiment, we analyze the role of habits in repeated brand purchases. The findings shed light on the effects of habit formation on consumer decision-making and have implications for marketers, retailers, policymakers, and individuals seeking behavioral change.

2 Introduction

One area of daily life that involves repetitive behavior and plays a vital role in our lives is grocery shopping. Grocery retail represents a substantial portion of consumer spending, with billions of dollars exchanged annually in this sector. Recognizing the factors that shape consumer decision-making in grocery stores is of utmost importance, given the substantial economic impact of the grocery retail sector. Researchers, marketers, retailers, and policymakers can benefit from a deeper understanding of consumer behavior in grocery shopping, as it provides valuable insights for optimizing strategies, improving customer satisfaction, and fostering economic growth.

In the everyday experience of grocery shopping, consumers are confronted with a wide array of brand choices spanning various product categories. Making purchasing decisions in such a context can be a complex and challenging process, given the large and diverse choice sets, as well as the influence of marketing factors such as prices and promotions. Consequently, consumers might rely on heuristics and habits as efficient strategies for faster and less effortful decision-making. Notably, these habits develop as a result of repetitive behavior occurring within the same context. Psychological theory suggests that habits tend to form when individuals repeatedly engage in rewarding behaviors within a consistent context (Verplanken and Aarts 1999, Wood and Rünger 2016). Over time, these contextual cues become associated with the behavior, triggering an automatic response (Orbell and Verplanken 2010). For a comprehensive overview, refer to the review by Gardner (2015b). Consequently, it is reasonable to anticipate that in product categories where consumers frequently make purchases in similar settings, such as shopping at the same store, their purchase decisions may be driven by habitual behavior.

Habitual behavior in shopping could be one explanation of the empirical regularity that purchase decisions exhibit substantial temporal-dependency or inertia (e.g., Carrasco *et al.* 2004, Dubé *et al.* 2010). Drawing on insights from psychology, we hypothesize that such repeated behaviors in stable contexts are often the result of slow-learning, fast-acting (i.e. System 1) processes (Mazar and Wood 2018, Wood and Rünger 2016). Of course, other varieties of inertia can also be present, such as learning and preference formation which are felicitously described through standard choice models. Here we aim to detect and estimate a role for habits specifically in explaining observed inertia in consumer purchase behavior. Hence, part of the contribution of this chapter is to provide robust and rigorous empirical evidence to bridge this gap and bring insights from the psychological definition to a setting of broad interest in marketing, economics, public policy, and decision science. This research contributes to the broader understanding of consumer behavior and provides insights into the mechanisms that drive habitual consumption. By shedding light on the influence of habits on consumer decision-making, this chapter has implications for marketing practitioners, policymakers, and individuals alike, offering opportunities to design effective strategies, promote healthy choices, and facilitate positive behavioral change.

2.1 Related Work

Habit formation, persistence, and disruption have garnered attention from various disciplines, each offering unique perspectives on this phenomenon. One aim of this chapter is to adapt a theory of habit developed in the psychology and consumer behavior literatures to one of the most widely-studied settings in quantitative marketing, in this process clarifying the causes of stylized facts in quantitative marketing and economics. So we first consider definitions and theories of habit in psychology and then empirical work in quantitative marketing.

The study of habits is a topic of ongoing debate and controversy within the research community. While there is no consensus on the precise definition, multiple competing accounts converge on the idea that habits involve a certain form of automaticity Berridge (2021), Wood and Rünger (2016). Wood *et al.* (2014) characterize habits as being activated by recurring contextual cues and being insensitive to short-term changes in goals. On this and related accounts, the repeated performance of a behavior in the presence of the same contextual cues creates an association between the context and the behavior, making performing it in that context automatic or proponent, thereby reducing deliberation prior to choice (Wood and Rünger 2016). Importantly, habits are relatively resistant to changes in goals or payoffs, meaning that even if the rewards associated with the behavior diminish or disappear, the habit may persist. This is reflected in experimental paradigms that test for habits (or the strength of habits) by examining persistence of the behavior even when the reward is absent or no longer desired (Adams and Dickinson 1981, Beshears *et al.* 2021, Dickinson 1985). In an influential empirical study related to the present work, Wood *et al.* (2005) had participants report their frequency of performing exercise and media consumption behaviors before and after transferring to a university, which changed the context for some of those behaviors for some participants. The results showed that behaviors in changed contexts shifted to align more closely with participants' self-reported goals after the transfer, highlighting the influence of contextual changes on habit formation.

This account of habits — in which deliberative processing of attributes of many available options is often absent — is consistent with studies that have probed visual attention of in-store shopping with mobile eye-tracking devices. For instance, in a study of grocery shoppers in Uruguay, Machín *et al.* (2020) found that 67% of shoppers have a tendency to directly select the product they are seeking without comparing it to other options. These shoppers quickly put the chosen product in their shopping basket, taking an average of just 7 seconds from the moment they grabbed the product. These findings support the idea that in-store shopping behavior can be highly habitual, where consumers quickly and automatically select familiar products without engaging in extensive deliberation or comparison. Building on this understanding, our hypothesis posits that when the context of shopping is disrupted, such as through the closure of a frequently-visited store, consumers will be prompted to engage in more thoughtful and deliberate decision-making processes. In these situations, they may consider a wider range of product options and make choices that align with their current, carefully considered preferences.

In quantitative marketing and economics, researchers have extensively studied the presence of strong autocorrelation or "inertia" in consumption patterns across various product categories (Cunningham 1956, Dubé *et al.* 2010, Guadagni and Little 1983, Keane 1997, Seetharaman *et al.* 1999). This inertia is often captured by a variable known as *state dependence* or loyalty, which represents the influence of past purchases on current decisions. The concept of state dependence was first introduced by Guadagni and Little (1983), who demonstrated that incorporating past purchases as a measure of loyalty in a choice model significantly improves the model's fit. Subsequent research has frequently employed a Markovian assumption, considering the immediate past purchase as a proxy for the consumer's current "state" (e.g., Dubé *et al.* 2010, Levine and Seiler 2021, Simonov *et al.* 2020). Consequently, the phenomenon of positive inertia or loyalty is often referred to as state dependence in the literature. It is worth noting that while negative state dependence or variety seeking has been observed in certain cases (McAlister 1982), it is generally less prevalent (Adamowicz and Swait 2013).

The state dependence literature has dedicated considerable attention to validating estimation procedures and employing flexible models to address unobserved heterogeneity in consumer preferences (Dubé *et al.* 2010). Overall, the existing literature shows a general consensus that structural state dependence exists even when employing complex models that account for potential preference heterogeneity (cf. Levine and Seiler 2021). However, the estimated magnitude of state dependence tends to diminish after controlling for factors such as consumer heterogeneity (Dubé *et al.* 2010, Keane 1997, Simonov *et al.* 2020). Notably, Dubé *et al.* (2010) found that the estimated state dependence vanishes when the order of shopping trips is randomly permuted. This observation supports the presence of genuine temporal state dependence rather than it being an artifact of model misspecification. If past decisions did not exert temporal influence on future decisions and the estimated state dependence solely arose from unobserved heterogeneity, it would have remained statistically significant even with random trip reordering. Additionally, Seetharaman *et al.* (1999) examined various product categories and discovered correlated state dependence among households, further suggesting that this phenomenon extends beyond model misspecification.

Inertia in purchases can have multiple underlying causes (Liu-Thompkins and Tam 2013). One typical explanation in the literature is psychological switching costs

(Dubé *et al.* 2010, Farrell and Klemperer 2007, Klemperer 1987), although this does not articulate underlying mechanisms which generate these mental costs. However, considering this established evidence for state dependence with respect to a psychological theory of habit, we may want to distinguish the associative, slow-learning, fast-acting process underlying habits from a deliberative, fast-learning, slow-acting process. In the next section, we describe how our empirical strategy is chosen for this purpose. Here we first consider some other decompositions of state dependence proposed in the quantitative marketing and economics literatures; see Thomadsen and Seetharaman (2018) for a review. Note that sometimes these have used the term "habit", but in ways that diverge both from our use and each other.

In a latent utility framework, Roy *et al.* (1996) incorporate both what they call "habit persistence" and "structural state dependence" in their model.[1] While structural state dependence only depends on realized past choices, "habit persistence" takes into account how prior propensities to choose a brand affect current choices (Heckman 1981). So if the household has high evaluation of brand j in trip t but purchases brand i, the high evaluation of brand j would persist in trip $t+1$ even though it was not purchased. Note that this might be interpreted as a reversal of terminology from the psychological account described above, which centrally features (multiple) prior choices (not just positive evaluations).

Moreover, Seetharaman (2004) builds on the work of Roy *et al.* (1996) and allows for more complicated forms of "habit persistence" in a utility-based framework. Seetharaman (2004) defines habit persistence type 1 as "serially correlated error terms in the random utility function". This form of habit accounts for persistence in choices for reasons unknown to the researcher such as long holidays or having guests which might require successive purchases of the same brands. Habit persistence type 2 is

[1] In this model, *structural state dependence* is defined as a direct boost in utility at time t coming from purchasing the same brand as time $t-1$, while "habit persistence" is modeled as the serial correlation between consequent choices in a Markov process that can be present even in the absence of structural state dependence. Similar to many other papers in this literature, Roy *et al.* (1996) impose a Markov assumption and consider the immediate past purchase as a proxy for past behavior. However, this modeling choice is questionable regarding modeling habits, which typically form and change slowly overtime.

then defined as "serial correlations between utility-maximizing alternatives on successive purchase occasions of a household", which accounts for temporal dependencies in successive brand choices due to unobserved information signals such as billboards or television advertisements. However, the link between the notion of habits in these papers and the psychological view as cue–response associations in memory is not clear. In particular, these models capture persistence around specific brands and implicitly assume the shopping environment to be fixed. As a result, they are silent on any changes in choices if the decision is being made in a quite different context with the same choice set.

While developing considered preferences for specific brands can contribute to overall inertia, so can consequences of repeated choices in a stable context. In settings from voting (Cantoni and Pons 2022) to food consumption (Privitera and Zuraikat 2014), empirical researchers have argued that context effects have a substantial impact on individuals choices (Amir and Levav 2008). The importance of contextual cues in triggering habits, despite conflict with current goals, has been extensively studied by social psychologists (Neal *et al.* 2011). Despite the conceptual distinction between brand loyalty and habits, the link between psychological measures of habits and the state-dependence literature remains unclear, and these two areas have largely evolved separately. While Tam *et al.* (2014) discuss the conceptual differences, empirical evidence supporting this distinction is lacking. In this chapter, we aim to bridge this gap by providing new empirical evidence that aligns with a significant role for habits in state dependent consumer purchases.

2.2 Overview

This chapter focuses on analyzing in-store purchase behavior by utilizing store closures as a means to disrupt the context in which consumers make their purchases. Through this approach, we aim to enhance our understanding of the role of habits in shaping repeated purchases.

We need to distinguish between habits — defined as an association in memory between purchase responses and contextual cues — and both any complementarity in repeated consumption and other learned preferences for particular products and brands. That is, in addition to distinguishing inertia from consumer heterogeneity (Pakes *et al.* 2021), in order to attribute inertia to habits, we should also distinguish it from other processes. Following the discussion by Tam *et al.* (2014), we propose using the term "state dependence" as an overarching concept that encompasses various mechanisms through which past purchases influence current choices. Within this framework, we distinguish between two key concepts: brand loyalty and shopping habits. Brand loyalty refers to the psychological disposition of consumers to evaluate a brand favorably, independent of the specific context of the purchase and when considering a choice among different brand options (Neal *et al.* 2006, Tam *et al.* 2014). It reflects the tendency to consistently prefer a particular brand due to established expectations or preferences, leading to higher expected utility for that brand regardless of other features of the purchase occasion, such as the store context. In contrast, shopping habits involve the repeated purchase of the same brands within a specific store context. These habits can transcend the nature of the purchased product and are influenced by contextual factors, such as the store layout or the placement of specific brands on shelves. Shopping habits can arise as a strategy to minimize search costs (Dong *et al.* 2020) or as a decision heuristic to conserve mental resources for more significant tasks (Macrae *et al.* 1994).

In this chapter, we aim to identify and measure how shopping habits can affect consumers' in-store decisions. To this end, we leverage store closures as a shock that disrupts part of the households' shopping behavior. The key idea is that each store closure can potentially force households to explore new store environments, where previous contextual cues are no longer present and consumers are engaged in a more thoughtful and deliberative decision-making process — driving them to explore other options that are normally ignored in a familiar store. This research design, in which a severe change of context is used to measure the strength of habits, is also related to the habit discontinuity hypothesis (Verplanken *et al.* 2008, Verplanken and Wood 2006). The idea is that when a habit is blocked or suspended due to a change of context, the person may need to search for information or advice, and be open to alternative options. Some examples of change of circumstances that can disrupt people's habits include: transitions from school to work (Busch-Geertsema and Lanzendorf 2017), residential relocation (Clark *et al.* 2016), changes in retail contexts (Figueroa *et al.* 2019, Poortinga *et al.* 2013), and lifestyle changes due to COVID-19 restrictions (Oblander and McCarthy 2022).

We argue that the use of the store closures increases the credibility of causal inference about habits. One might alternatively consider any two adjacent trips by the same household. Along these lines, Thomadsen (2016) finds some evidence that consumers exhibit higher levels of state dependence if the store they are shopping from is the same store they visited last time. However, the choice of the store could potentially confound the choice of brands, i.e., people might have chosen to go to a different store in the first place in order to buy a different brand. As a result, one cannot simply consider changes in purchase locations because the choice of the store could be correlated with brand choices. In our framework, the closure induces a relative increase in visits to new stores or at least newly shopping for a particular product category in a store. The set of exposed households is not all impacted equally by store closures. We posit each household's purchase behavior is primarily affected for the subset of categories which they used to buy from the closing store and the intensity of the effect increases with the frequency of visits. From this perspective, the set of household–category pairs for which the household had never purchased that product category in the closing store can be considered as the control group for a difference-in-differences (DID) causal identification strategy. Hence, our identification strategy is based on a combination of different households being exposed to store closures at different times, as well as within-household variation in how much that household is exposed to a closure for a particular product category.

Modeling brand choices involves a complicated multi-choice problem with varying choice sets. One conventional modeling approach in the literature considers various forms of latent utility choice models. However, these models are readily applicable only to a single product category, and researchers typically limit their analysis to a few frequently purchased brands and ignore changes in the households' choice sets. Instead, we conduct a "reduced-form" analysis where the outcome is a binary variable indicating whether the household is choosing their most frequent brand option (i.e. modal brand), following the approach in Larcom, Rauch, and Willems (2017). This simplification will allow us to detect changes in purchase patterns after the closures by simultaneously modeling households' purchases across multiple product categories, use regression machinery to estimate causal effects, and avoid the difficulties of estimating a model with a high-dimensional categorical outcome. Following this framework, we first show that the subset of households with a higher proxy for antecedent habits (i.e. higher frequency of visiting the closing store) experience a temporary disruption in shopping habits right after the closure. They then apparently form new habits over time in the newly visited stores. This observation is robust to accounting for unavailability of their favorite brands. Furthermore, the induced temporary disruption in shopping habits results in lasting changes in households' *modal* brands (brands most often purchased) suggesting that formation of shopping habits could lead to sub-optimal behavior.

These results augment our understanding of the state dependant consumers purchase behavior by demonstrating the importance of shopping habits, in addition to pure brand loyalty. Our findings have immediate implications for firms who could benefit from understanding (or discovery) of these shopping habits by incentivizing stores to keep the placement of their brands consistent inside the store. However, depending on the brands for which habits are formed (for any specific store), competing firms could have conflicting interests regarding keeping the product placements constant. For a less popular brand, the firm has an incentive to pay the store to change their product placement in order to disrupt existing habits. This would be most effective if it complements other marketing strategies such as providing free samples or different forms of advertisement.

The rest of this chapter is structured as follows: Section 3 describes the Nielsen scanner data and explains how closing stores and corresponding exposed households were identified. Section 4 discusses the problem formulation, in particular, how treatment exposure level and outcome variables used in the regression model are defined. Section 5 provides the results for various two-way fixed effects (TWFE) and event study models using different outcome measures. It also presents results for a Bacon decomposition analysis (Goodman-Bacon 2021) to explore any bias in TWFE estimation due to differential treatment timing. Section 6 concludes and discusses potential implications of our findings.

3 Data

We use Nielsen retail scanner and consumer panel data, containing detailed shopping information for more than 50,000 American households and 35,000 stores across the US between January 2006 and December 2018. We utilized the retail scanner data to identify closing stores, and the consumer panel data to detect changes in households' purchase decisions after one of their local stores closes.

3.1 Retail Scanner Data

The retail scanner data contains weekly pricing, volume, and store merchandising conditions generated by retail store point-of-sale systems. The data includes approximately 35,000 stores including grocery, drug, and mass merchandiser stores. The data is available from January 2004, but we only used from 2006 onward since we only needed the store closures relevant to panelists in the consumer panel data. All stores have unique anonymized identifiers, so we could track the sales of each store even if the retail chain changes, although more than 96% of the identified closing stores in the data operate under a single retailer.

3.2 Consumer Panel Data

The Consumer Panel Data represents a longitudinal panel of approximately 40,000–60,000 US households who use hand-held scanner devices to continually provide information to Nielsen about their purchases. Products include all Nielsen-tracked categories of food and non-food items, across all retail outlets in all US markets. Nielsen samples all states and major markets so panelists are geographically dispersed and demographically balanced. Importantly, the consumer panel data can be linked to the retail scanner data using unique store identifiers. Since we need to follow households' purchased brands, we do not use the "Magnet" data which includes non-barcoded products such as fresh fruit.

3.3 Store Closures

In order to identify closing stores, we compute the aggregate store weekly sales using retail scanner data.[2] Then, we single out the stores whose sales drop to zero at a certain time and remain zero afterward. We found 7,847 such permanent store closures during a 13-year period starting in 2006. We also investigated potential temporary store closures. Considering stores whose sales drops to zero and remains zero for at least the duration of a year, we found only 83 such cases. Varying the required zero sales duration would change the number slightly, but overall, there were very few temporary closures. Furthermore, there was no instance of multiple closures for any store in the data set. As a result, given the small number of temporary closures, we decided to drop them altogether and only consider permanent closures to avoid further complications in the causal analysis.

We expected that some of these stores simply stopped participating in the panel, while remaining open. To exclude such false closure identifications, we used the consumer panel data and ruled out any store for which there was a reported purchase trip after the closure date. After this correction, 3,243 closing stores remained. We could also identify a false closure if a certain store and all the related customers in the panel opt-out of reporting to Nielsen simultaneously. Even though there is a low chance of this incident happening, it will not affect our results because we only

[2]For computational simplicity we only compute the overall weekly sales of each store for top-5 purchased product categories: Refrigerated milk, refrigerated yogurt, fresh bakery bread, cereal, and canned soup.

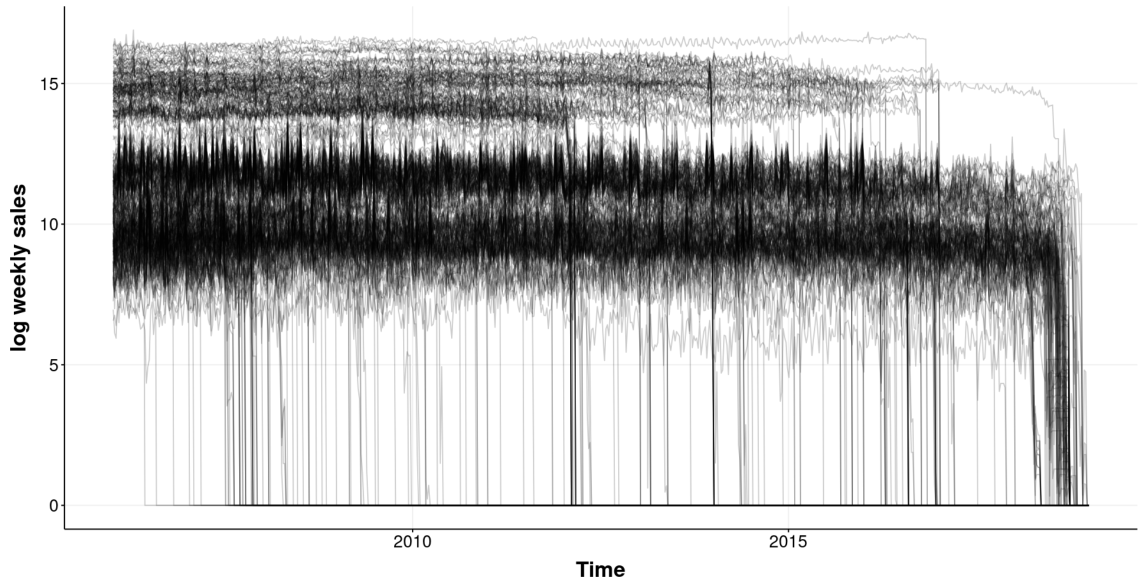


Figure 2.1: The figure shows log-weekly sales in the top-5 product categories by national purchase volume for the closing stores corresponding to the set of exposed households from January 2006 until December 2018.

consider households that are active both before and after their corresponding closure to measure the change in their behavior.[3]

Figure 2.1 shows log-weekly sales for the closing stores corresponding to the set of exposed households. Weekly sales by retailer, the distribution of closing stores over years, and their geographical dispersion are shown in Appendix A, Section 1.

3.4 Exposed Households

After matching the set of closing stores with the consumer panel data, we find 14,406 households who at some point in time visited one of the closing stores.[4] However, not all of these households were *exposed* to the exogenous closure shocks. Some of them might have visited the closing store months or years before it closed. We consider only households who were still shopping from the closing store near its closing time

[3]Finally, some households might stop going to a store around the time it drops out of the retail data, and therefore cause a false closure identification. Although this could potentially happen, it is less likely to happen for households who are frequently shopping from the closing store. And as we see in the following results, the main effect is driven by these more frequent visitors.

[4]In the final analysis, we only consider top-30 product categories and remove infrequent ones. This leaves us with 14,360 households.

to be *exposed* by the closure. Therefore, we marked households who had at least a shopping trip to the closing store within a 4-month interval prior to the corresponding closure date as the *exposed* set, which included 684 households with a total of 407,630 distinct shopping trips.

4 Framework and Definitions

In this section, first, we specify the treatment each household is receiving due to the store closures, and then define two distinct outcome variables which are used in our models.

4.1 Treatment Exposure

The set of exposed households are not all equally affected by a store closure. First, we posit each household's purchase behavior is primarily affected within the subset of categories that they used to buy from the closing store. For example, if someone frequently bought yogurt but not cereal from the closing store , we expect the closure affects their yogurt purchase behavior, with effects on cereal, if any, being much smaller. Second, we expect the effect to vary based on the prevalence of the shopping trips to the closing store. In order to capture both of these dimensions, we define the *treatment exposure level* ($e_{i,c}$) for household i and category[5] c as the relative fraction of household i shopping trips to the closing store in which a product in category c was purchased:

$$e_{i,c}(T_e) = \frac{\text{trips to the closing store by household } i, \text{ for category } c, \ T_e \text{ years before}}{\text{trips by household } i, \text{ for category } c, \ T_e \text{ years before}}. \tag{2.1}$$

The exposure level is a function of the pre-closure time period on which it is defined, T_e. The shorter we define this period, the better the fraction would capture

[5] Nielsen has a 3 level hierarchy for categorizing different products. There are 10 Departments, 125 product groups, and about 1100 product modules. For example, within the *Frozen foods* department, there are multiple groups including *frozen vegetables* or *frozen breakfast foods*. And within each group there could be multiple modules such as *frozen beans*, *frozen toaster items*, etc. Throughout the chapter by product category we mean the grouping at the product module level.

the true impact of the closure because someone could have a many shopping trips to the closing store many months before the closure but only a few such trips right before the closure. This household would be less likely affected by the closure. However, at the limit of $T_e \to 0$, we have zero observations to define the fraction, and with small values of T_e, measured exposure would be sensitive to a small number of recent trips. Since it is not obvious ex-ante what would be the optimal time period, we do the analysis for a range of values and show that the main results are robust to the choice of the T_e parameter. Results presented in the main text are for $T_e = 1$ year, and result for $T_e = 2$ and $T_e = \frac{1}{2}$ year are presented in Appendix A.

Figure 2.2(left) displays the treatment exposure levels for all household–category pairs. There are a total of 9,338 units (not shown) with zero exposure level; these household–category pairs can be used as the set of control units in a difference-in-differences framework with differential treatment timing.

Furthermore, to flexibly allow for potentially heterogeneous effects among household–category pairs, we partition treated pairs into four groups based on their exposure levels: $E_1 = \{(i,c)|e_{i,c} \leq 0.25\}$, $E_2 = \{(i,c)|0.25 < e_{i,c} \leq 0.5\}$, $E_3 = \{(i,c)|0.5 < e_{i,c} \leq 0.75\}$, $E_4 = \{(i,c)|0.75 < e_{i,c}\}$, and the control group is defined as $C = \{(i,c)|e_{i,c} = 0\}$. These groups are separated by dotted grey lines in Figure 2.2(left) where each point stands for a household–category pair. There are 1,636, 881, 641, 768 household–category pairs in E_1–E_4 correspondingly.

We expect household–category pairs with higher treatment exposure level to have more significant disruption in their purchasing behavior for two main reasons. First, according to psychological theory, more frequent trips make stronger shopping habits, and hence the disruption in brand choices could be more substantial (Wood *et al.* 2005). Second, the main channel through which the closure is affecting households purchasing behavior is the resulting forced exploration in visits to new store–category pairs[6] where old habitual cues are no longer present and people are susceptible to

[6] A purchase occasion at store s in category c is counted as *new store-category* visit for household–category (i,c) if household i purchases category c in store s during her first L trips after her corresponding closure date, while she was not purchasing any items in category c from store s during L

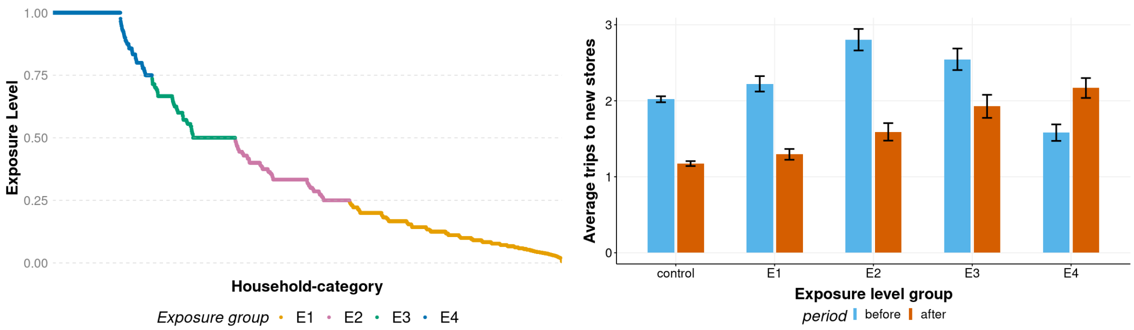


Figure 2.2: (left) Treatment exposure levels ($e_{i,c}$) for household–category pairs sorted in decreasing order, where each point shows a unique pair. Control group ($e_{i,c} = 0$) pairs are not shown in the figure. Dashed grey lines show how household–category pairs are categorized into four treatment exposure level groups. There are a total of 9,338 pairs in the control group, and 1,636, 881, 641, 768 pairs in E_1-E_4 correspondingly.
(right) The average number of new store–category pairs visited by each exposure group during an L trips before and after the corresponding closures. A purchase occasion at store s in category c is counted as *new store-category* visit for household–category (i, c) if household i purchases category c in store s during her first L trips after her corresponding closure date, while she was not purchasing any items in category c from store s during L trips prior to closure. Error bars show the 95% confidence intervals. The highest exposed group E_4 has relatively more new visits after the closure, which makes it more likely to observe a significant disruption effect in their purchasing behavior caused by the closure.

exploration and formation of new habits. As you can see in Figure 2.2(right), the average number of visits to new store–category pairs during the post-closure period increases with exposure level. A purchase occasion at store s is counted as *new store* visit for household–category (i, c) if household i purchases category c in store s during her first L trips after her corresponding closure date, while she had not purchased any items in category c from store s during L trips prior to closure. We compare these average new visits with the same quantity defined based on the period of L trips prior to each closure, while here a visit is counted as new if the customer had not purchased any item between $2L$ and L trips prior to closure. The comparison shows that relatively the new store visits increases only for E_4. This is another reason to expect effects of closures to be concentrated in the fourth exposure group.

trips prior to closure.

Finally, the timing of closures and whether a household is exposed to a closure at a given time is determined by households' own choices and hence is potentially endogenous. This motivates using differences-in-differences, whereby before–after closure changes for household–category pairs are compared with those changes in the control group of household–category pairs, which consists of households that did not purchase that category at the closing store. We will return to this issue later in Section 5.1.3 and show evidence consistent with parallel pre-treatment trends comparing each of the exposure groups with the control group.

4.2 Outcome Variables

Our goal is to detect changes in households' brand choice patterns that are indicative of habits and closure-induced search. This is a complicated multi-choice problem with varying choice sets. In order to simplify the problem and provide interpretable estimates, we follow the approach in Larcom, Rauch, and Willems (2017), and conduct a relatively "reduced-form" analysis where the outcome is a binary variable indicating whether the household is choosing their modal brand option. This simplification has multiple advantages. First, it allows us to use regression machinery to estimate causal effects and avoid the difficulties of estimating a model with a high-dimensional categorical outcome. Second, it allows us to readily pool information across different product categories, hence giving a more comprehensive view of shopping behavior. We separately define recent and baseline modal brands to capture different aspects of the changes in households' behavior.

4.2.1 Recent Modal Brand

We expect the effect of the habit disruption on purchase patterns to be, in some sense, temporary because customers will soon form new shopping habits in the new store environments they visit. Therefore, to measure the temporary effect of habit discontinuity on purchase decisions, we define the recent modal brand using a moving window, based on households' L-most-recent shopping trips for each category.[7]. More precisely, let $b_{i,t,c}$ be the brand purchase by household i at trip t in product category c. The L-*recent modal brand* for the triple (i, t, c) is defined as $\tilde{b}^r_{i,t,c} = \mod (b_{i,\tau,c})\big|^t_{\tau_c(L)}$, where $\tau_c(L)$ specifies L previous trips in which category c was purchased. Since different categories are purchased with different frequencies, the time duration in which modal brand is computed would be different for each category. Although time duration is not entirely irrelevant, habits are understood as persistent over time (Wood and Neal 2016). Therefore, we decided to define the modal brand based on the number of visits to each store because what matters most is the repetition and frequency of purchase behavior.[8]

As our first outcome variable, we define the recent modal brand indicator $y^r_{i,t,c}$ as a binary variable indicating buying the recent modal brand $\tilde{b}^r_{i,t,c}$, where i specifies the household, t the trip number, and c the corresponding category of the purchased product:

$$y^r_{i,t,c} = \begin{cases} 1, & b_{i,t,c} = \tilde{b}^r_{i,t,c} \\ 0, & b_{i,t,c} \neq \tilde{b}^r_{i,t,c} \end{cases}. \tag{2.2}$$

We hypothesize that when shopping habits are disrupted, households purchase decisions deviate more often from their modal options (so all the effects are expected to be negative) because the old contextual cues that used to trigger the behavior are no longer present. However, we expect the effect to be only temporary because after a while the recent modal brand is defined based on post-closure trips. Some deviations could also be caused by the unavailability of a household's modal brand in stores they visit after the closure; we return to this issue by conducting analyses that condition

[7]Results presented in the main text use $L = 20$, but they are robust to variations in L. More details can be found in Appendix A, Section 2

[8]This choice was primarily based on the frequency-in-context measure of habits (Labrecque and Wood 2015). Since people could have very different rates for visiting stores, considering same time frames could result in very different number of trips. In particular, habits are resilient to the passage of time and could be triggered even with the loss of memory (Bayley *et al.* 2005, Knowlton *et al.* 1996). As a result, we concluded the number of repetitions could matter more than the frequency over time. Although one could imagine the time passed between shopping trips could also play a role, this data set did not provide us with enough variation to study both of these phenomena simultaneously. This is an interesting research question that can be studied using carefully designed experiments.

on brand availability.

4.2.2 Baseline Modal Brand

Another interesting question to explore here is whether this disruption causes a lasting change in households' brand choices, or after doing some exploration they would return to their prior modal brands. To answer this question, we use all trips prior to closures to specify households' baseline modal brand and then measure deviations from that after the closures.[9]. For each household–category pair, the *baseline modal brand* $\tilde{b}^b_{i,c}$ is defined as the most frequently purchased brand during all trips before the household i's corresponding store closure date τ_i.[10] Note that the long-term modal brand is fixed for each household–category and independent of trip number, unlike the recent modal brand which is defined on a rolling basis. Similar to Equation 2.2, we define the long-term modal brand indicator $y^b_{i,t,c}$ as a binary variable indicating whether household i is buying her baseline modal brand in category c during trip t:

$$\tilde{b}^b_{i,c} = \quad \operatorname{mod}\,(b_{i,t,c})\Big|_{t=-\infty}^{\tau_i}\ , \quad y^b_{i,t,c} = \begin{cases} 1, & b_{i,t,c} = \tilde{b}^b_{i,c} \\ 0, & b_{i,t,c} \neq \tilde{b}^b_{i,c} \end{cases} . \tag{2.3}$$

4.3 Variation in Treatment Timing

Our identification strategy is based on a combination of different households being exposed to store closures at different times, as well as within-household variation in how much that household is exposed to a closure for a particular product category. Until recently, two-way fixed effects (TWFE) estimators would be the standard method for estimating treatment effects in such difference-in-differences (DID) set-

[9]Since the panel is not balanced, the number of pre-closure trips could be highly variable for different household–category pairs. The full panel, as well as the histogram of pre-closure trips can be found in Appendix A, Section 1. So in order to make analysis comparable across different units, we check the robustness of results using a fixed length of 40 trips to define the baseline modal brand; note that the average number of pre-closure trips is 42. All of the results are qualitatively the same as you can see in Appendix A, Section 3

[10]Households exposed to the same store closure might stop visiting the store on slightly different dates. So in practice, we set the corresponding closure date to be the last visit by that household.

tings. However, with variation in treatment timing (i.e. differential timing, staggered adoption), the estimated coefficient is more difficult to interpret and generally does not equal the average treatment effect (ATE), or the ATE on the treated; and need not be any weighted ATE either. Recent work has addressed this issue (Callaway and Sant'Anna 2020, Goodman-Bacon 2021, Imai and Kim 2019, Sun and Abraham 2021), including highlighting that, in some cases, this can make resulting estimates quite biased if interpreted as treatment effects (Baker *et al.* 2022).

Here we use TWFE as our primary estimator, but show that this choice is not so consequential. In particular, we use a decomposition (Goodman-Bacon 2021) of the estimates into a weighted average of individual 2×2 DID estimators with the weights proportional to group sizes and variance of treatment duration. The decomposition shows that the TWFE estimator consists of three comparisons and gives the corresponding weights for each: treated vs. untreated, lately-treated vs. early-treated, and early-treated vs. lately-treated.

5 Results

In this section, we present the results for many difference-in-differences and event study models using the outcomes defined in the previous section. We do the following analyses for each of the outcome variables discussed in the previous section:

1. Estimate a TWFE model to find the aggregated treatment effect.

2. Estimate a heterogeneous fixed effects model to explore possible heterogeneity in effects across household–category pairs in different exposure-level groups.

3. Estimate an event study model to examine the testable implication of the parallel trends assumption, and also examine how treatment effects change over time.

4. Estimate a conditional TWFE model to measure to what extent the post-closure unavailability of brands is driving the treatment effects.

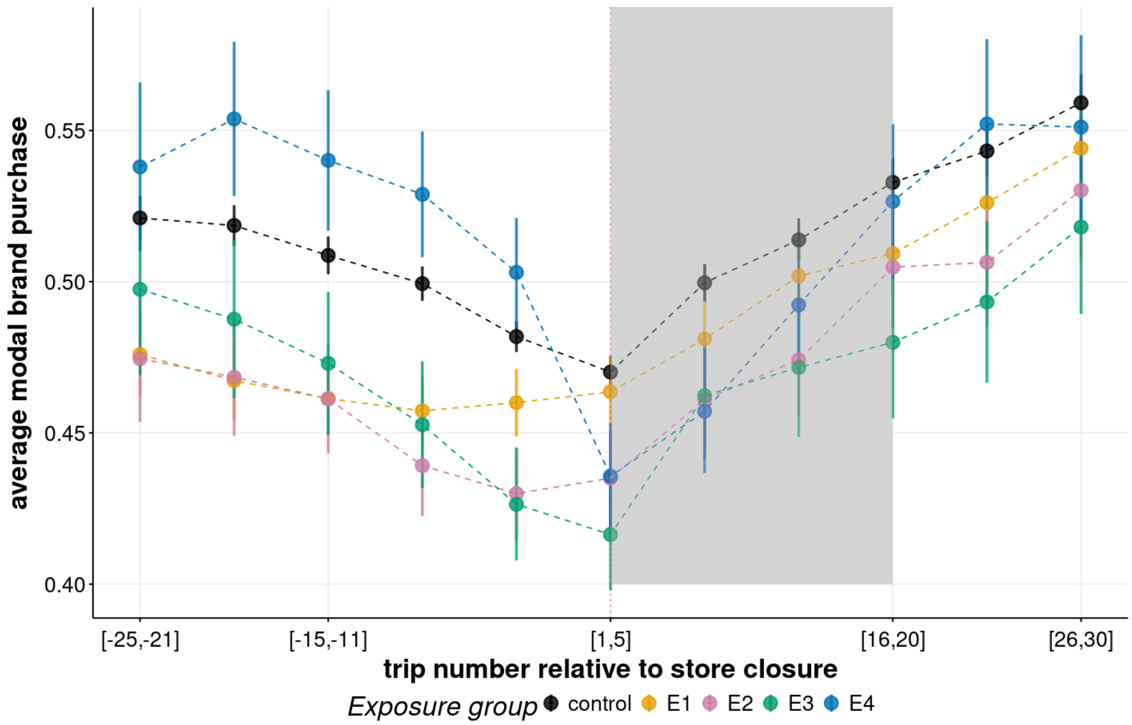


Figure 2.3: Dynamics of recent modal brand purchase rate for every 5 trips relative to closure date. The grey area shows the duration in which we are considering households to be treated. The subplot (a) compares all treated household–category pairs with the control group, and subplot (b) shows the averages separated by exposure level groups.

5. Use the Goodman-Bacon (2021) decomposition to find the extent of the bias in TWFE estimators.

5.1 Temporary Disruption in Shopping Habits

Here we use the recent modal brand indicator to measure the short-term effect of store closures on households' purchase behavior.

5.1.1 Descriptive Analysis

First, we examine the dynamics of modal brand choices by plotting the average recent modal brand purchase rates across all exposure groups. For each household, we consider the trip number relative to the corresponding store closure date and compute the average recent modal brand choice for blocks of 5 trips (Figure 2.3). There is a substantial drop in modal brand purchase rate for E_4 relative to the control group shortly after the closures happen, which shrinks over subsequent trips. This observation is consistent with our hypothesis that purchase behavior of higher exposed household–category pairs is more strongly affected. Moreover, the figure shows approximately parallel pre-trends between different exposure groups and the control group (maybe except for E_1). This analyses does not yet account for household–category or seasonal patterns in the panel data.

5.1.2 Difference-in-Differences

Here we use TWFE to estimate the effects of closures on modal brand purchase rate. Further, as explained previously, we separately estimate the effect for short-term and long-term treatment variables. The short-term treatment variable is active only for the first L trips after the closure for each category so that it can capture the temporary effect shortly after the closure. Moreover, the treatment intensity is equal to the corresponding household–category exposure level. As a result, the treatment vectors for household i in trip t and category c are defined as:

$$T^{r_1}_{i,t,c} = \begin{cases} e_{i,c}, & \tau_i \leq t \leq \tau_i + L \\ 0, & t < \tau_i,\ t > \tau_i + L \end{cases}, \quad T^{r_2}_{i,t,c} = \begin{cases} e_{i,c}, & t > \tau_i + L \\ 0, & t \leq \tau_i + L \end{cases}. \tag{2.4}$$

Using the recent modal brand outcome (Equation 2.2) and these treatment vectors, the fixed effects regression model can be formulated as:

$$y^r_{i,c,t} = \alpha_{i,c} + \gamma_t + X^T_{i,t}\theta + \beta_1 T^{r_1}_{i,c,t} + \beta_2 T^{r_2}_{i,c,t} + \epsilon_{i,c,t}, \tag{2.5}$$

where $\alpha_{i,c}$ are household–category fixed effects, and γ_t are the temporal (monthly) fixed effects. $X_{i,t}$ is a set of household covariates that are varying over time and could potentially affect shopping behavior; these include dummy variables for household income level, size, and composition.[11]

[11]Note that these covariates only include yearly changes and the Nielsen data does not provide more accurate temporal information on panelists.

Table 2.1: Estimation results for TWFE models with the recent modal brand indicator as the outcome variable. Columns 1 & 2 show the corresponding β parameter(s) in Equations 2.5, 2.6. These coefficients measure short-term and long-term rates of recent modal brand purchases, compared with control household–category pairs. Columns 3 & 4 contain the same parameters conditional on trips in which the recent modal brand was available. All standard errors are clustered at the closing store level, and all numbers are multiplied by 100.

	Dependent variable:			
	recent modal brand indicator (×100)			
	(1)	(2)	(3)	(4)
Overall, short-term	-4.616^{**}		-2.732^{***}	
	(1.514)		(0.714)	
Overall, long-term	0.128		-1.943^{*}	
	(2.210)		(0.971)	
E_1, short-term		0.849		-0.111
		(0.632)		(0.507)
E_2, short-term		1.072		-0.495
		(0.878)		(0.704)
E_3, short-term		-2.274		-2.117^{**}
		(1.397)		(0.772)
E_4, short-term		-6.350^{***}		-2.687^{***}
		(1.493)		(0.738)
E_1, long-term		1.634^{*}		-0.201
		(0.769)		(0.504)
E_2, long-term		3.629^{**}		0.153
		(1.361)		(0.960)
E_3, long-term		0.983		-0.879
		(2.176)		(1.059)
E_4, long-term		-3.686^{*}		-2.389^{**}
		(1.809)		(0.854)
Continuous treatment	✓		✓	
Conditioned on modal brand availability			✓	✓
Observations	887,544	887,544	535,313	535,313
R^2	0.279	0.279	0.149	0.149
Adjusted R^2	0.268	0.268	0.129	0.129
Note:		*p<0.05; **p<0.01; ***p<0.001		

Furthermore, we estimate a heterogeneous TWFE model to capture the treatment effect for each exposure level group. Note that, unlike the model in Equation 2.5 in which the treatment was defined proportional to the exposure level, in the following model the treatment is a binary indicator for the household–category pair belonging to each exposure group,

$$y^r_{i,c,t} = \alpha_{i,c} + \gamma_t + X^T_{i,t}\theta + \sum_{j=1}^{4} \beta_{1,j}\, \mathbb{1}_{i\in E_j}\, \mathbb{1}_{T^{r1}_{i,c,t}>0} + \sum_{j=1}^{4} \beta_{2,j}\, \mathbb{1}_{i\in E_j}\, \mathbb{1}_{T^{r2}_{i,c,t}>0} + \epsilon_{i,c,t}. \quad (2.6)$$

Table 2.1 summarizes the effects ($\hat{\beta}$s) estimated in Equations 2.5 and 2.6. All of the standard errors are clustered at the closing store level, thereby allowing for dependence among all household–category pairs associated with the same store closure. The estimated overall short-term effect (column 1) is negative and statistically significant, consistent with our hypothesis about the disruption caused by store closures. Furthermore, the estimates for different exposure groups (column 2) show that the effect magnitude is increasing for higher exposed household–category pairs and primarily driven by E_4, consistent with our expectation that household–category pairs in which the household frequently purchased from the closing store are most affected. There is no statistically significant long-term effect detected. Notice that there is an opposite significant short-term effect for E_1 and E_2. This is probably because the closing store did not play a major role in their modal brand purchases, but may have been causing deviations from their modal brands. So the closure makes them visit other stores that they already visited more often, which results in an increase in their modal brand purchase rates.

5.1.3 Event Study Analysis

A common robustness check for the TWFE model is an *event study analysis* (Granger 1969, Roth 2019, Sun and Abraham 2021). To this end, we estimate a model similar to Equation 2.5, but with lags and leads of the treatment variable:

$$y^r_{i,c,t} = \alpha_{i,c} + \gamma_t + X^T_{i,t}\theta + \sum_{\tau=1}^{4} \beta^{lead}_{\tau} T^r_{i,c,t-\tau} + \sum_{\tau=0}^{5} \beta^{lag}_{\tau} T^r_{i,c,t+\tau} + \epsilon_{i,c,t}. \tag{2.7}$$

If the lead estimates were statistically significant, it would be a violation of the parallel trends assumption because it would imply the cause is preceding the effect. Moreover, the estimated lagged effects are informative about the dynamics of post-treatment treatment effects. Similar to the analysis in Section 5.1.1, we estimate the leads and lags for every 5 trips grouped together. We added five leads and six lags, where the first and last include all trips whose relative trip number is less than -20 and more than 25. The fifth lead variable is the omitted baseline category in estimating the model.

The estimates for the event study model (Equation 2.7) with corresponding confidence intervals can be seen in Figure 2.4 (top). The lead parameters are not rejected at a 95% confidence level, which indicates the control and treatment units are indistinguishable prior to the treatment and is consistent with the parallel trends assumption. Furthermore, the point estimates are significantly negative for the next 15 trips after the closure with a diminishing magnitude. This is exactly what we expected since the recent modal brands are defined on a rolling basis (e.g., the point at $x = 4$, is entirely based on post-closure trips). Also, the last lagged variable which measures the long-term treatment effect is almost zero. Both of these observations support our hypothesis that store closures do not have a lasting effect on the degree to which households' purchases eventually concentrate into a model brand. The rate at which the effect goes to zero also gives us a sense of the required number of visits to new stores for that to happen.[12]

We do a similar event study analysis for heterogeneous effects across exposure groups to validate the parallel trends assumption for Equation 2.6 and also observe

[12]This is more like an upper limit for the required number of trips to form new habits because not of the trips in our data set are at new stores.

