8-2023

Empirical Studies on Global Product Reviews

Maneesh Reddy Ajjuguttu
majjugu@g.clemson.edu

Empirical Studies on Global Product Reviews

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Business Administration

by
Maneesh Reddy Ajjuguttu
August 2023

Accepted by:
Dr. Lawrence D. Fredendall, Committee Chair
Dr. Ahmet Colak
Dr. Craig Wallace
Dr. Blake Snider

Abstract

This dissertation comprises of three essays on global reviews on Amazon.com and their impact on future reviews. The first chapter examines the interplay between economic policy uncertainty (EPU) and global review sharing on consumer behavior and product reviews on Amazon.com. By analyzing data from Nike products available on the platform and considering EPU indices for the USA and China, econometric models are constructed to test various hypotheses. The study reveals that higher levels of USA EPU adversely affect product review rates, while China EPU positively influences review rates. Additionally, increased EPU volatility leads to a decrease in the sharing of non-English and English reviews by Amazon, impacting sales and review-sharing behavior.Moreover, the study investigates the effect of global reviews on customer decision-making for products at different stages of their lifecycle. For early-stage products, introducing global reviews influences customer perceptions and decision-making positively. However, for popular products with substantial review volumes, the addition of global reviews overwhelms customers, negatively impacting their decision-making. Furthermore, the research includes a counterfactual analysis to assess the potential implications of the absence of global reviews. The findings suggest that global review sharing positively impacts early-stage products but may adversely affect popular products. These insights shed light on how businesses can strategically leverage global reviews, considering economic policy uncertainties and product lifecycle stages, to enhance customer experiences on e-commerce platforms like Amazon.com. By understanding the nuanced effects of EPU and global review-sharing, companies can tailor review-sharing strategies for optimal customer engagement and improved business performance in a globalized marketplace.

The second chapter investigates the impact of product origin information on the review rate of products in online businesses, with a focus on Amazon.com as the platform of interest. The

study explores how references to "made-in USA/China" in customer reviews influence purchasing decisions and customer behavior. By understanding the relationship between product sources and customer behavior, businesses can optimize marketing strategies and enhance customer satisfaction. The literature review highlights the importance of the country of origin in consumer decision-making and its influence on product perception. Online reviews play a crucial role in guiding customer purchasing decisions, and product origin significantly affects the perception of product quality. While products made in the USA are positively perceived, products made in China have undergone a shift in perception due to increased awareness of China's manufacturing capabilities. Economic Policy Uncertainty (EPU) can also significantly impact consumer behavior and market dynamics. Using a comprehensive dataset, the study employs Ordinary Least Squares (OLS) methods to examine the direct effects of "made-in" mentions on changes in review rates and scores. The analysis includes control variables and product category fixed effects to ensure a robust analysis. The results indicate that "Made in China" mentions positively influence the review rate, while "Made in USA" mentions do not significantly impact the review rate. Economic uncertainty in both the USA and China leads to decreased review rates and scores, suggesting changes in consumer behavior during uncertain economic times. However, EPU volatility does not significantly impact mentions of products made in either country. This research provides valuable insights for businesses seeking to optimize marketing efforts and cater to consumer preferences effectively. The study underscores the significance of product origin information and its impact on customer behavior in the global marketplace.

The third chapter aims to explore how review language diversity, particularly English and Spanish reviews, influences customer behavior on the platform. The research addresses gaps in understanding the effects of non-English reviews on sales, providing insights for businesses to optimize their review strategies. Past research have addressed on how review language diversity can foster inclusivity and social proof, influencing the motivation of English and Spanish-speaking customers to post reviews. Language plays a crucial role in shaping customer behavior, and the study hypothesizes that the count of Spanish reviews affects English and Spanish-speaking customers differently. We employed econometric models to test the relationships between variables and their impact on review rates and scores accounting for different product categories. The results reveal that an increase in Spanish reviews positively affects future review scores for customers and encourages Spanish-speaking customers to contribute reviews. English reviews positively influence the review rate for both English and Spanish-speaking customers. The study concludes by emphasizing the importance of understanding language dynamics and their implications for customer behavior on Amazon.com, especially in the context of diverse language reviews. The research contributes to the field of consumer behavior, shedding light on the role of language in shaping purchase decisions and customer perceptions.