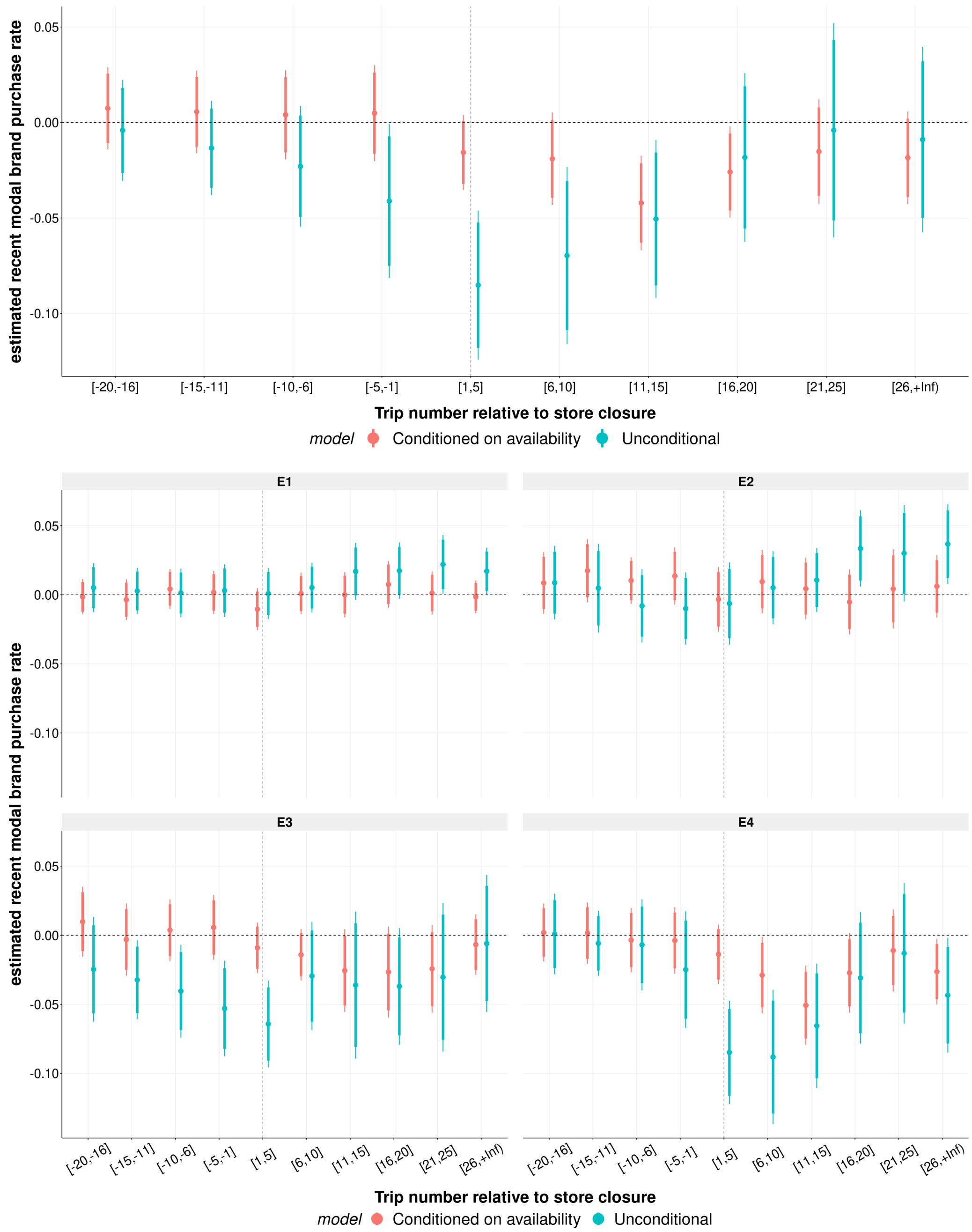


Figure 2.4: The plot shows the estimated lead and lag coefficients in Equations 2.7 & 2.8. Point estimates are computed for every 5 trips grouped together, the thin and thick error bars show the corresponding 95% and 90% confidence intervals. The blue color indicates the unconditional model, and red shows estimated coefficients conditional on recent modal brand availability. 5 lead and 6 lag variables are used, where the last lead (lag) includes all trips whose relative trip number is less than (more than) -20 (25). The fifth lead variable is used as the baseline in estimating the model and hence not shown in the figure.

the dynamics of the post-treatment effects,

$$y^r_{i,c,t} = \alpha_{i,c}+\gamma_t+X^T_{i,t}\theta+\sum_{j=1}^{4}\sum_{\tau=1}^{4}\beta^{lead}_{j,\tau}\,\mathbb{1}_{i\in E_j}\,\mathbb{1}_{T^r_{i,c,t-\tau}>0}+\sum_{j=1}^{4}\sum_{\tau=0}^{5}\beta^{lag}_{j,\tau}\,\mathbb{1}_{i\in E_j}\,\mathbb{1}_{T^r_{i,c,t+\tau}>0}+\epsilon_{i,c,t}. \tag{2.8}$$

The results are shown in Figure 2.4 (bottom). Except for leads coefficients of E_3, the rest of the leads are statistically non-significant, consistent with the parallel trends assumption for the heterogeneous fixed-effects model. The post-treatment trends also have a diminishing magnitude similar to the previous model, while the highest exposed group E_4 has the most significant and lasting effects.

5.1.4 Availability of Modal Brands

One potential source of the change in brand choices could be the lack of availability of prior modal brands in newly explored stores after the closure. In this section, we show that only part of the observed effect can be explained by unavailability of a modal brand on a given trip. Nielsen data does not directly provide information on all of the available brands in each store over time, so we need to infer that from retail scanner and consumer panel data. The list of stores in the retail scanner data does not have a full overlap with stores in the consumer panel data. We therefore used purchases from other households in the consumer panel data to identify available brands. For each week, we mark a modal brand as available in a store, if there is at least one purchase occasion by any household in the entire consumer panel data. In order to compare the availability among different exposure groups relative to control, we estimate event study models similar to Equations 2.7 and 2.8 using the availability indicator as the outcome variable.

As it might be expected, the overall percentage of available modal brands drops after the closure, both for the treated and control units. Figure A.7 (Appendix A, Section 2) shows the percentage of trips with available modal brands across exposure level groups and time periods. Nevertheless, these are averages over extended periods of time, and do not account for seasonal variations. To follow how modal brand avail-

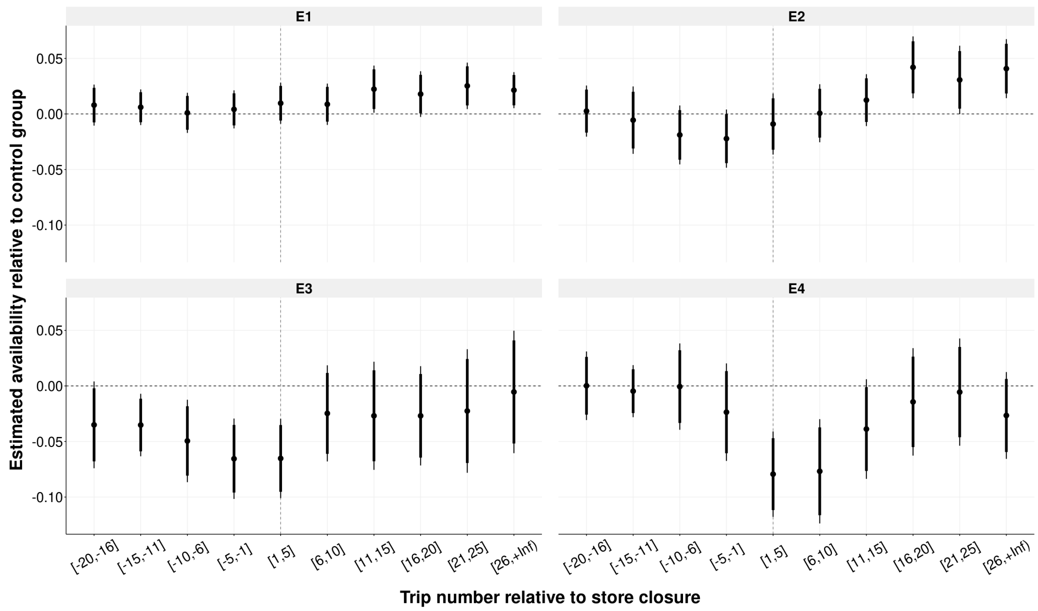


Figure 2.5: The plot shows the estimated lead and lag coefficients in Equation 2.7 with availability indicator as the outcome variable. Point estimates are computed for every 5 trips grouped together, the thin and thick error bars show the corresponding 95% and 90% confidence intervals. 5 lead and 6 lag variables are used, where the last lead (lag) includes all trips whose relative trip number is less than (more than) -20 (25). The fifth lead variable is used as the baseline in estimating the model and hence not shown in the figure.

ability varies in each exposure group over time, relative to the control, we estimate a similar event study model considering availability indicator as the outcome variable. The resulting coefficients can be seen in Figure 2.5. There are significant negative pre-trends for E_3, which can explain the negative lead coefficients in Figure 2.4. More importantly, the availability rates significantly drops (up to 8 percentage points) for E_4 right after the closure, which is expected since these household–category pairs are more likely to be purchasing in a new store (Figure 2.2, right sub-figure).

The decreased availability for E_4 could account for some or all of the significant short-term effect, so in order to be able to attribute the observed effect to disrupted habits we need to adequately account for that. To this end, we estimate similar DID models (Equations 2.5 and 2.6) conditional on the subset of trips in which the recent modal brand was available. The idea is that in trips where the households' modal brand in a certain category is available, increased average deviation from that brand would reveal the impact of the habits formed around the store environment.[13]

Results for the conditional model can be seen in Table 2.1, columns 3 and 4. We also estimate the same event study models (Equations 2.7 and 2.8) conditional on modal brand availability. As shown in Figure 2.4 (red points), the scarcity of modal brands can explain only part of the observed effect, and a substantial effect remains. For example, the short-term effect for E_4 is about a third of the unconditional effect. This remaining part could be attributed to what we called *shopping habits*, which happens due to the absence of previous contextual cues. Notice that in the long-term after the closure period, there is a significant increase in availability for E_1 and E_2. This can explain the positive long-term effect observed in the unconditional model for E_1 and E_2 in Table 2.1. This effect disappears as we account for brand availability, so there is no behavioral factor causing the effect. For lower exposed household–category pairs, by definition, the closing store plays a smaller role in their shopping. Hence the increased availability is plausibly because of the fact that they are increasingly often visiting the set of other store–category pairs they used to shop from — places where they bought their modal brands before the closure. Moreover, there is a marginal negative long-term effect for E_4, which could suggest there might be a lasting increase in variety-seeing. However, this effect is not robust to different choices of hyper-parameters (L and T_e), so we would not draw any conclusions based on that (see Appendix Afor more details).

5.1.5 Effects Without Temporary Unavailability

The previous analysis does not entirely rule out the impact of modal brand unavailability. Temporary unavailability of a certain brand could force the household to explore new ones which can cause increased information about alternative options leading to changed brand preferences. Such a process is consistent with our broad account of how disruptions to choice environments can have lasting consequences

[13]Note that, to keep estimates comparable, we define the short-term/long-term periods the same way using all trips. So the short-term coefficients would contain less than or equal to L trips. We could have alternatively used the first L trips in which the modal brand was available to estimate the short-term effect (which would indeed result in more substantial effect size), but then it would cover a longer period of time and one might worry about other factors affecting households' behavior.

by changing habits; however, we wish to also characterize whether some of this is attributable to changes in the store environment along, even if the modal brand remains continually available.

To this end, we now analyze subsets of household–category pairs in which we restrict any unavailability in the sequence of households' trips after the closure. So, for each household–category pair, we estimate the short-term effect for the first L_a trips in which the modal brand was always available, where $0 < L_a \leq L$. We include the remaining trips after L_a among the long-term effect. The long-term effect would not have a similar clear interpretation since it includes trips with unavailable brands, however, we still include them in the estimation because they help with estimation of fixed effects and hence improve the precision of the desired short-term parameter.

If we take the subset of the data for which $L_a > 0$, we are left with 6,084 household–category pairs and about 10,000 short-term trips; the full distribution of L_a can be seen in Appendix A, Figure A.8. Estimating Equation 2.6 gives the following short-term effect for the forth exposure group: $\hat{\beta}_{1,4} = -0.870$, $SE = 0.748$, $p = 0.245$. The estimate is not statistically significant, mostly because the condition requiring the full sequence of trips to have available modal brands is very restrictive and leaves us with very few observations.[14] For example, there could be many units for which there is only one trip where the modal brand is unavailable, and the information effect of exploration is minimal, but the previous condition would drop the sequence altogether. However, if we slightly loosen the conditions by allowing the sequence of short-term post-closure trips to have at most one trip with unavailable modal brand there are 7,825 household–category pairs with about 15,000 short-term trips remaining, and the estimated coefficient would be: $\hat{\beta}_{1,4} = -1.743$, $SE = 0.784$, $p = 0.027$. Here, we have included post-closure trips up to the point where the second trips with unavailable modal brand appears. This would minimize the effect of unavailable brands on households' information about alternative brands, while leaving enough data to be able to precisely estimate the parameter of interest. These

[14] Indeed, the same estimated parameter is statistically significant if we choose $T_e = 2$ years: $\hat{\beta}_{1,4} = -1.804$, $SE = 0.833$, $P = 0.031$, and $T_e = \frac{1}{2}$ year: $\hat{\beta}_{1,4} = -1.875$, $SE = 0.826$, $p = 0.024$.

results are broadly consistent with both an unavailability-driven mechanism and other effects of changing contexts for choices, even as modal brands remain largely available.

5.2 The Effect of Store Closures on Baseline Modal Brands

In this section, we utilize the baseline modal brand indicator as the outcome variable (Equation 2.3), and use TWFE to estimate effects of closures on consumers' rate of choosing their baseline modal brand. In particular, the difference of the current analysis is that it tells us whether people return to their prior modal brand options, or the disruption will lead to lasting changes in brand choices. As we discussed in the previous section, the disruption can be caused by various mechanisms including unavailability of brands and discontinuity in shopping habits. To capture the treatment effect on baseline choices, we define all post-closure trips to be treated because we want to see the overall effect and there is no reason to expect the effect to be temporary. The follow-up event study analysis will further justify this assumption. Note that this is different from how we defined the treatment vector for the recent modal brand (Equation 2.4). The baseline treatment variable is defined as:

$$T^b_{i,t,c} = \begin{cases} e_{i,c}, & \tau_i \geq t \\ 0, & t < \tau_i \end{cases}. \tag{2.9}$$

5.2.1 Descriptive Analysis

Again, we first plot the average modal brand purchase rates for different treated and control groups to compare the trends around store closures. Figure 2.6 shows the comparison for different exposure groups vs. control. Similar to the previous outcome, a differential change for E_4 is detectable even from comparing raw means without controlling for fixed effects. The average rate of purchasing modal brand for E_4 is always more than the control group prior to the treatment, while it drops below the control curve right after the closure time. Furthermore, pre-treatment trends are parallel for exposure groups and the control group, although we will later illustrate

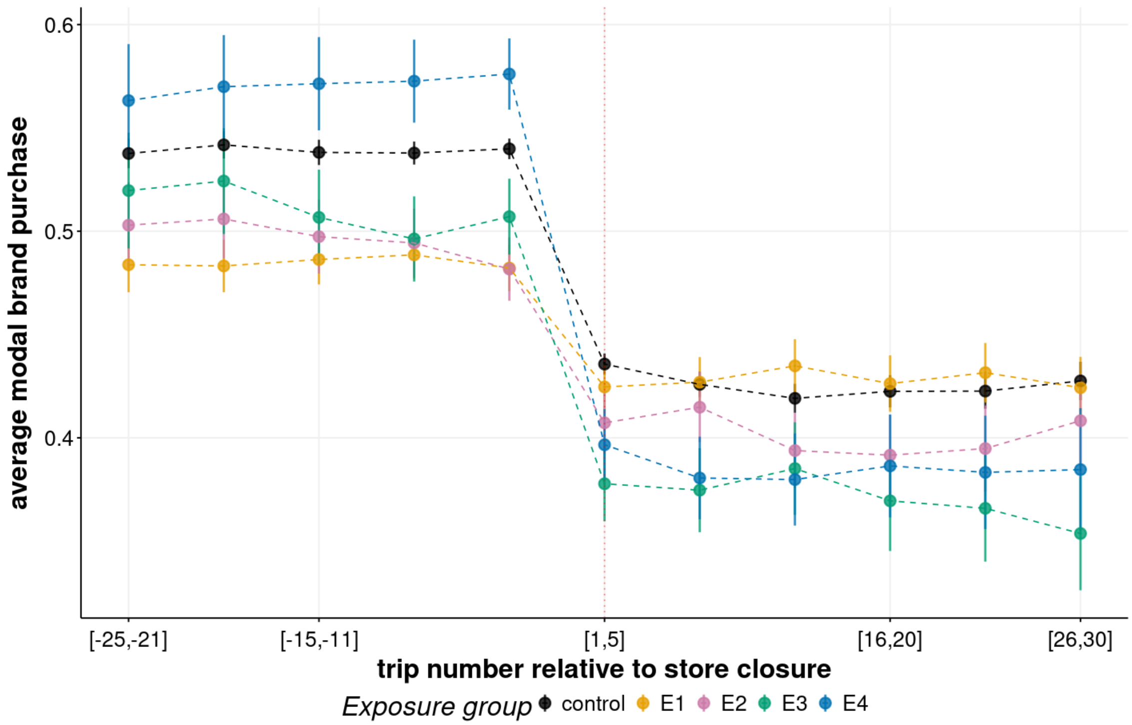


Figure 2.6: Dynamics of average long-term modal brand purchase rate for every 5 trips relative to closure date, across four exposure groups and the control group. The figure supports the parallel trends assumption between different groups and the control. It also suggests a constant lasting effect after the closure, unlike the temporary effect on recent modal brand choices.

this point more rigorously using event study analysis.

5.2.2 Difference-in-Differences

Similar to the recent modal brand analyses, we estimate two fixed-effects models to estimate the average and heterogeneous treatment effects. The outcome variable and treatment are defined differently as shown in Equations 2.3 and 2.9. The overall and heterogeneous TWFE models are:

$$y^b_{i,c,t} = \alpha_{i,c} + \gamma_t + X^T_{i,t}\theta + \beta\, T^b_{i,c,t} + \epsilon_{i,c,t} \tag{2.10}$$

$$y^b_{i,c,t} = \alpha_{i,c} + \gamma_t + X^T_{i,t}\theta + \sum_{j=1}^{4} \beta_j\, \mathbb{1}_{i \in E_j}\, \mathbb{1}_{T^b_{i,c,t} > 0} + \epsilon_{i,c,t} \tag{2.11}$$

Table 2.2: Estimation results for TWFE models with the baseline modal brand indicator as the outcome variable. Column 1 shows the estimated β parameter in Equation 2.10, and column 2 shows the same parameter where instead of a continuous treatment, a binary treatment indicator has been used. Column 3 displays corresponding β parameters in Equation 2.11. These coefficients measure the extent of deviation from baseline modal brands during the entire post-closures period, compared with entire pre-treatment period, relative to control household–category pairs. The fact that all coefficients are negative shows that disruption caused by the closure leads households to new brand options that are on average different from what they used to buy, and the effect becomes stronger for units with higher exposure. All standard errors are clustered at the closing store level, and all estimates are multiplied by 100.

	Dependent variable:		
	baseline modal brand indicator ($\times$100)		
	(1)	(2)	(3)
Overall	-14.415^{***}	-5.870^{***}	
	(2.021)	(0.894)	
E_1			-4.269^{***}
			(1.063)
E_2			-4.630^{***}
			(1.510)
E_3			-9.110^{***}
			(2.149)
E_4			-12.432^{***}
			(1.938)
Continuous treatment	✓		
Observations	895,035	895,035	895,035
R^2	0.343	0.342	0.342
Adjusted R^2	0.333	0.332	0.333
Note:		*p<0.05; **p<0.01; ***p<0.001	

where $\alpha_{i,c}$ are the household–category, and γ_t the temporal (monthly) fixed effects. $X_{i,t}$ is a set of household covariates that are varying over time and could potentially affect shopping behavior; these include dummy variables for household income level, size, and composition.

The estimation results for these equations are shown in Table 2.2. As it can be seen, the overall effect is negative and statistically significant. This result shows that the disruption due to store closures will cause a permanent change in modal brands for exposed household–category pairs, compared with the control units. So, although the disruption in habits is, in some sense, temporary in that people form new habits in their new environments, they converge to a set of brands that is on average different from what they used to buy. The second column shows the same parameter where instead of a continuous treatment, a binary treatment indicator has been used.[15] Finally, the third column contains heterogeneous treatment effects across exposure groups. The effect sizes are stronger for higher exposed groups, inline with our theoretical predictions and results from previous analysis. Note that although the identified temporary effect on recent modal brand was not statistically significant for lower exposed groups, the lasting effect on baseline brands is still statistically significant and substantial even for these lower exposed units.

5.2.3 Event Study Analysis

Next, we estimate two event study models for the baseline outcome both to test implications of the parallel trends assumptions used in difference-in-differences models and explore the dynamics of the treatment effect after closures. The estimated models are defined similarly to Section 5.1.3, but with two changes. First, the outcome and treatment variables are the baseline version defined in Equations 2.3 and 2.9. Second, we used the first lead coefficient as the reference point to estimate the model because we would like to know the difference in treated and control units' behavior compared with how it was right before the closure happens. The event study models for the overall and heterogeneous effects are as follows:

$$y^b_{i,c,t} = \alpha_{i,c} + \gamma_t + X^T_{i,t}\theta + \sum_{\tau=2}^{5} \beta^{lead}_{\tau}\, T^b_{i,c,t-\tau} + \sum_{\tau=0}^{5} \beta^{lag}_{\tau}\, T^b_{i,c,t+\tau} + \epsilon_{i,c,t}. \tag{2.12}$$

$$y^b_{i,c,t} = \alpha_{i,c}+\gamma_t+X^T_{i,t}\theta+\sum_{j=1}^{4}\sum_{\tau=2}^{5} \beta^{lead}_{j,\tau}\, \mathbb{1}_{i\in E_j}\, \mathbb{1}_{T^b_{i,c,t-\tau}>0}+\sum_{j=1}^{4}\sum_{\tau=0}^{5} \beta^{lag}_{j,\tau}\, \mathbb{1}_{i\in E_j}\, \mathbb{1}_{T^b_{i,c,t+\tau}>0}+\epsilon_{i,c,t}. \tag{2.13}$$

[15] We also included this result because the follow-up Bacon decomposition analysis (Goodman-Bacon 2021) does not work for continuous or heterogeneous treatments.

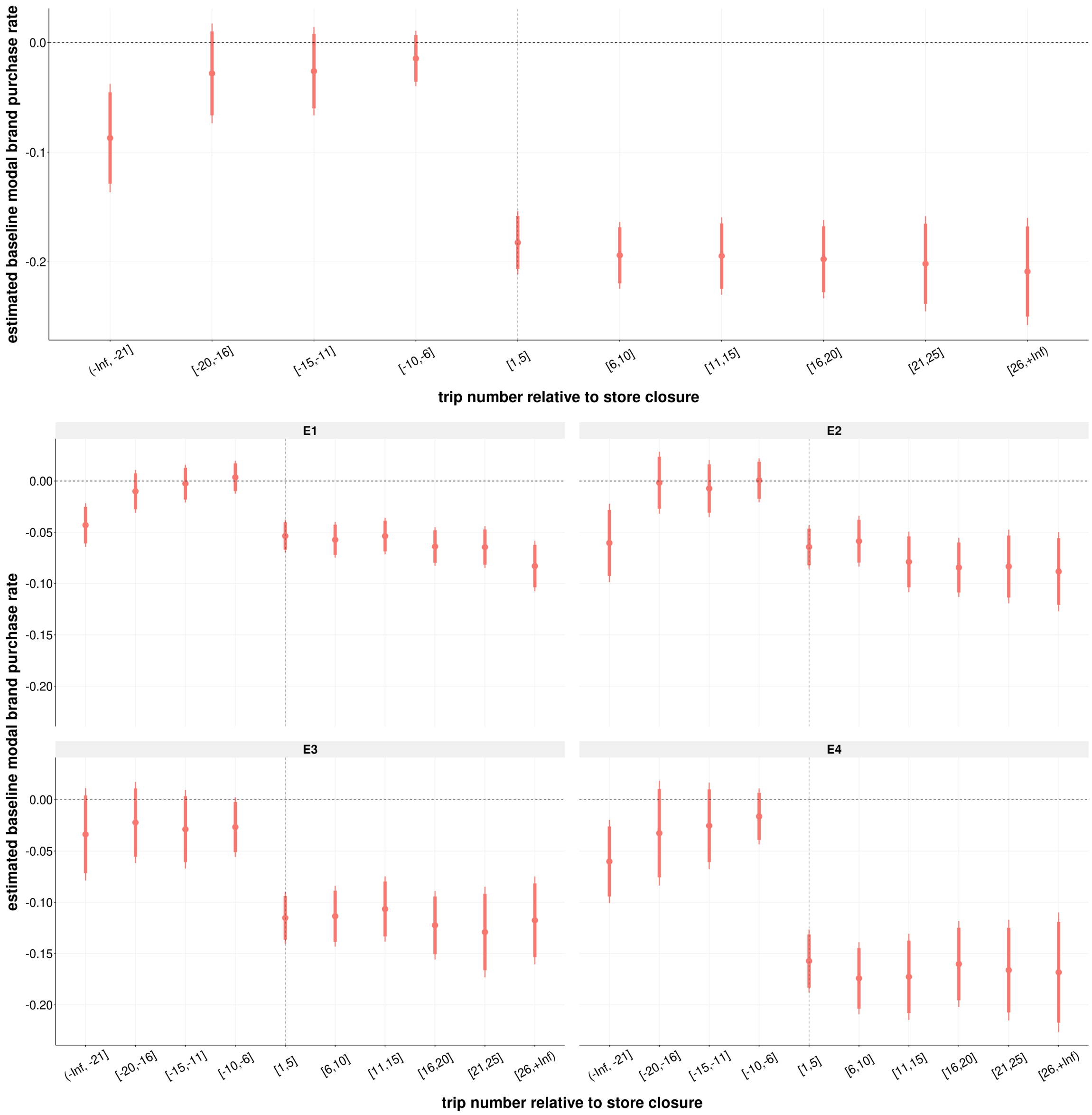


Figure 2.7: The plot shows the estimated lead and lag coefficients in Equations 2.12 & 2.13. Point estimates are computed for every 5 trips grouped together, the thin and thick error bars show the corresponding 95% and 90% confidence intervals. Ten lead and six lag variables are used, where the last lead (lag) includes all trips whose relative trip number is less than (more than) -21 (26). The first lead coefficient ($t = [-5, -1]$) is used as the reference level in estimating the parameters and hence not shown in the figure.

The results for these models can be seen in Figure 2.7. In the aggregate model, the first six lead parameters (up to trip 35 trips before closure) are not statistically significant, which means the treatment and control groups are behaving in parallel for a long time before closure in terms of purchasing baseline modal brands. However, coefficients for the last lead variable is non-zero showing that treated and control groups behave differently as we go further from the closure event. This is not unexpected since there is heterogeneity in the length of pre-closure panel across household-categories. The same observations is true for different exposure groups with some of them having parallel trends for a slightly longer period. Furthermore, the lag coefficients display a non-diminishing effect on baseline brands after the closures, unlike the recent modal brand model. The significance of the last lag coefficient ($t > 4$) shows there is a lasting effect on deviation from baseline modal brands, despite the disruption on shopping habits being temporary.

The fact that the event study model reveals non-zero and negative pre-trends long before the closures happens, could bias the results from the DID models towards zero, since those models compare the post-closure period with all pre-closure. That's is why the effect sizes in Table 2.2 are generally smaller than the coefficients in Figure 2.7.

5.3 Robustness: Bacon Decomposition

In this section, we use the decomposition proposed by Goodman-Bacon (2021) to access the extent of bias caused by variation in treatment timing (i.e. store closures). One can decompose the TWFE estimate into a weighted sum of individual 2×2 DID estimates (Goodman-Bacon 2021). Consider the group k of household–category pairs that face a store closure, and let u be the set of untreated pairs. An individual 2×2 DID estimate is the difference in means for between pre and post-treatment periods:

$$\hat{\beta}^{k}_{ku} = (\bar{y}_k^{post(k)} - \bar{y}_k^{pre(k)}) - (\bar{y}_u^{post(k)} - \bar{y}_u^{pre(k)}), \tag{2.14}$$

where the subscript shows comparison groups, and the superscript stands for the time of the treatment. For example, for two treated groups of early treated k and later

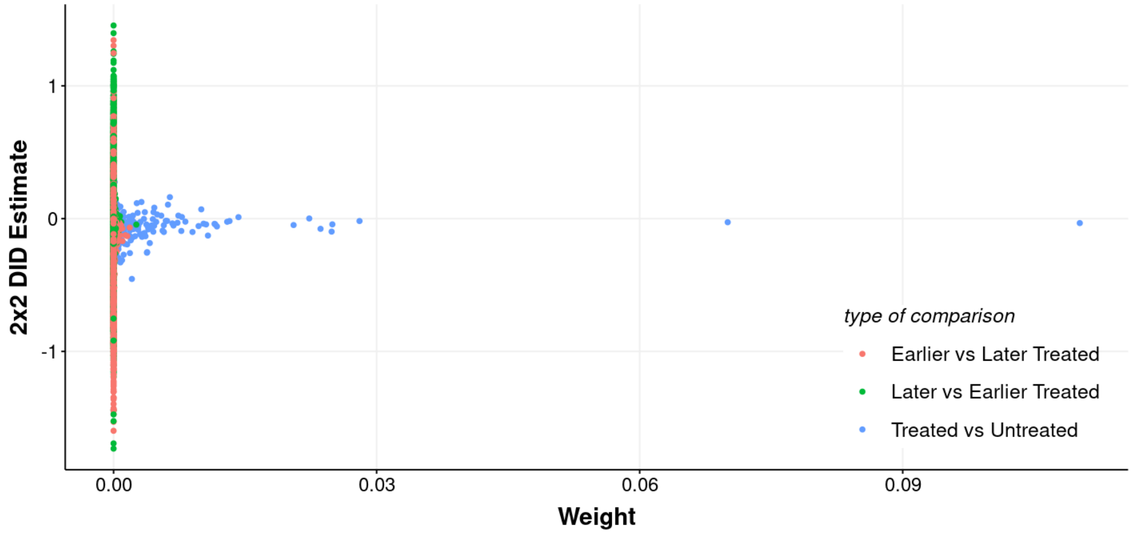


Figure 2.8: Individual 2×2 DID estimates and corresponding weights as defined in Equation 2.16. As you can see, most of the weights associated with earlier vs. later, and later vs. earlier treated units are close to zero.

treated l, $\hat{\beta}^k_{kl}$ is the 2×2 DID estimate between k and l when k get treated while l is still not affected, and $\hat{\beta}^k_{kl}$ captures the case where l get treated while k has already been treated before.

All household–category pairs facing a single store closure are gathered together in a group. If we have C store closures, there are C groups and one never treated group which consists of all household–category pairs that never face any closure. Therefore, there would be C 2×2 DID between treated and untreated units, and $C(C-1)$ 2×2 DID estimates between early treated and lately treated, and vice-versa. Following the notation in Goodman-Bacon (2021), the TWFE estimator can be decomposed as follows:

$$\hat{\beta}_{\text{TWFE}} = \sum_{k \neq u} s_{ku}\hat{\beta}^k_{ku} + \sum_{k \neq u}\sum_{l>k} s_{kl}\left[\mu_{kl}\hat{\beta}^k_{kl} + (1-\mu_{kl})\hat{\beta}^l_{kl}\right], \tag{2.15}$$

with the weights are shown in the following equations:

$$
\begin{aligned}
s_{ku} &= \frac{n_k n_u \bar{D}_k(1-\bar{D}_k)}{\widehat{\mathrm{Var}}(\tilde{D}_{i,t})} \\
s_{kl} &= \frac{n_k n_l(\bar{D}_k-\bar{D}_l)(1-(\bar{D}_k-\bar{D}_l))}{\widehat{\mathrm{Var}}(\tilde{D}_{i,t})} \\
\mu_{kl} &= \frac{1-\bar{D}_k}{1-(\bar{D}_k-\bar{D}_l)},
\end{aligned}
\tag{2.16}
$$

where n_k is the number of units in group k (i.e. the number of household–category pairs experiencing store closure k), n_u is the number of units in the control group, $\bar{D}_k$ is the fraction of the panel length group k is treated, and $\tilde{D}_{i,t}$ is the residual of the treatment vector after partialling out individual and time fixed effects.

Figure 2.8 show the $C^2 = 121,104$ separate 2×2 DID estimates and their weights. As it can be seen in the figure, most of the weight is contrasting earlier vs. later treated, while weight for the later vs. earlier treated contrast is close to zero. So we expect the bias to be small. Further, Table 2.3 shows the aggregated weights and corresponding average estimates for each comparison. This decomposition shows that there is a small, $\frac{(-5.78)-(-4.87)}{(-4.87)} \times 100 = 5.8\%$ bias in the TWFE estimator..

There are a number of implicit assumptions used in the decomposition results presented in the previous section. Note that the TWFE estimate here is different from the aggregate model in Table 2.2. The reason is that so far we have been using a continuous treatment vector where the treatment intensity was proportional to the exposure level as defined in Equation 2.1. However, Bacon decomposition only works for binary treatment variables, so in this section, we defined the treatment as an indicator that equals one in the post-closure period if the corresponding exposure level is not zero. The estimated parameter for a binary treatment is reported in Table 2.2 column 2.

Also, Bacon decomposition assumes a non-decreasing treatment assignment, which matches our baseline model in (Equation 2.10) but not the recent modal brand model (Equation 2.5). However, according to Equation 2.16, the weights for early and late

Table 2.3: The aggregated weights and associated average estimates for each comparison in the Bacon decomposition (Equation 2.15). The majority of the weight is attributed to the treated vs. untreated comparison. The middle comparison (Later vs. Earlier) is the one that can cause bias since it is comparing later treated units with already treated ones. The weighted average of the other two comparisons is -4.87 which is very close to the TWFE estimate -5.78 (Table 2.2 column 2) indicating small (5.8%) bias due to unwanted comparisons.

Type of comparison	Weight	Avg. estimate $\times$ 100
Earlier vs. Later Treated	0.131	-11.91
Later vs. Earlier Treated	0.109	-4.65
Treated vs. Untreated	0.759	-4.35

cross-comparisons s_{kl} depend on the differential amount of time units get treated $\bar{D}_k - \bar{D}_l$. While, in the recent modal brand model by definition (Equation 2.4), units are treated only temporarily, so the difference in duration of treatment timings are much less, although they are not exactly zero if the next closure happens before the L trips duration for the consumers of the previous closing store. As a result, we expect the bias inferred from Table 2.3 be an upper bound for any bias for the recent modal brand model.

6 Discussion

In summary, we showed that households on average choose different brands in product categories after facing a the closure of a store where they used to purchase in that category, thereby exposing them to new shopping contexts. This pattern is robust to accounting for the lack of availability of their prior modal brands. Furthermore, we show that the temporary disruption in shopping habits results in lasting changes in households' modal brands suggesting that formation of shopping habits could lead to sub-optimal behavior. These findings provide positive empirical evidence for the effect of shopping habits on consumers' in-store decision making, and that the measured inertia in the literature is not only due to brand loyalty, but also attributable to learned habits in the stable shopping contexts. We attribute the observed behavior to habits for two main reasons. First, the heterogeneous effect models show that the effect sizes increase with the frequency of prior purchases in that category at the closing store. This is in line with the prior literature on habits where it has been shown that frequent actions in a stable context lead to habit formation (Wood *et al.* 2005). Second, the observed effect is robust to accounting for the unavailability of brand options. Since store closures are not dependant on an individual consumer's tastes, there is no reason to believe consumers suddenly would have changed their preferred brands in the absence of the closure — and only in the categories purchased at that closing store. So different choices are a result of exposure to a new context where old habits are no longer present.

These results have consequences both for firms and customers. These kinds of behavioral factors in consumers' decision-making can have managerial implications for optimal pricing, advertising strategies, and allocation and location of goods inside stores. Strong consumption habits are at the core of marketing campaigns designed to attract consumers from other brands, e.g., by temporary price reductions or free sampling, in order to benefit from their choice inertia in the long-run. The average state dependence can affect firms' decisions about promotions or temporary price discounts. If people are very state-dependent, it would be an extra incentive for firms to attract customers sooner than later. So they have incentives to lower prices. This phenomenon has been studied in form of a dynamic game among firms to lower prices to exploit consumers' state dependence. The equilibrium of this game would depend on consumers' state dependence, and it can be shown that in some cases consumers can benefit from this competition (Klemperer 1987, Seetharaman and Che 2009). Therefore, correctly estimating the degree of consumer brand loyalty can be crucial for firms' decisions and biased estimates could adversely affect their total profit.

Having a better understanding of the psychological determinants of consumers shopping behavior can benefit both individuals and firms. On the consumer side, it can help us design more effective interventions to improve people's health, e.g., by nudging them to choose healthier options (Leonard 2008). Most habits are typically formed to help achieve particular goals (Aarts and Dijksterhuis 2000, Wood and Neal 2007). But these habits could continue to persist even after changes in outcome structure or reward devaluation (Adams and Dickinson 1981, Neal *et al.* 2011, Wood 2017). As a result, although in the short run habits could be advantageous by automating repetitive tasks and freeing up mental resources, in the long run they could lead to non-optimal behavior. For example, Larcom *et al.* (2017) show that a significant fraction of commuters on the London Underground used to take non-optimal routes, and an exogenous disruption such as a strike brought lasting behavior change in commuters' routing behavior. The same phenomenon could also happen for in-store purchase behavior where a shock causing extended periods of brand availability could bring lasting changes in brand choices (Figueroa *et al.* 2019). This inertia of habits could have adverse healthcare consequences when habits of buying unhealthy products are formed. As a result, it is important to understand the extent to which habit formation influences consumers' in-store shopping behavior. Being aware of the role of context can help those individuals seeking a change in their consumption habits. When they change the typical place they visit for shopping, they are less affected by exiting contextual cues and have an opportunity to start buying healthier products. Also, this could help policymakers design and implement more effective policies.

Finally, our findings shows a substantial role for shopping habits in people's purchasing decisions. This poses an immediate question for retailers regarding the effect of commonly practiced in-store re-arrangement of items on the store's aggregate sales and profits. One could imagine competing mechanisms in play here which could make these actions beneficial or damaging for the store. On the one hand, re-arrangement of items would nudge people to explore the store further and find things which could have been ignored previously due to existing shopping habits. On the other hand, these explorations are not free and require extra time and cognitive effort. These search costs could ultimately make people give up on buying an item. Perhaps future work will address this question through analysis of field experiments with changes to stores.

Chapter 3

Habits in Social Media Use: Entropy as a Habit Measure

1 Preface

This chapter investigates the relationship between users behavioral regularity and habits in the context of social media platforms. We introduce the concept of entropy as a measure of behavioral regularity and its potential as a proxy for habits in social media usage. Using statistical modeling techniques, we examine the predictive power of entropy and user engagement, controlling for factors such as average time spent and frequency of app use. Our findings contribute to understanding the dynamics of online social media behavior and offer insights into predicting user behavior in the digital era.

2 Introduction

Habits and routines play a crucial role in shaping our daily behaviors across various domains. They wield significant influence over our behaviors, impacting our health, well-being, and overall quality of life. In this chapter, we delve into the study of habits, defining them as the process by which well-learned associations between cues and behaviors prompt individuals to act (Rebar *et al.* 2018). Despite the importance of understanding habits, debates and controversies surround various aspects of their investigation. Contentious topics include the awareness of habits, distinguishing habits from mere behavior frequency, and disentangling the influence of habit from other motivational factors (Verplanken *et al.* 2018). Central to these debates is the issue of measurement and the validity of habit measures. Despite efforts, a definitive measure for "true habit" remains elusive in the habit literature (Rebar *et al.* 2018). The interplay between theory and measurement is essential in this pursuit, as theories inform the development of measurement schemes with strong construct validity, while empirical findings contribute to the refinement and evolution of theoretical frameworks. Overcoming the challenge of measurement in habit research is paramount to advancing our understanding of this phenomenon.

A 'good' psychological measure should have sound construct validity (Haynes *et al.* 1995, Messick 1990). Construct validity is a fundamental concept in psychological measurement, referring to the extent to which a measure accurately assesses the theoretical construct it intends to capture. It involves evaluating the degree to which a measure aligns with the underlying theoretical framework, providing evidence for the meaningfulness and coherence of the construct. Construct validity encompasses multiple facets, including predictive validity, convergent and discriminant validity, and reliability (Messick 1990). Predictive validity assesses the measure's ability to predict future outcomes or behaviors. Convergent validity examines the extent to which the measure correlates with other measures of similar constructs, while discriminant validity assesses its distinctiveness from unrelated constructs. Reliability is the degree to which the stability of observed scores conforms to theory (Kline 2013). A reliable measure should be responsive to gradual change while showing minimal assessment-to-assessment fluctuation in the absence of true change. Establishing construct validity is a rigorous process that involves theoretical justification, empirical evidence, and ongoing refinement of measures to ensure their accuracy and relevance. It enables researchers to confidently interpret and draw meaningful conclusions from their data, contributing to the advancement of scientific knowledge in psychology.

Various measurement approaches have been proposed to capture the essential constructs inherent in the definitions of habitual behavior. Although, none of these measures is perfect and each of them fails to satisfy certain aspects of construct validity individually. In early habit research, past behavioral frequency was commonly used as a proxy for habit, primarily due to its robust predictive validity (Aarts *et al.* 1998, Bagozzi 1981, Ronis *et al.* 1989). However, relying solely on past behavior lacks explanatory power as it fails to differentiate between habit and non-habitual influences that may regulate behavior, thereby lacking discriminant validity (Gardner *et al.* 2012). Meta-analysis conducted by Ouellette and Wood (1998) indicated that the association between past and future behavior was particularly strong for frequently and consistently executed behaviors within a stable context. Building on this understanding, Wood *et al.* (2005) introduced the concept of the frequency-in-context measure, which represents the product of behavior frequency and context stability. Although this measure demonstrates predictive validity in specific contexts such as travel mode choice (Friedrichsmeier *et al.* 2013), it does not capture other components of habits, such as behavior complexity. As shown by Verplanken (2006), simpler tasks are more likely to become habitual than complex tasks. Therefore, while the frequency-in-context measure can assess relative habits within similar tasks, it lacks applicability to diverse tasks without considering their complexity. Additionally, it does not adequately capture the automaticity observed in many habitual behaviors.

To address these shortcomings, researchers have proposed self-reported habit measures, including the Self-Reported Habit Index (SRHI) (Verplanken and Orbell 2003) and the Self-Reported Behavioral Automaticity Index (SRBAI) (Gardner *et al.* 2012). These measures require participants to reflect on different aspects of a behavior (e.g., "Behaviour X is something...") and answer corresponding questions about its automaticity, lack of awareness, and lack of control. The SRHI has demonstrated predictive validity across various domains (Gardner 2015a) and robust correlation with behavior frequency (Gardner *et al.* 2011). However, it is essential to acknowledge a limitation of these measures, as they rely on participants' subjective assessment of their own automaticity. This approach overlooks the possibility that people may not have complete recall of their unconscious decisions when reflecting on their actions (Hagger *et al.* 2015).

To mitigate the reliance on self-reports and enhance scalability, an emerging approach in habit measurement research involves the utilization of implicit measures. Implicit measures, commonly employed in cognitive and social psychology, provide indirect assessments that do not depend on participants' subjective evaluations (Gawronski and De Houwer 2014). By being indirect, implicit measures are less susceptible to response biases and less reliant on introspection compared to self-report measures (Greenwald *et al.* 2002). This avenue holds promise for advancing habit measurement towards a more objective and reliable assessment. Additionally, implicit behavioral measures can be more easily applied to large observational datasets, eliminating the need to individually probe participants with survey questions, which can be costly and impractical.

In this chapter, we make a novel contribution to the field of habit measurement by introducing entropy as an implicit measure of behavioral regularity. We thoroughly examine the predictive and convergent validity of the entropy measure and provide empirical evidence to support its effectiveness. Specifically, we apply the entropy measure to the context of interactions with mobile phones and the use of social media applications. Through rigorous statistical analysis, we demonstrate the utility and applicability of the entropy measure in capturing habitual behaviors related to mobile phone usage and social media engagement. By showcasing the predictive power and alignment with other established measures, we establish the credibility and value of entropy as a robust measure of behavioral regularity in the context of habit research.

2.1 Context and Application

The utilization of social media sites has escalated exponentially over the past two decades. Notably, in 2005, a meager 5% of American adults engaged with at least one of these platforms, a figure that had risen to 72% by 2021 (Pew Research Center, 2021). The average global user was documented to have spent over 145 minutes daily on social media in 2022 (We Are Social, DataReportal, Hootsuite 2022). The pervasive adoption of these platforms, coupled with the significant revenue garnered by social media corporations[1], underscores their critical role as subjects of policy, strategic business, and personal decision-making studies.

On a personal scale, excessive engagement with social media platforms has been associated with the potential development of addiction-like behaviors (Bányai *et al.* 2017), resulting in detrimental effects on individuals' mental health and overall well-being (Bekalu *et al.* 2019, Braghieri *et al.* 2022, Kross *et al.* 2013, Twenge *et al.* 2019). Multiple studies have exhibited a correlation between problematic social media use and heightened feelings of depression (Kelly *et al.* 2018, Lin *et al.* 2016) and loneliness (Dienlin *et al.* 2017). Comprehensive review articles by Pantic (2014) and Sun and Zhang (2021) offer more insights into this issue. In addition, there is mounting evidence suggesting that frequent social media usage could lead to poor sleep quality, thereby negatively affecting physical health (Chou and Edge 2012, Koc and Gulyagci 2013, Wolniczak *et al.* 2013).[2]

Delving into the understanding and measurement of habits is critical as it can guide the development of preventative measures designed to disrupt the possible progression from regular or habitual use to more problematic and addictive patterns. Engagement with social media platforms often elicits a series of cognitive rewards, including but not limited to, social validation precipitated by 'likes' and comments, entertainment satisfaction, or the acquisition of novel knowledge. Such rewards, when consistently experienced in response to particular cues, have the potential to forge cue-response relationships that underpin the emergence of habitual conduct (Anderson and Wood 2021). To evaluate whether habitual behavior manifests in the frequent utilization of social media, Anderson and Wood (2021) administered a survey to a

[1]In 2021, social network advertising revenues in the United States exceeded $50 billion (Hootsuite 2022)

[2]Conversely, a few studies, such as the one conducted by Orben and Przybylski (2019), propose a small inverse correlation between digital technology use and adolescent well-being. This implies that the negative impacts might be too marginal to warrant policy modifications. Nonetheless, as advocated by Götz *et al.* (2022), even slight effect sizes could produce significant implications when considered on a larger scale and over a prolonged duration.

cohort of Twitter and Facebook users acquired via MTurk. The study yielded evidence that the frequency of content posting was significantly correlated with the automaticity of the behavior, with correlation coefficients $r(124) = .46$, $p < .001$, for Twitter, and $r(60) = .50$, $p < .001$, for Facebook. To quantify the degree of automaticity, the Self-Report Habit Index (Gardner *et al.* 2012) was employed in the context of Facebook and Twitter usage. In alignment with these findings, Schnauber-Stockmann and Naab (2019) supplemented this body of evidence with a longitudinal study surveying students who began utilizing a sports-oriented application. Their data revealed that the strength of a habit is positively correlated with the number of behavioral repetitions. This realization accentuates the relevance of our research within this domain.

However, the existing body of literature has primarily focused on a limited subset of social media applications, thereby constraining our holistic understanding of the intricate dynamics of habitual social media usage in its entirety. The implicit measure we propose can be universally applied across the entire spectrum of applications habitually used by individuals on their mobile devices. This offers an opportunity to cast a broader investigative net that captures a richer and more comprehensive picture of habitual behavior in the context of social media usage.

2.2 Related work

Habitual behavior's influence on technology usage extends beyond the confines of contemporary social media applications. In fact, the broader sphere of technology habits subsumes social media habits, and this wider field has been a subject of academic scrutiny for more than a century (Bayer and LaRose 2018). Notably, the genesis of research on technology habits can be traced back to the pioneering work of William and Harter (1899), who explored the multifaceted stages involved in learning telegraphy. Their findings suggested the necessity of a complex hierarchy of psycho-physical habits for mastering the telegraphic language. Within the domain of technology habits, the habitual use of communication technologies bears the closest relevance to our topic of interest. The significance of habits in dictating media usage has been a longstanding theme in communication studies (Bae 2018, Du *et al.* 2020, Giannakos *et al.* 2013, LaRose 2010). For instance, Cooper and Tang (2009) offered compelling evidence that television users perceive their media selection to be predominantly habit-driven, underscoring the pivotal role of habitual behavior in technology utilization even prior to the emergence of modern digital platforms. Another study conducted by Bayer and Campbell (2012) revealed a robust correlation between habitual texting and both the sending and reading of text messages while driving. The conclusions derived from their work posit that such high-risk behaviors might often be enacted absent explicit intention or awareness, accentuating the potential perils associated with deeply ingrained habits. For a comprehensive overview of the current status of this research domain, we recommend referring to the review conducted by Bayer *et al.* (2022).

With the advent and subsequent ubiquity of the internet, academic focus has shifted from television habits to those associated with internet use. These habits have been identified as significant determinants of a variety of online behaviors, encompassing general internet usage, e-commerce, media file downloads, social networking, and online news consumption (Bayer and LaRose 2018). For instance, Oulasvirta *et al.* (2012) provide empirical support for the widely held assumption that mobile devices foster habit formation, and they elucidate how these habits contribute to the ubiquity of smartphone usage. The specific form of habit they identify is termed a 'checking habit', characterized by brief, repetitive scrutiny of dynamic content readily accessible on the device. From a more theoretical standpoint, LaRose and Eastin (2004) introduce an innovative model that amalgamates Bandura's Social Cognitive Theory (SCT) (Bandura 1999) with the Uses and Gratifications (U&G) theory (Katz *et al.* 1973) to provide a more comprehensive understanding of internet user behavior. They propose that internet usage is propelled by a series of anticipated outcomes, or gratifications, which are shaped by both social and individual factors. They argue that users actively pursue these outcomes, and their behavior is reinforced through a cyclical learning process encompassing attention, retention, reproduction, and motivation - the four cardinal constructs of SCT. This cyclical process is hypothesized to culminate in habitual internet use over time. From a measurement perspective, Araujo *et al.*

(2017) explore the biases and inaccuracies that can occur in self-reported measures of internet use. They combine automatic tracking data and survey data from the same participants to confirm low levels of accuracy and tendencies of over-reporting. Their results further underscore the need for including automatically collected behavioral data, a central advantage of the dataset we use and our proposed entropy measure.

Routines and habits are closely intertwined, with the key distinguishing factor being the emphasis on temporal structure in routine behavior. In the field of economics and marketing, researchers have developed several models and metrics to measure the extent of routineness in human behavior. These frameworks aim to capture the repetitive and consistent nature of routines, shedding light on the regular patterns and predictability in consumer choices and actions. In a seminal paper, Zhang *et al.* (2015) introduce an individual-level metric termed 'Clumpiness', defined as the extent of deviation from equal temporal spacing. If a user's historical events are densely clustered over time, it is considered highly clumpy usage, whereas the metric is minimized if there is equal spacing between consecutive events.[3] Utilizing various datasets, Zhang *et al.* (2015) empirically demonstrate the significant predictive power of the clumpiness metric for long-term user behavior and churn probability, controlling for baseline measures such as frequency and recency. In our study, we incorporate the clumpiness metric as a covariates, acknowledging its proven predictive power. This allows us to isolate the unique contributions of our proposed entropy measure while accounting for the influence of clumpiness.

In addition to straightforward metric-based methodologies, more intricate statistical models have been employed to encapsulate routine behavior. For example, Dew *et al.* (2021) utilize a hierarchical, Bayesian nonparametric Gaussian process model to discern customer-level routines, thereby enabling the dissection of a customer's behavior into routine and non-routine segments. When applied to a ride-sharing platform, their model revealed that individuals exhibiting a higher degree of "routineness" were correlated with increased future usage, decreased churn rates, and

[3]The clumpiness metric bears a close relationship to our metric, especially if we consider the entropy of inter-event times.

enhanced resilience to service disruptions. In the context of multidimensional longitudinal behavioral data, Eagle and Pentland (2009) propose an innovative approach to discerning the underlying structure in daily routines by employing dimensionality reduction techniques. They introduce the notion of 'eigenbehaviors', defined as the principal components of an individual's comprehensive behavioral dataset. Their findings indicate that an individual's behavior throughout a specific day can be approximated by a weighted sum of their primary eigenbehaviors. Remarkably, when these weights are computed halfway through a day, they can be utilized to predict the remainder of the day's behaviors with an accuracy of 79% for their test subjects.

In a more recent study that bears both conceptual and methodological similarities to our research, Buyalskaya *et al.* (2023) introduced a machine learning methodology predicated on LASSO regression. This approach was designed to discern which among numerous context variables are associated with habitual behavior and to infer the rate at which habits form. The authors applied this method to two extensive panel data sets with objective measures, in the contexts of gym attendance and hospital handwashing. In addition, they tested the association between habit formation and reward sensitivity in a random-assignment megastudy. Their findings suggest that gym attendees with more habitual tendencies are less responsive to interventions, implying a higher degree of reward insensitivity. Conceptually, they define habit formation through the extent of predictability of behavior, characterizing time series of behavior as habitual if the accuracy of prediction surpasses a certain threshold. However, their approach is not without limitations. It is plausible that mechanisms other than automaticity might account for the measured context-sensitive predictability. In contrast, our approach in this chapter diverges as our entropy measure estimates the extent to which circumstances are conducive to habit formation. This is achieved by gauging the context-stability, assessing the degree to which an action is consistently performed in an unchanging setting, thereby indicating the likely presence of a habit.

2.3 Overview

In this chapter, we present a novel contribution to the domain of habit measurement by introducing entropy as an implicit measure of behavioral regularity. We elucidate how entropy, a measure extensively employed to quantify uncertainty and randomness in probability distributions, serves as a metric for the stability of a behavior's context, thereby facilitating habit formation. While this metric could be applied to a multitude of settings involving regular and episodic behavior, such as gym visits, sleep patterns, running routines, etc., our focus herein is on mobile phone usage. Our analysis leverages a comprehensive and extensive dataset on mobile usage, which records timestamped entries each time a user engages with an app. While consumer panels based on desktop browsing data are more common (Chiou and Tucker 2012, Rao 2022), there a few recent studies based on phone usage panels (Agarwal *et al.* 2022, Horta Ribeiro *et al.* 2023). For the purposes of this chapter, our analyses are specifically concentrated on applications that fall under the category of social media. This choice is deliberate and is motivated by the recognition that social media behavior has been widely acknowledged to exhibit characteristics of habitual and addictive tendencies, rendering it particularly pertinent to our research objectives.

As previously explained, theories of habit formation in psychology posit that the repetition of a behavior in a specific context (e.g., checking your Facebook page upon arriving at the office in the morning) could engender the formation of cue-response associations (Anderson and Wood 2021). A plethora of contextual cues might consequently become associated with social media use, thereby triggering it automatically. Not all of these diverse triggers are observable in our dataset and are encapsulated by our entropy metric. Therefore, it is crucial to distinguish different mechanisms for triggering habits to illustrate precisely what is being measured by the entropy. As outlined by Anderson and Wood (2021), social media habits could be triggered by mood cues, design feature cues, technology cues, and activity cues. Firstly, the moods and emotions associated with social media use could potentially act as triggers for subsequent use. For instance, a study by Meier *et al.* (2016)

demonstrated that college students who felt uncertain about how to complete a school task were more likely to procrastinate by checking Facebook. Secondly, design features such as apps' push notifications could serve as cues automatically activating use. These features could also involve decreasing friction through the removal of stopping cues, or natural endpoints to limit media use. Many prominent social media sites such as Twitter, TikTok, Facebook, and Tumblr provide bottomless feeds devoid of natural endpoints. Thirdly, the technology consumers use to access social media, including computers, laptops, tablets, and smartphones, could provide cues for using social media. For example, one might pick up their phone to respond to a call or reply to a message, but then see the Facebook icon and decide to open it and engage with it. Finally, places, times of day, and activities can serve as cues triggering media use. A survey involving around 5,000 users of five prominent social media platforms found that Snapchat was the preferred app for socializing with friends, while Instagram, Facebook, and YouTube were commonly used at home or to pass time during waiting periods. Twitter, on the other hand, was frequently accessed during commuting (Anderson and Wood 2021). For individuals with habitual usage patterns, these locations, timings, and daily events can serve as cues that automatically trigger the use of specific social media platforms (e.g., instinctively opening Twitter while commuting on a train).

The entropy metric is intended to encapsulate the final form of these contextual cues, which are predominantly comprised of external cues since they are independent of the user's mood or the technology they engage with. To this end, entropy estimation is applied to quantify the uniformity of the distribution of app usage over a 24-hour clock time period. Clock time is selected as a pertinent contextual factor closely linked to daily routines and activities such as sleeping, eating, working, and commuting. It also serves as a proxy for other contextual cues like location. [4] By computing the entropy of the behavior distribution, we capture the level of regularity in behavior throughout the day. The entropy metric provides a quantifiable measure

[4] While we consider only time as our context variable, the same methods could be used if the data set captures more dimensions for the context of the behavior, such as location

of the concentration of behavior within a specific temporal context by evaluating the KL-divergence of the behavior's histogram from a uniform distribution. A low entropy indicates that individuals tend to repeat specific actions at fixed times, while a high entropy suggests more randomness in the timing of their behavior. Thus, entropy estimates habit indirectly by assessing the likelihood that habit has formed under conducive conditions.

Entropy estimation presents distinct challenges when dealing with data involving a continuous random variable, as most conventional entropy estimation methods are not directly applicable. A common approach is to employ a plug-in estimator. A continuous approximation of the distribution can be achieved by utilizing a density estimation method, followed by numerical integration to compute the entropy. In this study, we employ the Kernel-Density Estimation (KDE) method with a Gaussian kernel to estimate the distribution (Silverman 2018). We divide the data into two distinct periods: a baseline period and a prediction period. The baseline period, spanning three months, serves as the basis for estimating the distribution of daily usage and calculating the corresponding entropy. Additionally, during this baseline period, we calculate various control variables, including average frequency, recency, clumpiness, and total time spent on the app.

Subsequently, we conduct a series of analyses to establish the validity of entropy as a measure of habits. Given that learned habits are relatively impervious to changes in goal structures, we test the hypothesis that entropy, serving as a proxy for habit, can effectively predict long-term user behavior. Most importantly, we demonstrate its predictive validity by employing it as a tool for forecasting long-term social media behavior, while controlling for several other behavioral metrics that are anticipated to possess predictive capabilities for usage patterns. We employ a quasi-Poisson regression, which is suitable for count data exhibiting overdispersion. The results indicate that the estimated coefficient for entropy is negative and statistically significant. This suggests that higher entropy, indicating less regularity and weaker habit formation, is associated with less time spent and lower frequency of use in the long-term future, even after controlling for baseline covariates. To further substantiate this evidence, we utilize non-linear models that provide greater flexibility compared to traditional regression methods, as they avoid reliance on a fixed functional form. In our analysis, we utilize the binsreg implementation proposed by Cattaneo *et al.* (2019a), and accumulated local effect (ALE) plots (Apley and Zhu 2020, Hall *et al.* 2017) in conjunction with a random forest model. Both of these approaches corroborate the linear regression results, showing that high irregularity of behavior is associated with less time and frequency of usage in the future.

Moreover, we establish the face validity of this measure through an intra-application comparison. More specifically, we compare the average entropy values among popular social media applications with that of a clock application, which is presumed to be utilized regularly and thus exhibit low entropy. As anticipated, the results show that the average entropy for clock applications is significantly lower than that of social media apps. This finding supports the notion that entropy is capturing routines in app usage. We also establish the convergent validity of entropy by demonstrating its correlation with the frequency-in-context measure. However, since the exact context of engaging with each app is unknown, we develop a novel approach to compute this measure using unsupervised machine learning methods. Specifically, to identify stable contexts, we employ the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm (Ester *et al.* 1996), a popular density-based clustering algorithm, to identify clusters of concentrated usage patterns. Finally, we utilize variable importance plots to gain a comparative understanding of the significance of entropy compared to other predictors in driving social media usage. This comparison can shed light on the psychological mechanisms driving user engagement and provide insights into the specific role of habits driven by external cues.

By capturing the regularity of behavior in a specific context, the entropy metric provides a unique perspective on habit formation, offering a robust tool for predicting long-term user behavior. Furthermore, the flexibility of entropy metric allows it to be adapted to various settings involving regular and episodic behavior, extending its potential applications beyond social media usage. The findings of this study not only contribute to the existing literature on habit formation and internet usage but also have practical implications for the design of digital technologies and interventions aimed at managing online behavior. By understanding the role of habitual behavior in driving user engagement, designers and policymakers can develop more effective strategies to promote healthy internet usage habits. Future research could explore the application of the entropy metric in different contexts and for different types of apps, further expanding our understanding of online behavior and the mechanisms driving it.

3 Data

Our analysis is conducted using a comprehensive dataset on mobile usage, which was carefully compiled by a market intelligence firm based in the United States called Qrious.[5] This dataset consists of a representative panel of approximately 40,000 Android users who voluntarily opted in to participate. Panelists received a monthly fee from the firm, and their mobile usage data was collected through a mobile app. The data covers the period from October 2022 to June 2023 and captures anonymized, real-time mobile usage activity for each user within the panel.

App sessions data: This data records timestamped entries every time a user opens an app, providing valuable insights into the specific apps used by the panelists. As smartphones rely on individual apps for various functions like making/receiving calls, browsing the internet, or gaming, this dataset offers an unprecedented view of the digital footprint of a diverse set of users in the United States.

Demographics: In addition to app usage data, the dataset also includes detailed self-reported demographic information. This allows us to analyze various demographic factors such as gender, race, age, and more, providing valuable context and enabling a comprehensive understanding of the users involved in the study. Table 3.1 presents summary statistics of user characteristics within the panel, encompassing age, edu-

[5]Data source: https://www.qrious.co.nz

cation, gender, ethnicity, and income category.

This dataset is particularly well-suited for studying social media habits due to its ability to provide accurate and comprehensive large behavioral data. With a panel of approximately 40,000 users, it offers a substantial sample size and allows for reliable and generalizable findings. Covering a span of eight months, the longitudinal nature of the data enables the examination of social media habits over time. The inclusion of detailed app sessions data across multiple applications provides a holistic understanding of user engagement, while also facilitating comparative analyses and the exploration of emerging trends. This dataset is especially ideal for addressing our research questions because it aligns with the predominant trend of social media traffic originating from smartphones in the United States.[6] It is important to note that not all individuals in the panel remain active and utilize social media apps throughout the entire eight-month duration. In the subsequent section, we will elaborate on the data cleaning procedures employed and outline the criteria for selecting the subset of individuals for further analyses.

3.1 Social Media Applications

For the scope of this chapter, we have specifically focused our analyses on applications that fall under the category of social media.[7] This deliberate choice is motivated by the recognition that social media behavior has been widely acknowledged to possess characteristics of habitual and addictive tendencies, making it particularly relevant to our research objectives. It is worth noting that the analytical approaches and methodologies employed in this chapter can be readily extended and applied to other categories of applications, such as gaming or communication. The techniques utilized in our analyses can be adapted to investigate and examine the potential for habitual behavior within these alternative app categories as well. Therefore, the findings and

[6] https://www.statista.com/statistics/477368/us-social-media-visits-share/

[7] Despite YouTube not being traditionally categorized as social media, we have chosen to include it in our analysis due to its high usage volume and relevance to the topic of study. YouTube exhibits social media-like features, such as user interactions and comments, making it an important platform to consider.

Table 3.1: Demographic characteristics of all users involved in the panel

Characteristic	**N = 40,950**
Age	37 (31, 45)
Education	
Associate's Degree	4,434 (13%)
Bachelor's Degree	2,498 (7.4%)
Doctorate or Professional Degree	1,644 (4.8%)
High School Graduate	17,803 (52%)
Master's Degree	1,860 (5.5%)
Some Grade School	1,241 (3.7%)
Some High School	4,483 (13%)
Gender	
Female	21,465 (63%)
Male	12,745 (37%)
Ethnicity	
Asian/Pacific Islander	767 (2.3%)
Black/African-American	4,435 (13%)
Native American Indian	974 (2.9%)
Other	2,003 (5.9%)
White/Caucasian	25,899 (76%)
Income	
\$100,000 to \$249,999	2,652 (7.8%)
\$15,000 to \$24,999	5,710 (17%)
\$25,000 to \$34,999	4,520 (13%)
\$250,000 or more	577 (1.7%)
\$35,000 to \$49,999	3,742 (11%)
\$50,000 to \$74,999	3,105 (9.1%)
\$75,000 to \$99,999	2,060 (6.1%)
Less than \$15,000	11,662 (34%)

insights presented in this chapter offer a foundation for broader investigations into the formation of habits across different types of applications. By focusing on social media applications in this chapter, we aim to provide a comprehensive understanding of the role of habits in this prevalent context. However, we acknowledge the broader applicability of our methods and encourage future research to explore habits within diverse application categories to gain a holistic understanding of user behaviors in the

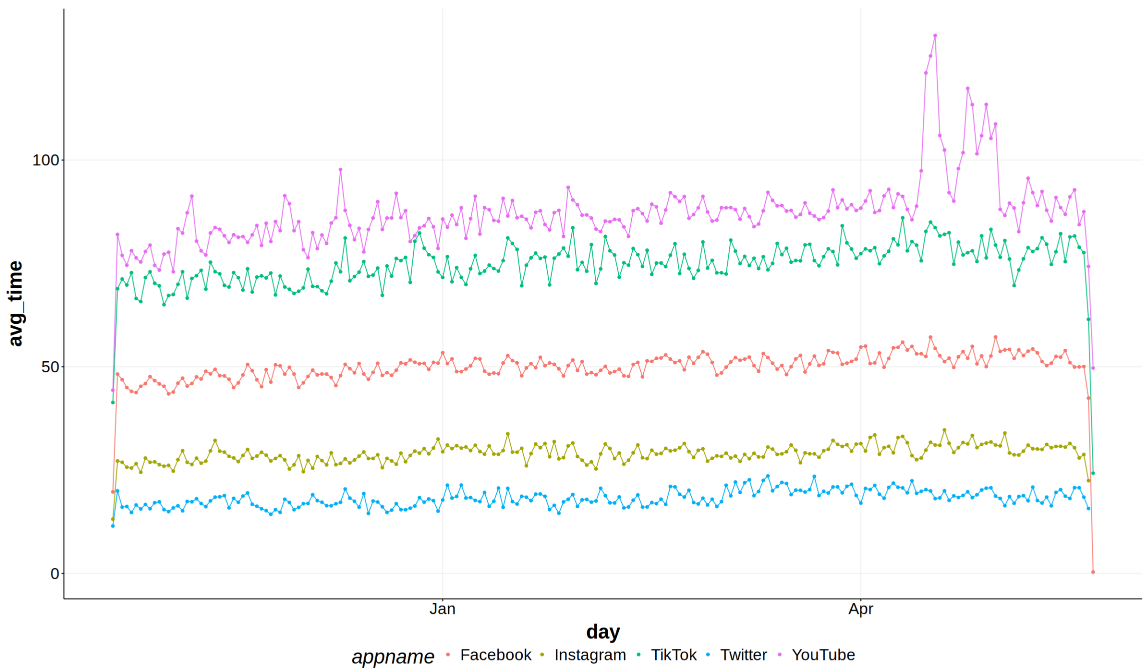


Figure 3.1: Daily usage pattern of top 5 social media application

digital realm.

There is a total 1,179 applications categorized as social media in our dataset. However, it is important to note that a significant portion of these applications are relatively lesser-known and have limited usage, resulting in only a few recorded sessions in the panel. To ensure an adequate number of observations for each application and facilitate meaningful subsequent analyses, we have made the decision to focus on the top 200 most frequently used apps. By selecting the top 200 applications based on the number of sessions recorded in our data, we capture a substantial portion of the overall usage patterns. These top apps account for more than 99% of the total observations, ensuring that our analysis includes a comprehensive representation of user behavior within the social media domain. Figure 3.1 shows the daily usage pattern of the top-5 mostly used applications throughout the panel horizon.

3.2 Sessionization

Sessionization involves the process of categorizing individual user interactions into meaningful sessions based on specified criteria. In order to ensure the meaningfulness of user behavior data for our analysis, we employ a sessionization procedure on the recorded data. While the data is initially captured with a precision of seconds, we acknowledge that this level of granularity may have limited relevance to our analysis. For instance, two consecutive sessions occurring within a few seconds of each other may not correspond to any significant change in the context of the behavior. To address this issue, we aggregate consecutive sessions that occur within a five-minute time window. By merging these sessions, we create a new session with a duration equal to the sum of the combined sessions. This approach enables us to capture more meaningful user interactions and contextualize app usage within a broader time frame.

Furthermore, as an additional step in refining the data, we exclude sessions that have a duration of less than five seconds after the merging process. This criterion ensures that we prioritize events where there is a reasonable level of interaction with the app, filtering out extremely brief sessions that may not reflect meaningful user behavior. By incorporating this criterion into our sessionization approach, we enhance the quality and relevance of the behavior data used in our analysis. This enables us to focus specifically on sessions that demonstrate a reasonable level of engagement and exclude irrelevant or minimal interactions, thereby providing a more accurate representation of user routines and habits.

4 Measures of Behavioral Regularity

Engagement with social media applications often involves regular and periodic behavior, which can be quantified using various approaches. In this section, we propose the entropy method as a measure of regularity, drawing on psychological theories of habit formation. According to these theories, habits are more likely to form when a rewarding behavior is repeatedly performed in the same context (Verplanken and Aarts 1999, Wood and Rünger 2016). This notion is supported by the cognitive rewards associated with social media engagement, such as social validation through likes and comments, and entertainment satisfaction, which increase the likelihood of habit formation. Therefore, having a measure that captures the stability of the behavior context can serve as a proxy for habit formation. The entropy metric is designed to capture the stability of these contextual cues, particularly those associated with the time of day. In the following sections, we will first explain how we define, estimate, and interpret the entropy metric, and then introduce the clumpiness metric, another widely used measure for assessing some form of regularity.

4.1 Entropy

Entropy, a fundamental measure of uncertainty and randomness in probability distributions, finds extensive application in various fields such as Information Theory and Physics. It quantifies the information necessary for describing or predicting outcomes of a random variable. Specifically, the entropy of a continuous probability distribution (also referred to as differential entropy), denoted as $P(x)$, is mathematically expressed by the formula:

$$H_P = -\int P(x) \log P(x) dx. \tag{3.1}$$

In our study, we calculate the entropy for the distribution of app usage across a 24-hour clock time for each user-app pair. This distribution is obtained by aggregating all observations over a period of three months into a single interval spanning from 0 to 24 hours. Consequently, the entropy derived from this distribution provides an indication of the regularity of app usage throughout the day. Clock time holds significant contextual relevance as it is intertwined with our daily routines, encompassing activities such as sleeping, eating, working, commuting, and more. Moreover, clock time can serve as a suitable proxy for other contextual cues, such as location. For instance, individuals are likely to be in their bedrooms at consistent times each night or at their office desks when they start work in the morning. It is important to note that entropy reaches its maximum value for a uniform distribution. Therefore, individuals who lack specific patterns of app usage over time would exhibit the highest entropy, whereas those who demonstrate concentrated and consistent usage during specific intervals of the day would exhibit lower entropy.

To illustrate this concept, let's consider the example of two users, User1 and User2. User1 has a specific pattern for checking her Facebook profile: she does so every time she arrives at her office in the morning around 9 am, and at the end of her work day before going home. On the other hand, User2 does not have a particular pattern for using Twitter and only checks it when she receives a notification. If we observe the time of use for these users over several days and plot its distribution, we can see clear differences in their usage patterns. For User1, the distribution of Facebook usage is concentrated around 9 am and 5 pm, indicating a regular and predictable pattern. In contrast, the distribution for User2 is more uniform across the day, with no clear concentration of usage at specific times.

Figure 1 illustrates a similar example for Facebook usage, where two panelists have the same number of sessions but exhibit distinct usage patterns with varying degrees of regularity. The left sub-figure shows the histogram and estimated probability distribution of Facebook usage for a specific panelist throughout the day. This panelist demonstrates more regular usage with concentrated patterns around 10 am, 3 pm, and 7 pm. On the other hand, the right sub-figure shows the usage patterns of the second panelist, which are more uniform throughout the day. This distinction in usage patterns is effectively captured by the differences in the estimated entropy for the two distributions. The first panelist on the left has a lower entropy value of 2.21, indicating a more predictable and less variable usage pattern. In contrast, the second panelist on the right has a higher entropy value of 2.89, indicating a less predictable and more variable usage pattern.

An alternative approach to interpreting entropy is through the perspective of Kullback-Leibler (KL) divergence. KL divergence serves as a measure of dissimilarity between two probability distributions by quantifying how one distribution diverges from a reference distribution. When considering two continuous probability distribu-

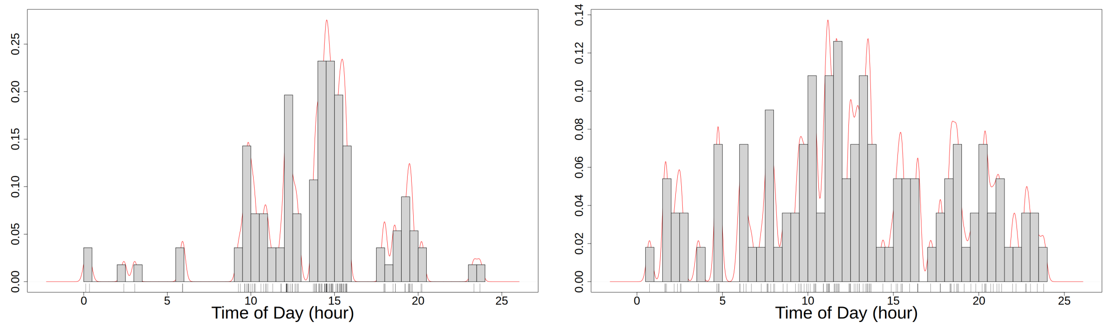


Figure 3.2: Comparison of Facebook usage patterns for two panelists with distinct levels of regularity. The left sub-figure shows the histogram and estimated probability distribution of Facebook usage for a panelist with concentrated patterns around 10 am, 3 pm, and 7 pm, resulting in lower entropy (2.21). The right sub-figure displays the usage patterns of a panelist with more uniform distribution throughout the day, resulting in higher entropy (2.89)

tions, P and Q, the KL divergence from P to Q can is defined as follows:

$$KL_{(P||Q)} = \int P(x)\log(\frac{P(x)}{Q(x)})dx. \tag{3.2}$$

If we consider uniform distribution as the reference distribution $U(x)$, then the KL-divergence is proportional to the negative entropy: $KL_{(P||U)} = \int_0^D P(x)log(\frac{P(x)}{\frac{1}{D}}))dx = -DH_P$, where D is the measure of the domain on which the integral is taken ($D = 24$ in our example if the scale is based on hours). From this perspective, our proposed entropy metric quantifies the deviation of the distribution of app usage from a uniform distribution such that a higher divergence (lower entropy) is indicative of more regularity in behavior, which could in turn facilitate the formation of associations between phone usage and contextual cues. Next, we will outline the estimation procedure used to calculate the differential entropy of the app usage distribution.

4.1.1 Estimation Methods

In the case of data involving a continuous random variable, most conventional entropy estimation methods are not directly applicable. However, a common approach to address this issue is to utilize a plug-in estimator. The simplest method involves approximating $P(x)$ by binning the x-axis and constructing an empirical histogram. By obtaining a discrete probability distribution, denoted as $P(x_i)$, the discrete entropy can be computed through a straightforward summation process.

$$\hat{H}(x) = \sum_i P(x_i) \log(P(x_i)/w_i) \tag{3.3}$$

where w_i is length of the i^{th} bin.

Alternatively, a continuous approximation of the distribution can be achieved by employing a density estimation method, followed by numerical integration of the integral in Equation 3.1 to compute the entropy. In this study, we utilize the Kernel-Density Estimation (KDE) method with a Gaussian kernel to estimate the distribution. Kernel density estimation utilizes a kernel function, denoted as K, along with a bandwidth parameter h, to estimate the density at a specific point. This estimation involves a weighted average calculation that incorporates nearby points from the sample. In this process, points in close proximity to the point of interest carry more weight and have a greater influence on the density estimation compared to points that are farther away. The estimated density $\hat{P}_h(x)$ could be computed using the following formula.

$$\hat{P}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{||x - x_i||}{h}) \tag{3.4}$$

The choice of the kernel function and the bandwidth parameter allows for adjustments to be made to the nature of this relationship, determining the level of contribution and influence from neighboring points. The flexibility of kernel density estimation enables researchers to customize the method according to the specific requirements of their problem. In essence, the selection of the bandwidth parameter is contingent upon the definition of an stable context for the behavior under investigation. In our case, considering that the duration of using social media apps typically spans a few minutes, we opt for a bandwidth in the order of 10 minutes. This choice is motivated by the desire to capture relevant patterns and variations in app usage within a reasonable temporal window. In Appendix B, we perform sensitivity analyses by varying the bandwidth and examine the robustness of our main regression results to different bandwidth selections.

4.2 Clumpiness

The notion of behavioral regularity has long been of interest in behavioral science and has prompted the development of various metrics to quantify this concept. Among these metrics, the *clumpiness* measure has emerged as a prominent tool for analyzing temporal patterns in individual-level event occurrences. Zhang *et al.* (2015) introduced the individual-level clumpiness metric, which gauges the extent to which event occurrences deviate from equal spacing over time. By assessing the degree of clustering or dispersion in an individual's temporal behavioral patterns, this metric offers valuable insights into the structure and clustering tendencies of user behavior. The clumpiness metric characterizes a highly clumpy usage pattern when events are densely concentrated or gathered closely together in time, signifying a pronounced lack of spacing between consecutive events. Conversely, minimal clumpiness is observed when there is a more uniform or equal spacing between events, indicating a more dispersed temporal pattern. By capturing patterns of temporal clustering or dispersion, the clumpiness metric sheds light on the regularity or randomness of event occurrences in individual users' behavioral trajectories. Mathematically, the clumpiness metric is defined as:

$$C = 1 + \frac{\sum_{i=1}^{N+1} \log(t_i) t_i}{N+1} \tag{3.5}$$

where t_i represents the inter-event time between two consecutive events, and N is the total number of events. Notably, the clumpiness measure is computed by scaling the inter-event times, rendering it conceptually invariant to the scaling of time units. However, in empirical applications, this invariance may not always hold true due to the discreteness and nonlinearity of C. Consequently, when the selected time unit is either too long or too short, the computed value may differ, potentially leading to the loss of information about the inter-event times or the risk of over-fitting patterns. In the context of our study, we opted to use minutes as the time scale, a choice that aligns with the nature of social media usage. Given that users can engage with social media platforms multiple times within a single hour, using minutes as the time unit allows for a granular examination of the temporal patterns of user behavior.

The original application of clumpiness lies within the domain of marketing, where it has been utilized to predict customer lifetime value and gain insights into individual-level temporal behaviors. The introduction of the Clumpiness metric has significantly impacted the field, offering a valuable tool for understanding and forecasting temporal patterns of behavior at the individual level. In this study, we make a notable contribution to this literature by evaluating the predictive efficacy of clumpiness in the context of social media behavior. To achieve this, we incorporate the clumpiness metric as a control variable in our analysis, acknowledging its well-established predictive power. By doing so, we can effectively isolate and examine the distinct contributions of our proposed entropy measure while accounting for the influence of clumpiness. This approach enables us to disentangle the effects of temporal clustering and regularity captured by clumpiness from the underlying temporal stability and routines represented by entropy. By extending the application of clumpiness to the realm of social media behavior, our study adds depth to the existing literature and sheds light on the unique predictive capacities of both clumpiness and entropy in this specific domain.

5 Validity of Entropy as a Habit Measure

In this section, we rigorously examine the validity of entropy as a measure of habits through a series of analyses. Our goal is to establish its credibility and utility in capturing regular behavioral patterns in the context of social media usage. To begin, we assess the face validity of entropy by conducting an intra-application comparison. Specifically, we compare the average entropy values among popular social media applications with those of a clock application, which is expected to exhibit low entropy due to its regular and predictable usage. Next, we investigate the convergent and discriminant validity of entropy by examining its correlation with the frequency-in-context measure and contrasting it with mere frequency. This analysis aims to showcase the unique information captured by entropy and how it differs from traditional frequency-based measures. Furthermore, we demonstrate the predictive validity of entropy by using it as a powerful tool for forecasting long-term social media behavior. To ensure a robust analysis, we control for several other behavioral metrics that are anticipated to possess predictive capabilities for usage patterns. Lastly, we quantify the significance of entropy in comparison to other predictors, elucidating its relative importance in driving social media usage. Throughout our investigation, we focus on two primary outcome variables:

- Total daily time spent on social media applications (measured in minutes).
- Frequency of sessions per day.

By systematically exploring the validity of entropy, we aim to provide compelling evidence for its suitability as a valuable and informative measure of habits in the realm of social media behavior.

5.1 Face Validity

Face validity refers to the extent to which a psychological measure appears to assess the construct or attribute it intends to measure based on its "face" or surface characteristics. It is often considered an initial and informal form of validity assessment. In this context, we aimed to establish the face validity of entropy as a measure of context stability and routines in app usage. To achieve this, we conducted a cross-app comparison between time-of-day clock app usage and social media applications. The clock app served as a relevant reference point due to its expected usage pattern, linked to individuals' sleeping routines.

Figure 3.3 illustrates the results, displaying the average entropy for the top 5 social media apps alongside two clock apps with the highest number of users. As expected, the average entropy for clock applications was significantly lower than that of social media apps. The significant difference in average entropy between clock applications

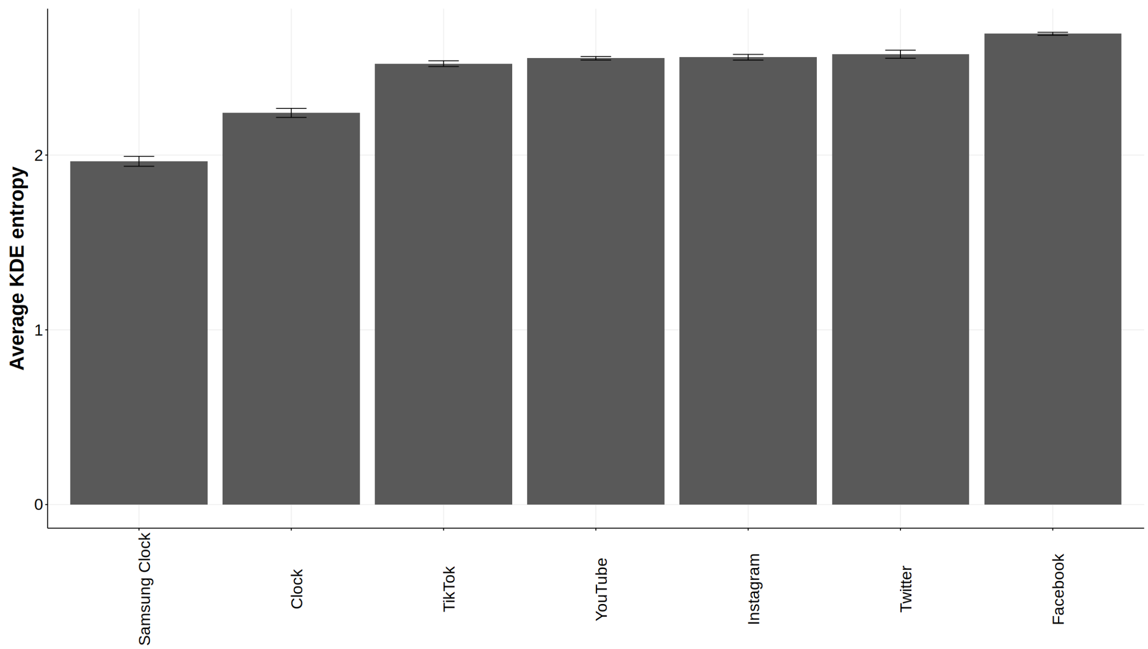


Figure 3.3: Comparison of Average Entropy between Popular Social Media Apps and Clock Apps

and social media apps provides preliminary support for the notion that entropy captures routines in app usage. Remarkably, this difference remains significant even though clock applications offer additional features, such as timers or stopwatches, which are not directly driven by routines like the clock feature. However, it is crucial to acknowledge that face validity alone does not offer strong empirical evidence for the measure's validity. Further steps will involve rigorous statistical analyses and comparisons with other relevant measures to thoroughly validate and enhance the measure's reliability and comprehensiveness.

5.2 Convergent and Discriminant Validity

Convergent validity is a crucial aspect of construct validity that evaluates how well a particular assessment or measure aligns with other measures that are theoretically expected to capture the same construct or concept (Messick 1990). This type of validity examines the extent to which different measures or indicators of the same construct yield similar or converging results. Strong convergent validity indicates that the measure effectively and consistently captures the intended construct, as evidenced by its correlations with other established measures of the same construct. On the other hand, discriminant validity assesses the distinctiveness of a measure from unrelated constructs. It ensures that the measure is not confounded with other unrelated variables, thereby demonstrating its ability to distinguish and measure the construct of interest independently from other constructs. In the domain of habit measurement, the absence of a definitive gold standard or directly observable true habit strength presents certain challenges. However, researchers have developed and validated a few habit measures that are widely accepted in the literature. Two commonly used measures are the frequency-in-context and the self-reported habit index (SRHI). The SRHI relies on self-reported survey responses, which unfortunately are not available in our dataset, rendering it unobservable to us. In the absence of the SRHI, we can leverage the available frequency-in-context measure as a valid proxy for habit strength in our analysis.

In prior studies, the frequency-in-context measure has been limited to specific and well-defined contexts, carefully selected to capture habitual behavior in particular situations (Neal *et al.* 2011). However, while the frequency component of the frequency-in-context measure can be readily computed from the observed data, the context component remains unobservable. To overcome this limitation, we develop a novel approach to approximate the stable context from observed behavioral data using unsupervised machine learning methods. Our approach focuses on identifying stable contexts for daily routines, where users consistently engage in app usage during specific intervals, such as lunchtime or bedtime. These stable contexts are characterized by a higher concentration of app usage compared to other times of the day when usage is more random and not linked to any specific context. Even if these dense usage intervals are not directly associated with other activities, the time of day itself can serve as a relevant context that may influence behavior due to its repetitive nature.

5.2.1 DBSCAN Algorithm

To identify these stable contexts, we employ the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm (Ester *et al.* 1996), which is a popular density-based clustering algorithm. The DBSCAN algorithm operates based on the assumption that clusters are dense regions in space, separated by regions of lower density. It groups together data points that are closely located to form a cluster, while data points located in less dense regions are considered as noise or outliers. The DBSCAN algorithm relies on two parameters: minPts and ϵ. The parameter minPts represents the minimum number of points required to form a cluster in a region for it to be considered dense, while ϵ serves as a distance measure, determining the neighborhood of any given point. The algorithm's procedure is as follows:

1. **Inputs:** Dataset D, Minimum number of points $minPts$, Distance threshold ε
2. **Outputs:** Clusters and noise points
3. Mark all points in D as unvisited
4. For each unvisited point p in D, do the following:
   (a) Mark p as visited
   (b) Find all points within distance ε of p and store them in $neighborhood$
   (c) If $|neighborhood| \geq minPts$, then:
      i. Create a new cluster
      ii. Add p to the cluster
      iii. Expand the cluster with (p, $neighborhood$, ε, $minPts$)
   (d) Else, mark p as noise

For a more detailed understanding of the algorithm's procedure, you can refer to the pseudo code provided in Appendix B, Algorithm 1.

By applying the DBSCAN algorithm to our observed events data, we can uncover distinct patterns of app usage throughout the day and identify clusters of concentrated usage patterns. The algorithm considers the density of app usage events in different time periods and creates clusters based on their proximity and density. These clusters serve as proxies for stable contexts associated with habitual behavior. Using DBSCAN for clustering offers several advantages. It enables us to automatically detect and define stable contexts without relying on pre-defined time intervals or external context information. This flexibility and data-driven approach expand the applicability of the frequency-in-context measure to settings where diverse and unobservable contexts may exist.

5.2.2 Convergent Validity

As discussed earlier, the DBSCAN algorithm relies on two hyperparameters: ϵ and minPts. The parameter ϵ represents the radius of the interval where high-density points are located, determining the length of the context. In our study, we set ϵ to 15 minutes to facilitate a relevant comparison with the bandwidth used for entropy estimation. On the other hand, the parameter minPts indicates the minimum number of points required to form a dense region or cluster. Since the number of observations (sessions) can vary significantly among different user-app pairs, applying a fixed minPts value is not appropriate. To address this issue, we consider density as a relative concept and set minPts to twice the baseline density. The baseline density is defined as the fraction of points relative to the number of clusters obtained when breaking the 24-hour interval into regions of 2ϵ intervals. This approach accommodates variations in sample sizes and effectively identifies regions with relatively high density as stable clusters.

Indeed, the DBSCAN algorithm is designed to identify core clusters, where data points are densely grouped together, and separate outlying points in less dense regions as noise or outliers. By computing the fraction of data points belonging to the core clusters, we effectively estimate the frequency in context. This measure quantifies the relative frequency of app usage occurring within time-of-day intervals that have been

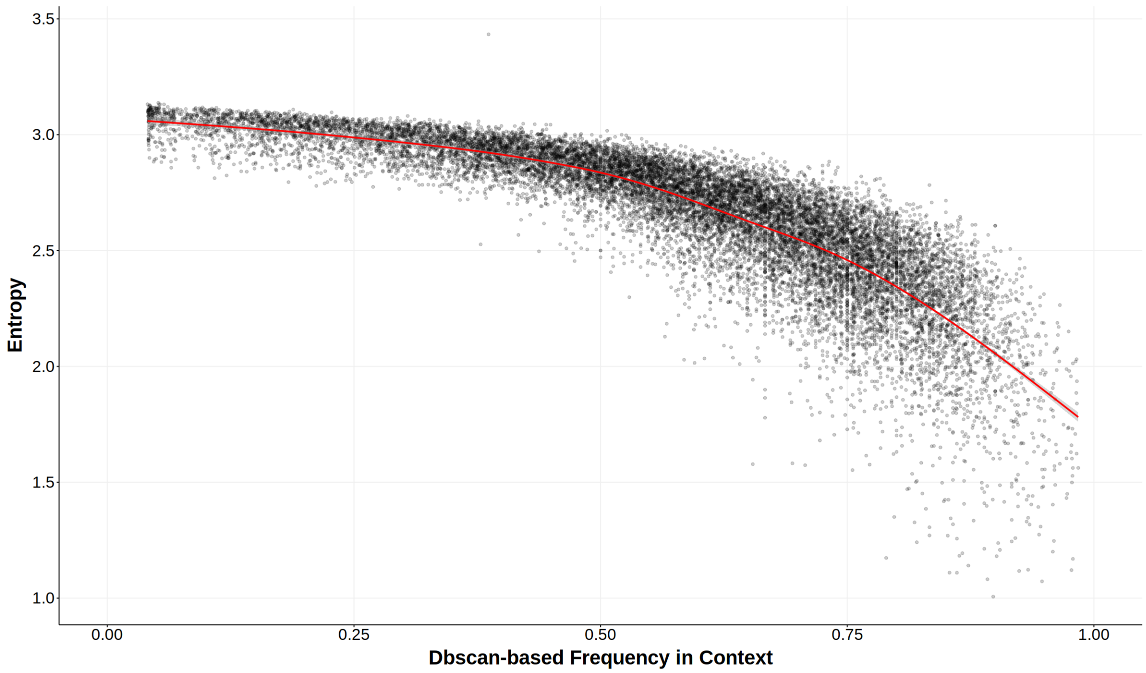


Figure 3.4: The plot demonstrates the correlation between the frequency in context, estimated using the DBSCAN algorithm, and the entropy measure obtained through kernel density estimation. As the frequency in context increases, the entropy decreases, indicating convergent validity of the entropy measure. The plot highlights the relationship between temporal regularity captured by entropy and the occurrence of app usage in high-density intervals identified by DBSCAN.

identified as high-density regions by the DBSCAN algorithm. In other words, the frequency in context represents the proportion of app usage events that fall within the stable and repetitive time intervals detected by the algorithm. These intervals correspond to periods of dense app usage, which are indicative of potential habitual behaviors or routines. Figure 3.4 illustrates the correlation between DBSCAN-based frequency in context and entropy measured by kernel density estimation. The results indicate that as the frequency in context increases, the entropy decreases, providing evidence for the convergent validity of the entropy measure. This finding reinforces the idea that higher levels of behavioral regularity are associated with more stable and repetitive contexts, as reflected in lower entropy values.

5.2.3 Discriminant Validity

As explained in the introduction, behavioral frequency and habit measurements represent distinct constructs and cannot be equated. Habits go beyond mere frequency and involve automatic and repetitive behaviors that are triggered by contextual cues, whereas frequent behaviors may not necessarily exhibit habitual characteristics (Ajzen 2002, Rebar *et al.* 2018). Therefore, it is essential for a habit measure to discriminate between true habits and mere behavioral frequency. In our study, we explore the relationship between our entropy measure and the frequency of checking social media apps. Figure 3.5 illustrates how entropy varies with the logarithm of the average daily frequency of app use where each point represents a user-app pair. As depicted, higher entropy (indicating lower regularity) is associated with higher frequency of app usage. This finding supports the notion that the entropy measure captures a distinct construct from mere frequency, and indicates that our entropy measure is successfully capturing the temporal irregularities and variations in app usage that go beyond the frequency of checking apps.