Dedication

I dedicate this dissertation to my parents, whose unwavering support and encouragement allowed me to pursue my dreams. Their belief in me has been a constant motivation throughout my academic journey. I also want to express my deepest gratitude to my wife, who stood by my side during the most challenging phase of my career. Her love, understanding, and support gave me the strength and determination to overcome obstacles and achieve my goals.

Acknowledgments

Completing my dissertation has been made possible with numerous individuals' invaluable support and contributions. I want to express my deepest gratitude to all of them.

First and foremost, I extend my heartfelt appreciation to my advisor, Prof.Lawrence Fredendall. His belief in me and the opportunity to work under his guidance have shaped my research journey. I am grateful for the freedom he provided me and his constant mentorship, which empowered me to become an independent researcher. I will forever be indebted to him for his guidance throughout my dissertation.

I would also like to express my profound gratitude to Prof.Ahmet Colak, another esteemed committee member. His unwavering support and assistance at every research step have been invaluable. His guidance in refining my research and programming skills has been vital to my dissertation journey. I am grateful for his constant mentorship for the three years I have worked with him. Without him the journey would have been impossible to finish.

My heartfelt thanks also go to Prof. Craig Wallace and Prof. Blake Snider, whose contributions as committee members have been crucial in completing my dissertation. I would like to extend my gratitude to Prof. Yann Ferrand and Prof. Bernardo F. Quiroga for their support and guidance throughout my graduate school journey. I am also grateful to all the faculty members I worked with and the administrative staff at Clemson University for their assistance and cooperation. Special thanks go to Prof.Janis Miller at Clemson University for her support in helping me secure a scholarship and a teaching award. Her assistance was instrumental in furthering my academic pursuits.

I would like to thank Prof.Ganesh Janakiraman at the University of Texas at Dallas for inspiring me to pursue a doctoral degree. His guidance and encouragement have played a significant role in shaping my aspirations. I also want to thank Prof. Bharadwaj Kadiyala for helping through my Ph.D. applications and guiding me in understanding what a doctoral program is.

I want to acknowledge my peers in graduate school, Daniel Nielubowicz, Brandon Lee, Min Kyung Lee, Remi Charpin, and Niratcha "Grace" Tungtisanont, Jonathan Preedom, Hongli Ye, for their company and support. Their presence made my Ph.D. journey comfortable, and I am grateful for their willingness to lend an ear whenever needed and guide me in the right direction. I would also like to thank my friends Shahil, Aradhana, Apoorva, Parthiva, Hruday, Akhil, and Hari Keshav for their presence and assistance. They have consistently been there for me, offering help and support whenever needed. I am grateful to my peers and friends for their willingness to support me throughout my academic pursuits.

A special mention of gratitude goes to my friends Adithya Krishnamurthy and Dr. Shilpi Mukherjee. Their unwavering support and guidance, both technically and emotionally, have been invaluable. I am deeply grateful for their presence during a critical juncture in my career. Without their generous assistance, my Ph.D. journey may have ended abruptly. From the bottom of my heart, I thank them for standing by me and instilling self-belief when I needed it the most. I also want to thank my friends Ravi, Mounika, and Aashish for encouraging me to pursue my dream of a doctoral degree.

Lastly, but most importantly, I am forever grateful to my parents for their support and smiles that have accompanied every decision I made. I am also thankful to my sister for being a constant motivator throughout my journey. And to my wife, I extend my deepest appreciation for her unconditional love and care, particularly during the final year of my Ph.D. Her support made an already challenging journey more manageable.

In conclusion, I extend my heartfelt gratitude to everyone who played a role in realizing my dream. I have learned that a person's achievements are not solely the result of their hard work but

also the support of many helping hands. Thank you all for being a part of my journey.