5.3 Predictive Validity

In this section, we employ various statistical modeling frameworks to establish the predictive validity of the entropy metric as a proxy for habits. Each model has its own strengths and limitations, which together offer a comprehensive understanding of the relationship between entropy and future behavior. The results obtained from these models are crucial in demonstrating the validity of entropy as a proxy for habits. Operationalizing habits through their predictive power is a valuable approach, especially when direct measures of automaticity of behavior are not observable in the data. This approach aligns with the vast literature in social psychology and behavioral science, where predictive models have been widely used to study and validate habit formation and related constructs (Buyalskaya *et al.* 2023, Ouellette and Wood 1998). By establishing the predictive validity of entropy, we contribute to the growing body of research on behavioral regularity and its impact on future behaviors in the context

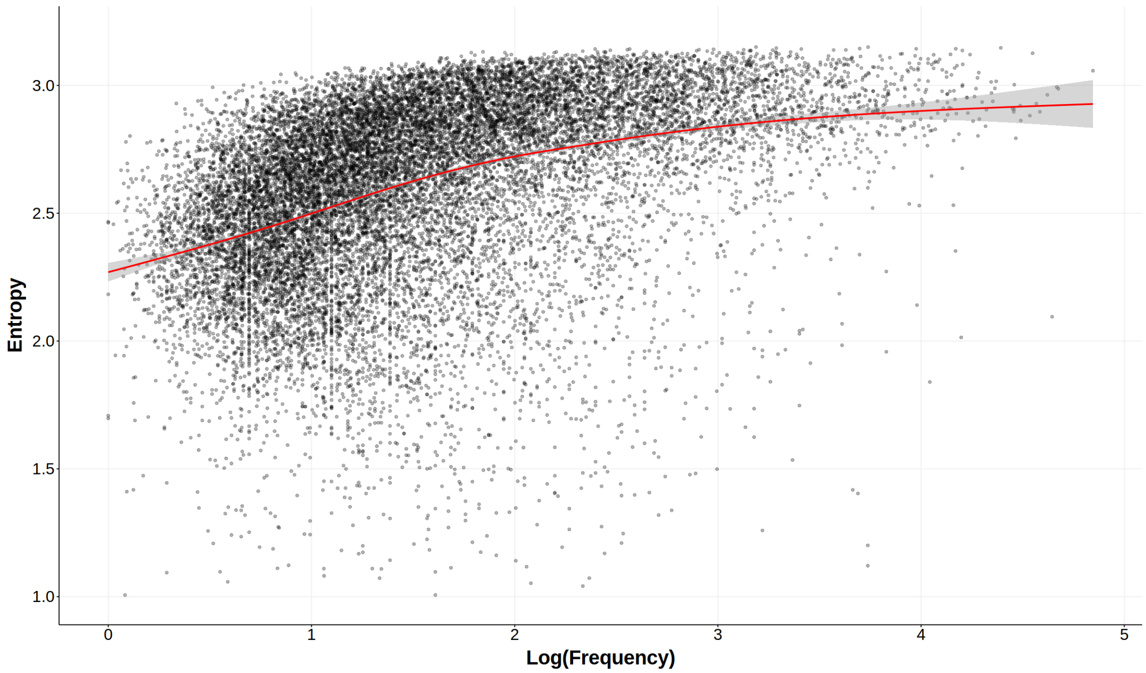


Figure 3.5: The plot shows how entropy varies with the logarithm of the average daily frequency of app use. Each point represents a user-app pair. Higher entropy values (indicating lower regularity) are associated with higher frequency of app usage, demonstrating the distinct construct captured by the entropy measure compared to mere frequency.

of social media usage.

Control Variables: To ensure a robust analysis, we account for several other behavioral metrics that are expected to have predictive capabilities for app usage patterns. This approach allows us to isolate the unique contributions of our proposed entropy measure while considering the influence of other factors affecting future engagement. Given the complexity of this prediction problem and the diverse psychological mechanisms driving future behavior, selecting a comprehensive set of control variables is challenging. Nonetheless, we draw on the vast literature in marketing about RFM (Recency, Frequency, Monetary) models, used for customer segmentation and predicting customer lifetime value (CLV) (Berger and Nasr 1998). From a theoretical perspective, Fader *et al.* (2005) demonstrated the statistical sufficiency of the RFM method under certain assumptions. Furthermore, Zhang *et al.* (2015) extended the traditional RFM approach by incorporating the clumpiness component.

In the context of social media behavior, where purchase behavior is not involved, we include the parameters involved in the RFM model by replacing the monetary value with the time spent on the app, which serves as a proxy for user engagement and value attribution. Additionally, we include the total number of days users actively engage with the app to account for varying panel involvement lengths. The set of control variables, defined at the user-app pair level during the baseline period, includes the following:

- **Frequency**: The average number of times users interact with the app.
- **Time Spent**: The average time spent on the app in minutes.
- **Recency**: The number of days since the user's last engagement with the app.
- **Days Active**: The number of unique days that the user engages with the app.
- **Clumpiness**: Refer to section 4.2 for a complete description.

It is important to note that, given the lack of random variation in entropy, none of the following results provide causal evidence for the impact of entropy on app usage when controlling for other factors. Nevertheless, controlling for these predictive variables offers additional evidence that entropy may capture nuanced aspects of behavior distinct from simple factors like frequency, or even more complex patterns like clumpiness. Indeed, future research with randomized interventions to manipulate entropy while controlling for frequency and total time spent on social media apps could provide valuable insights into the impact of regularity on future usage behavior. This would allow for isolating the effect of regularity from other confounding factors, providing more conclusive evidence on the role of habits in shaping social media usage patterns.

5.3.1 Quasi-Poisson Regression

The first model we employ is a linear quasi-Poisson regression. Quasi-Poisson regression is a statistical modeling technique designed for count data that exhibit overdispersion, where the variance exceeds the mean (Wedderburn 1974). This approach is particularly suitable for modeling data that follow a Poisson-like distribution but display greater variability than what can be accounted for by a standard Poisson regression model. Overdispersion can arise from various sources of variation, such as unobserved heterogeneity or excess zeros in the data. In our specific context, the panel dataset includes numerous users who do not engage with different apps on many days, resulting in a significant number of zero values in the outcome variables (i.e., time spent and frequency). Quasi-Poisson regression allows us to handle such data with overdispersion while accounting for the excess zeros, making it a well-suited model for our analysis.

Quasi-Poisson regression offers several advantages: it enables the modeling of count data with overdispersion without assuming a specific distribution, and it relaxes the stringent assumption of equidispersion (where the variance is equal to the mean) made by standard Poisson regression models. By introducing an additional dispersion parameter, quasi-Poisson regression accommodates the increased variability commonly observed in count data, making it a more flexible and applicable approach for diverse scenarios (Cameron and Trivedi 2013, Ver Hoef and Boveng 2007). The relaxation of the equidispersion assumption is particularly relevant in our analysis of social media behavior, where the mean and variance of the outcome variables (daily time spent and daily frequency) exhibit notable discrepancies. For instance, the mean daily time spent on social media apps is 28.8 minutes, while the variance is as high as 5968. Similarly, the mean daily frequency is 4.2, with a variance of 103. The presence of such substantial overdispersion underscores the significance of using quasi-Poisson regression to effectively model these count data with greater variability, ensuring a more accurate representation of the underlying patterns and relationships between the predictor variables, including entropy, and the outcome variables.

Estimation Procedure: For our estimation purposes, we employ the "fixest" package (Bergé 2018) in the R programming language, which facilitates the estimation of Generalized Linear Models (GLMs) with various link functions, including quasi-Poisson regression. This package is particularly advantageous as it efficiently handles scenarios involving a large number of dummy variables, ensuring computational efficiency and accuracy in our analyses. The quasi-Poisson regression equation used for our model is as follows:

$$\log(\mathbb{E}[y_{i,t,a}|e_{i,a}, \boldsymbol{W}_{i,a}]) = \alpha_t + \alpha_a + \alpha_{n(i,a)} + \beta\, e_{i,a} + \gamma^T \boldsymbol{W}_{i,a} + \epsilon_{i,t,a} \qquad (3.6)$$

where the outcome variable ($y_{i,\mathrm{t},\mathrm{a}}$) represents either the time or frequency associated with user i on day t for application a. To account for temporal and application-specific factors, we incorporate fixed effects at the daily level (α_t) and app level (α_a), respectively. Additionally, we address potential biases in entropy estimation due to variations in sample size by including sample size indicators ($\alpha_n(i,a)$) in our analysis. These indicators help mitigate any potential biases that may arise and ensure a more robust and accurate estimation of entropy, considering the impact of sample size on our results. The estimated time-of-day entropy ($e_{i,a}$) is calculated for user i and application a based on the first three months of the panel. To control for other influential factors, we incorporate a vector of control variables ($\boldsymbol{W}_{i,a}$). These predictors include average time, average frequency, recency, number of active days, and clumpiness, all derived from the same initial three months of the panel.

Results: Table 3.2 presents the estimated regression models, where standard errors are clustered at the panelist level to account for correlated observations across different applications used by the same person. As expected, the coefficients for average time and frequency are positive, indicating that more prior engagement increases the likelihood of future app use. However, the parameter of particular interest to us is entropy. The coefficient for entropy is found to be negative and statistically significant, suggesting that higher entropy, reflecting less regularity and weaker habit formation, is associated with less time spent and lower frequency of use in the long-term future, even after controlling for baseline covariates. The regression results suggest that, among users with equal covariates such as frequency and total time spent, those with more regular usage patterns are more likely to spend more time in the future. This alignment with our expectations from psychological theory provides further support for interpreting entropy as a meaningful proxy for habit and its impact on future social media behavior.

The results also shed light on the influence of clumpiness on future time and frequency of social media usage. Lower clumpiness, reflecting patterns with more evenly spaced inter-session times, is associated with higher future time spent and frequency of app usage. Interestingly, our findings contrast with those reported by Zhang *et al.* (2015) in the context of shopping behavior. In their study, customers with higher clumpiness, indicating more concentrated shopping behavior, exhibited a higher likelihood of retention and future shopping. However, in our analysis of social media usage, we observed the opposite pattern, where higher clumpiness was associated with lower future usage. This disparity suggests the presence of distinct mechanisms at play in the context of social media usage compared to shopping behavior. The nature of social media platforms, with their continuous availability, diverse content, and varied social interactions, may shape user behavior differently than the more transactional nature of shopping. These contextual differences highlight the need to consider the unique characteristics and dynamics of different domains when studying user behavior and the underlying mechanisms driving it. Moreover, it underscores the importance of domain-specific investigations to gain a comprehensive understanding of behavior in various contexts.

5.3.2 Binscatter Plots

Conventional regression-based methods estimate a log-linear association between the outcomes (time or frequency) and the predictor of interest (entropy), potentially disregarding non-linear patterns in the data. However, such an assumption may not fully capture the underlying dependence present in the data. Although one can estimate the model by applying a non-linear transformation to entropy, this approach still relies on predetermined functional form assumptions, limiting its ability to accurately model complex relationships. An alternative approach to gain more insight into the relationship between the outcome variables and entropy is to visualize their association using scatter plots. However, traditional scatter plots lack the inherent capacity

Table 3.2: The table presents the estimated coefficients from the quasi-Poisson regression models, Equation 3.6, for Total Time (left) and Frequency (right) as the outcome variables. Standard errors are reported in parentheses. The fixed effects for day, application, and sample size are included in the models. All standard errors are clustered at the panelist level

	Total Time (1)	Frequency (2)
Entropy	-0.2484***	-0.1587***
	(0.0521)	(0.0489)
Clumpiness	-4.437***	-3.884***
	(0.1568)	(0.1730)
log(Average Time)	0.8608***	0.0012
	(0.0143)	(0.0135)
log(Frequency)	0.1452***	0.9099***
	(0.0544)	(0.0616)
Recency	-0.0107***	-0.0109***
	(0.0028)	(0.0035)
Days Active	-0.0024*	-0.0046***
	(0.0014)	(0.0014)
Observations	1,313,336	1,313,336
Squared Correlation	0.48177	0.54301
Daily fixed effects	✓	✓
Application fixed effects	✓	✓
Sample size fixed effects	✓	✓

to control for other covariates or variables of interest, hindering a comprehensive analysis of their joint impact on the outcomes.

Binned scatter plots, also referred to as binscatters, offer an effective solution to this issue and have gained popularity as a valuable tool in applied microeconomics for visualizing relationships between two variables (Starr and Goldfarb 2020). This approach involves dividing the range of one variable into a small number of bins and representing each bin with a single point, showing the average outcome for observations within that bin. By employing binscatters, we achieve greater flexibility in capturing the underlying relationship between entropy and the outcome variables, without the need for rigid functional form assumptions commonly employed in traditional regression methods. To implement our analysis, we employ the binsreg implementation proposed by Cattaneo *et al.* (2019a). This implementation not only facilitates visualizing the data with confidence intervals in each bin but also enables rigorous testing of shape restrictions. A key advantage of this approach is its ability to address the limitations of residual-based methods and effectively account for control variables (Cattaneo *et al.* 2019b). Moreover, the binsreg implementation extends the estimation procedure to generalized linear models, allowing us to utilize quasi-Poisson link functions that are particularly suitable for our specific application. The formulated model is presented below.

$$\log(\mathbb{E}[y_{i,\tau,a}|e_{i,a}, \boldsymbol{W}_{i,a}]) = \alpha_\tau + \alpha_a + \alpha_{n'(i,a)} + \mu(e_{i,a}) + \gamma^T \boldsymbol{W}_{i,a} + \epsilon_{i,\tau,a} \tag{3.7}$$

where the outcome variable ($y_{i,\tau,a}$) represents either the average time or frequency associated with user i during week τ for application a. To address potential biases in entropy estimation due to variations in sample size, we incorporate sample size bin indicators ($\alpha_n'(i,a)$). Additionally, fixed effects at the weekly level (α_τ) and app level (α_a) are included to account for temporal and application-specific factors, respectively. The estimated time-of-day entropy ($e_{i,a}$) is calculated for user i and application a based on the first three months of the panel. We also incorporate a vector of control variables ($\boldsymbol{W}_{i,a}$) that includes average time, average frequency, recency, number of active days, and clumpiness. These covariates are derived from the same initial three months of the panel.

The key distinction from the previous model (Equation 3.6) lies in the representation of entropy, which enters the model in a fully non-parametric manner through the function $\mu()$. Unlike traditional approaches, the binscatter regression method allows for the discovery of the underlying shape of the function $\mu(.)$ and computes the standard errors of each bin, without imposing specific parametric assumptions. By employing the binscatter regression approach, we gain greater flexibility in capturing the relationship between entropy and the outcome variables, enabling us to explore non-linear patterns in the data and providing valuable insights into the association between entropy and future social media usage.

Estimation Procedure: When estimating Model 3.7 with daily and sample size fixed effects, it becomes necessary to handle large matrices of control variables. In our case, we have 150 days in the prediction period and 1000 unique sample sizes, resulting in a substantial number of dummy variables. Additionally, there are 181 dummies for application-specific fixed effects, further increasing the number of variables to control for. Dealing with such a large number of variables can pose computational challenges. To address these challenges and make the computations more feasible, we implemented some simplifications. Firstly, we aggregated the data at the weekly level, which effectively reduced the number of time fixed effects and the number of rows in the dataset. This aggregation helps in managing the computational complexity associated with a large number of variables. Additionally, we binned the sample size into 100 equal intervals and included dummy variables for each associated sample size bin. In total, this will result in 332 dummy controls. This approach allowed us to capture the variation in sample sizes while reducing the dimensionality of the data. By employing these simplifications, we aimed to strike a balance between computational feasibility and capturing the relevant effects in the model estimation process.

Results: Figure 3.6 displays the results of the binscatter regression models for time (top) and frequency (bottom) as outcome variables. Each point estimate represents the expected value of the outcome within the corresponding bin, while controlling for the covariates. The blue curve in the figure represents a quadratic function fitted to the point estimates, capturing the overall trend. The results reveal a more nuanced relationship between entropy and the outcomes compared to the linear quasi-Poisson regression. While the linear model suggests a significant negative association between time and frequency with entropy, the binscatter models demonstrate a non-monotonic relationship. Visually, the curve appears relatively flat for lower entropy values, while there is a substantial drop for higher values. To provide a more rigorous assessment of the functional form, we conducted formal hypothesis tests to examine the monotonicity and concavity of the relationship, as proposed by Cattaneo *et al.* (2019b). For the top figure (time as the outcome), we find that monotonicity is not rejected with $P = 0.154$, and concavity is also not rejected with $P = 0.870$. Similarly, for frequency as the outcome, monotonicity is not rejected with $P = 0.732$, and concavity is not rejected with $P = 0.298$. These results further support the visual intuition, indicating that high entropy (reflecting high irregularity of behavior) hinders the formation of habits, while within the range of smaller entropy (indicating more regularity), the dependence between regularity and future use becomes weaker. The non-monotonic pattern in the relationship highlights the complexity of habit development and its dependence on different levels of regularity.

Finding a non-monotonic relationship between entropy and future app usage presents an opportunity to conduct further analysis by dividing the data into different subsets based on entropy values. Subsequently, we can estimate the linear model separately for each subset. This data segmentation approach allows for a more detailed examination of the relationship between entropy and future app usage, as it enables us to explore how the strength and direction of the association vary across different levels of irregularity in behavior. Table 3.3 presents the results of the model 3.6 estimated separately for user-app pairs with entropy higher or lower than 2.6. The table demonstrates that the estimated coefficient for entropy is not significant for pairs with $e_{i,a} \leq 2.6$, while it is significant and nearly three times larger for pairs with entropy greater than 2.6. These results reaffirm the findings from the binscatter plot and provide further evidence of the non-linear relationship between entropy and future app usage. The lack of significance for pairs with lower entropy indicates that at lower levels of irregularity, the impact of entropy on future behavior is limited. However, as entropy increases and the behavior becomes more irregular, its effect on future usage becomes more pronounced. Furthermore, this analysis highlights the importance of considering the non-linear relationship between entropy and future behavior and demonstrates how the binscatter approach can assist in model selection for standard regression analysis. By allowing for a more flexible representation of the relationship, binscatter enables a better understanding of the underlying patterns and enhances the robustness of our findings.

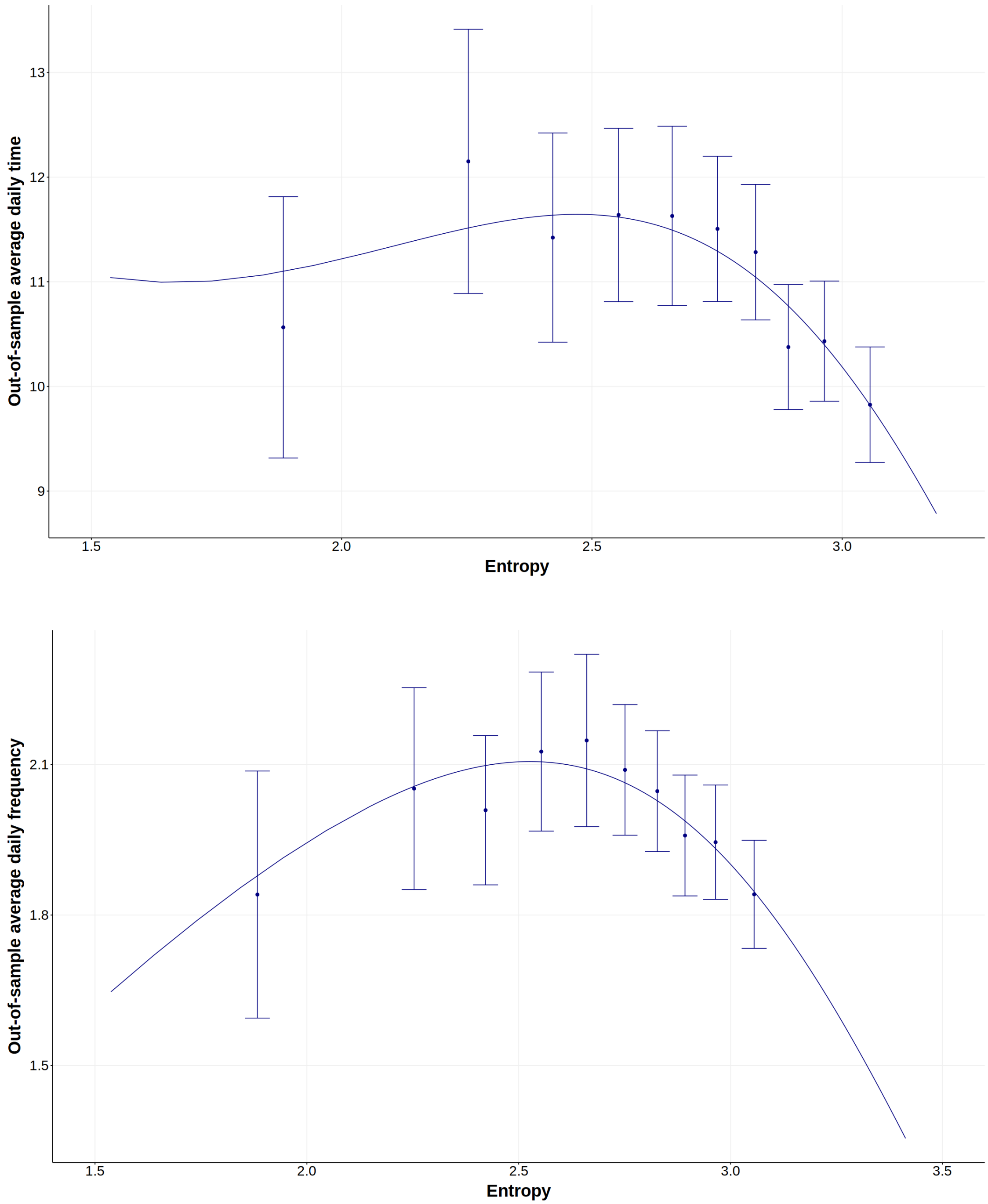


Figure 3.6: Binned scatter plots illustrating the relationship between entropy and the expected values of average daily time (top) and frequency (bottom). Each bar represents the corresponding 95% confidence interval. The blue curve represents a fourth-order polynomial fit to the point estimates.

Table 3.3: This table presents the results of the model 3.6 estimated separately for user-app pairs with entropy higher or lower than 2.6. The table demonstrates that the estimated coefficient for entropy is not significant for pairs with $E \leq 2.6$, while it is significant and nearly three times larger for pairs with entropy greater than 2.6.

	Total Time		Frequency	
	($e_{i,a} \leq 2.6$)	($e_{i,a} > 2.6$)	($e_{i,a} \leq 2.6$)	($e_{i,a} > 2.6$)
Entropy	0.0191	-0.5378***	0.0916	-0.4421***
	(0.1138)	(0.0944)	(0.0933)	(0.0926)
Clumpiness	-3.394***	-4.738***	-3.463***	-4.018***
	(0.2958)	(0.1836)	(0.2944)	(0.2100)
log(Frequency)	1.133***	0.8501***	1.102***	0.8060***
	(0.0850)	(0.0846)	(0.0779)	(0.0896)
log(Average Time)	0.7250***	0.8305***	0.0438	-0.0103
	(0.0408)	(0.0230)	(0.0351)	(0.0219)
Recency	-0.0122***	-0.0094***	-0.0137***	-0.0093**
	(0.0033)	(0.0034)	(0.0037)	(0.0041)
Days Active	0.0005	-0.0061***	-4.18×10^{-5}	-0.0070***
	(0.0039)	(0.0020)	(0.0034)	(0.0021)
Observations	526,879	823,449	526,879	823,449
Squared Correlation	0.24005	0.48318	0.29334	0.53422
Daily fixed effects	✓	✓	✓	✓
Application fixed effects	✓	✓	✓	✓
Sample size fixed effects	✓	✓	✓	✓

5.3.3 Partial Dependence and Accumulated Local Effect Plots

While our binscatter regression model with various fixed effects is capable of capturing different functional forms, it does impose a linear structure on the other set of covariates. To further examine the robustness of our analysis, we employ a fully non-parametric approach inspired by machine learning techniques. Specifically, we train a random forest model using entropy and the other control variables, similar to our previous approach. This trained model allows us to generate two types of plots: partial dependence plots (PDP) and accumulated local effect (ALE) plots (Apley and Zhu 2020, Hall *et al.* 2017).

Both PDP and ALE plots provide visual representations of the relationship between entropy and the predicted outcomes, while considering the influence of other covariates in a non-parametric manner. However, they differ in some key aspects. Partial dependence plots resemble the stratified mean of the outcome at different values of entropy, but with fine-grained control variables defining the strata. It holds all other variables at fixed values, disregarding whether the value of entropy is meaningful for all data instances. Thus, PDP may encounter potential issues when there is a high correlation between entropy and the control variables. ALE plots serve the same purpose as PDP in terms of visualizing the relationship between the feature of interest (entropy) and the predicted outcomes. However, ALE plots incorporate the correlation structure of the covariates and offer a clearer interpretation. The ALE plot is centered at zero, making it easy to interpret each point on the curve as the difference from the mean prediction. This allows us to understand the relative effect of changing the feature (entropy) on the prediction, conditional on a given value of the feature. In summary, both PDP and ALE plots provide non-parametric visualizations of the relationship between entropy and the predicted outcomes, but they differ in terms of how they handle the influence of other covariates. By utilizing these approaches, we aim to gain a deeper understanding of the relationship between entropy and the outcomes, considering both the main effect and potential interactions with other variables.

Figure 3.7 presents the partial dependence plot (PDP) (top) and accumulated local effect (ALE) plot (bottom) generated from the two random forest models trained on time and frequency. These plots provide insights into the relationship between entropy and the outcomes of interest, namely the average daily time spent and frequency of app use. The ALE plot aligns with the findings from the binscatter plots, revealing that higher entropy (indicating higher irregularity) is associated with lower levels of time spent and frequency of app use in the future. On the other hand, lower and moderate levels of entropy correspond to increased engagement with the apps, and the dependence between regularity and future use becomes weaker.

In contrast, the PDP plot displays a slightly different pattern, showing a rising trend for higher entropy values. As it was noted before, the PDP approach may not

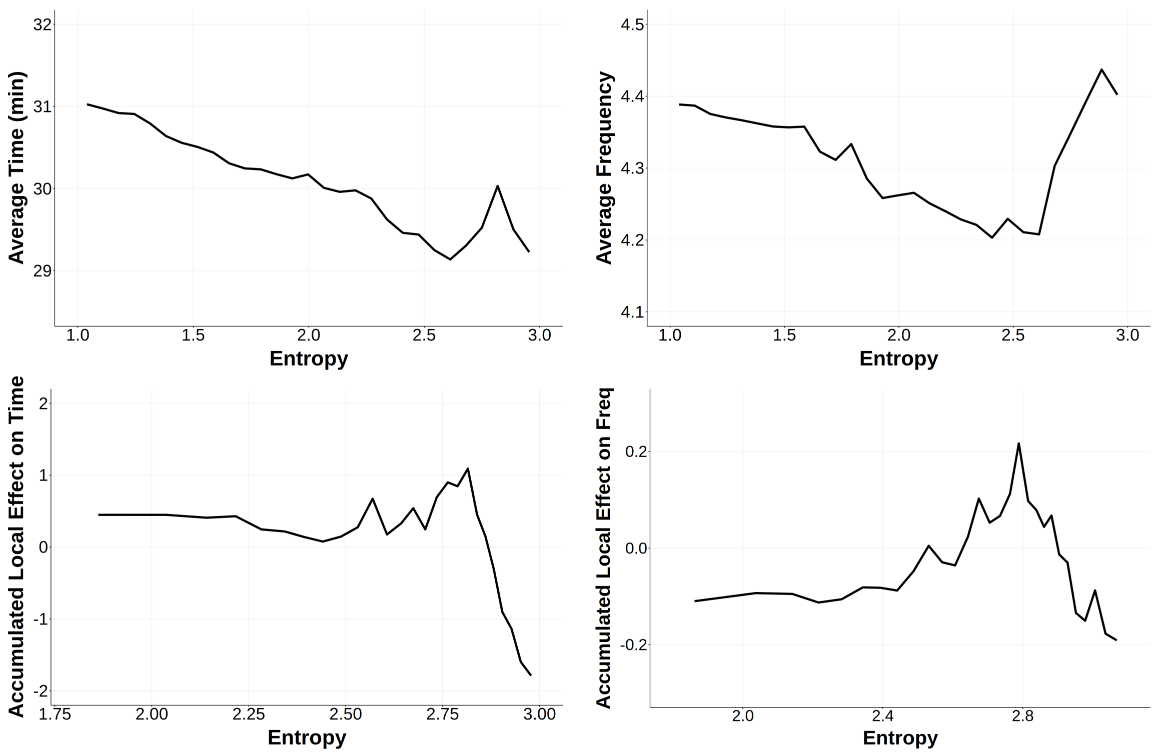


Figure 3.7: PDP (top) and ALE (bottom) plots for average time (left) and frequency (right). ALE aligns with binscatter plots, showing how higher entropy is associated with lower engagement. PDP displays a different pattern, possibly due to unaccounted correlation with covariates.

fully account for the correlation between entropy and other covariates, which could explain the discrepancy with the ALE plot. The PDP plot assumes fixed values for the other covariates, potentially disregarding their influence on the relationship between entropy and the outcomes. Considering these limitations, the ALE plot provides a more reliable representation of the relationship between entropy and the outcomes, as it properly incorporates the correlation structure of the covariates.

Disadvantages: While the random forest model offers the advantage of capturing fully non-parametric and non-linear relationships between the predictive variables, it does come with certain disadvantages compared to the regression approach. It is important to consider these limitations when interpreting the results. Firstly, one of the key drawbacks of random forest models is the inability to perform statistical inference. Unlike regression models, random forests do not provide confidence intervals or p-values for the estimated relationships. Therefore, the points in the partial dependence (PDP) and accumulated local effect (ALE) plots do not have an associated statistical measure of uncertainty. Secondly, random forest models are not readily applicable to panel data structures, where observations are correlated within individuals over time. These models do not explicitly account for the temporal dependence between observations for the same user. While there are approaches available to address this issue, such as random forest with fixed effects or mixed effects models, these methods can be more complex and computationally demanding.

To circumvent this issue, a common simplification is to aggregate the data over time, resulting in a cross-sectional dataset. However, this simplification can lead to a loss of information regarding the temporal dynamics and individual-level variations. It is important to be aware of these limitations when using random forest models and interpreting the results. While random forests offer flexibility and can capture complex relationships, they may not provide the same level of statistical inference and consideration of temporal dependencies as regression models. As a result, care should be taken to interpret the results within the context of these limitations and explore alternative approaches when necessary.

6 Quantifying Importance of Regularity for Social Media Use

In the preceding section, we have presented a multitude of evidential findings elucidating the relationship between entropy and long-term user behavior on social media platforms, thereby indicating that entropy serves as a pertinent proxy for time-of-day behavioral regularity. However, in order to gain a comprehensive and comparative understanding of the significance of entropy in relation to other predictors, it becomes imperative to quantify the relative importance of entropy versus other regularity measures such as clumpiness. This comparative analysis holds the potential to illuminate the psychological mechanisms governing user engagement and offer profound insights into the specific role played by entropy in this context. To achieve this, we employ variable importance plots as a valuable analytical tool. Variable importance plots, widely used in the realm of machine learning, offer a profound understanding of the contributions made by individual input variables (also known as features) towards the predictive performance of a machine learning model. These plots serve the purpose of evaluating the relative importance or significance of each feature in influencing the model's predictions. While multiple methods exist for assessing variable importance, the fundamental concept remains consistent across these approaches. Typically, variable importance is determined either by evaluating the impact of feature perturbations or by analyzing the feature's influence on the model's performance metrics.

In our specific context, we employ the same random forest model discussed earlier. To estimate variable importance in random forests, we adopt a common approach known as permutation importance. This method involves measuring the decrease in the model's accuracy or mean squared error (MSE) when a specific feature is randomly perturbed. The process entails randomly permuting the values of a single feature while keeping the other variables unchanged. By evaluating the resulting decrease in model performance (i.e., increase in MSE), we can assess the importance of each feature. Features that lead to a substantial decrease in model performance when perturbed are considered more important, as they exert a stronger influence on the predictions made by the model.

The variable importance plots for random forest models trained on average time (left) and frequency (right) are displayed in Figure 3.8. These plots illustrate the percentage increase in mean squared error (MSE) resulting from permuting each associated variable. Notably, the importance of entropy is evident in both plots, providing further support for the findings of previous models. However, it is worth noting that the relative importance of entropy is comparatively lower than that of other predictors. This suggests that the formation of habits driven by external contextual cues, such as time and location, may not be the primary factor influencing social media behavior. As discussed in the introduction, the triggers for habits can extend beyond external cues to include internal cues, such as mood and emotions, which might be

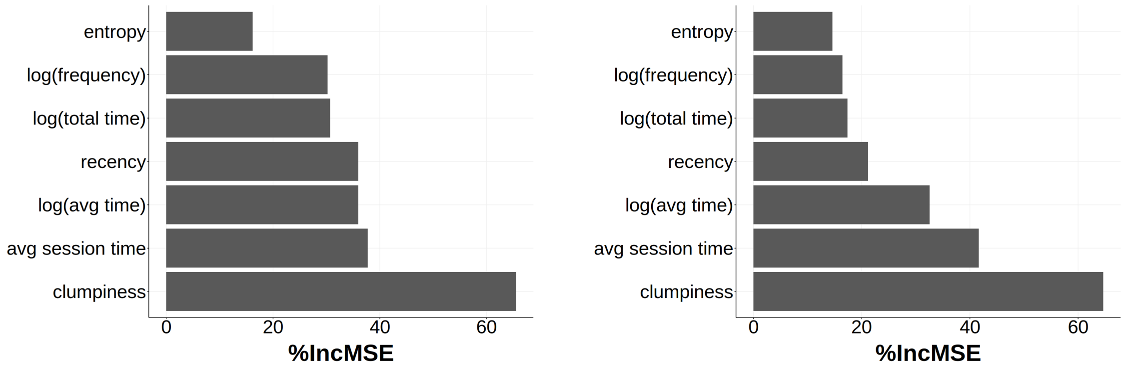


Figure 3.8: The variable importance plots for random forest models trained on average time (left) and frequency (right). These plots illustrate the percentage increase in mean squared error (MSE) resulting from permuting each associated variable.

playing a more significant role in shaping social media usage patterns. These results underscore the complexity of habitual behavior and encourage further exploration of the diverse mechanisms contributing to user engagement in the social media context.

Another intriguing finding is the significant relative importance of clumpiness. This observation indicates that the regularity of inter-session times is a substantial predictor of long-term frequency and time spent on social media platforms, even surpassing the importance of the baseline frequency and time variables themselves. This suggests that the pattern of clustering in app usage, with more evenly spaced inter-session times, plays a crucial role in shaping future user behavior. Further research is warranted to delve deeper into the specific psychological mechanisms underlying the relationship between clumpiness and future usage in the context of social media. Such investigations could provide valuable insights into the complex interplay of user behaviors and contextual factors in different settings, shedding light on the underlying drivers of user engagement and habit formation. Understanding these mechanisms can have practical implications for app design and personalized interventions aimed at fostering positive usage patterns and habit formation among users.

7 Discussion

The current chapter introduces a novel approach to assess habitual behavior in the context of social media usage, employing entropy as an implicit measure of behavioral regularity. Our findings emphasize the significance of considering the specific context in which behavior occurs when studying habit formation. The entropy metric, by capturing the regularity of behavior within a distinct context, offers a unique perspective on habit formation, aligning with psychological theories that underscore the role of stable contextual cues in triggering habitual behavior.

Through a series of comprehensive analyses, we establish the validity of entropy as a measure of habits. The correlation of entropy with the frequency-in-context measure demonstrates its convergent validity, and its distinction from mere frequency of use confirms how it captures more complex behavioral patterns. Additionally, we demonstrate the predictive validity of entropy by showcasing its ability to forecast long-term user behavior, even after accounting for other behavioral metrics. This reinforces the potential usefulness of the entropy metric in predicting social media usage patterns. Our results indicate that higher entropy values, reflecting lower time-of-day behavioral regularity, are linked to weaker habit formation, leading to reduced time spent and lower frequency of app usage in the long-term future. Remarkably, the use of binscatter regression uncovers a non-monotonic relationship between entropy and future app usage, revealing that lower entropy values exhibit a positive association with future usage, while excessively high entropy values result in a significant decline in future usage. This observation suggests that while a minimum level of regularity is necessary for habit formation, excessive irregularity impedes the development of habits. These findings contribute to a deeper understanding of habit formation and provide valuable insights into the dynamics of user behaviors in the context of social media platforms.

The findings from our analyses carry several implications, both for advancing the understanding of habitual behavior and for practical applications in the design of digital technologies and interventions. Firstly, our research contributes to the academic understanding of habitual behavior by shedding light on the pivotal role of behavioral regularity in shaping long-term engagement with social media platforms. The demonstrated association between entropy and habit formation provides valuable insights into the underlying mechanisms driving user behavior in this context. Secondly, an intriguing application of our findings is in the design of self-control apps or browser extensions. Currently, these tools primarily utilize total time or frequency as metrics to set usage limits (Grüning *et al.* 2023). However, our research suggests that incorporating the regularity of time of usage as an additional dimension may be beneficial. For example, individuals could set limits in a way that restricts app usage if they have been using it excessively around the same time of day recently. This approach could potentially disrupt the formation of problematic habitual behavior and promote more balanced usage patterns. The potential of utilizing entropy as a treatment in intervention methods presents an exciting avenue for future research. Randomized experiments could be employed to test different intervention strategies, including those that incorporate varying entropy. Such studies could assess the effectiveness of these interventions in preventing the formation of problematic usage habits and promoting healthier digital behaviors. By investigating the impact of interventions targeting behavioral regularity, researchers and practitioners can contribute to the development of evidence-based strategies for managing digital behaviors and cultivating healthier online habits.

It is essential to acknowledge that despite the robust support for using entropy as a measure of habits, there are certain limitations to this approach. While our analyses demonstrate the convergent validity of the entropy metric with the frequency-in-context measure, it is worth noting that we were unable to test its convergent validity with self-reported habit measures, such as the Self-Report Habit Index (SRHI), due to the absence of self-report data in our dataset. Therefore, future research could benefit from incorporating self-report measures to further validate the entropy metric against established habit assessment tools. Additionally, while our study highlights the distinct nature of entropy from mere frequency, a comprehensive measure of habit should also demonstrate discriminant validity from motivational constructs. Addressing this limitation could enhance the reliability and validity of the entropy metric in capturing habitual behavior accurately. Moreover, it is crucial to recognize that the entropy metric primarily captures the influence of external cues on behavior. As a result, it may not fully encapsulate the impact of internal cues, such as mood or motivation, on habit formation. Overall, these limitations serve as valuable avenues for future research to refine and expand upon our findings.

Furthermore, future research could expand the application of the entropy metric to various contexts and diverse types of applications, allowing for a more comprehensive understanding of online behavior and the underlying mechanisms driving it. Exploring how the entropy metric can be adapted to capture the influence of different contextual cues, such as location or social context, presents a promising avenue. One potential approach is to estimate the joint probability distribution of app usage across multiple contextual dimensions and subsequently compute entropy. However, it is essential to acknowledge that Kernel Density Estimation may encounter challenges in higher dimensions. In such cases, alternative methods like clustering could be explored to handle the increased complexity. By generalizing entropy to multiple context parameters, researchers can gain deeper insights into the factors shaping internet usage habits and contribute to a better understanding of digital routines and behavioral patterns.

Chapter 4

Predicting and Shaping Climate Change Attitudes

1 Preface

This fourth chapter delves into the complex relationship between individual stances on climate change and the capacity to comprehend and anticipate the mental states of others. Employing Theory of Mind and persuasion dynamics as theoretical guides, this chapter explores how personal viewpoints might affect an individual's accuracy in predicting others' shifts in opinion after exposure to relevant news articles. The core objective is to discern whether climate change deniers or believers exhibit varying levels of accuracy in predicting the persuasive impact of news articles emphasizing the gravity of climate change[1]. Our hypothesis posits that those sharing similar beliefs may exhibit heightened predictive accuracy, potentially stemming from an empathetic understanding of peers' perspectives and an ability to simulate likely reactions to new information. Through a well-constructed survey and randomized experiment, we test these hypotheses to uncover the intricate dynamics of persuasion and understanding in the context of climate change discourse.

[1]Throughout this chapter we use the terms climate denier and believer based on the classification explained in section 3.3. We do not assign a value or any particular connotation to either stance.

2 Introduction

We live an age of unparalleled digital revolution that has radically transformed the ways in which we create, share, and interact with information (Bennett and Iyengar 2008). Due to the emergence of digital platforms, particularly social media, news and information are no longer solely disseminated through traditional outlets like print media or broadcasting networks. These innovative media forms have introduced a more interactive, personalized, and pervasive mode of information dissemination, fostering an era marked by engagement, participation, and user-generated content (Bakshy *et al.* 2012). In this evolving participatory culture, individuals are not mere consumers of information, but also active producers and distributors. This empowers them to influence, persuade, and shape public opinion on a mass scale (Conover *et al.* 2011).

Understanding the dynamics of information sharing in the digital era is crucial for unraveling the complex mechanisms underlying belief and attitude formation, and its societal implications such as political polarization. Scholars concur that political polarization has been on the rise since the 1970s (Boxell *et al.* 2022, McCarty *et al.* 2003), although the causes of this trend continue to be a subject of debate (Winkler 2019). The Internet and social networks have frequently been criticized for promoting either voluntary or algorithmic readership that aligns with existing attitudes (Adamic and Glance 2005, Kubin and von Sikorski 2021, Pariser 2011).

In the act of disseminating news, users intrinsically depend on a mental model of how others might react to the given information. A growing body of neuroscientific research substantiates the notion that communicators' perceptions about the mental state of receivers influence the value they attribute to information sharing (Baek *et al.* 2017). Moreover, psychological studies have underscored that adopting the perspective of others can bolster communication effectiveness (Traxler and Gernsbacher 1993). In the same vein, neural studies have indicated that successful persuaders tend to engage brain regions associated with understanding others' minds more effectively than unsuccessful persuaders (Dietvorst *et al.* 2009, Falk *et al.* 2013). For an in-depth review on this topic, refer to Falk and Scholz (2018). This cognitive ability to ascribe mental states – such as beliefs, desires, and intentions – to others is a key aspect of human social cognition, commonly referred to as Theory of Mind (ToM) or mentalizing (Happé *et al.* 2017). A well-developed ToM enables individuals to empathize with others and engage in effective communication by considering their perspectives and mental states (Baron-Cohen 1997). Consequently, ToM offers a valuable insights for understanding how individuals share, interpret, and respond to information on social media platforms, as well as its implications for political persuasion.

Inspired by the theoretical underpinnings of Theory of Mind (ToM), this chapter aims to systematically explore the complex interplay between an individual's stance on a politically contentious issue and their capacity to anticipate and comprehend the cognitive processes of others. The central impetus for investigating this matter stems from the notion that a fundamental incentive for disseminating information to those who hold divergent views lies in the desire to sway their perspective, with the aspiration of converging it towards one's own (Hsu *et al.* 2021). However, a challenge arises as personal convictions might inadvertently impair one's ability to adequately understand the perspectives of others. This might precipitate a failure in identifying the appropriate evidence or arguments that would effectively influence the counterpart, especially when there are deep-seated disagreements on critical issues. In this study, we take a novel approach and address this problem through the lens of prediction accuracy of opinion shift. We specifically measure how variations in perspectives might impact an individual's capacity to accurately predict the shifts in others' stances following their exposure to pertinent news articles.

The subject chosen for this investigation is climate change, a highly polarized and pivotal topic of our era with the potential for profound impacts on both human society and the natural world (Dunlap *et al.* 2016). Despite an overwhelming scientific consensus acknowledging the reality and severity of climate change, there persists significant political polarization surrounding the issue. For instance, a survey conducted by Leiserowitz *et al.* (2020) revealed that 83% of Democrats, 56% of Independents, and a mere 22% of Republicans reported that global warming should be a high priority for the President and Congress. Such polarization frequently stems from divergent ideologies, disparate economic interests, and conflicting priorities among various stakeholders (McCright *et al.* 2016, Wong-Parodi and Feygina 2020). Regrettably, this polarization can obstruct efforts to address climate change by causing delays or even outright prevention of essential policy changes. Therefore, comprehending the intricate mechanisms that foster this increasing polarization is of critical importance. By exploring these dynamics, we aim to shed light on the factors that obstruct or facilitate consensus-building around climate change, thereby contributing valuable insights towards the development of more effective strategies for information dissemination and persuasion on this vital issue.

2.1 Related Literature

Our research is deeply intertwined with the extensive and diverse literature on political persuasion. This field encompasses a myriad of theories and methodologies, all centered around the concept of persuasion - the act of altering others' attitudes or behaviors through the strategic dissemination of information (Perloff 2020). Historically, scholars have presented differing viewpoints on the implications of persuasion resulting from communication by motivated agents (DellaVigna and Gentzkow 2010). Some have portrayed persuasion as a predominantly negative force, suggesting that citizens and consumers are easily swayed by those wielding political or economic power (Lippmann 1965, Robinson 1933). This perspective paints a picture of a society vulnerable to manipulation, where power dynamics heavily influence the flow and impact of information. Conversely, other scholars adopt a more optimistic stance, viewing even motivated communications as a form of information dissemination that can ultimately enhance efficiency (Bernays 1928, Stigler 1961). From this viewpoint, persuasion is seen as a tool that, when used responsibly, can foster informed decision-making and contribute to the efficient functioning of society.

The study of persuasion is indeed multidisciplinary, with various approaches drawn from fields like economics, political science, psychology, and communications. Economic theories of persuasion often incorporate aspects of strategic communication and information disclosure, where an informed sender aims to influence the actions of a less-informed receiver. In these models, persuasion is viewed as a game-theoretic problem that centers on the sender's ability to optimally design the disclosure of information to maximize their expected utility (Kamenica and Gentzkow 2011). In the realm of political science, persuasion is frequently examined in the context of mass communication, public opinion, and political behavior. Research such as the study conducted by Druckman and Nelson (2003) delves into the pivotal role of framing and priming in molding public attitudes. This investigation emphasizes the integral part played by media and political campaigns in directing public sentiment. Conversely, the discipline of psychology scrutinizes persuasion through a lens centered on the individual, accentuating cognitive processes and behavioural responses. A paramount model in this regard is the Elaboration Likelihood Model, as proposed by Petty *et al.* (1986). This model delineates central and peripheral routes to persuasion, underscoring the impact of variables such as the quality of the message, the credibility of the source, and the personal relevance of the information. In a similar vein, recent research by **?** demonstrates that sharing similar non-political features with the source of a message can facilitate persuasion on political topics.

Within the domain of climate change, the task of persuading individuals, while plausible, presents significant challenges and preceding scholarship demonstrates relative stability in public opinion. In a review of empirical studies and polling data related to American public sentiment towards climate change, Egan and Mullin (2017) uncover a conspicuous absence of considerable long-term fluctuations in mass opinion. In their extensive meta-analysis of experimental studies on climate change persuasion, Rode *et al.* (2021) report that interventions on average yielded a small yet significant positive influence on attitudes. Nonetheless, the difference between treatment and control groups in terms of attitudes was observed to be less pronounced for policy support as compared to belief in climate change. This suggests a higher degree of resistance to influence in the realm of policy attitudes than in beliefs concerning climate change. Consequently, interventions must be exceptionally persuasive and thoughtfully tailored to align with an audience's values to augment supportive attitudes towards climate change and ultimately sway policy support (Druckman and McGrath 2019). Furthermore, Rode *et al.* (2021) present evidence for an asymmetry between positive and negative interventions. A generic message downplaying the effects of climate change may sufficiently instigate skepticism, whereas a generic message underscoring the pervasive effects of climate change may not significantly enhance positive attitudes. Such a phenomenon could potentially be attributed to individual incentives to disbelieve in climate change, given the potential constraints and costs such beliefs could impose, including limitations on transportation modes, energy consumption, and the burden of various energy taxes.

At the conceptual level, our study intersects with several psychological concepts deeply embedded within the persuasion literature. One such concept is the cognitive bias commonly referred to as the "curse of knowledge." This bias, particularly relevant in the context of complex issues such as climate change, significantly impacts the effectiveness of communication and persuasion. The curse of knowledge manifests when individuals possessing extensive knowledge on a subject grapple with considering the perspective of those less informed (Camerer *et al.* 1989). This struggle can lead to potential miscommunication and misunderstanding. For instance, a study by Birch and Bloom (2007) demonstrated that adults' knowledge of an event's outcome could compromise their ability to reason about another person's false beliefs regarding that event. In the realm of climate change, this phenomenon could potentially impede effective communication between experts and deniers. Experts, well-versed in the scientific evidence and implications of climate change, may unconsciously assume that deniers share the same level of understanding or acceptance of the science. This assumption could lead to the creation of messages that fail to resonate with deniers, or worse, inadvertently intensify polarization. Moreover, the curse of knowledge can lead to a miscalibration of explanatory insight, where an abundance of knowledge can foster overconfidence in one's ability to explain concepts to others (Fisher and Keil 2016). This could further complicate communication between climate change believers and deniers, as believers may overestimate their ability to articulate the science and implications of climate change in a manner that is comprehensible and persuasive to deniers.

A notable aspect of the curse of knowledge phenomenon is the disparity in expertise levels. However, a similar phenomenon could occur due to differing viewpoints, independent of expertise levels. It is important to note that individuals' opinions are often shaped more by factors such as partisan polarization, misinformation, and elite influence, rather than a thorough examination of scientific evidence (Druckman *et al.* 2013, Zollo and Quattrociocchi 2018). Furthermore, in certain ethical scenarios, attributing right and wrong to either side of a disagreement can be challenging. Our study represents a departure from the conventional focus of the "curse of knowledge" literature, which typically centers on the communication dynamics between experts and non-experts. Instead, we turn our attention to ordinary individuals who are not necessarily experts. Our aim is to explore how these individuals, with their diverse perspectives and varying levels of understanding, navigate the challenge of comprehending the viewpoint of those on the opposite side of a contentious issue such as climate change.

Another psychological concept that intersects with our study is the "framing effect," a phenomenon extensively studied for its impact on the effectiveness of persuasive messages (Chong and Druckman 2007). Framing refers to the process where minor alterations in the presentation of an issue or event can elicit substantial shifts in opinion. For instance, a study by Sniderman and Theriault (2004) vividly demonstrated this effect. Participants were asked whether they would favor or oppose allowing a hate group to hold a political rally. The researchers found that when the question was prefaced with the phrase, "Given the importance of free speech," 85% of participants expressed favor. In contrast, when the question was introduced with, "Given the risk of violence," favorability dropped to 45%. In the context of climate change, the framing effect emerges as a potent tool. A study by Fielding *et al.* (2020) underscored how tailoring interventions to align with audience values could bolster their effectiveness. The researchers presented messages about a carbon tax policy to Republican and Democrat participants, varying the promotion of the policy based on Republican or Democrat values. The results revealed that participants were more inclined to engage in policy-supportive behavior when the climate change policy was endorsed by members of their own political group, rather than the opposing group. The framing effect has been harnessed in numerous studies aiming to augment engagement with climate change by reframing it as a moral issue. Notably, reframing climate change in terms of conservative morality appears to be an effective strategy for engaging conservatives in climate change discourse (Feinberg and Willer 2013, Wolsko *et al.* 2016).

The final crucial aspect that complements this investigation is the concept of motivated beliefs and reasoning, which asserts that beliefs often fulfill essential psychological needs beyond mere information provision for decision-making (Kunda 1990). According to this perspective, people's preferences, goals, and motivations subtly guide the stages of belief formation, from initial evidence interpretation to the construction of biased yet subjectively unbiased beliefs (Epley and Gilovich 2016). Consequently, as individual preferences play a pivotal role in shaping held beliefs, a lack of comprehension or an incorrect assessment of a belief's value could impede attempts to persuade individuals toward alternative viewpoints. Empirical support for motivated reasoning in the climate change context has also emerged (Hart and Nisbet 2012), often fueled by diverse influences. One significant factor is the inclination for individuals to align their climate change opinions with those of like-minded individuals within their political party or ideological sphere (Palm *et al.* 2017). A concise overview of literature linking motivated beliefs and reasoning to climate change can be found in Rode *et al.* (2021).

These findings underscore the importance of understanding the audience's perspective and values in crafting persuasive messages, a theme central to our investigation. If climate change believers fail to empathize with and understand the perspective of climate change deniers, they may not only fall short in creating effective persuasive messages but could also inadvertently exacerbate polarization. Indeed, persuasion attempts sometimes lead to a backfire or boomerang effect which can exacerbate political polarization, a burgeoning issue in contemporary society. For instance, Hart and Nisbet (2012) demonstrated this phenomenon in a study where participants were exposed to news stories about potential health impacts of climate change on different groups. The results revealed that exposure to these messages significantly increased support for climate mitigation policies among Democrats compared to a control group. However, for Republicans, message exposure significantly decreased support for such policies. Consequently, there was a significant increase in opinion polarization between Democrats and Republicans regarding climate mitigation following exposure to messages designed to highlight the health risks of climate change. At the platform level, prior research offers some evidence that increasing exposure to content from ideologically opposing sources can sometimes backfire, leading to an intensification of polarization (Bail *et al.* 2018). These findings illuminate the potential pitfalls of persuasion attempts that do not adequately consider the perspectives and beliefs of the target audience. They underscore the necessity for a nuanced understanding of audience perspectives in crafting messages. Such messages should not only inform but also resonate with the audience's existing beliefs and values. Our research aims to shed light on this phenomenon by investigating how differences in stance can lead individuals to select inappropriate information for persuading those on the opposing side, potentially driving them further away from their viewpoint. This exploration is crucial in our endeavor to navigate the complex dynamics of persuasion in the context of climate change discourse.

2.2 Overview

The primary objective of this chapter is to investigate the potential impact of an individual's stance on a contentious political issue, such as climate change, on their ability to predict the persuasive impact of news on others. Within the context of climate change discourse, this study delves into the question of whether individuals who identify as climate change believers and deniers display differential levels of accuracy when forecasting the persuasive outcomes of news articles that accentuate the importance of climate change matters. Drawing inspiration from the theory of mind literature, a central hypothesis to be examined is whether individuals who share congruent beliefs possess an enhanced ability for accurate prediction due to their heightened capacity for perspective-taking. This cognitive faculty is posited to stem from their empathetic grasp of the perspectives held by their peers, particularly those belonging to the same stance (i.e., climate deniers), coupled with their proficiency in simulating potential reactions to novel information. Conversely, it is conjectured that individuals with opposing beliefs (i.e., climate believers) might encounter difficulties in adequately apprehending the anticipated responses of others, thereby potentially impeding their predictive accuracy. Furthermore, this inquiry also contemplates an alternative scenario wherein the biases inherent in deniers' viewpoints may hinder their objective assessment of an article's influence, while the comprehensive knowledge possessed by believers regarding the topic equips them with better predictive capabilities.

To empirically investigate this phenomenon, a comprehensive research approach encompassing two surveys and a randomized experiment was undertaken. The initial survey was meticulously designed to curate a collection of news articles predominantly emphasizing the significance of climate change and the urgency for collective action. The methodology commenced by extracting a sample of 1351 authentic news articles addressing climate change. Subsequent refinement involved the inclusion criterion of articles featuring "climate change" or "global warming" in either the first three sentences or the title, a measure intended to mitigate the inclusion of extraneous and unrelated content. Subsequently, 798 participants were recruited from the Prolific platform, comprising an evenly distributed representation of 402 Democrats and 396 Republicans. The participants' task involved meticulously reading and categorizing the articles. The assessment encompassed evaluating the article's relevance to climate change, determining the article's stance on the issue, and gauging the participants' personal stance on the subject. This survey yielded 550 articles, with the majority of evaluators concurring that the prevailing theme of these articles revolved around climate change. An essential component of this survey involved gathering participants' predictions on how a hypothetical reader, with attitudes ranging from supportive to indifferent and opposing, might alter their opinions after reading a specific news article. This aspect aimed to gauge the participants' expectations of persuasive impact on others. However, it is important to note that the survey's structure introduced limitations, stemming from variations in assessments among labelers appraising the same articles. This divergence emerged due to the stipulation that respondents identifying the central theme as climate change-related would proceed to the final question, inadvertently leading to an uneven distribution of labels across the articles.

To address the previously noted limitation and to further enhance the accuracy of estimated predicted opinion shifts, a subsequent follow-up survey was meticulously designed. This survey utilized a carefully curated subset of articles, chosen specifically based on unanimous labeling as climate change related and possessing persuasive qualities. A deliberate endeavor was made to achieve a more equitable distribution of respondents across various stance groups, a departure from the naturally skewed distribution observed in the first survey on the Prolific platform. This strategic adjustment was pivotal as it sought to ensure equitable precision in estimating opinion shifts across divergent stance groups. Given the skewed representation towards climate change believers, a systematic over-sampling of climate change deniers was undertaken in an iterative manner to achieve this desired balance. In this follow-up survey, participants encountered the same fundamental survey question concerning the prediction of persuasive impact stemming from news articles. Notably, the survey results illuminated a marked incongruity between the expectations of individuals who believe in climate change and those of climate change deniers. Specifically, individuals who hold the belief in climate change did not envisage a substantial capacity to persuade climate change deniers through the selected news articles. On the contrary, climate change deniers exhibited an alternative perspective, anticipating a potential backfire effect. According to this anticipation, fellow deniers were envisioned to become even more resolutely opposed to climate change policies following exposure to articles accentuating the gravity of the climate change phenomenon.

To rigorously examine these divergent predictions, a meticulously designed randomized experiment was executed, aimed at empirically probing the impact of news articles on participants' viewpoints. Specifically, a cohort comprising climate change deniers was randomly assigned to read the curated selection of articles, affording a controlled framework for gauging the extent of opinion shift following exposure. The experimental outcomes unveiled a noteworthy and statistically significant positive shift in the convictions of climate change deniers with respect to the gravity of climate change. This pivotal discovery underscores the potency of the chosen stimuli in eliciting a substantive shift in perspective, thereby underscoring the potential flexibility of deniers' viewpoints. Nevertheless, in contrast, the experiment yielded null effects when assessing measures related to policy support, personal actions, and contributions to environmental causes. This intriguing pattern suggests that the endeavor to reshape deeply entrenched policy stances and induce tangible behavioral transformations might be more resistant to the mechanisms of persuasion. These findings harmoniously align with antecedent research on climate change persuasion, corroborating the prevailing notion that instigating changes in attitudes is a multifaceted and intricate process, shaped by an array of factors beyond mere exposure to persuasive discourse. Nonetheless, the discerned positive persuasion effect pertaining to belief alteration is promising, as shifts in the perception of climate change's severity represents the critical inaugural stride towards engendering subsequent modifications in policy endorsement and individual conduct.

A noteworthy outcome of this study is the direct contradiction it presents to the hypothesis that climate change deniers might exhibit a heightened ability to predict opinion shifts within their own ranks, underpinned by an elevated capacity for perspective-taking arising from shared beliefs. Additionally, our findings highlight an intriguing trend whereby climate change believers tend to underestimate their potential for persuading deniers. Our study thus imparts significant insights into the intricate interplay among individual beliefs, theory of mind, and the intricate mechanisms of persuasion, all within the domain of climate change discourse. As society confronts the pressing imperatives of environmental challenges, these insights hold the potential to shape the design of targeted communication strategies and interventions, with the overarching goal of fostering consensus and catalyzing meaningful actions in response to the exigent issue of climate change.

3 Phase one: Broad Survey

The primary intent of this survey was to categorize a specific collection of articles centered on climate change and global warming. To this end, our starting point was a set of 3.6 million authentic news articles, sourced from Factiva, a leading news provider[2]. To guarantee the pertinence of the chosen articles, we instituted specific criteria: an article must include the terms "climate change" or "global warming" within its title or initial three sentences (Barberá *et al.* 2021). This procedure was aimed at minimizing the incorporation of articles not directly pertaining to the designated subject. In addition, we confined our sample to articles with fewer than 500 words to mirror the type of content commonly disseminated on platforms like Twitter, known for their character constraints. In this era of dwindling attention spans, individuals are less likely to interact with extensive content, especially if it clashes with their pre-existing beliefs. This criterion considered not only the preferences of digital users but also aimed to reduce potential fatigue and distraction during the labeling task. After applying these filters, we were left with a corpus of 1,351 news articles spanning the period from 1990 to 2020.

3.1 Survey Design

In the initial stages of the survey, respondents were queried on a range of demographic factors, such as age, educational attainment, and gender, enabling a detailed understanding of their respective backgrounds. This was followed by inquiries concerning their political orientation. Next, we sought to discern their viewpoints on various aspects of climate change through a series of targeted questions, assessing their concern level, acceptance of anthropogenic climate change, and confidence in mainstream

[2]Factiva: https://professional.dowjones.com/factiva/

climate science.

In order to validate respondents' proficiency in comprehending and critically evaluating news articles, two comprehension check questions were incorporated. Despite all participants residing in the United States and having English as their mother tongue, their abilities to accurately discern the main theme of a news article may differ, or some may provide responses randomly. Consequently, two concise articles were presented: one directly addressing climate change issues in the United States, the other unrelated to the subject matter. Respondents were required to successfully pass this comprehension check before proceeding to assess a randomly selected set of 12 articles for which pertinent labels were collected. Within this set, two articles (the 2nd and the 11th) bore ground truth labels and functioned as attention checks, facilitating the identification of respondents whose accuracy may diminish over the course of the survey. Notably, we intentionally withheld the source of these articles from the respondents, displaying only the title and publication date. This design decision aimed to eliminate the influence of source credibility, thus enabling a focus solely on content.

For each of the 12 articles, respondents were first tasked with determining whether the article predominantly addressed climate change. If the response was negative, they were directed to the subsequent article. However, if the response was affirmative, additional labels were gathered concerning the article's stance on the gravity of climate change, its argumentative style (e.g., reliance on empirical data and/or statistics or anecdotal narratives), and whether the information was perceived as deceptive or misleading. Lastly, respondents were asked a trio of questions related to persuasion prediction. Specifically, they were asked to project their expectations regarding how a hypothetical reader with supportive, indifferent, or opposing attitudes towards climate change policies might alter their viewpoint post-reading the respective news article. The comprehensive set of survey questions is detailed in Appendix C.

3.2 Respondents

For the task of article labeling, we recruited a total of 798 respondents via the Prolific platform. The sample comprised 402 individuals identifying as liberals and 396 as conservatives. After removing respondents who did not pass the comprehension or attention check questions[3], the final tally stood at 777 respondents, including 385 liberals and 392 conservatives. Table 4.1 outlines the demographic traits of the participants engaged in the initial survey, encapsulating age, education, gender, and self-identification as either liberal or conservative. It is crucial to acknowledge that the demographic composition of the participants was not intended to mirror the wider U.S. population. Rather, the goal was to maintain a fairly balanced representation of users across diverse political spectra, given the potent correlation between political affiliations and attitudes toward climate change. To accommodate a wide array of perspectives and curtail potential bias, each article was assessed by approximately four respondents, comprising an equal distribution of liberals and conservatives. This methodology aimed to account for the potential influence of the participants' political inclinations on their perception and interpretation of the articles. In the end, a total of 6,893 labels were collected for the articles.

3.3 Respondents' Climate Stance

A pivotal subsequent step involves quantifying the labelers' stances on the severity of climate change and categorizing them accordingly. This is of paramount importance given our project's focus on cross-stance comparisons in persuasion prediction. As elucidated earlier in the survey design section, participants were tasked with predicting the potential shift in opinion of a hypothetical reader with supportive, indifferent, or opposing attitudes towards climate change policies. To ensure the relevance of these comparisons, we also categorized respondents' stances into three distinct groups. Initially, we employed questions adapted from a study by Sunstein *et al.* (2016) to construct an index - the Sunstein score - by aggregating responses to the following three

[3] Overall, 9 respondents failed the comprehension check questions.

Table 4.1: Characteristics of respondents involved in the first survey. The liberal vs conservative dimension is rated on a scale from 1 (Strongly Liberal) to 7 (Strongly Conservative).

Characteristics	**N = 777**
Age	37 (28, 52)
Education	
College	398 (51%)
Elementary	3 (0.4%)
Grad School	131 (17%)
High-School	245 (32%)
Gender	
Female	404 (52%)
Male	357 (46%)
Non-binary	16 (2.1%)
Liberal/Conservative	
1	179 (23%)
2	155 (20%)
3	50 (6.4%)
4	21 (2.7%)
5	120 (15%)
6	147 (19%)
7	105 (14%)

[1] Median (IQR); n (%)

questions probing participants' views on climate change and environmental issues:

- Q_1. I consider myself an environmentalist.
- Q_2. I believe that man-made climate change is occurring.
- Q_3. The United States was right to rejoin the Paris Agreement in 2021 to reduce greenhouse gas emissions.

Each question was rated on a five-point scale, resulting in the Sunstein score spanning from 3 to 15. Subsequently, based on their Sunstein score, individuals were classified into three groups. Those scoring below 9 were identified as *Deniers*, while those with scores exceeding 11 were classified as *Believers*[4]. Respondents with scores ranging between these two extremes were designated as *Neutrals*[5]. The rationale for these thresholds is derived from the structure of the index. It's important to emphasize that responses to the survey questions were collected on a 5-point scale: [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree]. deniers are characterized by an average response score of less than 3, whereas believers have an average response score of 4 or higher. Consequently, when posed with a general question about the gravity of climate change, deniers are typically within the "strongly disagree" to "disagree" range, while believers are within the "agree" to "strongly agree" range. Figure 4.1 illustrates the distribution of respondents' Sunstein scores. It is evident that deniers are relatively fewer in number, despite our endeavor to balance political affiliation. This disparity was the primary driver for conducting a second survey, details of which will be discussed in the ensuing sections.

3.4 Limitations

As previously elucidated, the primary objective of this survey was to select a suitable set of stimuli for the final experiment, wherein we assess the precision of opinion shift predictions. However, the structure of the survey presented certain limitations, particularly concerning disagreement among labelers who assessed the same article. As detailed in Section 3.1, respondents were required to perceive the main topic of an article as climate change related to proceed with the survey. This requirement was put in place to ensure the relevance of inquiring about potential persuasion effects consequent to exposure to the news. As a result, for the first question concerning whether an article is "about climate change," every article was evaluated by at least four individuals, comprising two conservatives and two liberals. However, the primary challenge emerged for questions succeeding the "about climate change" question, including the pivotal persuasion prediction query. For instance, if the two conservatives

[4]Throughout the chapter, we use the terms "deniers" and "believers" solely based on this classification. We do not assign any particular connotation or value judgment to either stance.

[5]We employ the terms denier, neutral, and believer to distinguish this classification from the prediction questions where respondents are asked about individuals with supportive, indifferent, or opposing attitudes towards climate change policies.

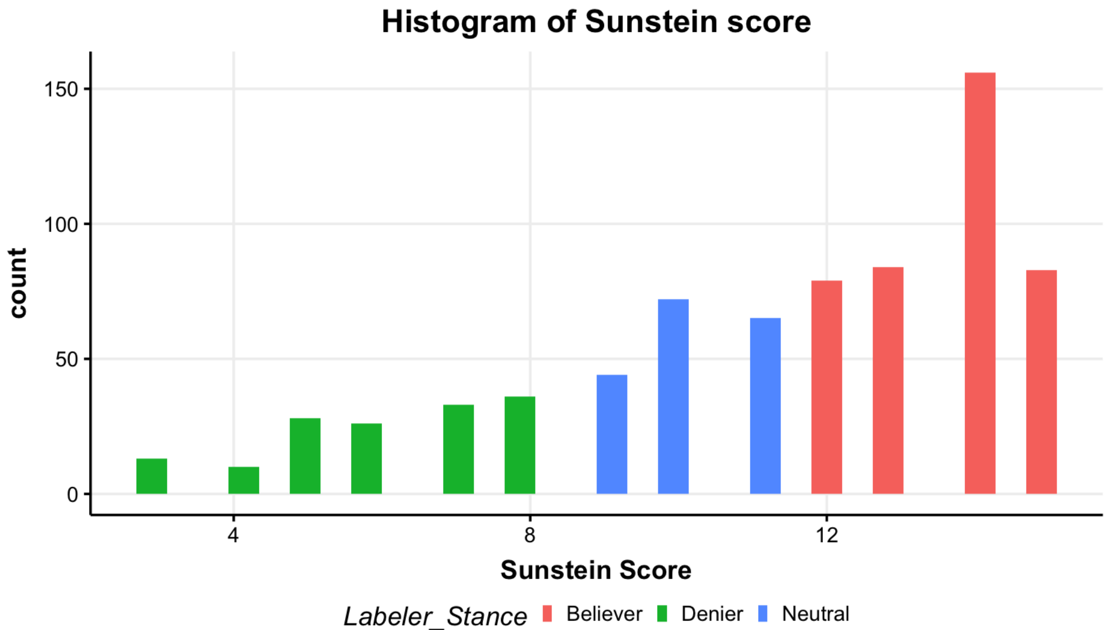


Figure 4.1: Distribution of respondents' Sunstein scores. The Sunstein score, calculated based on responses to three questions regarding people's opinions on climate change and the environment, is used to classify individuals into deniers, believers, and neutrals. deniers exhibit a score below 9, believers have a score higher than 11, while those in between are classified as neutrals.

who evaluated article A deemed it unrelated to climate change, they did not proceed to answer the concluding survey questions. As a result, we were left with only two labels from liberals. This situation led to a reduction in precision and an imbalance in the stances of respondents. Furthermore, given the overall underrepresentation of deniers, several articles remained unlabeled by any deniers.

4 Phase 2: In-depth Classification of News Articles

To surmount the constraints of the initial survey, a subsequent survey was conducted. The aim of this second survey was to collect additional labels, enabling more precise estimates and ensuring full balance across various stance groups. By collecting more data through the second survey, we aimed to enhance the reliability and accuracy of stimuli selection for the final experiment. These additional labels would offer a more comprehensive and balanced perspective, thereby facilitating more informed decisions concerning the stimuli chosen for the final experiment.

Initially, to ensure relevance, we constrained the set of articles to a subset of 550, wherein the majority of labelers concurred that the main subject pertained to climate change. Following this, we leveraged the labels acquired from the first survey to judiciously select a set of 60 articles for the second phase of data collection. The opinion shift survey queries presented to participants in the first survey adhered to a specific format such as: "Reader 1 believes that climate change is not a problem and is opposing national and international actions to combat climate change. After reading the article, Reader 1 will...". Participants were then given three response options that encapsulated the potential direction of opinion shift. These options were: becoming less favorable, remaining unchanged, or becoming more favorable towards actions intended to combat climate change. Similar questions were posed regarding hypothetical individuals who were indifferent or opposed to climate change. The comprehensive set of questions can be found in Appendix C.

The responses from participants were encoded as (-1, 0, +1), and the average predicted persuasion score was computed at the article level by averaging these scores. Based on these average scores, the articles were ranked to identify the ones predicted to be the most persuasive for each stance group. Given that conceptually persuasion could only occur among those who are against or indifferent to climate policies, as the opinions of supporters could only experience reinforcement as opposed to persuasion, we aimed to include 30 articles for each of these two groups. To counteract biases originating from labelers' stances, each set of 30 articles consisted of the top 10 persuasive articles chosen by each stance group. Consequently, we had a total of 60 stimuli that met the following criteria:

- 10 articles predicted to be most persuasive for individuals *opposing* climate policies, as determined by *believers*.
- 10 articles predicted to be most persuasive for individuals *indifferent* to climate policies, as determined by *believers*.

- 10 articles predicted to be most persuasive for individuals *opposing* climate policies, as determined by *neutrals*.
- 10 articles predicted to be most persuasive for individuals *indifferent* to climate policies, as determined by *neutrals*.
- 10 articles predicted to be most persuasive for individuals *opposing* climate policies, as determined by *deniers*.
- 10 articles predicted to be most persuasive for individuals *indifferent* to climate policies, as determined by *deniers*.

Practically, the process of selecting the most persuasive articles necessitated consideration of overlapping articles and the need for balance within the stimuli set. For instance, when it came to believers' selected articles, the top 10 persuasive articles for those who opposed climate policies and those who were indifferent were not mutually exclusive, leading to overlaps. To counteract this, we utilized a modified approach. Initially, the sets of articles chosen by each labeler stance for both categories of target individuals (those against or indifferent to climate policies) were combined. Subsequently, taking into account the persuasion rankings of the articles, we adjusted the ranking threshold to include a total of 20 articles. This approach was designed to ensure balanced representation across different labeler stances and foster diversity within the experiment's stimuli set.

4.1 Survey Design

The main structure of the second survey remained largely consistent with that of the first survey, albeit with several enhancements to refine the data collection process. To start, in an effort to bolster survey credibility and mitigate participant fatigue, each respondent was assigned to review only three articles. The selection of articles provided balanced representation from the different stance groups, with one article selected by believers, one by deniers, and one by neutrals. The sequence in which the articles were presented was randomized to counteract any potential order bias. A significant modification was made to the final persuasion prediction question to obtain more nuanced data and bolster statistical power. Rather than providing discrete options, we introduced a continuous scale using a slider that ranged from -10 to 10 (Broockman *et al.* 2017). Verbal descriptors indicating less favorable or more favorable positions were provided at the extremes, with the midpoint set at zero, to represent an unchanged stance.

4.2 Respondents

For the article labeling process in the second survey, a total of 794 respondents were recruited via the Prolific platform. One significant divergence from the first survey was the emphasis placed on achieving a nearly balanced distribution of respondents across different stance groups. This breakdown comprised 308 believers, 259 deniers, and 227 neutrals. Given the lower prevalence of deniers in the general population, we intentionally over-sampled individuals who identified as Republican to ensure a more balanced representation. In an effort to maintain balance at the article label level, our aim was to gather approximately 10 labels from each stance group for each article. Figure 4.2 demonstrates the distribution of labels assigned to articles across the different stance groups. As evident in the figure, slight variation exists in the number of labels allocated to different articles due to the iterative labeling process that was conducted in separate batches to ensure adequate participation from deniers and neutrals. Despite these minor variations, nearly all articles received the minimum target of 10 labels per stance group. This balance in stance representation led to improved prediction accuracy compared to the results of the first survey.

4.3 Respondents' Climate Stance

The process of classifying respondents' stance on climate change in the second survey followed a methodology similar to that employed in the first survey. However, in an attempt to encompass additional facets of belief regarding the scientific consensus and concern about climate change, we expanded the set of questions which were originally based on Sunstein *et al.* (2016). Two more questions were included in the survey, to which respondents provided their answers on a 5-point Likert scale. The additional questions were as follows:

- Q_4. In general, how much do you trust the science on global warming?
- Q_5. How worried are you about climate change on a scale from 1 to 5?

The categorization into the three stance groups, namely deniers, believers, and neutrals, retained the rationale used in the previous study. Deniers were characterized by an average response score of less than 3, indicating a lesser degree of trust and concern with respect to climate change. In contrast, believers exhibit an average response score of 4 or higher, signifying a greater level of trust and concern. Given that this survey incorporated a total of five questions, the new climate score ranged from 5 to 25. As a result, respondents who scored below 15 were classified as deniers, those scoring 20 or above were categorized as believers, and the remaining participants were identified as neutrals.

4.4 Results

4.4.1 Articles' Stance

In order to verify the alignment between the selected articles and the perception of climate change as a serious concern, we conducted an analysis based on the article stance. Participants were asked to judge the degree of concern expressed in each article about climate change, using response options that ranged from "Not a problem at all" to "A very serious problem." In instances where disagreements among labelers occurred, the majority vote was used to finalize the label. Figure 4.3 displays the distribution of the article stances, divided by the labelers' own stance on climate change. Interestingly, none of the articles were labeled as indicating that climate change is not a problem or only a minor problem, even by those participants identifying as deniers. This outcome validates our initial selection process and ensures that the selected articles passed a basic relevance check.

Additionally, we noticed a compelling pattern that reflects how pre-existing beliefs can influence individuals' perception of the articles' stance. Among the believers, a

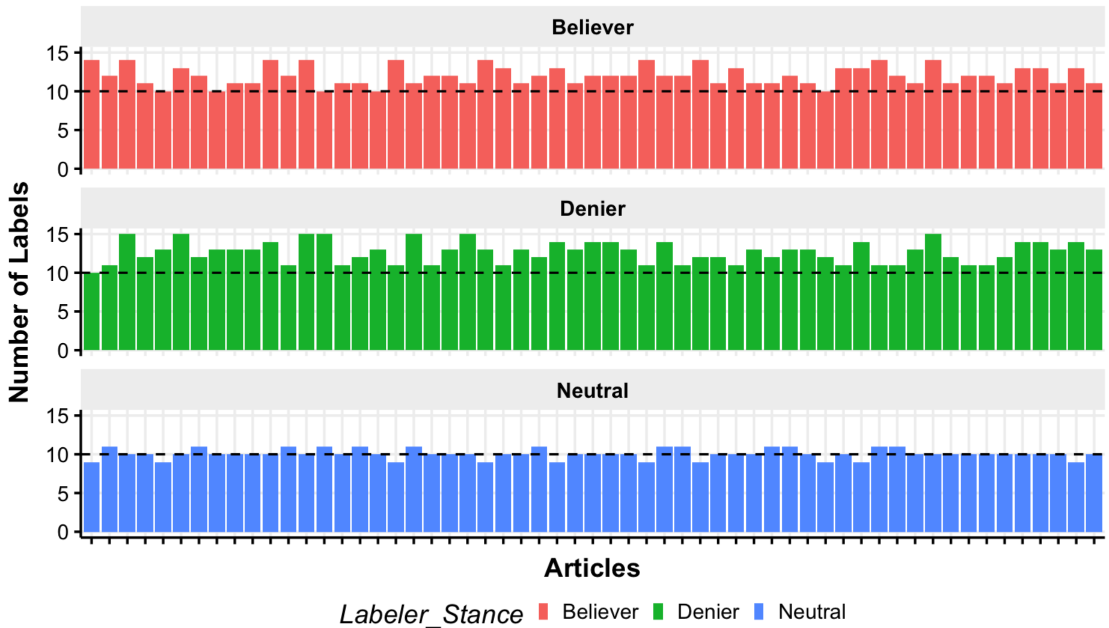


Figure 4.2: The figure displays the number of of labels received by all articles, categorized by different stance groups. The aim was to maintain balance at the article level, with approximately 10 labels from each stance group per article. The figure shows slight variability in the number of labels assigned to different articles, which is a result of the iterative labeling process conducted in separate batches to ensure adequate participation from deniers and neutrals. Notably, nearly all articles have achieved the minimum requirement of 10 labels per stance group, leading to more precise predictions compared to the previous survey

majority perceived the articles as signaling that climate change is a "Very Serious Problem." Neutrals, on the other hand, showed a somewhat less severe perception, with most falling into the "Serious" category. In contrast, deniers predominantly interpreted the articles as indicating that climate change is "A Problem," a category comprising more than twice as many articles as perceived by believers. These findings accentuate how the same rating scales can lead to varied interpretations, depending on individuals' viewpoints on the topic. They underscore the importance of ensuring a balanced representation of different stance groups in the data collection process. This analysis offers crucial insights into the relationship between article selection, personal viewpoints, and the perceived stance on climate change, contributing to a holistic understanding of our dataset.

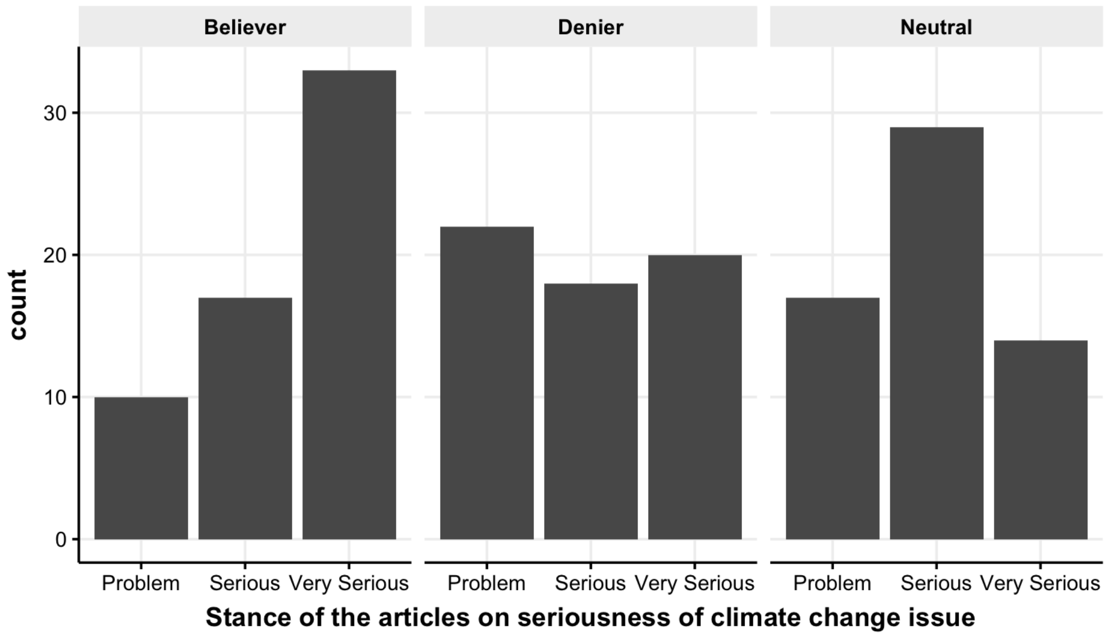


Figure 4.3: This figure illustrates the distribution of article stances, as determined by labelers' responses to the question, "According to the article, climate change is or could soon be..." The article stances range from "Not a problem at all" to "A very serious problem". The figure demonstrates how participants' perspectives on climate change influence their interpretation of the articles' stance. Believers predominantly perceive the articles as indicating that climate change is a "Very Serious Problem," while deniers perceive most articles as suggesting that climate change is only "A Problem".

4.4.2 Average Predicted Opinion Shift

To gain insights into the differences in expected persuasion among different stance groups, we utilized the opinion shift survey questions. The responses were collected on a scale ranging from -10 to 10, where -10 indicated stronger opposition to climate policies and +10 indicated stronger support for them. To align the scores with the previous survey, each score was divided by 10 and normalized to a range of [-1, +1].

The average predicted persuasion score, denoted as $PP_{i,j,k}$, represents the expected opinion shift of individuals in stance group i (believers, neutrals, deniers) for others who are (against, indifferent to, supportive of) climate policies after reading article k. To compute $PP_{i,j,k}$, we averaged the (approximately) 10 labels that met the relevant criteria. However, it should be noted that the scores at the article level are relatively noisy due to the limited number of labels available to estimate each $PP_{i,j,k}$. To obtain a more precise estimate of the predicted persuasion for the group of articles, we averaged over k. Since each respondent labeled three distinct articles, the labels are not entirely independent. Therefore, to account for potential correlation among responses from the same individual, we clustered all standard errors at the respondent level.

Figure 4.4 presents the average predicted persuasion scores for believers, neutrals, and deniers, categorized by the stance of the target group regarding climate policies (against, indifferent, supportive). The results indicate that, on average, all stance groups expect a reinforcement of existing beliefs for individuals who are already supportive of climate policies when exposed to articles emphasizing the severity of the issue. For the indifferent target group, all three stance groups anticipate some level of persuasion, with believers being the most optimistic, neutrals having a moderate expectation, and deniers expecting a smaller effect size. However, the most notable observation pertains to the target group who are against climate policies. While believers and neutrals, on average, do not expect significant persuasion to occur for this group, deniers anticipate a backfire effect, wherein exposure to these articles would intensify their opposition to climate policies. This divergence in the expected direction of the opinion shift forms one of the primary phenomena investigated in the subsequent follow-up experiment.

Although the study does not directly capture the underlying mechanisms influencing these predictions due to certain limitations, an analysis of the average prediction scores across different article stances reveals an intriguing pattern. Figure 4.5 illustrates the average prediction scores by respondents categorized as deniers, neutrals, and believers, specifically focusing on individuals who hold an opposing stance toward climate policies (referred to as the "Against" column in Figure 4.4). The findings indicate that believers and neutrals generally anticipate no significant persuasion to occur when individuals with an opposing stance are exposed to articles emphasizing

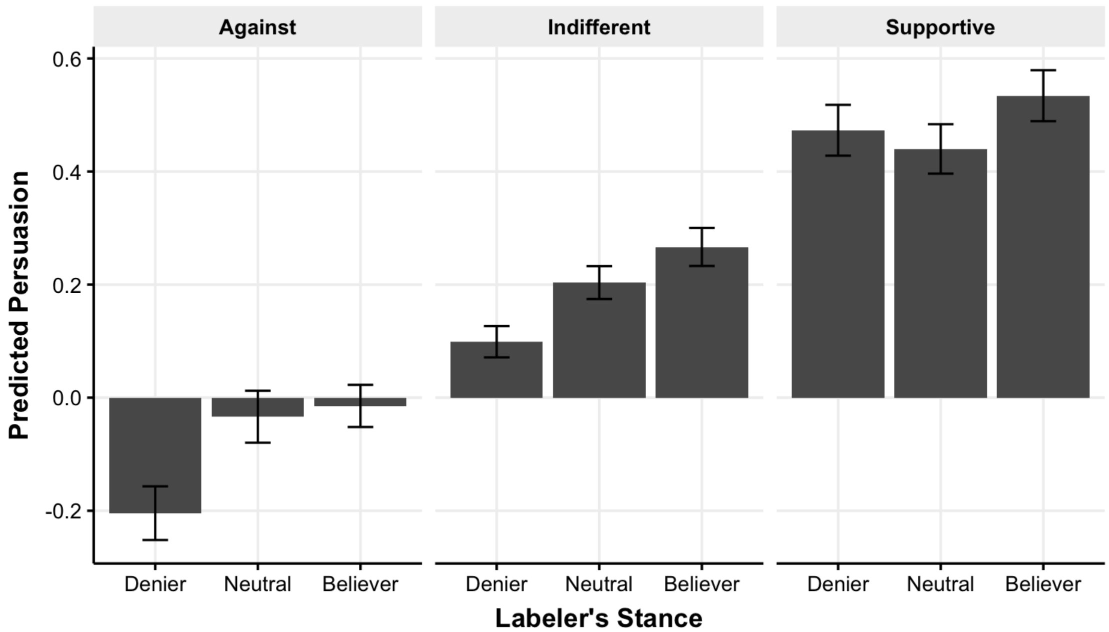


Figure 4.4: This figure illustrates the average predicted persuasion scores for believers, neutrals, and deniers across all 60 articles, categorized by the stance of the target group (against, indifferent, supportive) regarding climate policies. The results highlight distinct expectations among the groups, with believers and neutrals anticipating no persuasion for the *against climate policy* group, while deniers expect a potential backfire effect.

climate change as *A Problem* or worse. However, they predict a reinforcement of opposition when these individuals encounter articles that they perceive as indicating *No Problem* or *A Small Problem*. It should be noted that although the majority vote in Figure 4.4 does not indicate any articles labeled as *No Problem* or *A Small Problem*, a minority of cases exist with such labels. Consequently, standard errors tend to be larger for the first two columns in each group. For the purpose of this analysis, we included each prediction alongside the associated article stance perceived by the same respondents, recognizing that the prediction depends on their perception of the article rather than the consensus.

In contrast, deniers expect minimal change in individuals' stance when exposed to articles characterized as "No" or "Small Problem." However, they anticipate a backfire effect, resulting in a reinforcement of opposition, when individuals encounter more

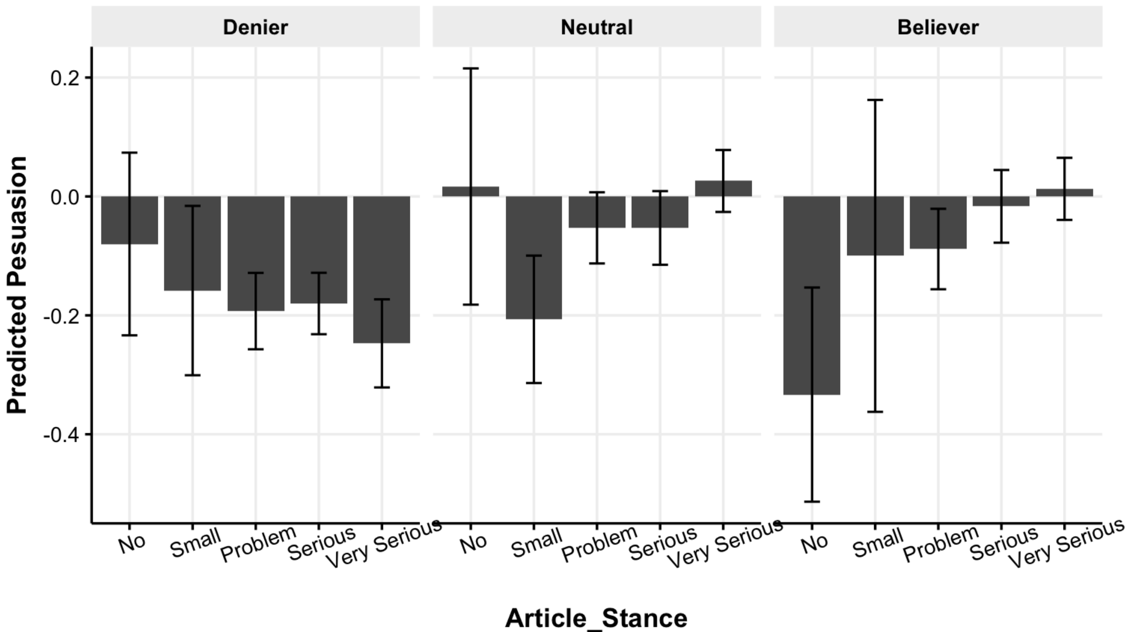


Figure 4.5: Average Persuasion Predictions by denier, neutral, and believer respondents for individuals against climate policies, categorized by articles' stance on severity of climate change.

extreme articles. Although the observed differences in predictions are not statistically significant, a slight increasing trend can be observed in terms of the expected backfire effect size as the severity of the article stance intensifies. This observation suggests a potential relationship between the emphasis on the severity of climate change and the occurrence of unwanted backfire effects. To gain deeper insights into the reasons underlying these variations in persuasion predictions, further research is warranted.

4.4.3 Heterogeneity of Predicted Persuasion

The variation in article stances on the gravity of climate change, along with different expectations of respondents, has led to a substantial heterogeneity in the average predicted persuasion scores $PP_{i,j,k}$. Delving into this heterogeneity could provide us with more profound insights into this complex issue, as well as guide us to generate more nuanced hypotheses to test through our subsequent randomized experiment. While a three-dimensional plot would be needed to display the full distribution of these scores due to the fact that $i \in \{\text{denier}, \text{neutral}, \text{believer}\}$, $j \in \{\text{against}, \text{indifferent}, \text{supportive}\}$, and $k \in \{1, \ldots, 60\}$, we have chosen to present the results in two dimensions for clearer illustration. This is primarily because we are particularly interested in understanding the differences between deniers' and believers' predictions.

Figure 4.6 illustrates the distribution of $PP_{i,j,k}$ for all the articles. The x-axis denotes the average prediction made by deniers, while the y-axis represents those made by believers. Each article k appears thrice on the plot—once for each target group (i.e., against, indifferent, or supportive)—distinguished by different colors. The gray line indicating the $y = x$ axis aids in visualizing the degree of disagreement between deniers and believers regarding the potential persuasive effect of each article. Points closer to this gray line signal higher agreement, while those farther away indicate higher levels of disagreement.

For the purposes of this study, we are particularly interested in the differing predictions for those who are against climate policies, represented by the color green in the Figure 4.6. As demonstrated earlier in Figure 4.4, deniers generally predict a backfire effect for like-minded individuals, while believers anticipate no significant effect. This pattern can also be observed in Figure 4.6; however, this latter figure shows a considerable heterogeneity among different articles in terms of the predictions made by believers. In particular, there is a total of 21 articles for which believers anticipate a positive persuasion effect, while deniers expect a backfire effect. This set of articles falls in the upper left quadrant of the figure, and we refer to it as set M since the predictions of deniers and believers are misaligned. There is also a set of 35 articles in the lower left quadrant for which the predictions are aligned, and both groups anticipate a backfire effect. We refer to this set as A.