Table of Contents

List of Tables

List of Figures

Chapter 1

Do Global Review Sharing Impact Online Business? Evidence from Nike Products on Amazon

1.1 Introduction

The proliferation of online reviews has surged in recent years, driven by the rapid growth of online commerce (Howarth, 2022). This surge has prompted businesses to recognize the significance of online reviews in consumer decision-making (Podium, 2017). A report by Podium.com revealed that a substantial majority of consumers (93%) are influenced by online reviews when making purchasing decisions. Concurrently, globalization has gained momentum, compelling businesses to establish a global presence and cater to diverse markets (Federico, 2016). In response to this trend, many companies have adopted a review-sharing strategy, whereby reviews from one market are disseminated to customers in different markets. Examples of this strategy include Nike sharing reviews from their distributor's website and featuring them on Nike.com and Amazon sharing reviews from multiple countries on Amazon.com and vice versa. Amazon, which operates in over 20 countries

(Mbabazi, 2022), offers products on both Amazon.com and its international websites. Each country-specific website accumulates reviews for the products available in that particular market. Starting in January 2020, Amazon has implemented a practice of sharing reviews posted on one country's website with other country-specific websites, thereby expanding the pool of visible reviews for customers.

Simultaneously, macroeconomic factors, such as policy uncertainties, profoundly influence customers (Handley and Limão, 2017), thereby impacting businesses. In this study, we identify economic policy uncertainty (referred to as EPU hereafter) as a potential factor that can significantly affect sales on Amazon.com. Studies have demonstrated that EPU amplifies domestic prices and diminishes consumer income (Handley and Limão, 2017). Moreover, research has shown that increasing uncertainty reduces consumer spending, highlighting the consequential impact of EPU on businesses (Tamara Charm, 2022; Lai, 2022). Baker et al. (2013) further contribute to this understanding by establishing a negative correlation between increased EPU and investment, employment, and output. Their subsequent work Baker et al. (2016) reveals that EPU correlates negatively with stock market returns and corporate investment. Furthermore, an upsurge in EPU results in declining consumer confidence, ultimately reducing consumer spending (Ozdemir et al., 2021). While traditional macroeconomic variables, such as GDP and GDP per capita, are undoubtedly important, EPU captures the uncertainty and unpredictability arising from changes in government policies and regulations. Such uncertainty often prompts cautious behavior among businesses and individuals, resulting in reduced investment and spending. Consequently, EPU emerges as a critical variable that merits consideration in economic research, particularly during political and economic turmoil. Recognizing the significance of EPU, we have chosen it as the focal variable of interest in our study. Importantly, the specific impacts of varying EPU conditions on businesses still need to be answered. Our study seeks to address this knowledge gap by investigating the influence of EPU on the domestic review rate of products on Amazon.com. Additionally, we aim to explore the factors that influence Amazon's review-sharing behavior and assess the impact of the global review-sharing program on the domestic product review rate. We investigate three main research questions as a part of this study: (i) How does EPU affect the product review rate on Amazon.com? (ii) What factors influence Amazon's review-sharing behavior? (iii) How does global review sharing impact product review rate? Prior studies have investigated the impact of product reviews on sales, but to our knowledge, this study has yet to explore the impact of foreign reviews on domestic sales.

To address the research questions, we have developed three main hypotheses. The first set of hypotheses addresses the impact of EPU on the domestic product review rate. Specifically, we hypothesize that USA and China EPU has a negative direct effect on domestic product sales. The second set of hypotheses focuses on the variables that may impact Amazon's review-sharing behavior. We hypothesize that EPU positively affects review-sharing behavior, such that higher EPU leads to higher review-sharing. Finally, in the third set of hypotheses, we examine how global review sharing affects the domestic review rate. In this hypothesis, the direction of the effect needs to be clarified. Higher levels of sharing will positively affect the domestic review rate because customers can see more reviews. However, at the same time, review manipulation will have adverse effects. Overall, we hope to provide valuable insights into the complex relationship between EPU and domestic review rate and to offer practical implications for businesses operating in a globalized market by investigating these hypotheses.

To address our research questions, we required a dataset encompassing reviews from multiple markets available on a single platform. Amazon is the world's largest online retailer (Angelovska, 2019), and Nike is one of the largest sportswear companies with a presence in over 170 countries. We selected Amazon as our primary focus and Nike as the brand of interest. Amazon's practice of sharing reviews from various countries with domestic customers on Amazon.com made it an ideal platform for our study. We conducted web scraping to collect data on 2,197 Nike products, capturing 174,172 unique reviews. Among these reviews, 117,516 originated from domestic customers in the USA, while the remaining reviews were shared global reviews. From the Amazon.com website, we gathered information such as ratings, review texts, and other relevant variables, enabling the construction of variables like review rate, recent five reviews rating, cumulative review rating, and more.