For each of these sets, we can compute the average predicted persuasions in a similar fashion to what was done in Figure 4.4. The result of these average predictions for the 'against' group is shown in Figure 4.7. The left figure shows the average predicted persuasion across the set A, while the right figure does the same for the

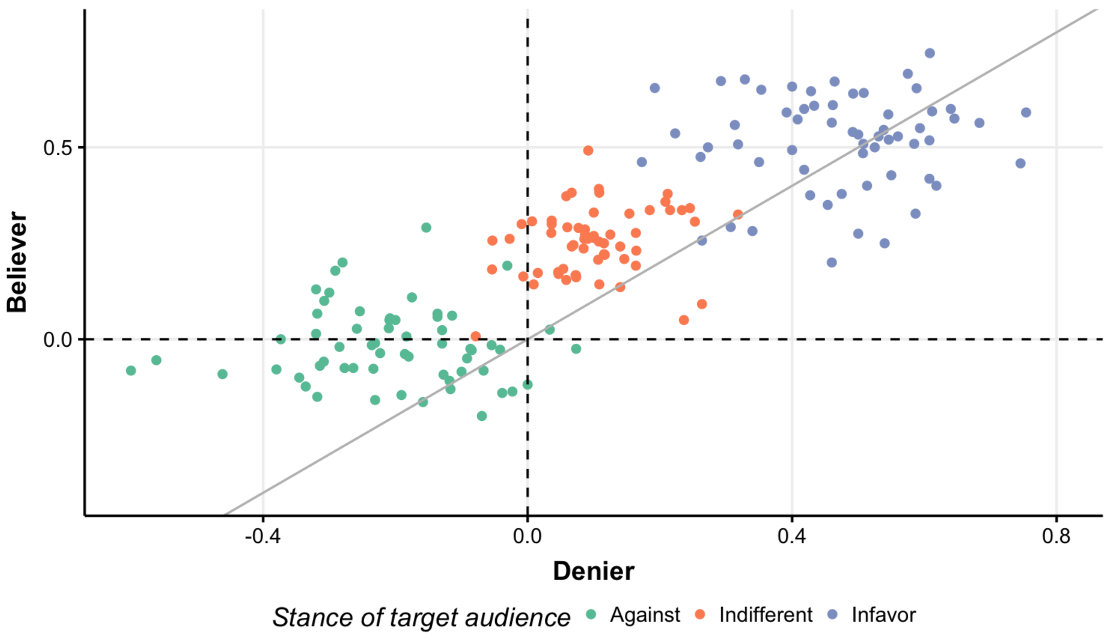


Figure 4.6: Two-dimensional scatter plot of the average predicted persuasion scores for all articles. The x-axis represents the average prediction by deniers, while the y-axis corresponds to the average prediction by believers. Each article is represented three times in different colors, corresponding to different target groups (against, indifferent, or supportive). The gray line indicates the $y = x$ axis. Points closer to this line signify higher agreement between the two groups, whereas points farther from the line indicate more significant disagreement.

set M. As can be seen, on average there is a statistically significant disagreement for the misaligned set, with believers predicting a significant positive persuasion effect while deniers anticipate a backfire effect. Furthermore, even for the aligned set, the magnitude of predictions is significantly different, with deniers predicting a larger backfire effect. These predictions form the basis of the secondary hypotheses which will be tested in the subsequent randomized experiment.

5 Phase 3: Randomized Survey Experiment to Assess Actual Shifts in Climate Change Attitudes

In the subsequent phase of our study, we conducted a randomized survey experiment to empirically test the predictions obtained in the second survey. Through this exper-

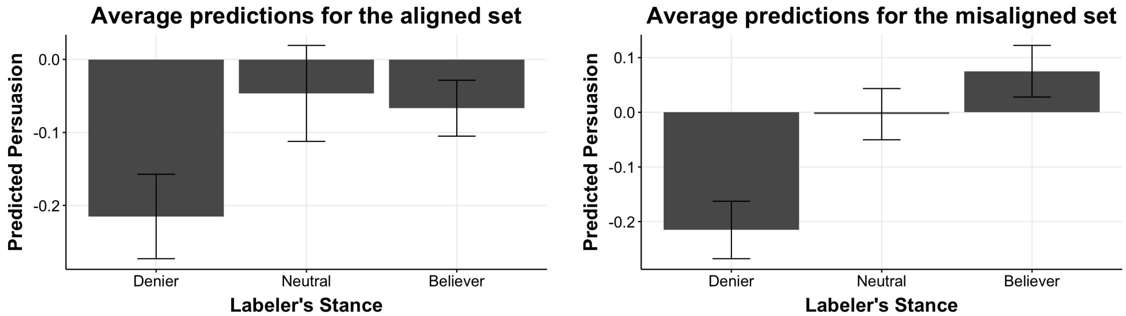


Figure 4.7: Comparisons of Average Predicted Persuasions for Misaligned and Aligned Sets. The left panel represents the average predicted persuasion effect for the set of articles with aligned predictions (A), while the right panel represents the same for the misaligned set (M). The discrepancies between believers' and deniers' predictions are apparent, with a notable difference in the direction of predicted persuasion effects, particularly in the misaligned set.

iment, we aimed to compare the predicted opinion shifts, obtained from the second survey, with the observed changes in climate change attitudes resulting from exposure to the news articles. For this experiment, we used the same set of 60 news articles that were used in the second survey to collect the predictions. All of these articles were carefully chosen to emphasize the severity and consequences of climate change. Additionally, we included a few articles unrelated to climate change in the control group. These control articles were selected to ensure that they were non-polarizing and did not invoke any partisan biases[6]. It is important to note that the experiment was pre-registered, and the pre-registration plan, as well as the stimuli used in the experiment, are available on the Open Science Framework (OSF) for reference[7]. The pre-registration ensured that the research plan and hypotheses were established prior to conducting the experiment, preventing any potential biases or data-driven decisions during the analysis phase.

[6]Specifically, the control group articles covered topics such as NASA's project on Mars and the importance of investment in education.

[7]https://rb.gy/kj8ty

5.1 Experiment Design

The experiment was divided into two main parts. In the initial phase, we employed a similar approach to the previous survey to identify climate change deniers. Participants were presented with a series of demographic questions, followed by a set of questions aimed at understanding their stance on climate change and global warming. In addition to the questions used in the previous survey, we included one more question to measure the extent of their belief in the urgency of climate change and the necessity of national and international actions to combat it. This question was carefully framed to mirror the same prediction question used in prior surveys, allowing us to capture their pre-treatment stance accurately. The complete set of questions used in the survey experiment can be found in Appendix C. After the initial phase, participants' stances were determined using the procedure described in Section 4.3. Those classified as climate change deniers were given the option to participate in the second task for an additional bonus.

In the second part of the experiment, participants were asked to read one news article randomly chosen from either the set of 60 climate change articles or a control group of unrelated news articles. Approximately one-third of the participants were assigned to the control condition, while the remaining two-thirds received the treatment condition with climate change-related articles. After reading the article, participants were asked to respond to a set of questions aimed at eliciting their attitudes towards climate change. These questions covered three different dimensions: posterior beliefs in climate change, support for specific climate policies, and intention to engage in private actions that mitigate climate change. We adapted certain policy and action questions from the study conducted by Dechezleprêtre *et al.* (2022), but we made modifications to better align these questions with our specific stimuli. Participants were instructed to rate their level of agreement with each statement on a scale from 1 to 10, where 1 represented "fully disagree" and 10 indicated "fully agree." The statements included the following:

- **Belief**
  - If nothing is done to limit climate change, there will be dire consequences for humanity in the not-distant future.
  - We urgently need national and international actions to combat climate change.
  - I am worried about climate change.
  - Human-caused climate change is real and it is occurring.
- **Policy Support**
  - Subsidize the insulation of buildings to make homes more energy efficient.
  - Subsidize the development and use of low-carbon technologies (e.g., renewable energy, capture and storage of carbon, etc.).
  - Impose a carbon tax on all products proportional to the amount of CO2 emitted for producing them.
  - Increase fuel duty, the tax motorists pay for petrol and diesel.
- **Personal Actions**
  - Increase walking, cycling, or using public transport instead of driving.
  - Use only green electricity, that is electricity produced by renewable energy, even if it costs more.
  - Vote for a candidate who is vocal about climate change issues.
  - Make a significant donation to an environmental cause.

By incorporating a diverse set of questions, we were able to comprehensively assess various aspects of climate change denial and gauge the potential impact of the news articles on respondents' attitudes and behaviors. To enhance the statistical power of the experiment and detect meaningful effects, we presented each set of questions related to posterior beliefs, support for climate policies, and intention to engage in private actions on separate pages to the participants. This allowed us to gather detailed and focused responses for each attitude dimension.

To complement the self-reported survey responses on participants' opinions and beliefs about climate change, we also incorporated a behavioral measure to capture more practical effects of the treatment articles. In this measure, we assessed participants' willingness to donate a flexible share of their bonus, received for participating in the survey, to one of four non-governmental organizations (NGOs). Among the four NGOs, two were in support of climate change mitigation, while the other two were against it. By including this behavioral measure, we aimed to explore potential disparities between participants' expressed attitudes and their actual actions regarding climate change mitigation. This approach provides valuable insights into the alignment or divergence between individuals' stated beliefs and their willingness to take concrete steps to address the issue. The behavioral measure offers a more objective and real-world perspective on the impact of the treatment articles on participants' climate-related decision-making.

5.2 Respondents

The respondents were recruited from Prolific online research platform[8]. Since we only needed the responses for those who classified as a denier based on our classification of climate stance, we had to set some filters to sample users who most likely satisfied this condition. In the end, 1123 respondents started the survey, 677 were classified as a deniers and had the choice to proceed to the second task for an extra bonus, from which 644 chose to continue. We excluded any participant who fulfills any of the following: voluntarily declare that they did not pay attention to the questions, perform straight-lining on at least 2 blocks of the main survey, read the article too fast or spent 2 standard deviations less than the mean time to complete the whole task, and answer nonsensical text in open-ended questions. 30 respondents were removed because of these constraints which leaves us with 614 responses from deniers.

[8]In accordance with the pre-registration plan, our initial goal was to secure a sample size of 1000 responses from deniers. Nonetheless, the findings presented in this chapter draw from approximately half of this intended sample size, reflecting the accumulated responses up to this point.

5.3 Analysis

We estimate the average treatment effect (ATE) of being randomly assigned to the treatment group with the following regression model:

$$Y_i = \beta\, T_i + \alpha_1 Y_{i,0} + \delta_1 S_i + \gamma_1\, \boldsymbol{C}_i + \epsilon_i \tag{4.1}$$

The main independent variable is T_i, which is a binary treatment indicator that takes the value of 1 if the participant was assigned to the treatment group (exposed to the selected news articles about climate change) and 0 if the participant was in the control group (exposed to unrelated news articles). To account for participants' pre-treatment attitudes towards climate change policies, the variables $Y_{i,0}$ and S_i are included. The variable S_i represents the climate score defined in Section 4.3, and $Y_{i,0}$ is the additional climate change belief question asked before the treatment. These variable help control for any pre-existing differences in attitudes among participants. The vector $\boldsymbol{C}_i$ includes additional covariates, such as age, gender, education level, race, employment status, urbanity, social media activity, and political affiliations (democratic/republican scale and conservative/liberal scale). The inclusion of these covariates is not strictly necessary for model identification since the treatment assignment is randomized. However, including them can help improve the precision of the inference and account for any potential imbalances across the control and treatment groups. The error term ϵ_i represents the random variability or unexplained factors in the outcome.

The dependent variable, denoted as Y_i, represents the outcome of interest, which could be various measures related to respondents' attitudes towards climate change. To ensure reliable and robust measurements of respondents' attitudes, we created standardized indices for each outcome measure, including belief, policy, and actions. This was achieved by combining multiple items related to each attitude dimension into an index. By using this approach, we increased the stability of our survey measures, resulting in more precise estimates and greater efficiency in capturing true effects

(Broockman *et al.* 2017). To standardize the indices, we computed the average of the corresponding components, subtracted the mean value in the control group, and divided the result by the standard deviation in the control group (Kling *et al.* 2007). This transformation allowed us to compare and interpret the results on a common scale, ensuring consistency across the different attitude dimensions. For the donation outcome, participants were asked to make a two-step decision. First, they selected one of four NGOs, two of which supported climate change mitigation and two that were against it. Next, they decided the amount they wanted to donate from their bonus. To capture this behavior, we encoded the donation amount as a continuous variable, taking negative values for donations to NGOs that act against climate change, 0 if no donation was made, and positive values for donations to NGOs that support climate change mitigation. Finally, to account for potential correlation among responses from the same news articles, we clustered the standard errors at the news article level. This approach allows us to properly handle any dependencies in the data and obtain more accurate statistical inference.

Moreover, an additional model will be employed to determine the average treatment effects for two distinct subsets of articles characterized by aligned and misaligned predictions, as elaborated in Section 4.4.3. This model aims to further elucidate the comparative predictive accuracy of deniers and believers. As previously elucidated, the first subset encompasses 39 articles wherein both deniers and believers foresee a backfire effect following exposure to treatment, signifying concurrence in their predictions of news article influence. Conversely, the second subset comprises 21 articles characterized by misaligned predictions, with believers anticipating persuasion and deniers still projecting a backfire effect. In pursuit of this analysis, the subsequent regression model will be employed.

$$Y_i = \beta_1 TA_i + \beta_2 TM_i + \alpha_2 Y_{i,0} + \delta_2 S_i + \gamma_2 \boldsymbol{C}_i + \epsilon_i \tag{4.2}$$

where the binary indicators TA_i and TM_i differentiate respondents assigned to articles with aligned and misaligned predictions, respectively. Our primary focus lies on estimating the coefficient β_1, as it serves as a crucial metric to ascertain which group's predictions more accurately correspond to the actual treatment effect in cases where deniers and believers hold differing views. A statistically significant and negative β_2 would indicate that deniers' predictions better align with the observed treatment effect during instances of prediction disagreement. Conversely, a statistically significant and positive β_2 would suggest that believers' expectations more accurately predict the actual treatment effect in such scenarios.

5.4 Results

The outcomes of the regression models yield the Average Treatment Effects (ATEs), detailed in Table 4.2. The first row of the table showcases the ATEs for the entire set of articles across all outcomes. Remarkably, the ATE for the 'belief' outcome presents a statistically significant and positive effect. The substantial and statistically significant positive shift in perspectives concerning the severity of climate change attests to both the effectiveness of the selected stimuli and the malleability of deniers' convictions. Importantly, this effect size, when compared with analogous studies on persuasion experiments, demonstrates not only statistical significance but also considerable magnitude. For instance, in the meta analysis by Rode *et al.* (2021), an effect size of $g = 0.08$ is found by analyzing 396 effect sizes derived from 76 distinct experiments.

Contrarily, the estimated persuasion effects remain statistically non-significant for the remaining outcomes. The lack of effects on policy support, personal actions, and donation outcomes underscores the formidable challenge of altering these specific attitudes. These findings resonate with existing research on climate change persuasion, highlighting the heightened resistance to influence in the domain of policy attitudes (Rode *et al.* 2021). These outcomes were in anticipation, given the focal point of the chosen articles which primarily emphasize the consequences of climate change and global warming, such as escalating sea levels and temperatures. Additionally, these articles are succinct, lacking the in-depth exploration of policy measures to mitigate the impacts of climate change. Consequently, the observed null effects across these

Table 4.2: Estimated Average Treatment Effects on Main Outcomes.

	Dependent variable							
	Belief		Policy Support		Private Actions		Donations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ATE	0.268***		0.058		0.066		−0.021	
	(0.049)		(0.084)		(0.073)		(0.123)	
ATE Aligned		0.329***		0.096		0.050		−0.003
		(0.054)		(0.094)		(0.081)		(0.139)
ATE Misaligned		0.184**		0.010		0.053		0.021
		(0.061)		(0.106)		(0.092)		(0.157)
Observations	608	581	607	580	607	580	605	578
R^2	0.822	0.821	0.521	0.517	0.487	0.482	0.197	0.205
Adjusted R^2	0.788	0.784	0.429	0.417	0.388	0.376	0.042	0.040

Notes: Significance is denoted as follows: *** p<0.001, ** p<0.01, and * p<0.05.

outcomes are congruous with these characteristics. Furthermore, the absence of a significant effect on donation behavior logically aligns with the unchanged self-reported policy support among respondents, rendering it unlikely for substantial behavioral actions to stem from this modest intervention.

The ATEs for both aligned and misaligned articles, presented in the second and third rows, reveal positive and statistically significant effects for the 'belief' outcome, while other outcomes exhibit null effects. Although no statistically significant distinction is evident between subsets of articles with aligned and misaligned predictions, the estimated effect is more substantial and precise for the aligned set. This outcome further underscores the insufficiency of both deniers and believers in predicting the persuasive impact of news articles, as this is the set for which both groups predicted a backfire effect. Nevertheless, the notable positive and significant effect of misaligned articles suggests a relatively more accurate prediction ability among believers, particularly concerning the anticipation of effects on deniers for this subset of articles.

The central crux of this study lies in the comparison between persuasion effects and the average predictions offered by distinct stance groups. This juxtaposition provides crucial insights into the disparities between individuals' anticipations and the factual impacts stemming from exposure to climate-related news articles. For

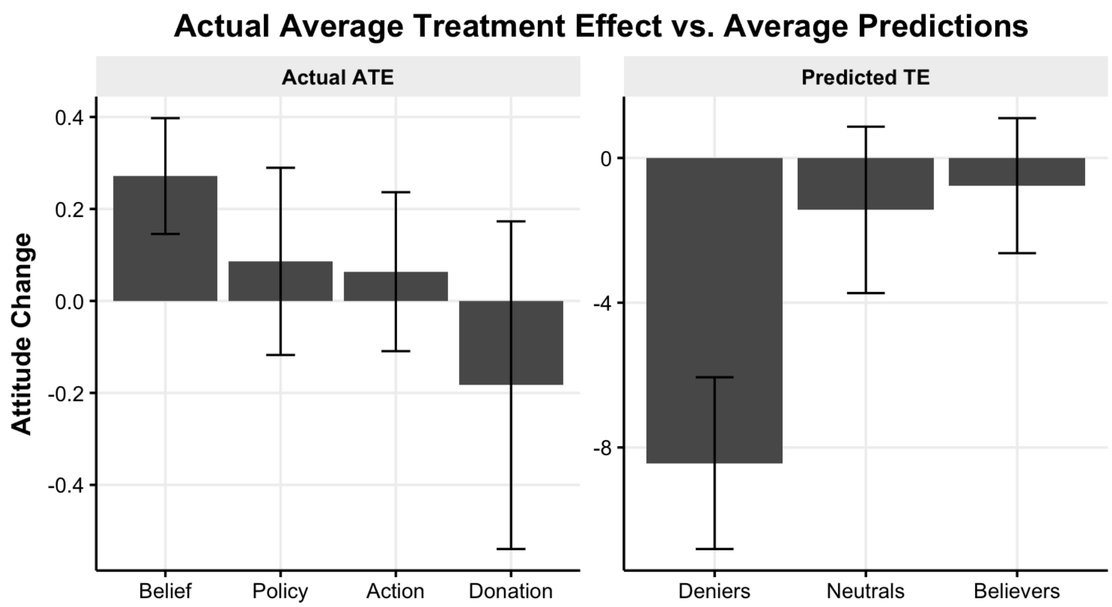


Figure 4.8: A comparison between the average treatment effect for attitude outcomes on deniers and the average prediction by climate stance.

enhanced clarity of this comparison, Figure 4.8 visually presents the average treatment effect for attitude outcomes alongside the average prediction made by each stance group. It is important to note that, in order to render the comparison more robust, the average predictions have been normalized by the standard deviation of the complete set of responses for each stance group, considering the absence of a control group in the prediction survey. Strikingly, the deniers' average prediction stands in stark contrast to the observed effects, thereby contradicting the hypothesis that deniers possess an advantage for accurately predicting opinion shifts among their like-minded peers, fostered by an augmented capacity for perspective-taking due to shared beliefs. Quite to the contrary, these findings underscore that deniers' preexisting biases on the subject impede their ability to presage the consequences of exposure to news articles effectively.

The forecasts articulated by both believers and neutrals also display a noticeable divergence from the actual effects, carrying a certain degree of pessimism, particularly when contrasted with the treatment effect on belief. However, a constellation of factors could potentially elucidate these disparities. A potential factor lies in the formulation of prediction queries presented to participants during the second survey phase, which encompassed both belief and policy dimensions. Given that the estimated Average Treatment Effect (ATE) for policy outcomes did not attain statistical significance, it becomes conceivable that believers' predictions might not be overly distant from actual outcomes if their considerations extended to policy support, in tandem with their stance on the gravity of climate matters. Moreover, it is imperative to acknowledge that respondents' mental conception of an individual who perceives climate change as "not a problem" could potentially entail a higher degree of denial than what has been delimited as the cohort of deniers within the experiment.

To scrutinize this conjecture, we embarked on an exploration of treatment effect heterogeneity across weak and strong deniers. Strong deniers were operationally defined as individuals holding a climate score below 10, while weak deniers were identified as those possessing a climate score ranging between 10 and 15, given the original classification of deniers encompassing a climate score between 5 and 15 (see Section 4.3 for comprehensive elucidation). In order to quantify these effects, an indicator for weak/strong denial was created and interacted with the treatment variable within the regression model. The outcomes of this analysis are expounded in Table 4.3. As discerned from the results, the average treatment effect for the belief outcome is primarily propelled by the persuasion effect on weak deniers, where a notably substantial, statistically significant positive effect is detected. In contrast, the treatment effect on strong deniers remains statistically insignificant.[9] Notably, the treatment effect on policy support for weak deniers registers as positive and statistically significant. These findings imply that for individuals who staunchly deny climate change, even effecting a change in their belief system proves to be a formidable undertaking—a trend harmonious with the predictions posited by believers and neutrals. Consequently, this suggests that the distinction between the actual effect and

[9]The difference between a coefficient that holds statistical significance and one that does not is not inherently statistically significant in itself. Nonetheless, the discrepancy between coefficients for individuals categorized as weak deniers and those classified as strong deniers does exhibit statistical significance, as indicated by a P-value of 0.024.

Table 4.3: Heterogeneity of Treatment Effects on Main Outcomes for Weak and Strong Deniers

	Dependent variable:			
	Belief	Policy Support	Personal Actions	Donations
	(1)	(2)	(3)	(4)
MTE, Weak Deniers	0.354***	0.216*	0.026	−0.002
	(0.061)	(0.106)	(0.092)	(0.157)
MTE, Strong Deniers	0.133	−0.192	0.129	−0.051
	(0.077)	(0.133)	(0.116)	(0.196)
Observations	608	607	607	605
R^2	0.824	0.526	0.488	0.197
Adjusted R^2	0.790	0.434	0.387	0.040
Note:				*p<0.05; **p<0.01; ***p<0.001

the average prediction may potentially stem from believers taking strong deniers into consideration while responding to prediction queries.

6 Discussion

This research study delves into the dynamics of attitude and behavioral changes triggered by exposure to information that challenges preexisting beliefs. Furthermore, it endeavors to determine whether these shifts can be accurately predicted by individuals with either congruent or opposing viewpoints. To tackle these inquiries, we undertake a comprehensive series of online surveys focusing on the contentious topic of climate change, involving respondents from the United States. The initial phase involves curating a substantial collection of actual newspaper articles sourced from a diverse array of international media outlets. Through collaboration with participants from the Prolific platform, these articles are assessed and rated based on the severity of their depiction of the climate change issue. Participants are also tasked with predicting the potential change in stance that an average climate change denier would undergo upon reading each article. Subsequently, we select a subset of the top

60 articles with the highest predicted persuasiveness scores. This selection forms the foundation for a survey experiment targeted exclusively at climate change deniers. Utilizing random assignment, participants are assigned to read either one of the pre-selected climate change articles or an analogous editorial piece on an unrelated subject. Following exposure to the chosen article, we gauge participants' revised stance on climate change, their professed endorsement of public policies aimed at addressing the issue, their intent to undertake individual actions to mitigate climate change, and their willingness to allocate a portion of their survey rewards to a non-governmental organization either supportive of or opposed to climate change mitigation efforts.

Our study yields robust evidence indicating that the exposure of climate change deniers to newspaper articles centered on the subject effectively leads to a revision of their initial beliefs. However, this effect does not align with concurrent modifications in either stated or observed preferences related to actively combating climate change, at least within the scope of our experiment. Nonetheless, the discerned positive persuasion effect pertaining to belief alteration holds substantial promise. The transformation of perceptions regarding the gravity of climate change serves as an essential foundational step in catalyzing potential shifts in both policy endorsement and individual actions. This fundamental change in belief has the potential to lay the groundwork for subsequent adjustments in attitudes and behaviors. Moreover, our findings offer robust evidence that in the case of weak deniers the policy support can also be influenced. This result holds significant promise, revealing that even concise articles can effectively alter reported endorsements of climate policies among individuals with less entrenched resistance to climate change. This finding underscores the criticality of acknowledging heterogeneity in persuasive endeavors. Neglecting such nuances and amalgamating a diverse range of stances under a singular "climate denier" category could potentially overlook opportunities to influence those with less steadfast viewpoints on climate change.

In synthesizing the findings from prediction survey and the randomized experiment, a comprehensive view emerges, revealing the general inadequacy of individuals' predictive capabilities regarding the persuasive impact of climate change articles on others. Interestingly, this deficiency is further accentuated among deniers, who, counter to our original hypothesis grounded in the perspective-taking concept, exhibit a notably poorer accuracy in predicting opinion changes among fellow deniers. Conversely, believers, while not immune to prediction errors, display a relatively higher degree of accuracy in certain instances. In particular, they aptly predict positive persuasion effects for a subset of articles. Nevertheless, believers still exhibit a tendency to underestimate the potential for persuasion, particularly evident in their predictions concerning the other subset of articles for which they foresee a significant backfire effect, contrary to the observed positive actual effect.

These results carry important societal and policy implications. Firstly, our findings challenge initial assumptions by revealing that deniers' beliefs possess a greater degree of malleability than previously envisaged, as they can indeed be influenced to recognize the gravity of the climate change issue. Secondly, our study underscores a noteworthy trend: climate change believers frequently underestimate their potential to persuade deniers. This incongruity between anticipated and actual outcomes extends beyond the confines of our investigation, resonating particularly within the realm of communicative interactions, such as those prevalent on social media platforms. Given that climate change believers are primed to initiate such persuasive endeavors, this underestimation could potentially discourage their engagement in socially constructive actions aimed at tempering the escalating polarization. Such hesitancy may arise from a perception of ineffectiveness, thus impacting the cultivation of productive discourse and collective efforts to address climate-related divisions.

The insights garnered from this chapter are situated within the domain of climate change, a topic that is both significant and polarizing in our societal discourse. However, the implications of our findings extend beyond this specific context and hold relevance for a broader spectrum of issues characterized by divergent perspectives. This includes subjects of comparable societal importance such as immigration, escalating income inequality, and racial justice. By adopting a broader perspective, our results could potentially illuminate dynamics relevant to these domains, shedding light on the interplay between differing viewpoints and the potential for persuasion.

Chapter 5

Conclusion

The chapters in this dissertation contribute significantly to the evolving research areas of habit formation and political persuasion. Central to all these chapters is the examination of everyday repetitive behaviors in diverse contexts such as shopping, social media usage, and interaction with news articles.

In the initial chapter, a comprehensive exploration of the psychological theories underpinning the research is presented. Notably, the discussion revolves around habit formation and theory of mind. Habits exhibit distinctive features that motivate our investigations. Firstly, the consistent repetition of behaviors within similar contexts transitions behavioral control from being goal-dependent to context-dependent. Consequently, contextual cues trigger the associated behavior automatically, reducing the need for conscious decision-making. Secondly, established habits tend to persist over time due to their resistance to change, exerting substantial influence over individuals' actions. Furthermore, the concept of theory of mind, a cognitive skill enabling the comprehension of others' cognitive processes, furnishes a valuable framework for scrutinizing persuasion dynamics. This aligns with neuroscientific findings indicating that the value of information sharing is shaped by communicators' considerations of recipients' mental states. This theoretical backdrop forms the foundation for the hypothesis tested in the last study – the notion that individuals who share similar beliefs possess an advantage in accurately predicting shifts in opinion among their peers, attributed to enhanced perspective-taking driven by shared beliefs.

The second chapter delves into the influence of habits on recurring purchase decisions within the domain of grocery store shopping. While the concepts of repeated purchases and consumer inertia have received extensive attention in quantitative marketing and economics, the underlying psychological mechanisms driving this inertia have remained insufficiently elucidated. In particular, the empirical examination of the role of habits within this specific context has been notably absent. To address this gap, the chapter sets out to investigate the impact of shopping habits on in-store choices, utilizing store closures as a natural disruption that affects households' shopping behaviors. The central premise rests on the notion that each store closure presents an occasion for households to engage with new store environments devoid of the familiar contextual cues that trigger established habits. The findings of the study disclose a substantial role played by habits, manifesting as households with stronger habits encountering temporary upheavals in their shopping routines following store closures. With the passage of time, these households appear to form new habits within alternative stores, leading to lasting shifts in brand preferences. This underscores the notion that the formation of shopping habits can result in suboptimal consumer behavior.

These findings carry substantial managerial implications that extend to optimal pricing strategies, advertising tactics, and the allocation and arrangement of products within stores. The insights gained from understanding or discovering these shopping habits can prove instrumental for firms seeking to fine-tune their operations. A pivotal strategy lies in incentivizing stores to maintain or modify the placement of their products, depending on the nature of the habits formed. This interplay between competing brands adds complexity to the equation. For less popular brands, there exists an incentive to induce stores to alter their product placement, thus disrupting existing habits and potentially boosting sales. Conversely, popular brands within a store would aim to ensure consistent placement to leverage established habits. Simultaneously, this mechanism presents a challenge for retailers contemplating the effects of common in-store item rearrangements on overall sales and profits. Such rearrangements could prompt shoppers to explore the store more extensively, potentially leading to the discovery of items that had been previously overlooked due to ingrained shopping habits. However, this exploration carries associated search costs, which could eventually deter individuals from making a purchase altogether. Future research, both theoretical and empirical, could delve into the intricate interplay between rival brands and retailers resulting from the impact of shopping habits. Employing a game-theoretic framework could offer insights into the equilibrium that retailers could strive to establish in terms of negotiating the consistency of product placements with different brands within their stores. This line of inquiry holds the potential to uncover strategies that optimize profits while addressing the nuanced dynamics of consumer behavior shaped by habit formation.

Chapter three contributes to the existing literature on habit measurement by introducing entropy as an implicit measure of behavioral regularity. In this chapter, entropy is harnessed to assess the distribution of behaviors across a 24-hour clock time, effectively quantifying the degree of regularity present in behavior patterns throughout the day. At its core, entropy functions as a metric to measure the stability of contextual factors that surround behaviors, thereby providing an indirect estimation of habit formation facilitated by conducive conditions. The key advantage of this approach is its ability to gauge habit development using observational behavioral data, mitigating the reliance on self-reports and enhancing the scalability of the method. To showcase the practicality of this measure, it is applied to the realm of social media usage.

Through a series of rigorous analyses, the validity of entropy as a measure of habits is rigorously established. The findings highlight that higher entropy values, indicative of lower levels of time-of-day behavioral regularity, correspond to weaker habit formation. This, in turn, translates to diminished time spent on social media platforms and reduced usage frequency over the long-term future. The results not only underscore the utility of entropy as a valuable metric for habit measurement but also offer valuable insights into the dynamics of habit development in the context of technology engagement. The insights gleaned from our analyses carry multifaceted implications, spanning both platform management and the practical implementation of digital technologies and interventions. Platforms can strategically leverage the identified mechanism to foster consistent user engagement through targeted interventions, such as sending push notifications on particular times of the day. Furthermore, this understanding can be translated into the design of interventions that facilitate self-control and mindful technology usage. While current self-control tools often rely on metrics like total time or frequency to set usage limits, our research underscores the potential benefits of integrating the regularity of time of usage as an additional dimension.

It is crucial to acknowledge the limitations inherent in our analysis. While we demonstrate robust support for using entropy as a measure of habits, our focus on clock time as a contextual factor may not capture the full spectrum of influences on online behavior. Contextual factors like location, environment, and internal cues such as mood and emotion could also serve as a trigger for learned habits. Expanding our approach to encompass a broader range of contextual factors could enhance its applicability. Addressing these constraints provides a valuable potential for future research to embrace a wider array of contexts and applications, offering a more comprehensive understanding of online behavior and the underlying driving forces. Furthermore, our findings offer a foundation for future research that employs randomized interventions to manipulate entropy while controlling for frequency and total time spent on social media apps. This approach would yield valuable insights into how altering regularity impacts subsequent usage behavior.

In this fourth chapter, we delve into the intricate interplay between individual attitudes towards climate change and their capacity to grasp and anticipate shifts in others' opinions. Given the urgency of the climate crisis, the potential disparities between actual and predicted persuasive effects of news articles take on significant importance, particularly in the context of exacerbating polarization. Through a meticulously designed series of surveys, we systematically collect predictions regarding the impact of climate change-related news articles on climate change deniers. Surprisingly, our findings illuminate starkly discordant predictions: deniers envision a backfire effect among their peers, while respondents with neutral or supportive stances anticipate negligible impact. Employing a rigorous randomized survey experiment involving deniers, we put these predictions to the test, ultimately unveiling an unforeseen positive shift in belief in severity of climate change issue after exposure to the articles. It is noteworthy that this effect does not manifest as discernible changes in either stated or revealed support for climate change actions among the average denier. However, in the case of weak deniers, there is evidence of a positive influence on policy support as well.

By integrating the outcomes of both the prediction survey and the randomized experiment, a comprehensive picture emerges, underscoring the inherent limitations of individuals' predictive capacities concerning the persuasive influence of climate change articles on others. These findings challenge our initial hypothesis, which postulated that deniers might possess a heightened ability to anticipate persuasion due to their shared perspective with fellow deniers. Moreover, while believers fare relatively better in their predictive accuracy, they too display a tendency to underestimate the potential for persuasion among deniers. This underestimation could carries significant implications for society. Given that climate change believers are likely to initiate such persuasive endeavors, their underestimation could potentially dissuade them from participating in socially constructive actions aimed at ameliorating escalating polarization. Such hesitancy may stem from a perceived lack of effectiveness, which in turn could hinder the fostering of productive discourse and collective efforts to address the divisions exacerbated by climate-related discussions. However, it is important to note that our research does not directly capture the consequences of this phenomenon and its extent in a real-world social network remains an avenue for future investigation. Subsequent research could delve into these dynamics within a more naturalistic context, providing a deeper understanding of the complex interplay between predictive accuracy, belief polarization, and effective communication strategies.

Appendix A

Robustness Results (Chapter 2)

1 Additional Statistics

In this appendix, we include some additional figures to illustrate the data. Figure A.1 shows the number of identified closing stores over years, and Figure A.2 shows the geographic dispersion of closing stores over the US at the county level. Figure A.3 displays log-weekly sales in the top-5 product categories by national purchase volume for the closing stores corresponding to the set of exposed households from January 2006 until December 2018, separated based on retailers.

2 Recent modal brand replication results

In this appendix we replicate the results for all DID and event study models for different values of L and T_e. L determines the length of the window for used to define the recent modal brand (Equation 2.2), and T_e specifies the pre-closure duration used to define the exposure levels (Equation 2.1).

2.1 Changing T_e

Here, we change the value of the period on which treatment exposure is defined,T_e, to show that results are qualitatively robust to choice of this parameter in a range of

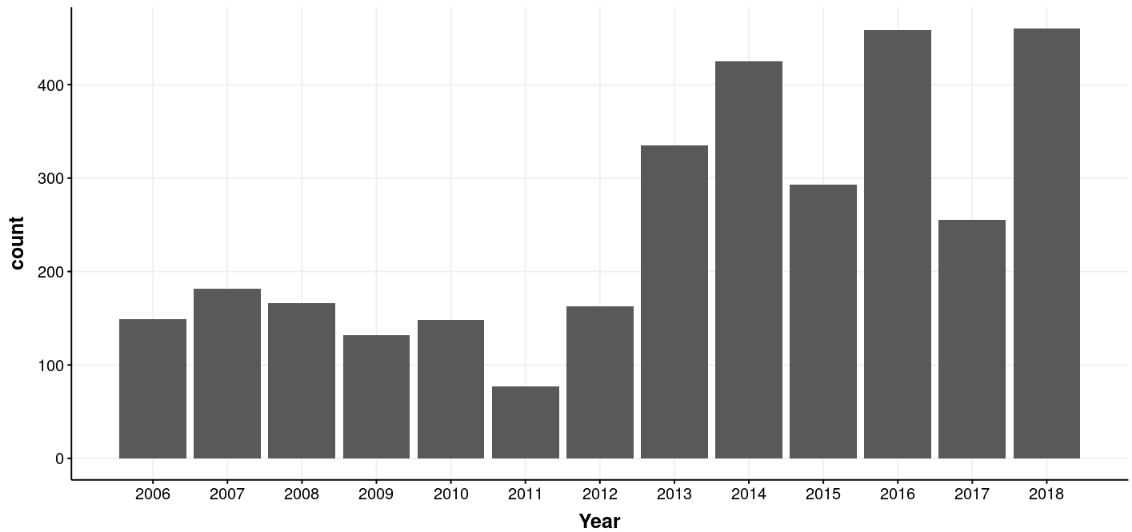


Figure A.1: Number of closing stores in each year. The fact that there are more identified closings later in the timeline is not necessarily indicator of more closures, but could rather be due to change in the Nielsen data sample over time.

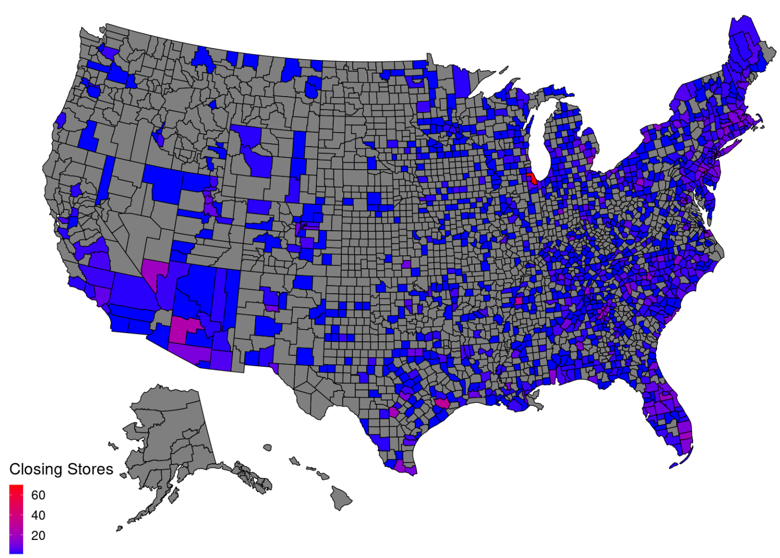


Figure A.2: The geographic dispersion of closing stores in the US at the county level.

values. In the main text, we used $T_e = 1$ year. Here, we present results for $T_e = 1/2$ year, and $T_e = 2$ years.

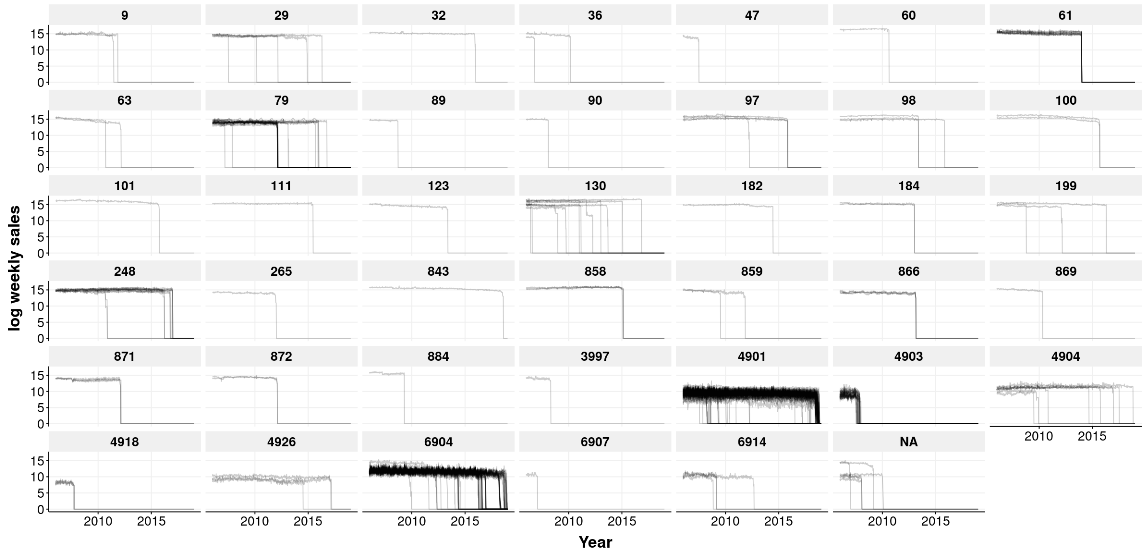


Figure A.3: The figure shows log-weekly sales in the top-5 product categories by national purchase volume for the closing stores corresponding to the set of exposed households from January 2006 until December 2018, separated based on retailers. Numbers on top of each facet show the corresponding retailer code in the Nielsen data.

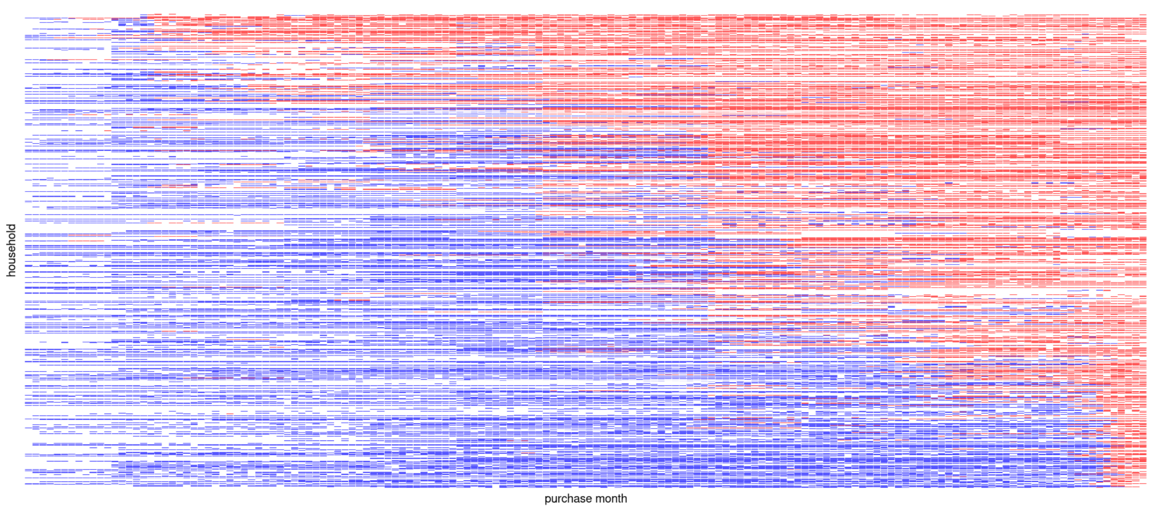


Figure A.4: Each row stands for a household, and columns corresponds to different purchase months. Blue rectangles show pre-closure and red rectangles show post-closure purchase occasions. Note that each household could be both in treatment and control groups, based on the household–category exposure level. Closures are staggered over time, and due to Bacon decomposition, units for which the closure happens in the middle of the panel have a higher weight in the TWFE fixed effects estimate.

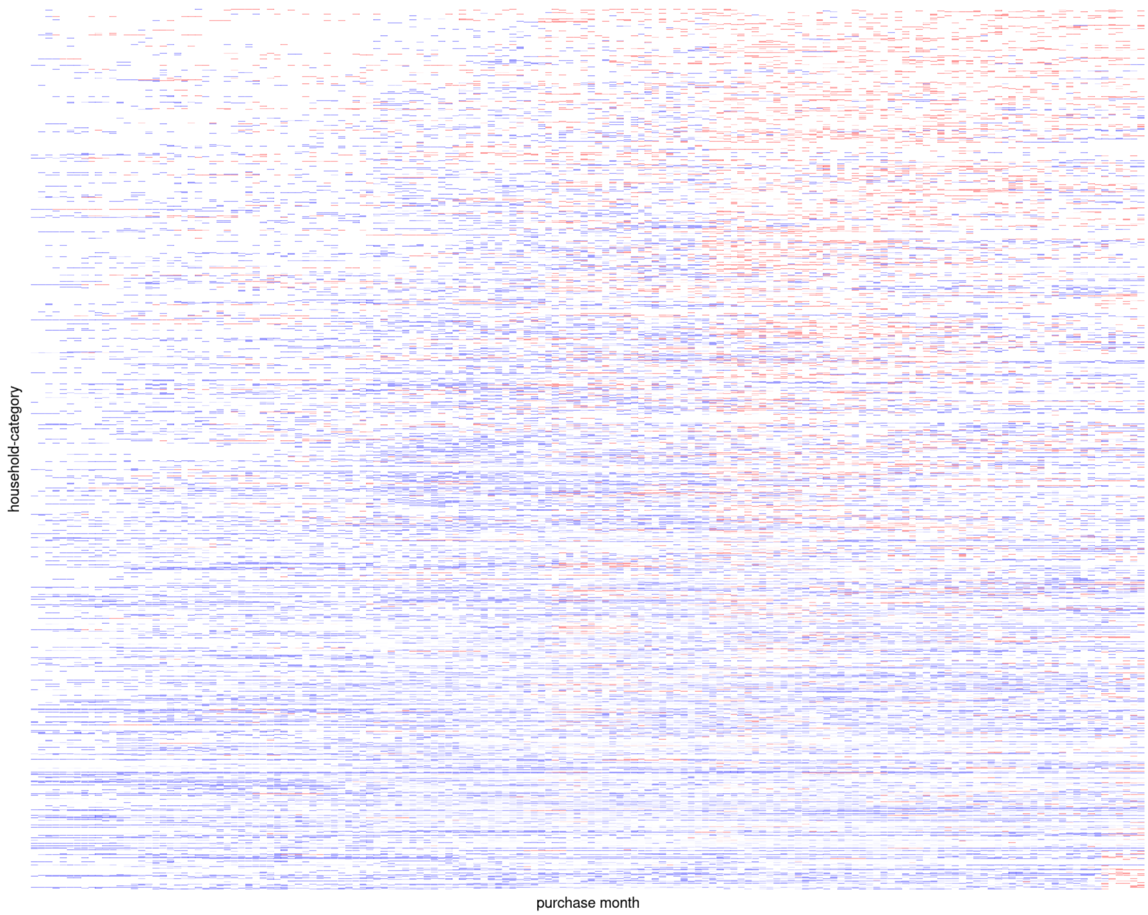


Figure A.5: In this figure each row corresponds to a household–category pair, and that is why is looks sparse. Also, in each category, only the first L trips after the closure are marked as treated (red) to reflect the short-term β_1 in Equation 2.4. Control units (household–category pairs with zero exposure level) are not shown in the figure.

2.2 Changing L

Note that L enters the analysis in two ways, first through the definition of the L-recent modal brand, and second by the short-term treatment definition (Equation 2.4). The former specifies the width of the moving window based on which the modal brand is defined, and the latter determines the short-term period after closure for which we expect habits to be disrupted. One could imagine using two different L values, but we considered a symmetric case for simplicity of presenting results. Also, note that the dynamics of short-term treatment effect is estimated more systematically with the event study models, so what really matters is the width of the moving window.

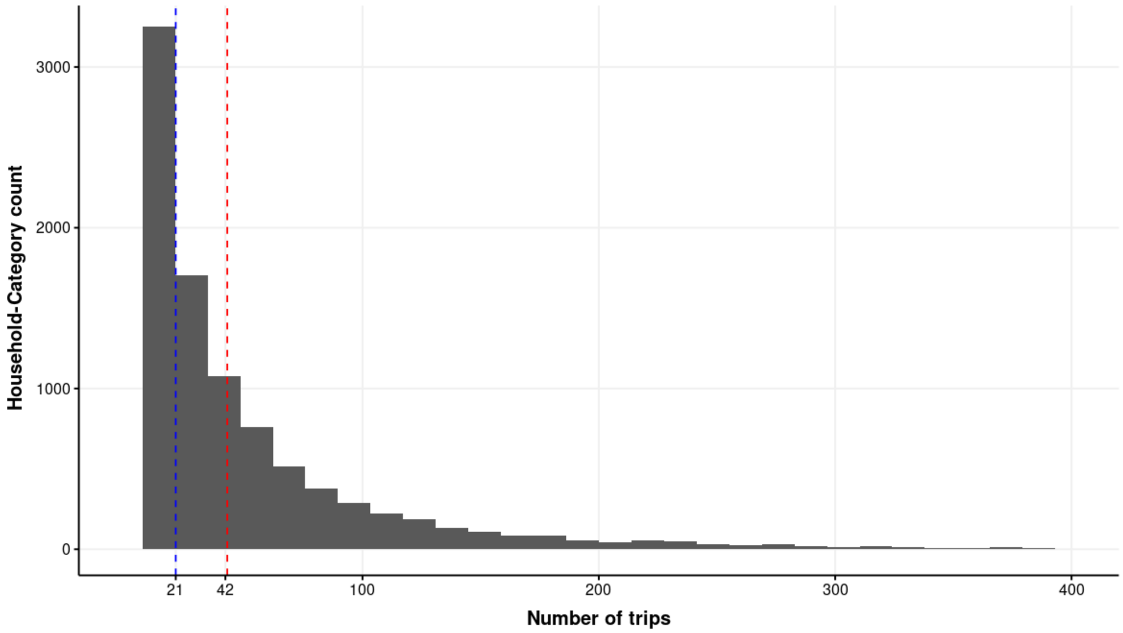


Figure A.6: The figure shows the histogram of number of shopping trips for all household–category pairs, before the corresponding store closure. Mean (red) of the distribution is 42 and the median (blue) equals 21.

3 Baseline modal brand replication results

In this appendix we replicate the results for the the baseline modal brand analysis where instead of all pre-closure period shopping trips, the baseline modal brand is defined based on the prior 40 trips in each category. Since the panel is not balanced, the number of pre-closure trips could be highly variable for different household–category pairs. So in order to make analysis comparable across different units, we use a fix length of 40 trips to define the baseline modal brand. Note that the average number of pre-closure trips is 42.

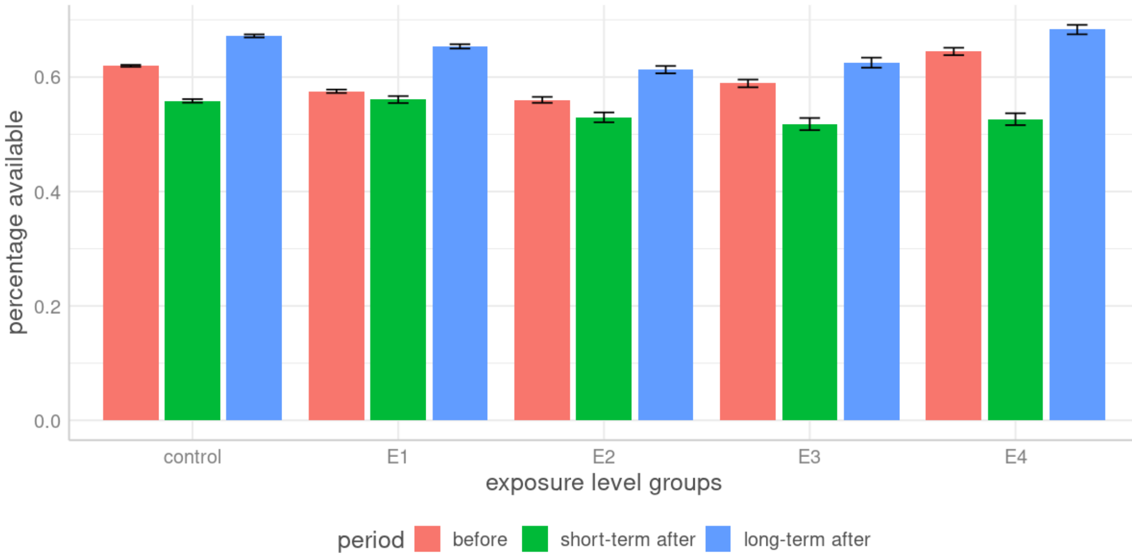


Figure A.7: Different periods are defined based on the treatment definitions in Equation 2.4. The *before* period contains the whole pre-closure trips, *short-term after* period includes the first L trips for each category (where $T^{r_1}_{i,t,c} \neq 0$), and *long-term after* marks the rest of the trips. Note that since periods are defined at the household–category level, for a certain household a trip could be in the *short-term after* period for category a, while in the *long-term after* period for category b. For each week, we mark a modal brand as available in a store if there is at least one purchase occasion by any household in the entire consumer panel data in the same store.

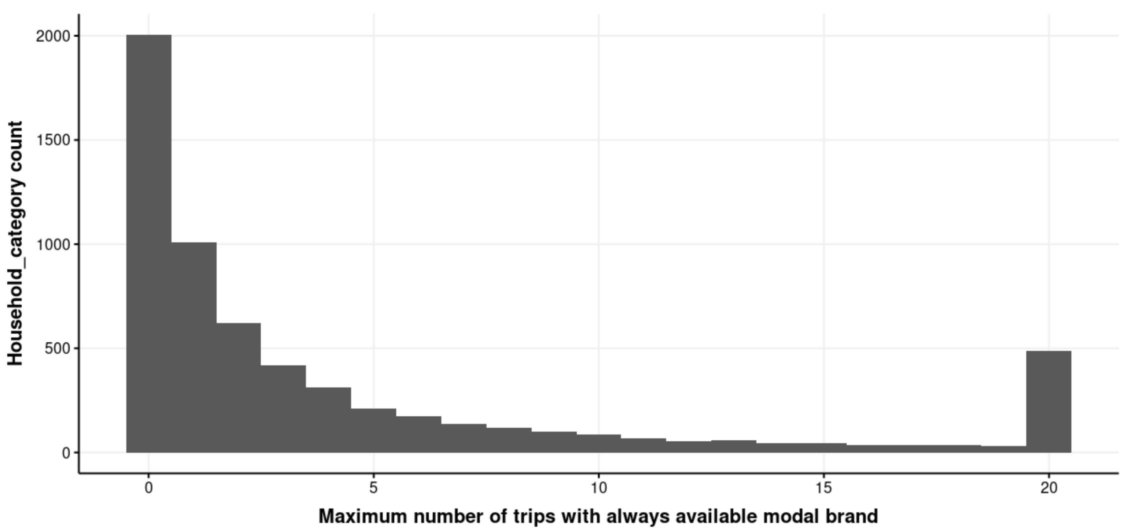
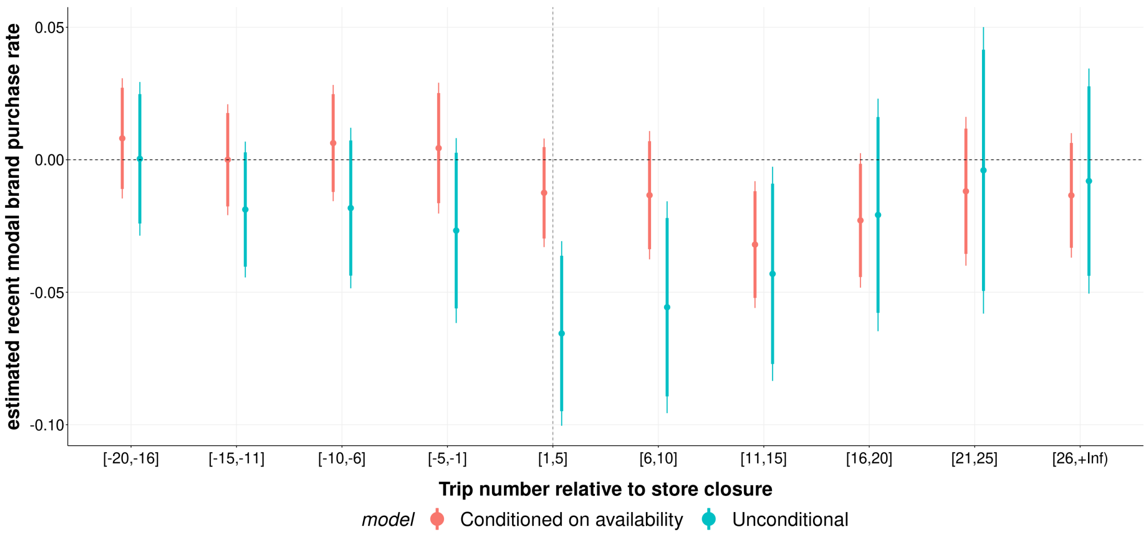


Figure A.8: Distribution of the maximum number of trips where modal brand is always available in a sequence, over all household–category pairs.



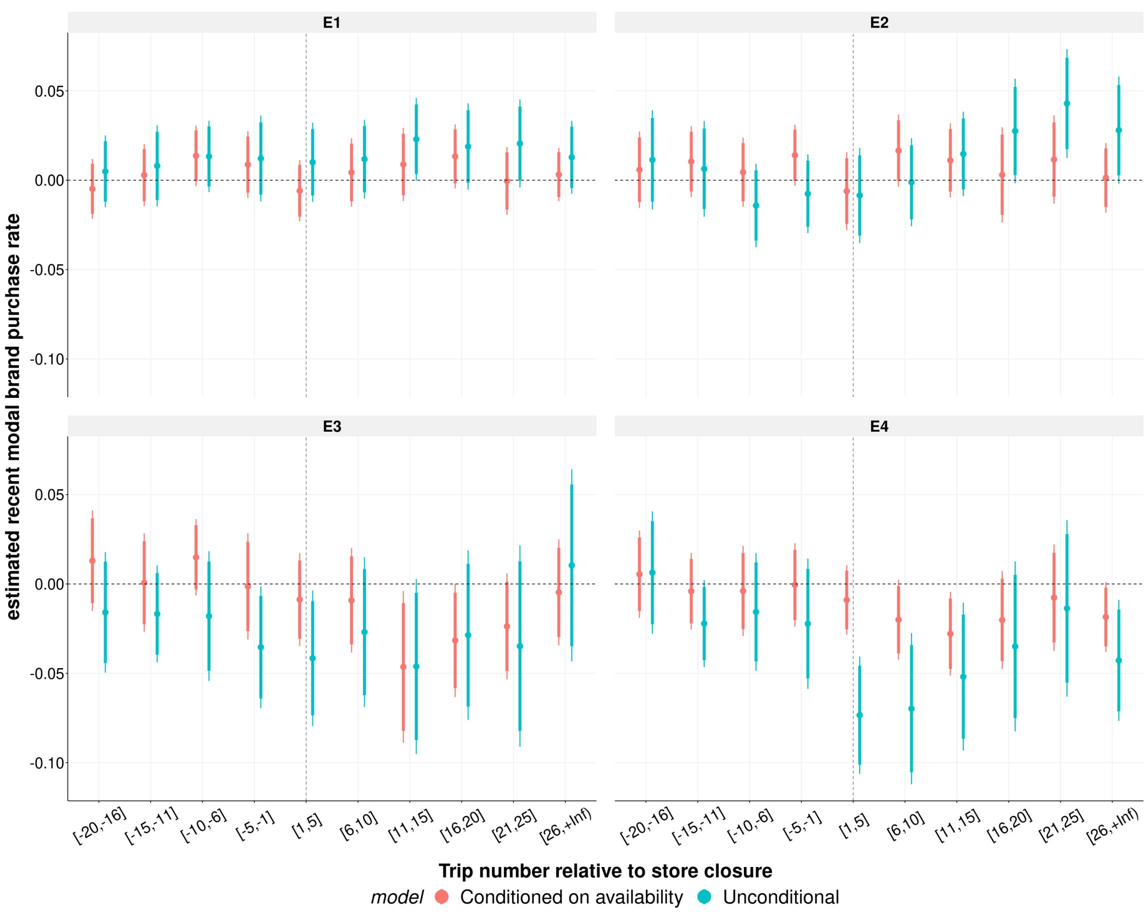
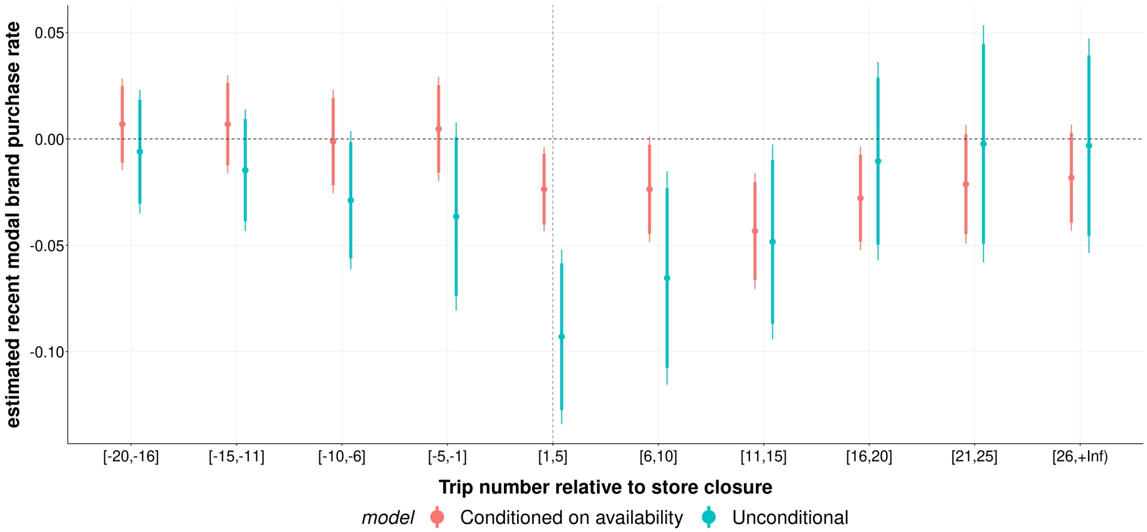


Figure A.9: Replication of the results in Figure 2.4 for $T_e = \frac{1}{2}$ year, and $L = 20$.



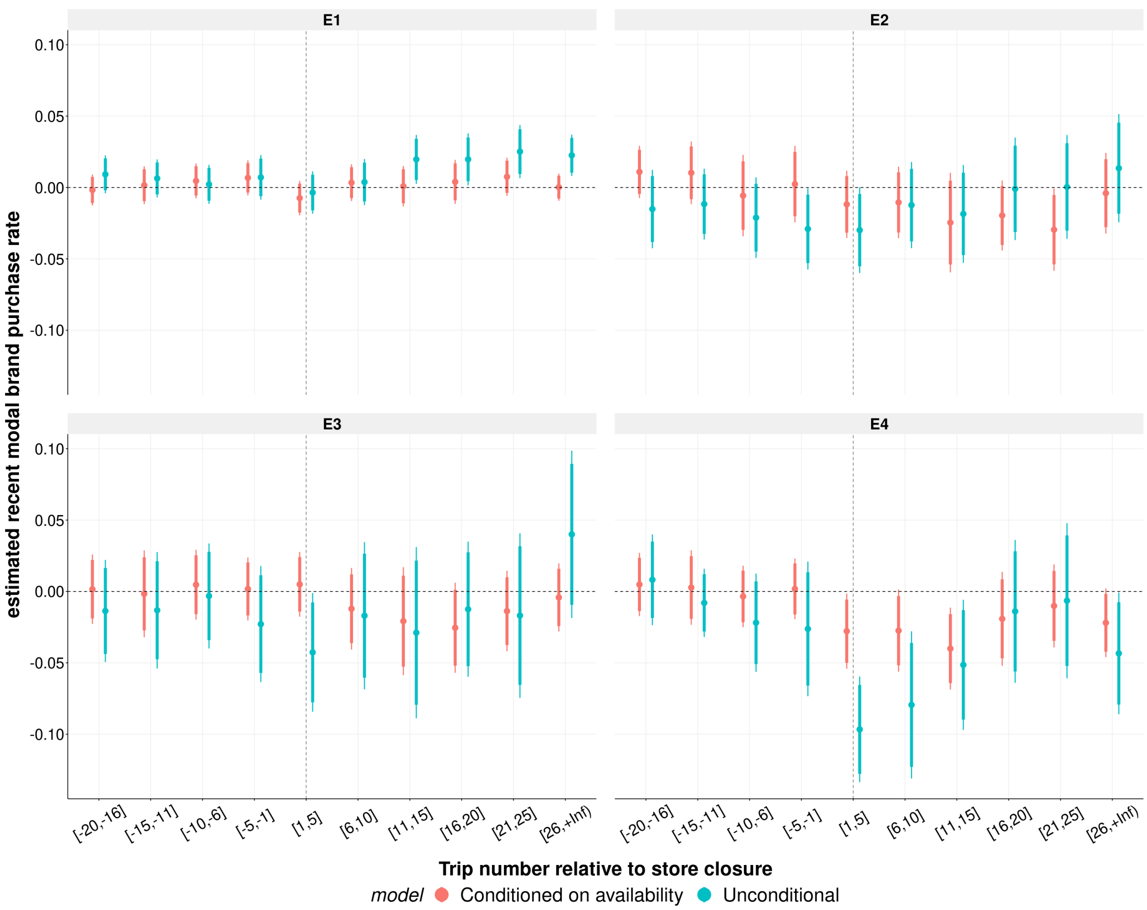


Figure A.10: Replication of the results in Figure 2.4 for $T_e = 2$ years, and $L = 20$.

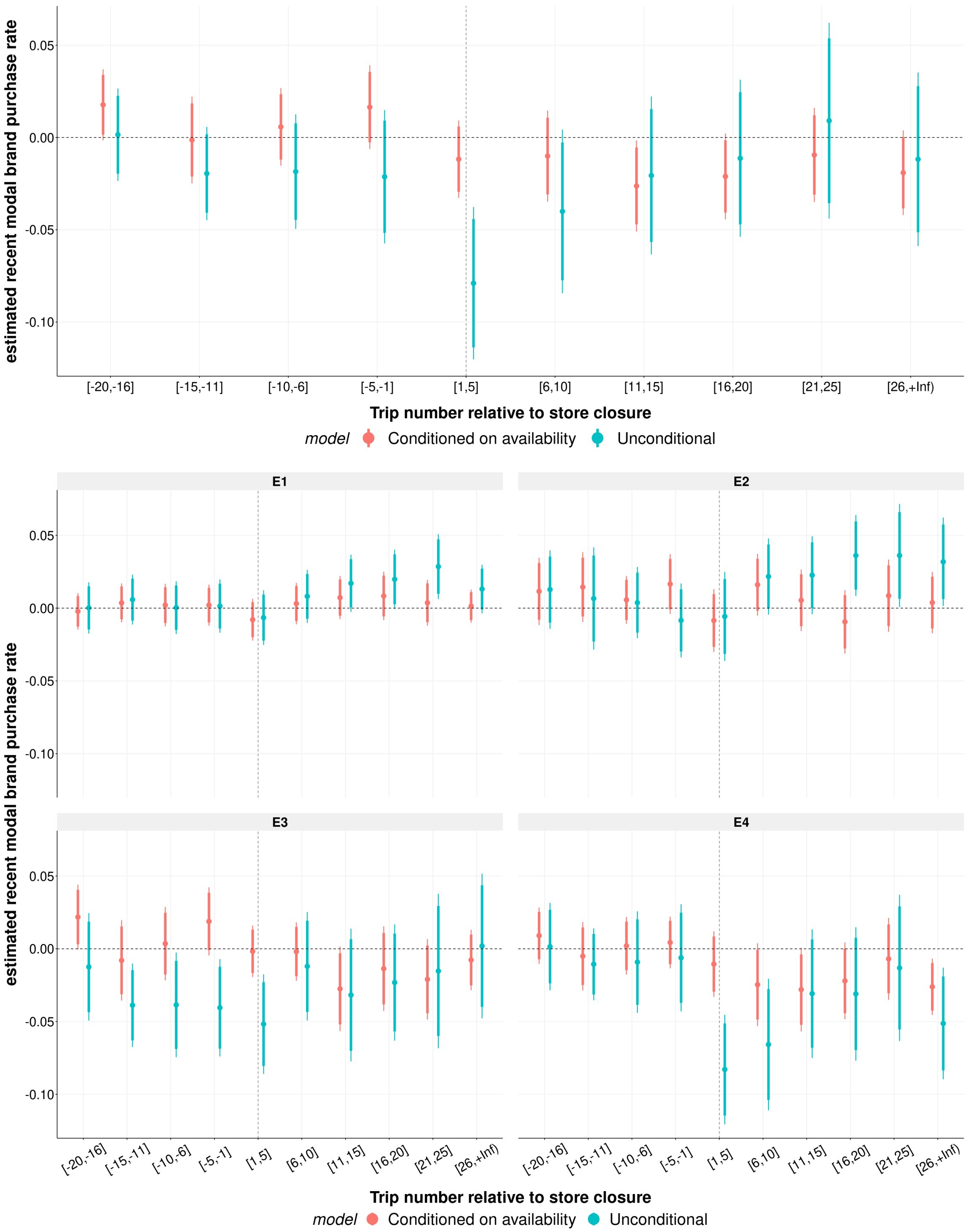
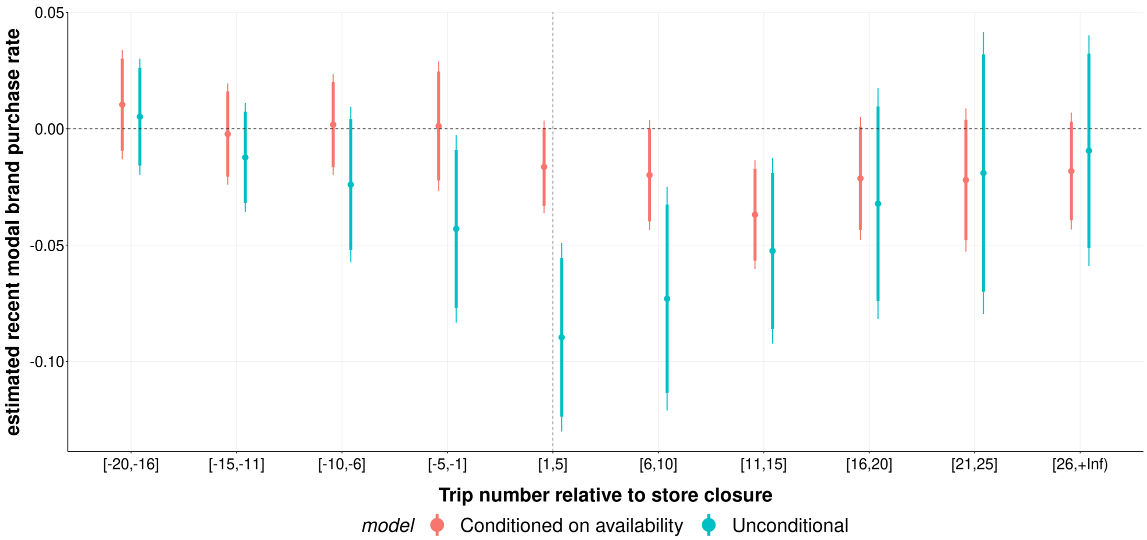


Figure A.11: Replication of the results in Figure 2.4 for $T_e = 1$ year, and $L = 10$.



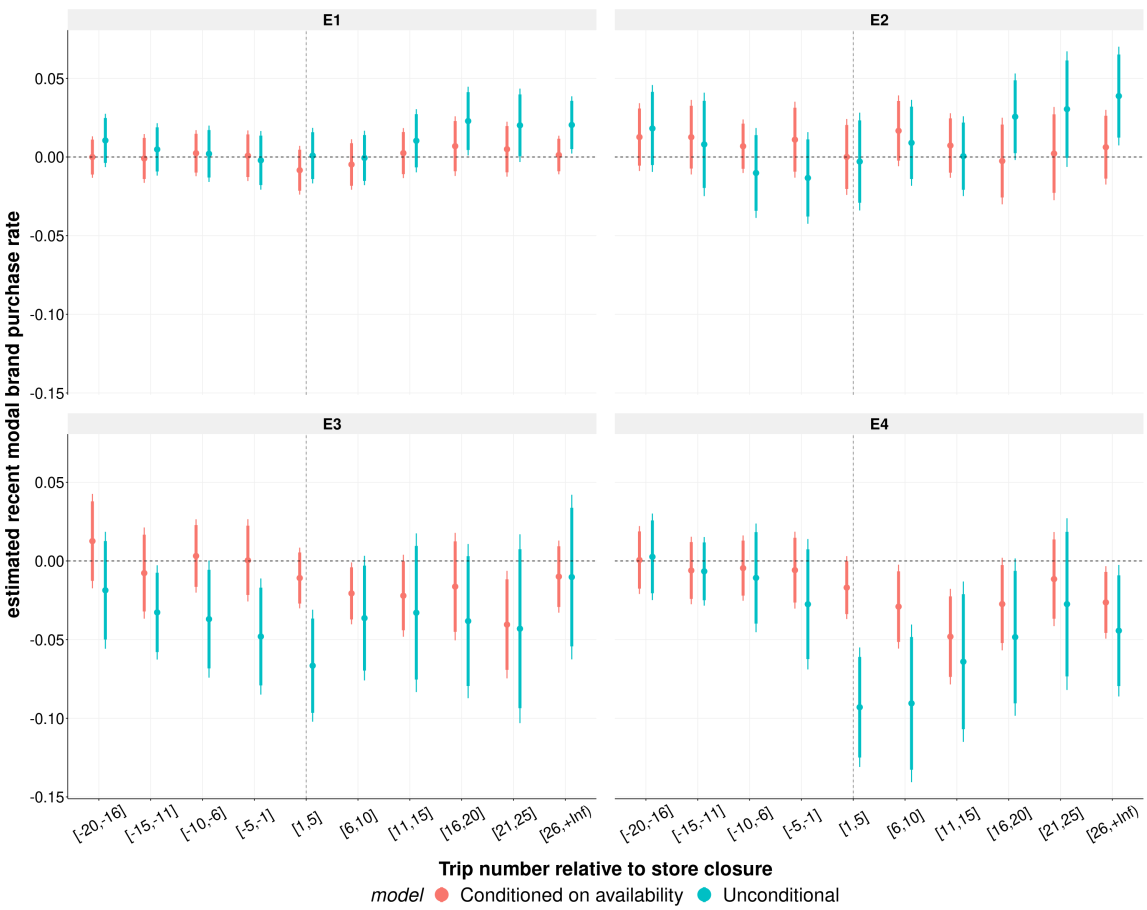


Figure A.12: Replication of the results in Figure 2.4 for $T_e = 1$ year, and $L = 30$.

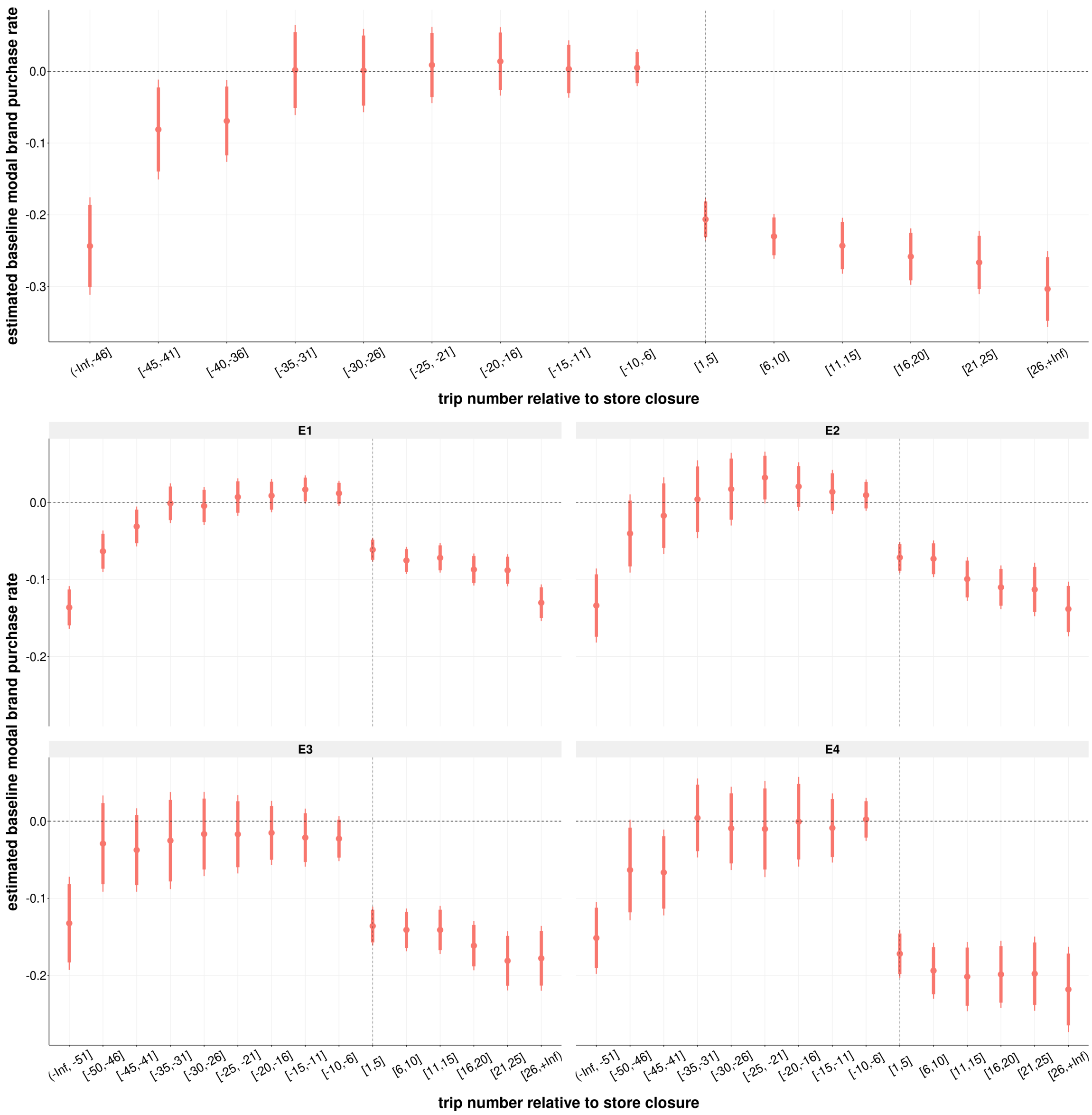


Figure A.13: Replication of the results in Figure 2.7, with the baseline modal brand defined based on only 40 trips prior to each closure.

Table A.1: Replication of the results in table 2.1 for $T_e = \frac{1}{2}$ year and $L = 20$.

	Dependent variable:			
	recent modal brand indicator (×100)			
	(1)	(2)	(3)	(4)
Overall, short-term	-3.867^{**}		-2.207^{**}	
	(1.330)		(0.676)	
Overall, long-term	-0.006		-1.402	
	(1.922)		(0.911)	
E_1, short-term		1.308		0.292
		(0.772)		(0.619)
E_2, short-term		0.757		0.055
		(0.874)		(0.699)
E_3, short-term		-2.553		-2.650^{**}
		(1.611)		(0.998)
E_4, short-term		-5.055^{***}		-1.742^{*}
		(1.276)		(0.715)
E_1, long-term		0.982		0.052
		(0.911)		(0.640)
E_2, long-term		3.022^{*}		-0.065
		(1.325)		(0.884)
E_3, long-term		1.613		-0.726
		(2.487)		(1.236)
E_4, long-term		-3.368^{*}		-1.648^{*}
		(1.414)		(0.712)
Continuous treatment	✓		✓	
Conditioned on modal brand availability			✓	✓
Observations	846,032	846,032	520,216	520,216
R^2	0.279	0.279	0.144	0.144
Adjusted R^2	0.269	0.269	0.126	0.126
Note:		*p<0.05; **p<0.01; ***p<0.001		

Table A.2: Replication of the results in table 2.1 for $T_e = 2$ years and $L = 20$.

	Dependent variable:			
	recent modal brand indicator (×100)			
	(1)	(2)	(3)	(4)
Overall, short-term	-4.568^{**}		-3.102^{***}	
	(1.715)		(0.764)	
Overall, long-term	0.582		−1.953	
	(2.300)		(1.037)	
E_1, short-term		0.621		−0.136
		(0.533)		(0.406)
E_2, short-term		−0.572		-1.767^{*}
		(1.039)		(0.826)
E_3, short-term		−2.084		−1.467
		(1.811)		(0.984)
E_4, short-term		-5.705^{***}		-2.855^{**}
		(1.621)		(0.869)
E_1, long-term		1.992^{**}		−0.076
		(0.644)		(0.400)
E_2, long-term		2.045		−0.775
		(1.604)		(1.183)
E_3, long-term		3.943		−0.340
		(2.677)		(1.059)
E_4, long-term		−3.579		-2.197^{*}
		(1.905)		(0.961)
Continuous treatment	✓		✓	
Conditioned on modal brand availability			✓	✓
Observations	899,487	899,487	547,614	547,614
R^2	0.279	0.279	0.145	0.145
Adjusted R^2	0.268	0.268	0.125	0.125
Note:		$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$		

Table A.3: Replication of the results in Table 2.1 for $T_e = 1$ year, and $L = 10$.

	Dependent variable:			
	recent modal brand indicator ($\times$100)			
	output100			
	(1)	(2)	(3)	(4)
Overall, short-term	-4.900^{**}		-1.624^{*}	
	(1.626)		(0.789)	
Overall, long-term	-0.341		-2.457^{**}	
	(1.892)		(0.844)	
E_1, short-term		0.071		-0.235
		(0.677)		(0.509)
E_2, short-term		0.846		-0.153
		(0.953)		(0.679)
E_3, short-term		-1.589		-0.797
		(1.254)		(0.848)
E_4, short-term		-6.813^{***}		-1.844^{*}
		(1.551)		(0.893)
E_1, long-term		1.499^{*}		0.168
		(0.708)		(0.449)
E_2, long-term		3.047^{*}		-0.246
		(1.210)		(0.753)
E_3, long-term		0.967		-1.620
		(1.991)		(0.965)
E_4, long-term		-3.989^{*}		-2.588^{***}
		(1.553)		(0.711)
Continuous treatment	✓		✓	
Conditioned on modal brand availability			✓	✓
Observations	883,237	883,237	544,180	544,180
R^2	0.279	0.279	0.141	0.141
Adjusted R^2	0.268	0.268	0.121	0.121
Note:				*p<0.05; **p<0.01; ***p<0.001

Table A.4: Replication of the results in Table 2.1 for $T_e = 1$ year and $L = 30$.

	Dependent variable:			
	recent modal brand indicator ($\times 100$)			
	(1)	(2)	(3)	(4)
Overall, short-term	-4.188^{*}		-2.364^{**}	
	(1.731)		(0.750)	
Overall, long-term	-0.400		-1.830	
	(2.206)		(1.066)	
E_1, short-term		1.123		0.118
		(0.684)		(0.518)
E_2, short-term		1.392		-0.191
		(1.025)		(0.818)
E_3, short-term		-2.393		-1.999^{*}
		(1.620)		(0.863)
E_4, short-term		-6.065^{***}		-2.495^{**}
		(1.615)		(0.783)
E_1, long-term		1.643		-0.119
		(0.866)		(0.572)
E_2, long-term		3.836^{**}		0.459
		(1.473)		(1.069)
E_3, long-term		0.662		-1.049
		(2.275)		(1.147)
E_4, long-term		-4.531^{*}		-2.393^{*}
		(2.013)		(0.981)
Continuous treatment	✓		✓	
Conditioned on modal brand availability			✓	✓
Observations	883,237	883,237	533,618	533,618
R^2	0.281	0.281	0.149	0.149
Adjusted R^2	0.270	0.270	0.129	0.129
Note:	*p<0.05; **p<0.01; ***p<0.001			

Table A.5: Replication of the results in Table 2.2, with the baseline modal brand defined based on only 40 trips prior to each closure.

	Dependent variable:		
	baseline modal brand indicator (×100)		
	(1)	(2)	(3)
Overall	-25.07^{***}	-8.09^{***}	
	(1.74)	(0.98)	
E_1			-5.47^{***}
			(1.07)
E_2			-8.10^{***}
			(1.59)
E_3			-11.62^{***}
			(2.20)
E_4			-15.99^{***}
			(1.78)
Continuous treatment	✓		
Observations	895,035	895,035	895,035
R^2	0.343	0.342	0.342
Adjusted R^2	0.333	0.332	0.333
Note:			*p<0.05; **p<0.01; ***p<0.001

Appendix B

Robustness Results (Chapter 3)

1 Quasi-Poisson replication results

In this section, we investigate the robustness of our results to the choice of bandwidth used in kernel density estimation to compute entropy[1]. To verify the stability of our findings, we replicate the regression analyses for the Quasi-Poisson model, as described in Equation 3.6, while varying the bandwidth values. The following tables display the estimated coefficients for entropy for different bandwidth values. As evident from the results, the coefficients for entropy remain relatively stable and consistent across various bandwidth settings. This robustness of the results demonstrates that the relationship between entropy and long-term user behavior remains significant and unaffected by changes in the bandwidth hyper-parameter. These findings further support the validity and reliability of our conclusions and lend additional credibility to the use of entropy as a reliable proxy for time-of-day behavioral regularity.

2 DBSCAN Algorithm

A detailed description of the DBSCAN algorithm's pseudo code can be found in Algorithm 1.

[1]The main results presented in Chapter 3 use a bandwidth of 10 minutes for kernel density estimation

Algorithm 1: DBSCAN Algorithm

```
Input  : D - Dataset of points
         eps - Maximum distance between points to be considered neighbors
         minPts - Minimum number of points to form a dense region
Output: Clusters - List of clusters
Procedure DBSCAN(D, eps, minPts)
    Mark all points in D as unvisited;
    Create an empty list Clusters;
    for each point p in D do
        if p is visited then
            continue;
        end
        Mark p as visited;
        neighbors ← getNeighbors(p, eps);
        if size(neighbors) < minPts then
            Mark p as noise;
        end
        else
            Create a new cluster C;
            expandCluster(p, neighbors, C, eps, minPts);
            Add C to Clusters;
        end
    end
    return Clusters;
Procedure expandCluster(p, neighbors, C, eps, minPts)
    Add p to cluster C;
    for each point q in neighbors do
        if q is not visited then
            Mark q as visited;
            neighbors_q ← getNeighbors(q, eps);
            if size(neighbors_q) ≥ minPts then
                Add neighbors_q to neighbors (merge the sets);
            end
        end
        if q is not yet a member of any cluster then
            Add q to cluster C;
        end
    end
Procedure getNeighbors(p, eps)
    neighbors ← empty set;
    for each point q in D do
        if distance(p, q) ≤ eps then
            Add q to neighbors;
        end
    end
    return neighbors;
```

Table B.1: Replication of the result in Table 3.2, with bandwidth equal to 20 minutes.

	Total Time (1)	Frequency (2)
Entropy	-0.2664***	-0.1888***
	(0.0542)	(0.0528)
Clumpiness	-4.387***	-3.855***
	(0.1572)	(0.1712)
log(Average Time)	0.8645***	0.0025
	(0.0147)	(0.0134)
log(Frequency)	0.1280**	0.9044***
	(0.0540)	(0.0607)
Recency	-0.0103***	-0.0112***
	(0.0027)	(0.0035)
Days Active	-0.0026*	-0.0046***
	(0.0014)	(0.0014)
Observations	1,352,823	1,352,823
Squared Correlation	0.48165	0.54384
Daily fixed effects	✓	✓
Application fixed effects	✓	✓
Sample size fixed effects	✓	✓

Table B.2: Replication of the result in Table 3.2, with bandwidth equal to 5 minutes.

	Total Time (1)	Frequency (2)
Entropy	-0.2370***	-0.1627***
	(0.0489)	(0.0488)
Clumpiness	-4.426***	-3.897***
	(0.1660)	(0.1715)
log(Average Time)	0.8647***	-0.0023
	(0.0152)	(0.0138)
log(Frequency)	0.0944	0.8354***
	(0.0626)	(0.0650)
Recency	-0.0109***	-0.0122***
	(0.0030)	(0.0040)
Days Active	-0.0032**	-0.0062***
	(0.0016)	(0.0015)
Observations	1,224,693	1,224,693
Squared Correlation	0.48305	0.54563
Daily fixed effects	✓	✓
Application fixed effects	✓	✓
Sample size fixed effects	✓	✓

Appendix C

Survey Designs (Chapter 4)

1 Prediction Survey

1.1 Demographics

- Q1.**age:** What is your age? {Optional}

- Q2.**agegroup:**You choose not provide your age. Please provide your age group. [18-20, 21-25, 26-30, 31-35, 36-40, 41-45, 46-50, 51-55, 56-60, 61-65, 66-70, 71-75, 76-80, 80+] {If Q1 is not answered}

- Q3.**gender:** What is your gender? [Male, Female, Non-binary]

- Q4.**race:** Do you identify with any of the following races/ethnic groups? Select all that apply. [White, African American, Latino, Asian, American Indian, Alaska Native, Native Hawaiian, Pacific Islander]

- Q5.**education:** What is the highest education level that you have completed? [None, Elementary, High-School, College, Grad School]

- Q6.**employment:** What is your employment status? [Unemployed, Self-employed, Employed, Retired]

- Q7.**urban:** Would you describe the area where you currently live as mostly rural or urban? [Rural, Sub or Ex-urban, Urban]

- Q8.**socialmedia:** How much time do you spend on social media? [[never, I never use them], [rarely, I rarely use them], [somewhat, I am a somewhat active user], [very, I am a very active user]]

- Q9.**hardship:** Hardship is a condition that causes difficulty or suffering. In the course of your life, would you say that you have experienced hardship? *(Examples are being without a job or enough money)* [Yes, No, Prefer not to say]

1.2 Political orientation.

- Q1.**followpol:** On a scale from 1 to 7, where 1 means "not at all" and 7 means "very closely," how closely do you follow US politics? [1,2,3,4,5,6,7]

- Q2.**demrep:** On a scale from 1 to 7, where 1 means "strong Democrat" and 7 means "strong Republican," where do you position yourself? [1,2,3,4,5,6,7]

- Q3.**libcons:** On a scale from 1 to 7, where 1 means "very liberal" and 7 means "very conservative," where do you position yourself? [1,2,3,4,5,6,7]

- Q4.**libertarian:** *Libertarianism is a political philosophy and movement that upholds liberty as a core principle. Libertarians seek to maximize autonomy and political freedom, emphasizing free association, freedom of choice, individualism and voluntary association.* Do you consider yourself a libertarian? [Yes, No]

- Q5.**candidate:** Which candidate did you support in the 2020 election? [Joe Biden, Donald Trump, Other, None]

- Q6.**othercandidate:** Please say the name of the other candidate that supported in the 2020 election. {If Q5=Other}

1.3 Views about climate change and the environment.

- Q1.**climate_me:** I consider myself an environmentalist. [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree] {Optional}

- Q2.**climate_manmade:** I believe that man-made climate change is occurring. [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree] {Optional}

- Q3.**climate_paris:** The United States was right to rejoin in 2021 the Paris Agreement to reduce greenhouse gas emissions. [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree] {Optional}

1.4 Views about the impact of climate change.

- Q1.**climate_worried:** How worried are you about climate change on a scale from 1 to 5, where 1 means "not worried at all" and 5 means "very worried"? [1,2,3,4,5] {Optional}

- Q2.**climate_impact_us:** How much do you think global warming will harm people in the United States? [Not at all, Only a little, A moderate amount, A great deal] {Optional}

1.5 Actions to combat climate change.

- Q1.**climate_demonstration:** Have you ever participated in any environmental public demonstration? [Yes, No] {Optional}

- Q2.**climate_donation:** During the last 12 months, have you donated time, money, or in-kind to an environmental organization? [Yes, No] {Optional}

1.6 Views about the science of climate change.

- Q1.**climate_science_consensus:** Which of the following statements comes closer to your point of view? *(Select only one)* [[happening, Most scientists think global warming is happening.], [disagreement, There is a lot of disagreement among scientists about whether or not global warming is happening.], [not happening, Most scientists think global warming is not happening.]] {Optional}

- Q2.**climate_science_trust:** In general, how much do you trust the science on global warming? [Not at all, A little, A moderate amount, A lot, A great deal] {Optional}

1.7 Attention Check Questions

- Q1.**abouttopic:** Would you say that climate change is the **main** topic of the article above? *Answer **"Yes"** if the article discusses climate change **at length**, either in real or fictional terms or if the article denies the existence of climate change. Topics that the article may discuss include but are not limited to: environmental, socio-economic, geopolitical causes or consequences of climate change. Answer **"No"** if the article is completely unrelated to climate change, or it mentions climate change only in passing.* [No, Yes]

- Q2.**aboutus:** Would you say that it discusses climate change in the context of **the United States**? *Answer **"Yes"** if the article contains references to places, persons, institutions, products, or ideas that are likely to be familiar to an **average** person living the United States. Answer **"No"** if the article does not specifically discuss climate change in the context of the United States, or if it mentions the United States only in passing.* [No, Yes] {If Q1=Yes}

- Q3.**topicprob:** According to the article, climate change is or could soon be: [Not a problem at all, A very serious problem] {If Q1=Yes}

1.8 Label

- Q1.**abouttopic:** Would you say that climate change or global warming is the **core** topic of the article above? *Answer **"Yes"** if the article discusses climate change **at length**, either in real or fictional terms or if the article denies the existence of climate change. Topics that the article may discuss include but are not limited to: environmental, socio-economic, geopolitical causes or consequences of climate change. Answer **"No"** if the article is completely unrelated to climate change, or it mentions climate change only in passing. Answer **"I am not sure"** if you are not sure.* [No, Yes, I am not sure]

- Q2.**aboutus:** Would you say that it discusses climate change in the context of **the United States**? *Answer **"Yes"** if the article contains references to places, persons, institutions, products, or ideas that are likely to be familiar to an **average** person living the United States. Answer **"No"** if the article does not specifically discuss climate change in the context of the United States, or if it mentions the United States only in passing. Answer **"I am not sure"** if you are not sure.* [No, Yes, I am not sure] {If Q1= Yes or I am not sure}

- Q3.**topicprob:** According to the article, climate change is or could soon be: [[No, Not a problem at all], [Small, A small problem], [Problem, A problem], [Serious, A serious problem], [Very Serious, A very serious problem]] {If Q1= Yes or I am not sure}

- Q4.**feedback_problem:** Please motivate your choice to the previous question. *You can give examples or citations from the article text.* {If Q1= Yes or I am not sure}

- Q5.**arguments:** What kinds of arguments does the article use to discuss climate change? *Answer **"Facts and Statistics"** if the article cites expertsór scientific evidence. Answer **"Personal Stories"** if the article focuses on real or fictional experiences. Answer **"Morality"** if the arguments rely on moral or ethical principles.* [Facts and Statistics, Personal Stories, Morality] {If Q1= Yes or I am not sure}

- Q6.**actions:** Does the article invite the reader to take actions with regards to climate change? [No, Yes] {If Q1= Yes or I am not sure}

- Q7.**like:** Did you like the article? [Strongly disliked, Disliked, Was indifferent, Liked, Strongly liked] {If Q1= Yes or I am not sure}

- Q8.**truthful:** Do you think the article contains any misleading, false, or inaccurate claims? [No, Yes, I am not sure] {If Q1= Yes or I am not sure}

- Q9.**guess:** Consider now three different types of readers from the United States in 2021. Each reader holds a different view on climate change and what to do about it. **How do you think their views towards climate change will change after having read the article?** {If Q1= Yes or I am not sure}

  - Q1.**guess_against:** Reader 1 believes that climate change is **not a problem** and is **opposing** national and international actions to combat climate change. After reading the article Reader 1 will: [[less favorable, Be even more opposed to actions to combat climate change.], [no change, Not change his/her beliefs.], [more favorable, Be less opposed to actions to combat climate change.]]