In our study, we categorized our variables into international and domestic levels. To address the language barrier in shared global reviews, we employed the fastText package (Bojanowski et al., 2016; Facebook, 2016; Bojanowski et al., 2017a,b; Mouselimis, 2022) in R to classify them into English and non-English reviews. We then calculated the number of reviews, review scores, and review rates separately for English and non-English reviews. Additionally, we constructed variables based on the last five reviews to explore the potential impact of recency bias. We used five reviews to test the recency bias as Zhou (2023) presented the significant influence of having five reviews on a product's selling behavior.

Furthermore, a report by Bizrate.com highlighted that most customers read at least five reviews before making a purchase decision. To measure the EPU indexes, a critical variable in our study, we utilized the index developed by Baker et al. (2016). Baker et al. (2016) evaluates the USA EPU, which incorporates three components: newspaper coverage of policy-related economic uncertainty, federal tax code provisions, and the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. Baker et al. (2013) developed the index by using the frequency count of articles related to government policies in the South China Morning Post newspaper.

Our estimation strategy comprised three main components. Firstly, we aimed to assess the impact of EPU. For this analysis, we employed an Ordinary Least Squares (OLS) approach to examine the effects of USA EPU and China EPU on the future review rates of the product. The domestic review rate served as a proxy for sales. Review rate indicates the speed at which reviews are posted on Amazon.com, which is associated with sales performance Spool (2009). Secondly, we utilized a two-stage regression model to explore the factors influencing the number and score of reviews shared under the review-sharing strategy. In the first stage, we employed predictive modeling to estimate the global review variables, such as the number of reviews and the average score, based on the product's past performance. Subsequently, in the second stage, we used the residual values from the first stage as an independent variable to predict the impact of the global review-sharing program on the product's review rate. Multiple dependent variables, including the next review rate and score, were tested against different models to examine the impact of review sharing. Additionally, we conducted a counterfactual analysis to assess the impact of review sharing. This analysis allowed us to explore the potential effects on the product's review rate under different scenarios and shed light on the significance of the review-sharing strategy in influencing customer behavior. Overall, our estimation strategy involved a combination of OLS regression, two-stage modeling, and counterfactual analysis to comprehensively investigate the impacts of EPU and global review sharing on the product's review rates and associated factors.

Our findings reveal intriguing insights regarding the impact of EPU and the review-sharing program on Amazon.com. Specifically, we observed contrasting effects between USA EPU and China EPU on product review rates and scores. The USA EPU negatively influenced review rates and scores, suggesting that higher levels of EPU in the USA context adversely affect customer perceptions and evaluations on Amazon.com. On the other hand, the China EPU impacted review rates and scores positively, implying that increased EPU in China may generate positive evaluations on the platform. Regarding the review-sharing program, our analysis yielded mixed results. Shared English reviews negatively affected the future scores, indicating that Amazon.com customers might perceive English reviews from a different country negatively. However, the non-English review count positively affected future review scores. These results suggest that shared reviews do not elicit the same positive response from Amazon.com customers in the USA, potentially due to customers perceiving foreign reviews negatively. Overall, our findings highlight the complex interplay between EPU, the review-sharing program, and customer behavior on Amazon.com. The contrasting effects of USA EPU and China EPU and the differential responses to English and non-English reviews underscore the importance of considering cultural and contextual macroeconomic factors in understanding customer reactions and the effectiveness of review-sharing strategies.