  - Q2.**guess_indifferent:** Reader 2 is **undecided** whether climate change is a problem or not and is **indifferent** to national and international actions to combat climate change. After reading the article Reader 2 will: [[less favorable, Become opposed to actions to combat climate change.], [no change, Not change his/her beliefs.], [more favorable, Become supportive of actions to combat climate change.]]

  - Q3.**guess_infavor:** Reader 3 believes that climate change is **a problem** and **supports** national and international actions to combat climate change. After reading the article Reader 3 will: [[less favorable, Be less supportive of actions to combat climate change.], [no change, Not change his/her beliefs.], [more favorable, Be even more supportive of actions to combat climate change.]]

- Q10.**feedback_guess:** Think about your answers to the previous question. What features of the article do you think could determine the three readers above to change or not change their beliefs about climate change? {If Q1= Yes or I am not sure}

1.9 Feedback

Thank you for participating. This was a pilot of the main study, therefore we are very interested in hearing your **feedback** about the following points:

1. Was the task too long or too short?
2. Did you feel you could express your opinion?
3. Did you find any question unclear or uncomfortable?
4. Did you feel that the survey was balanced, or rather biased towards the left or right?
5. Did you experience any technical difficulty?
6. How can we improve the study?

2 Experiment

2.1 Demographics

1. What is your age? [Optional. Numeric > 18]

2. You chose not to provide your age. Please provide your age group. [Age bins of 3 years]

3. What is your gender? [Male, Female, Non-binary]

4. Do you identify with any of the following races/ethnic groups? Select all that apply. [White, African American, Latino, Asian, American Indian, Alaska Native, Native Hawaiian, Pacific Islander]

5. What is the highest education level that you have completed? [None, Elementary, High-School, College, Grad School]

6. What is your employment status? [Unemployed, Self-employed, Employed, Retired]

7. How would you describe the area where you currently live? [Rural, Sub or Ex-urban, Urban]

8. How much time do you spend on social media? [I never use them, I rarely use them, I am a somewhat active user, I am a very active user]

2.2 Politics

1. On a scale from 1 to 7, where 1 means "not at all" and 7 means "very closely," how closely do you follow US politics? [1-7]

2. On a scale from 1 to 7, where 1 means "strong Democrat" and 7 means "strong Republican," where do you position yourself? [1-7]

3. On a scale from 1 to 7, where 1 means "very liberal" and 7 means "very conservative," where do you position yourself? [1-7]

4. Do you consider yourself a libertarian? Libertarianism is a political philosophy and movement that upholds liberty as a core principle. Libertarians seek to maximize autonomy and political freedom, emphasizing free association, freedom of choice, individualism and voluntary association. [Yes, No]

5. Which candidate did you support in the 2020 election? [Joe Biden, Donald Trump, Other, None]

6. Please say the name of the other candidate that you supported in the 2020 election. [If Other]

2.3 Climate Change Stance

1. I consider myself an environmentalist. [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree]

2. I believe that man-made climate change is occurring. [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree]

3. The United States was right to rejoin in 2021 the Paris Agreement to reduce greenhouse gas emissions. [I strongly disagree, I disagree, Neither agree nor disagree, I agree, I strongly agree]

4. How worried are you about climate change on a scale from 1 to 5, where 1 means "not worried at all" and 5 means "very worried"? [1-5]

5. How much do you think global warming will harm people in the United States? [Not at all, Only a little, A moderate amount, A great deal]

6. Have you ever participated in any environmental public demonstration? [Yes, No]

7. During the last 12 months, have you donated time, money, or in-kind to an environmental organization? [Yes, No]

8. Which of the following statements comes closer to your point of view? (Select only one) [Most scientists think global warming is happening, There is a lot of disagreement among scientists about whether or not global warming is happening, Most scientists think global warming is not happening]

9. In general, how much do you trust the science on global warming? [Not at all, A little, A moderate amount, A lot, A great deal]

10. On a scale from 1 to 10, where 1 is "fully disagree" and 10 is "fully agree," how much do you agree with the following statement: Climate change is a pressing problem and we urgently need national and international actions to combat it. [1-10]

2.4 Beliefs

On a scale from 1 to 10, where 1 is "fully disagree" and 10 is "fully agree," how much do you agree with the following statements:

- If nothing is done to limit climate change, there will be dire consequences for humanity in the not-distant future. [1-10]
- We urgently need national and international actions to combat climate change. [1-10]
- I am worried about climate change. [1-10]
- Human-caused climate change is real and it is occurring. [1-10]

2.5 Policies

1. On a scale from 1 to 10, where 1 is "fully disagree" and 10 is "fully agree," what is your level of support for the following governmental policies to mitigate climate change?

- Subsidize the insulation of buildings to make homes more energy efficient. [1-10]
- Subsidize the development and use of low-carbon technologies (e.g., renewable energy, capture and storage of carbon, etc.). [1-10]
- Impose a carbon tax on all products proportional to the amount of CO2 emitted for producing them. [1-10]
- Increase fuel duty, the tax motorists pay for petrol and diesel. [1-10]

2.6 Personal Actions

On a scale from 1 to 10, where 1 is "I would never do it" and 10 is "I would certainly do it," what is your level of support for the following climate actions?

- Increase walking, cycling, or using public transport instead of driving. [1-10]
- Use only green electricity, that is electricity produced by renewable energy, even if it costs more. [1-10]
- Vote for a candidate who is vocal about climate change issues. [1-10]
- Make a significant donation to an environmental cause. [1-10]

2.7 Commitment

Before proceeding to the next set of questions, we want to ask for your feedback about the responses you provided so far. It is vital to our study that we only include responses from people who devoted their full attention to this study. This will not affect in any way the payment you will receive for taking this survey. In your honest opinion, should we use your responses, or should we discard your responses since you did not devote your full attention to the questions so far? Please answer:

- (Yes) I have devoted full attention to the questions so far, and I think you should use my responses for your study.

- (No) I have not devoted full attention to the questions so far, and I think you should not use my responses for your study.

2.8 Donation

A decision about the bonus reward: In this additional task, you receive a bonus of [BONUS AMOUNT] USD. On the next page, you will have the opportunity to contribute part of your bonus to a non-governmental organization (NGO) that either supports or is against climate actions in the US and worldwide. In case you decide to contribute, we will transfer the amount you choose to the selected organization. The rest of the bonus, together with the base pay of the task, will be paid out to you via Prolific. Note: we evaluate whether your responses are rushed, inattentive, or otherwise negligent, and this might affect your earnings. Once the study has concluded, contribution receipts will be displayed on the webpage of our research group (webpage link) under study ID 1334C4CH.

- Which non-governmental initiative do you want to support? Note: some organizations are advocating for climate actions, while others are not.
  - (SUPPORTS climate actions) The Sierra Foundation promotes climate solutions combining strategic philanthropy and grassroots advocacy. More info.
  - (SUPPORTS climate actions) Earth Justice is a public interest environmental law organization fighting for climate goals in court. More info.
  - (AGAINST climate actions) Americans for Prosperity believes freedom and opportunity are the keys to unleashing prosperity for all. More info.
  - (AGAINST climate actions) Heartland aims to develop and promote free-market solutions to social and economic problems. More info.
  - (I do not wish to make a donation.)
- How much of the bonus reward of [BONUS AMOUNT] USD do you want to contribute? Please move the slider to your preferred contribution amount. Your

contribution will be given to the initiative of your choice. The rest will go to you.

- Please confirm your contribution. By clicking the "Yes, I confirm my contribution of [CHOSEN DONATION] USD to the Americans for Prosperity." button below, you authorize Dr. Anca Balietti (Heidelberg University, anca.balietti@awi.uni-heidelberg.de) to transfer the specified amount to the selected organization. If you click "No, I do not consent," we will not be able to make the transfer in your name.
  - (Yes, I confirm my contribution of [CHOSEN DONATION] USD to [CHOSEN NGO].)
  - (No, I do not consent.)

2.9 Emotions

Please take a moment to reflect on your emotions during this task. How strongly or weakly did you experience each of the following emotional states? (Provide your answer on a scale from 1 to 7, where 1 is "not at all" and 7 is "very much")

- Upset
- Hostile
- Alert
- Ashamed
- Inspired
- Nervous
- Determined
- Attentive

- Afraid
- Active

2.10 Feedback

Thank you for participating. We are very interested in your feedback about the following points:

1. Was the task too long or too short?
2. Did you feel you could express your opinion?
3. Did you find any question unclear or uncomfortable?
4. Did you feel that the survey was balanced or rather biased towards the left or right?
5. Did you experience any technical difficulty?
6. How can we improve the study? (at least 50 characters)

Bibliography

Aarts H, Dijksterhuis A (2000) Habits as knowledge structures: automaticity in goal-directed behavior. *Journal of Personality and Social Psychology* 78(1):53.

Aarts H, Paulussen T, Schaalma H (1997) Physical exercise habit: on the conceptualization and formation of habitual health behaviours. *Health education research* 12(3):363–374.

Aarts H, Verplanken B, Van Knippenberg A (1998) Predicting behavior from actions in the past: Repeated decision making or a matter of habit? *Journal of applied social psychology* 28(15):1355–1374.

Adamic LA, Glance N (2005) The political blogosphere and the 2004 us election: divided they blog. *Proceedings of the 3rd international workshop on Link discovery*, 36–43.

Adamowicz WL, Swait JD (2013) Are food choices really habitual? Integrating habits, variety-seeking, and compensatory choice in a utility-maximizing framework. *American Journal of Agricultural Economics* 95(1):17–41.

Adams CD, Dickinson A (1981) Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B* 33(2b):109–121.

Agarwal S, Ananthakrishnan UM, Tucker CE (2022) Deplatforming and the control of misinformation: Evidence from parler. *Available at SSRN* .

Ajzen I (2002) Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and social psychology review* 6(2):107–122.

Allcott H, Gentzkow M, Song L (2022) Digital addiction. *American Economic Review* 112(7):2424–63.

Amir O, Levav J (2008) Choice construction versus preference construction: The instability of preferences learned in context. *Journal of Marketing Research* 45(2):145–158.

Anderson IA, Wood W (2021) Habits and the electronic herd: The psychology behind social media's successes and failures. *Consumer Psychology Review* 4(1):83–99.

Apley DW, Zhu J (2020) Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82(4):1059–1086, URL http://dx.doi.org/https://doi.org/10.1111/rssb.12377.

Araujo T, Wonneberger A, Neijens P, de Vreese C (2017) How much time do you spend online? understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods and Measures* 11(3):173–190.

Astington JW, Jenkins JM (1995) Theory of mind development and social understanding. *Cognition & Emotion* 9(2-3):151–165.

Bae M (2018) Understanding the effect of the discrepancy between sought and obtained gratification on social networking site users' satisfaction and continuance intention. *Computers in human behavior* 79:137–153.

Baek EC, Scholz C, O'Donnell MB, Falk EB (2017) The value of sharing information: a neural account of information transmission. *Psychological science* 28(7):851–861.

Bagozzi RP (1981) Attitudes, intentions, and behavior: A test of some key hypotheses. *Journal of personality and social psychology* 41(4):607.

Bail CA, Argyle LP, Brown TW, Bumpus JP, Chen H, Hunzaker MF, Lee J, Mann M, Merhout F, Volfovsky A (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115(37):9216–9221.

Baker AC, Larcker DF, Wang CC (2022) How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics* 144(2):370–395.

Bakshy E, Rosenn I, Marlow C, Adamic L (2012) The role of social networks in information diffusion. *Proceedings of the 21st international conference on World Wide Web*, 519–528.

Bandura A (1999) Social cognitive theory: An agentic perspective. *Asian journal of social psychology* 2(1):21–41.

Bányai F, Zsila Á, Király O, Maraz A, Elekes Z, Griffiths MD, Andreassen CS, Demetrovics Z (2017) Problematic social media use: Results from a large-scale nationally representative adolescent sample. *PloS one* 12(1):e0169839.

Barasch A, Berger J (2014) Broadcasting and narrowcasting: How audience size affects what people share. *Journal of Marketing Research* 51(3):286–299.

Barberá P, Boydstun AE, Linn S, McMahon R, Nagler J (2021) Automated text classification of news articles: A practical guide. *Political Analysis* 29(1):19–42.

Baron-Cohen S (1997) *Mindblindness: An essay on autism and theory of mind* (MIT press).

Bayer JB, Anderson IA, Tokunaga R (2022) Building and breaking social media habits. *Current Opinion in Psychology* 101303.

Bayer JB, Campbell SW (2012) Texting while driving on automatic: Considering the frequency-independent side of habit. *Computers in Human Behavior* 28(6):2083–2090.

Bayer JB, LaRose R (2018) Technology habits: Progress, problems, and prospects. *The psychology of habit: Theory, mechanisms, change, and contexts* 111–130.

Bayley PJ, Frascino JC, Squire LR (2005) Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature* 436(7050):550–553.

Bekalu MA, McCloud RF, Viswanath K (2019) Association of social media use with social well-being, positive mental health, and self-rated health: disentangling routine use from emotional connection to use. *Health Education & Behavior* 46(2_suppl):69S–80S.

Bennett WL, Iyengar S (2008) A new era of minimal effects? the changing foundations of political communication. *Journal of communication* 58(4):707–731.

Bergé L (2018) Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discussion Papers* (13).

Berger PD, Nasr NI (1998) Customer lifetime value: Marketing models and applications. *Journal of interactive marketing* 12(1):17–30.

Bernays EL (1928) *Propaganda* (Ig publishing).

Berridge KC (2021) Comment on vandaele and ahmed: Rethinking habits in addiction. *Neuropsychopharmacology* 46(4):687–688.

Beshears J, Lee HN, Milkman KL, Mislavsky R, Wisdom J (2021) Creating exercise habits using incentives: The trade-off between flexibility and routinization. *Management science* 67(7):4139–4171.

Birch SA, Bloom P (2007) The curse of knowledge in reasoning about false beliefs. *Psychological Science* 18(5):382–386.

Birnbaum MH (2004) Human research and data collection via the internet. *Annu. Rev. Psychol.* 55:803–832.

Boxell L, Gentzkow M, Shapiro JM (2022) Cross-country trends in affective polarization. *Review of Economics and Statistics* 1–60.

Braghieri L, Levy R, Makarin A (2022) Social media and mental health .

Broockman DE, Kalla JL, Sekhon JS (2017) The design of field experiments with survey outcomes: A framework for selecting more efficient, robust, and ethical designs. *Political Analysis* 25(4):435–464.

Busch-Geertsema A, Lanzendorf M (2017) From university to work life – Jumping behind the wheel? Explaining mode change of students making the transition to professional life. *Transportation Research Part A: Policy and Practice* 106:181–196, ISSN 09658564, URL http://dx.doi.org/10.1016/j.tra.2017.09.016.

Buyalskaya A, Ho H, Milkman KL, Li X, Duckworth AL, Camerer C (2023) What can machine learning teach us about habit formation? evidence from exercise and hygiene. *Proceedings of the National Academy of Sciences* 120(17):e2216115120.

Callaway B, Sant'Anna PH (2020) Difference-in-differences with multiple time periods. *Journal of Econometrics* .

Camerer C, Loewenstein G, Weber M (1989) The curse of knowledge in economic settings: An experimental analysis. *Journal of political Economy* 97(5):1232–1254.

Cameron AC, Trivedi PK (2013) *Regression analysis of count data*, volume 53 (Cambridge university press).

Cantoni E, Pons V (2022) Does context outweigh individual characteristics in driving voting behavior? Evidence from relocations within the united states. *American Economic Review* 112(4):1226–72.

Carrasco R, Labeaga JM, David López-Salido J (2004) Consumption and Habits: Evidence from Panel Data. *The Economic Journal* 115(500):144–165, ISSN 0013-0133, URL http://dx.doi.org/10.1111/j.1468-0297.2004.00963.x.

Cattaneo MD, Crump RK, Farrell MH, Feng Y (2019a) Binscatter regressions. *arXiv preprint arXiv:1902.09615* .

Cattaneo MD, Crump RK, Farrell MH, Feng Y (2019b) On binscatter. *arXiv preprint arXiv:1902.09608* .

Ceylan G, Anderson IA, Wood W (2023) Sharing of misinformation is habitual, not just lazy or biased. *Proceedings of the National Academy of Sciences* 120(4):e2216614120.

Chetty R (2015) Behavioral economics and public policy: A pragmatic perspective. *American Economic Review* 105(5):1–33.

Chiou L, Tucker C (2012) How does the use of trademarks by third-party sellers affect online search? *Marketing Science* 31(5):819–837.

Chong D, Druckman JN (2007) Framing theory. *Annu. Rev. Polit. Sci.* 10:103–126.

Chou HTG, Edge N (2012) "they are happier and having better lives than i am": The impact of using facebook on perceptions of others' lives. *Cyberpsychology, behavior, and social networking* 15(2):117–121.

Clark B, Chatterjee K, Melia S (2016) Changes to commute mode: The role of life events, spatial context and environmental attitude. *Transportation Research Part A: Policy and Practice* 89:89–105.

Clutterbuck RA, Callan MJ, Shah P (2023) Socio-demographic and political predictors of theory of mind in adulthood. *Plos one* 18(5):e0284960.

Conover M, Ratkiewicz J, Francisco M, Gonçalves B, Menczer F, Flammini A (2011) Political polarization on twitter. *Proceedings of the international aaai conference on web and social media*, volume 5, 89–96.

Converse PE (1964) The nature of belief systems in mass publics. *Critical review* 18(1-3):1–74.

Cooper R, Tang T (2009) Predicting audience exposure to television in today's media environment: An empirical integration of active-audience and structural theories. *Journal of Broadcasting & Electronic Media* 53(3):400–418.

Cunningham RM (1956) Brand loyalty — What, where, how much. *Harvard Business Review* 34(1):116–128.

Danner UN, Aarts H, De Vries NK (2008) Habit vs. intention in the prediction of future behaviour: The role of frequency, context stability and mental accessibility of past behaviour. *British Journal of Social Psychology* 47(2):245–265.

de Wit S, Watson P, Harsay HA, Cohen MX, van de Vijver I, Ridderinkhof KR (2012) Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *Journal of Neuroscience* 32(35):12066–12075.

Dechezleprêtre A, Fabre A, Kruse T, Planterose B, Chico AS, Stantcheva S (2022) Fighting climate change: International attitudes toward climate policies. Technical report, National Bureau of Economic Research.

DellaVigna S, Gentzkow M (2010) Persuasion: empirical evidence. *Annu. Rev. Econ.* 2(1):643–669.

Dew R, Ascarza E, Netzer O, Sicherman N (2021) Detecting routines in ride-sharing: Implications for customer management. *Available at SSRN 3982612* .

Dickinson A (1985) Actions and habits: The development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 308(1135):67–78.

Dickinson A, Balleine B (1994) Motivational control of goal-directed action. *Animal Learning & Behavior* 22(1):1–18.

Dienlin T, Masur PK, Trepte S (2017) Reinforcement or displacement? the reciprocity of ftf, im, and sns communication and their effects on loneliness and life satisfaction. *Journal of Computer-Mediated Communication* 22(2):71–87.

Dietvorst RC, Verbeke WJ, Bagozzi RP, Yoon C, Smits M, Van Der Lugt A (2009) A sales force–specific theory-of-mind scale: Tests of its validity by classical methods and functional magnetic resonance imaging. *Journal of Marketing Research* 46(5):653–668.

Dong X, Morozov I, Seiler S, Hou L (2020) Estimation of preference heterogeneity in markets with costly search. Technical report, Working paper.

Druckman JN, McGrath MC (2019) The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change* 9(2):111–119.

Druckman JN, Nelson KR (2003) Framing and deliberation: How citizens' conversations limit elite influence. *American journal of political science* 47(4):729–745.

Druckman JN, Peterson E, Slothuus R (2013) How elite partisan polarization affects public opinion formation. *American political science review* 107(1):57–79.

Du J, van Koningsbruggen GM, Kerkhof P (2020) Spontaneous approach reactions toward social media cues. *Computers in Human Behavior* 103:101–108.

Dubé JP, Hitsch GJ, Rossi PE (2010) State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics* 41(3):417–445, ISSN 07416261, URL http://dx.doi.org/10.1111/j.1756-2171.2010.00106.x.

Dunlap RE, McCright AM, Yarosh JH (2016) The political divide on climate change: Partisan polarization widens in the us. *Environment: Science and Policy for Sustainable Development* 58(5):4–23.

Eagle N, Pentland AS (2009) Eigenbehaviors: Identifying structure in routine. *Behavioral ecology and sociobiology* 63:1057–1066.

Egan PJ, Mullin M (2017) Climate change: Us public opinion. *Annual Review of Political Science* 20:209–227.

Einav L, Levin J (2014) The data revolution and economic analysis. *Innovation Policy and the Economy* 14(1):1–24.

Epley N, Gilovich T (2016) The mechanics of motivated reasoning. *Journal of Economic Perspectives* 30(3):133–40, URL http://dx.doi.org/10.1257/jep.30.3.133.

Ester M, Kriegel HP, Sander J, Xu X, *et al.* (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. *kdd*, volume 96, 226–231.

Fader PS, Hardie BG, Lee KL (2005) Rfm and clv: Using iso-value curves for customer base analysis. *Journal of marketing research* 42(4):415–430.

Falk E, Scholz C (2018) Persuasion, influence, and value: Perspectives from communication and social neuroscience. *Annual review of psychology* 69:329–356.

Falk EB, Morelli SA, Welborn BL, Dambacher K, Lieberman MD (2013) Creating buzz: the neural correlates of effective message propagation. *Psychological science* 24(7):1234–1242.

Farrell J, Klemperer P (2007) Coordination and lock-in: Competition with switching costs and network effects. *Handbook of Industrial Organization* 3:1967–2072.

Fechner GT (1860) *Elemente der psychophysik*, volume 2 (Breitkopf u. Härtel).

Feinberg M, Willer R (2013) The moral roots of environmental attitudes. *Psychological science* 24(1):56–62.

Fielding KS, Hornsey MJ, Thai HA, Toh LL (2020) Using ingroup messengers and ingroup values to promote climate change policy. *Climatic Change* 158:181–199.

Figueroa C, Musalem A, Noton C (2019) Earthquakes and brand loyalty: Beyond the short-term effects of stockouts. *Available at SSRN 3325923* .

Fisher M, Keil FC (2016) The curse of expertise: When more knowledge leads to miscalibrated explanatory insight. *Cognitive science* 40(5):1251–1269.

Foerde K (2018) What are habits and do they depend on the striatum? a view from the study of neuropsychological populations. *Current opinion in behavioral sciences* 20:17–24.

Friedrichsmeier T, Matthies E, Klöckner CA (2013) Explaining stability in travel mode choice: An empirical comparison of two concepts of habit. *Transportation Research Part F: Traffic Psychology and Behaviour* 16:1–13.

Frith CD, Frith U (2006) The neural basis of mentalizing. *Neuron* 50(4):531–534.

Gardner B (2015a) A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. *Health Psychology Review* 9(3):277–295, ISSN 1743-7199, 1743-7202, URL http://dx.doi.org/10.1080/17437199.2013.876238.

Gardner B (2015b) A review and analysis of the use of 'habit'in understanding, predicting and influencing health-related behaviour. *Health psychology review* 9(3):277–295.

Gardner B, Abraham C, Lally P, de Bruijn GJ (2012) Towards parsimony in habit measurement: Testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *International Journal of Behavioral Nutrition and Physical Activity* 9(1):102, ISSN 1479-5868, URL http://dx.doi.org/10.1186/1479-5868-9-102.

Gardner B, de Bruijn GJ, Lally P (2011) A systematic review and meta-analysis of applications of the self-report habit index to nutrition and physical activity behaviours. *Annals of Behavioral Medicine* 42(2):174–187.

Gardner B, Lally P (2013) Does intrinsic motivation strengthen physical activity habit? modeling relationships between self-determination, past behaviour, and habit strength. *Journal of behavioral medicine* 36:488–497.

Gawronski B, De Houwer J (2014) Implicit measures in social and personality psychology. .

Giannakos MN, Chorianopoulos K, Giotopoulos K, Vlamos P (2013) Using facebook out of habit. *Behaviour & Information Technology* 32(6):594–602.

Gilbert PA, Khokhar S (2008) Changing dietary habits of ethnic groups in europe and implications for health. *Nutrition reviews* 66(4):203–215.

Goodman-Bacon A (2021) Difference-in-differences with variation in treatment timing. *Journal of Econometrics* .

Götz FM, Gosling SD, Rentfrow PJ (2022) Small effects: The indispensable foundation for a cumulative psychological science. *Perspectives on Psychological Science* 17(1):205–215.

Granger CW (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 424–438.

Greenwald AG, Banaji MR, Rudman LA, Farnham SD, Nosek BA, Mellott DS (2002) A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological review* 109(1):3.

Grüning DJ, Riedel F, Lorenz-Spreen P (2023) Directing smartphone use through the self-nudge app one sec. *Proceedings of the National Academy of Sciences* 120(8):e2213114120.

Guadagni PM, Little JDC (1983) A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science* 27(1):29–48, ISSN 0732-2399, 1526-548X, URL http://dx.doi.org/10.1287/mksc.1070.0331.

Hagger MS, Rebar AL, Mullan B, Lipp OV, Chatzisarantis NL (2015) The subjective experience of habit captured by self-report indexes may lead to inaccuracies in the measurement of habitual action. *Health Psychology Review* 9(3):296–302.

Hall P, Gill N, Kurka M, Phan W (2017) Machine learning interpretability with h2o driverless ai. *H2O. ai* .

Happé F, Cook JL, Bird G (2017) The structure of social cognition: In (ter) dependence of sociocognitive processes. *Annual review of psychology* 68:243–267.

Hart PS, Nisbet EC (2012) Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication research* 39(6):701–723.

Haynes SN, Richard D, Kubany ES (1995) Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological assessment* 7(3):238.

Heckman JJ (1981) Statistical models for discrete panel data. *Structural Analysis of Discrete Data with Econometric Applications* 114:178.

Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? *Behavioral and brain sciences* 33(2-3):61–83.

Hootsuite (2022) Social network advertising revenues in the united states from 2017 to 2021 (in billion u.s. dollars) [graph]. in statista. retrieved march 07, 2023. https://www.statista.com/statistics/271259/advertising-revenue-of-social-networks-in-the-us/.

Horta Ribeiro M, Hosseinmardi H, West R, Watts DJ (2023) Deplatforming did not decrease parler users' activity on fringe social media. *PNAS nexus* 2(3):pgad035.

Hsu CC, Ajorlou A, Jadbabaie A (2021) Persuasion, news sharing, and cascades on social networks. *2021 60th IEEE Conference on Decision and Control (CDC)*, 4970–4975 (IEEE).

Huddy L, Sears DO, Levy JS (2013) *The Oxford handbook of political psychology* (Oxford University Press).

Hussam R, Rabbani A, Reggiani G, Rigol N (2017) Habit Formation and Rational Addiction: A Field Experiment in Handwashing. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.3040729.

Imai K, Kim IS (2019) When should we use unit fixed effects regression models for causal inference with longitudinal data? *American Journal of Political Science* 63(2):467–490.

Iyengar S, Kinder DR, *et al.* (1987) News that matters: Agenda-setting and priming in a television age. *News that Matters: Agenda-Setting and Priming in a Television Age* .

James W (1890) *The Principles of Psychology* (Macmillan London).

Kahneman D, Tversky A (1979) Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2):263–291, URL https://ideas.repec.org/a/ecm/emetrp/v47y1979i2p263-91.html.

Kamenica E, Gentzkow M (2011) Bayesian persuasion. *American Economic Review* 101(6):2590–2615.

Katz E, Blumler JG, Gurevitch M (1973) Uses and gratifications research. *The public opinion quarterly* 37(4):509–523.

Keane MP (1997) Modeling Heterogeneity and State Dependence in Consumer Choice Behavior. *Journal of Business & Economic Statistics* 15(3):310–327, ISSN 0735-0015, 1537-2707, URL http://dx.doi.org/10.1080/07350015.1997.10524709.

Kelly Y, Zilanawala A, Booker C, Sacker A (2018) Social media use and adolescent mental health: Findings from the uk millennium cohort study. *EClinicalMedicine* 6:59–68.

Klemperer P (1987) Markets with consumer switching costs. *The Quarterly Journal of Economics* 102(2):375–394.

Kline P (2013) *Handbook of psychological testing* (Routledge).

Kling JR, Liebman JB, Katz LF (2007) Experimental analysis of neighborhood effects. *Econometrica* 75(1):83–119.

Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273(5280):1399–1402.

Koc M, Gulyagci S (2013) Facebook addiction among turkish college students: The role of psychological health, demographic, and usage characteristics. *Cyberpsychology, Behavior, and Social Networking* 16(4):279–284.

Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences* 110(15):5802–5805.

Kosinski M, Wang Y, Lakkaraju H, Leskovec J (2016) Mining big data to extract patterns and predict real-life outcomes. *Psychological methods* 21(4):493.

Kross E, Verduyn P, Demiralp E, Park J, Lee DS, Lin N, Shablack H, Jonides J, Ybarra O (2013) Facebook use predicts declines in subjective well-being in young adults. *PloS one* 8(8):e69841.

Kubin E, von Sikorski C (2021) The role of (social) media in political polarization: a systematic review. *Annals of the International Communication Association* 45(3):188–206.

Kunda Z (1990) The case for motivated reasoning. *Psychological bulletin* 108(3):480–98, URL http://dx.doi.org/10.1037/0033-2909.108.3.480.

Labrecque JS, Wood W (2015) What measures of habit strength to use? Comment on Gardner (2015). *Health Psychology Review* 9(3):303–310.

Larcom S, Rauch F, Willems T (2017) The benefits of forced experimentation: Striking evidence from the London Underground network. *The Quarterly Journal of Economics* 132(4):2019–2055, ISSN 0033-5533, 1531-4650, URL http://dx.doi.org/10.1093/qje/qjx020.

LaRose R (2010) The problem of media habits. *Communication Theory* 20(2):194–222.

LaRose R, Eastin MS (2004) A social cognitive theory of internet uses and gratifications: Toward a new model of media attendance. *Journal of broadcasting & electronic media* 48(3):358–377.

Lazer D, Pentland A, Adamic L, Aral S, Barabási AL, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M, *et al.* (2009) Computational social science. *Science* 323(5915):721–723.

Leiserowitz A, Maibach E, Rosenthal SA, Kotcher J, Bergquist P, Ballew MT, Goldberg MH, Gustafson A (2020) Climate change in the american mind: November 2019 .

Leonard TC (2008) Richard h. thaler, cass r. sunstein, nudge: Improving decisions about health, wealth, and happiness.

Levine J, Seiler S (2021) Identifying state dependence in brand choice: Evidence from hurricanes, CEPR Discussion Paper No. DP16632.

Lin LY, Sidani JE, Shensa A, Radovic A, Miller E, Colditz JB, Hoffman BL, Giles LM, Primack BA (2016) Association between social media use and depression among us young adults. *Depression and anxiety* 33(4):323–331.

Lippmann W (1965) Public opinion. 1922. *URL: http://infomotions.com/etexts/gutenberg/dirs/etext04/pbp nn10. htm* .

Liu-Thompkins Y, Tam L (2013) Not all repeat customers are the same: Designing effective cross-selling promotion on the basis of attitudinal loyalty and habit. *Journal of Marketing* 77(5):21–36.

Machín L, Curutchet MR, Gugliucci V, Vitola A, Otterbring T, de Alcantara M, Ares G (2020) The habitual nature of food purchases at the supermarket: Implications for policy making. *Appetite* 104844, ISSN 01956663, URL http://dx.doi.org/10.1016/j.appet.2020.104844.

Macrae CN, Milne AB, Bodenhausen GV (1994) Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of personality and Social Psychology* 66(1):37.

Madrian BC (2014) Applying insights from behavioral economics to policy design. *Annu. Rev. Econ.* 6(1):663–688.

Mazar A, Wood W (2018) Defining habit in psychology. *The Psychology of Habit*, 13–29 (Springer).

McAlister L (1982) A dynamic attribute satiation model of variety-seeking behavior. *Journal of consumer research* 9(2):141–150.

McCarty N, Poole KT, Rosenthal H (2003) Political polarization and income inequality. *Available at SSRN 1154098* .

McCright AM, Marquart-Pyatt ST, Shwom RL, Brechin SR, Allen S (2016) Ideology, capitalism, and climate: Explaining public views about climate change in the united states. *Energy Research & Social Science* 21:180–189.

Meier A, Reinecke L, Meltzer CE (2016) "facebocrastination"? predictors of using facebook for procrastination and its effects on students' well-being. *Computers in Human Behavior* 64:65–76.

Messick S (1990) Validity of test interpretation and use. .

Mullainathan S, Thaler RH (2000) Behavioral economics.

Neal DT, Wood W, Quinn JM (2006) Habits—a repeat performance. *Current directions in psychological science* 15(4):198–202.

Neal DT, Wood W, Wu M, Kurlander D (2011) The pull of the past: When do habits persist despite conflict with motives? *Personality and Social Psychology Bulletin* 37(11):1428–1437.

Oblander ES, McCarthy DM (2022) Persistence of consumer lifestyle choices: Evidence from restaurant delivery during COVID-19, available at https://dx.doi.org/10.2139/ssrn.3836262.

Orbell S, Verplanken B (2010) The automatic component of habit in health behavior: Habit as cue-contingent automaticity. *Health Psychology* 29(4):374–383, ISSN 1930-7810, 0278-6133, URL http://dx.doi.org/10.1037/a0019596.

Orben A, Przybylski AK (2019) The association between adolescent well-being and digital technology use. *Nature human behaviour* 3(2):173–182.

Ouellette JA, Wood W (1998) Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological bulletin* 124(1):54.

Oulasvirta A, Rattenbury T, Ma L, Raita E (2012) Habits make smartphone use more pervasive. *Personal and Ubiquitous computing* 16:105–114.

Pakes A, Porter JR, Shepard M, Calder-Wang S (2021) Unobserved heterogeneity, state dependence, and health plan choices. Technical report, National Bureau of Economic Research.

Palm R, Lewis GB, Feng B (2017) What causes people to change their opinion about climate change? *Annals of the American Association of Geographers* 107(4):883–896.

Pantic I (2014) Online social networking and mental health. *Cyberpsychology, Behavior, and Social Networking* 17(10):652–657.

Pariser E (2011) *The filter bubble: How the new personalized web is changing what we read and how we think* (Penguin).

Perloff RM (2020) *The dynamics of persuasion: Communication and attitudes in the twenty-first century* (Routledge).

Petty RE, Cacioppo JT, Petty RE, Cacioppo JT (1986) *The elaboration likelihood model of persuasion* (Springer).

Poortinga W, Whitmarsh L, Suffolk C (2013) The introduction of a single-use carrier bag charge in Wales: Attitude change and behavioural spillover effects. *Journal of Environmental Psychology* 36:240–247, ISSN 02724944, URL http://dx.doi.org/10.1016/j.jenvp.2013.09.001.

Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? *Behavioral and brain sciences* 1(4):515–526.

Privitera GJ, Zuraikat FM (2014) Proximity of foods in a competitive food environment influences consumption of a low calorie and a high calorie food. *Appetite* 76:175–179, ISSN 01956663, URL http://dx.doi.org/10.1016/j.appet.2014.02.004.

Rabin M (1998) Psychology and economics. *Journal of economic literature* 36(1):11–46.

Rao A (2022) Deceptive claims using fake news advertising: The impact on consumers. *Journal of Marketing Research* 59(3):534–554.

Rebar AL, Gardner B, Rhodes RE, Verplanken B (2018) The measurement of habit. *The psychology of habit: Theory, mechanisms, change, and contexts* 31–49.

Rebar AL, Gardner B, Verplanken B (2020) Habit in exercise behavior. *Handbook of sport psychology* 986–998.

Robinson J (1933) Imperfect competition. *An Essay on Marxian Economics*, 73–81 (Springer).

Rode JB, Dent AL, Benedict CN, Brosnahan DB, Martinez RL, Ditto PH (2021) Influencing climate change attitudes in the united states: A systematic review and meta-analysis. *Journal of Environmental Psychology* 76:101623.

Ronis DL, Yates JF, Kirscht JP (1989) Attitudes, decisions, and habits as determinants of repeated behavior. *Attitude structure and function* 213:39.

Roth J (2019) Pre-test with caution: Event-study estimates after testing for parallel trends. *Department of Economics, Harvard University, Unpublished manuscript* .

Roy R, Chintagunta PK, Haldar S (1996) A framework for investigating habits,"the hand of the past," and heterogeneity in dynamic brand choice. *Marketing Science* 15(3):280–299.

Salganik MJ (2019) *Bit by bit: Social research in the digital age* (Princeton University Press).

Schnauber-Stockmann A, Naab TK (2019) The process of forming a mobile media habit: Results of a longitudinal study in a real-world setting. *Media Psychology* 22(5):714–742.

Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: a meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews* 42:9–34.

Schwartz C, Scholtens PA, Lalanne A, Weenen H, Nicklaus S (2011) Development of healthy eating habits early in life. review of recent evidence and selected guidelines. *Appetite* 57(3):796–807.

Sears DO (1987) Political psychology. *Annual Review of Psychology* 38(1):229–255.

Seetharaman P (2004) Modeling multiple sources of state dependence in random utility models: A distributed lag approach. *Marketing Science* 23(2):263–271.

Seetharaman P, Ainslie A, Chintagunta PK (1999) Investigating household state dependence effects across categories. *Journal of Marketing Research* 36(4):488–500.

Seetharaman P, Che H (2009) Price competition in markets with consumer variety seeking. *Marketing Science* 28(3):516–525.

Sheth J (2020) Impact of Covid-19 on consumer behavior: Will the old habits return or die? *Journal of Business Research* 117:280–283, ISSN 01482963, URL http://dx.doi.org/10.1016/j.jbusres.2020.05.059.

Silverman BW (2018) *Density estimation for statistics and data analysis* (Routledge).

Simon HA (1955) A behavioral model of rational choice. *The quarterly journal of economics* 99–118.

Simonov A, Dubé JP, Hitsch G, Rossi P (2020) State-dependent demand estimation with initial conditions correction. *Journal of Marketing Research* 57(5):789–809.

Sniderman PM, Theriault SM (2004) The structure of political argument and the logic of issue framing. *Studies in public opinion: Attitudes, nonattitudes, measurement error, and change* 133–65.

Spearman C (1904) "general intelligence," objectively determined and measured. *The American Journal of Psychology* 15(2):201–292, ISSN 00029556, URL http://www.jstor.org/stable/1412107.

Starr E, Goldfarb B (2020) Binned scatterplots: A simple tool to make research easier and better. *Strategic Management Journal* 41(12):2261–2274.

Stigler GJ (1961) The economics of information. *Journal of political economy* 69(3):213–225.

Stigler SM (1992) A historical view of statistical concepts in psychology and educational research. *American Journal of Education* 101(1):60–70.

Sun L, Abraham S (2021) Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225(2):175–199.

Sun Y, Zhang Y (2021) A review of theories and models applied in studies of social media addiction and implications for future research. *Addictive Behaviors* 114:106699.

Sunstein CR, Bobadilla-Suarez S, Lazzaro SC, Sharot T (2016) How people update beliefs about climate change: Good news and bad news. *Cornell L. Rev.* 102:1431.

Tam L, Wood W, Ji MF (2014) Brand loyalty is not habitual. *Handbook of Brand Relationships* 65–84.

Thaler RH (2016) Behavioral economics: Past, present, and future. *American economic review* 106(7):1577–1600.

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

Thomadsen R (2016) The Impact of Switching Stores on State Dependence in Brand Choice. *SSRN Electronic Journal* ISSN 1556-5068, URL http://dx.doi.org/10.2139/ssrn.2759868.

Thomadsen R, Seetharaman PB (2018) The strategic effects of state-dependent consumer preferences: The roles of habits and variety seeking. *The Psychology of Habit*, 131–150 (Springer).

Thorndike EL (1898) Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements* 2(4):i.

Thrailkill EA, Bouton ME (2015) Contextual control of instrumental actions and habits. *Journal of Experimental Psychology: Animal Learning and Cognition* 41(1):69.

Tomasello M (2018) How children come to understand false beliefs: A shared intentionality account. *Proceedings of the National Academy of Sciences* 115(34):8491–8498.

Traxler MJ, Gernsbacher MA (1993) Improving written communication through perspective-taking. *Language and Cognitive Processes* 8(3):311–334.

Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG (2019) Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of abnormal psychology* 128(3):185.

Van Bavel JJ, Pereira A (2018) The partisan brain: An identity-based model of political belief. *Trends in cognitive sciences* 22(3):213–224.

Vandaele Y, Ahmed S (2021) Habit, choice, and addiction. *Neuropsychopharmacology* 46(4):689–698.

Ver Hoef JM, Boveng PL (2007) Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology* 88(11):2766–2772.

Verplanken B (2006) Beyond frequency: Habit as mental construct. *British Journal of Social Psychology* 45(3):639–656, ISSN 01446665, URL http://dx.doi.org/10.1348/014466605X49122.

Verplanken B, Aarts H (1999) Habit, attitude, and planned behaviour: is habit an empty construct or an interesting case of goal-directed automaticity? *European review of social psychology* 10(1):101–134.

Verplanken B, Orbell S (2003) Reflections on Past Behavior: A Self-Report Index of Habit Strength [1]. *Journal of Applied Social Psychology* 33(6):1313–1330, ISSN 00219029, URL http://dx.doi.org/10.1111/j.1559-1816.2003.tb01951.x.

Verplanken B, Verplanken B, Ryan (2018) *Psychology of habit* (Springer).

Verplanken B, Walker I, Davis A, Jurasek M (2008) Context change and travel mode choice: Combining the habit discontinuity and self-activation hypotheses. *Journal of Environmental Psychology* 28(2):121–127, ISSN 02724944, URL http://dx.doi.org/10.1016/j.jenvp.2007.10.005.

Verplanken B, Wood W (2006) Interventions to Break and Create Consumer Habits. *Journal of Public Policy & Marketing* 25(1):90–103, ISSN 0748-6766, 1547-7207, URL http://dx.doi.org/10.1509/jppm.25.1.90.

Wallas G (1921) *Human nature in politics* (Constable).

Watson P, de Wit S (2018) Current limits of experimental research into habits and future directions. *Current opinion in behavioral sciences* 20:33–39.

We Are Social, DataReportal, Hootsuite (2022) Daily time spent on social networking by internet users worldwide from 2012 to 2022 (in minutes) [graph]. in statista. retrieved march 07, 2023. https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/.

Wedderburn RW (1974) Quasi-likelihood functions, generalized linear models, and the gauss—newton method. *Biometrika* 61(3):439–447.

Welle PD, Graf HM (2011) Effective lifestyle habits and coping strategies for stress tolerance among college students. *American Journal of Health Education* 42(2):96–105.

William LB, Harter N (1899) Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological review* 6(4):345.

Winkler H (2019) The effect of income inequality on political polarization: Evidence from european regions, 2002–2014. *Economics & Politics* 31(2):137–162.

Wolniczak I, Cáceres-DelAguila JA, Palma-Ardiles G, Arroyo KJ, Solís-Visscher R, Paredes-Yauri S, Mego-Aquije K, Bernabe-Ortiz A (2013) Association between facebook dependence and poor sleep quality: a study in a sample of undergraduate students in peru. *PloS one* 8(3):e59087.

Wolsko C, Ariceaga H, Seiden J (2016) Red, white, and blue enough to be green: Effects of moral framing on climate change attitudes and conservation behaviors. *Journal of Experimental Social Psychology* 65:7–19.

Wong-Parodi G, Feygina I (2020) Understanding and countering the motivated roots of climate change denial. *Current Opinion in Environmental Sustainability* 42:60–64.

Wood W (2017) Habit in Personality and Social Psychology. *Personality and Social Psychology Review* 21(4):389–403, ISSN 1088-8683, 1532-7957, URL http://dx.doi.org/10.1177/1088868317720362.

Wood W, Labrecque JS, Lin PY, Rünger D (2014) Habits in dual process models. Sherman JW, Gawronski B, Trope Y, eds., *Dual Process Theories of the Social Mind*, 371–385 (Guilford Press New York, NY).

Wood W, Neal DT (2007) A new look at habits and the habit-goal interface. *Psychological Review* 114(4):843–863, ISSN 1939-1471, 0033-295X, URL http://dx.doi.org/10.1037/0033-295X.114.4.843.

Wood W, Neal DT (2016) Healthy through habit: Interventions for initiating & maintaining health behavior change. *Behavioral Science & Policy* 2(1):71–83.

Wood W, Quinn JM, Kashy DA (2002) Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology* 83(6):1281–1297, ISSN 1939-1315, 0022-3514, URL http://dx.doi.org/10.1037/0022-3514.83.6.1281.

Wood W, Rünger D (2016) Psychology of Habit. *Annual Review of Psychology* 67(1):289–314, ISSN 0066-4308, 1545-2085, URL http://dx.doi.org/10.1146/annurev-psych-122414-033417.

Wood W, Tam L, Witt MG (2005) Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology* 88(6):918–933, ISSN 1939-1315, 0022-3514, URL http://dx.doi.org/10.1037/0022-3514.88.6.918.

Yadamsuren B, Erdelez S (2011) Online news reading behavior: From habitual reading to stumbling upon news. *Proceedings of the american society for information science and technology* 48(1):1–10.

Yarkoni T, Westfall J (2017) Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science* 12(6):1100–1122.

Yim O, Ramdeen KT (2015) Hierarchical cluster analysis: comparison of three linkage measures and application to psychological data. *The quantitative methods for psychology* 11(1):8–21.

Zhang Y, Bradlow ET, Small DS (2015) Predicting customer value using clumpiness: From rfm to rfmc. *Marketing science* 34(2):195–208.

Zollo F, Quattrociocchi W (2018) Misinformation spreading on facebook. *Complex spreading phenomena in social systems: Influence and contagion in real-world social networks* 177–196.