On e-commerce platforms, reviews have gained tremendous significance, shaping consumer perceptions and influencing purchasing decisions. As brands increasingly position themselves as global entities, the opportunity to gather reviews from various countries where their products are sold has expanded. However, devising and implementing effective strategies for utilizing these reviews effectively on domestic customers is crucial. Merely increasing the number of product reviews may not suffice; platforms must ensure that the shared reviews benefit customers. Our research emphasizes critical findings that provide insights into when managers should consider sharing reviews from foreign markets with domestic customers. Managers should exercise caution during volatile periods of policy uncertainty before embracing a global brand identity and implementing review-sharing strategies. Understanding the impact of EPU is crucial in this regard. By examining the effects of EPU on review rates and scores, our work provides valuable guidance for managers in leveraging reviews from foreign markets to enhance customer experiences. Our research contributes to developing effective review-sharing strategies by highlighting the importance of context and timing. Managers must be mindful of policy uncertainty conditions and consider the potential implications on customer perceptions and behaviors before implementing review-sharing initiatives. By doing so, platforms can leverage reviews to add value to customers rather than increase product review counts. Overall, our work underscores the importance of strategic review sharing and the need for managers to align their decisions with the prevailing policy uncertainty conditions. By adopting a thoughtful approach, platforms can harness the power of reviews to enhance customer satisfaction and drive business success in the global marketplace.

1.2 Literature and hypothesis development

This section focuses on the impact of global review sharing on customers influenced by economic policy uncertainty (EPU) and develops hypotheses related to several aspects of the phenomenon. Firstly, we investigate how reviews can serve as a source of information for customers, and secondly, how global the product's domestic performance influences review sharing. We also examine the effect of shared review language on the review rate. Our work is rooted in three main research streams: (i) online reviews, (ii) economic uncertainty, and (iii) information sharing across multiple markets. Although the study of reviews and their impact on business has been well-established in marketing, information systems, and economics literature, operations academia has only recently begun to focus on the topic (Lei et al., 2022; Xu et al., 2021; Ko et al., 2019). The importance of customer reviews in business has been well-established in previous research. Operations management researchers have observed a positive effect of reviews on business (Xu et al., 2021; Deshpande and Pendem, 2022). Reviews help optimize search algorithms for hotels (Ghose et al., 2014) and pricing strategies for sellers (Moreno and Terwiesch, 2014). Service quality can be assessed from reviews using text mining techniques (Mejia et al., 2021; Xu et al., 2021). Theoretical work has also shown that customers delay their purchase decisions unless they see reviews from peers (Papanastasiou and Savva, 2017). In our study, we contribute to the literature by empirically investigating how reviews from a different country impact local customers and examining the phenomenon of review-sharing to understand when and why Amazon shares reviews.

1.2.1 Effect of EPU on product review rate

1.2.2 Economic Policy Uncertainty

The first hypothesis aims to investigate the potential effect of EPU in the USA on online business. Previous studies have explored the impact of economic uncertainty on various economic variables. During high uncertain times, firms follow more conservative policies (Çolak et al., 2017) by delaying investments (Jens, 2017) and reduce capital spending (Gulen and Ion, 2015). (Hu and Liu, 2021) has shown that EPU negatively affected China's gross exports. On the other hand, EPU's impact on customers is studied extensively. Handley and Limão (2017) showed that policy uncertainty increases domestic prices and reduces consumers' income. Furthermore, multiple corporate reports have indicated that uncertainty reduces customer spending (Tamara Charm, 2022; Lai, 2022), which in turn reduces customer consumption (Romer, 1990; Gudmundsson, 2012). Prior literature has shown that uncertainty affects household consumption (Bernanke, 1983) and delays non-essential purchases(Eberly, 1994). During high uncertainty, spending becomes less attractive to households (Bloom, 2009). This reduced consumption again affects production (Bloom et al., 2007). Also, Yu et al. (2017); Phan et al. (2018) noted the volatility clustering around the index of EPU. Goodell et al. (2021) notes that the volatility of EPU is more economically significant than the EPU itself. Using this literature, we use the volatility of EPU as our measure to predict the product review rate. We use a six-month EPU volatility measure which we discuss more in the data section of this paper. With the literature mentioned, as EPU increases, the sales in an economy decrease. As we are using a product review rate to see how more reviews are coming into the system, we hypothesize that:

Hypothesis 1 (a): Increased USA's EPU negatively affects the domestic review rate.

At the same time, uncertainty in one country can have spillover effects on another country's economy, especially under the economic conditions (Kelly et al., 2016). Depending on the exposure of a business in a country, the EPU of one country can have a negative spillover effect on another country.(Ghirelli et al., 2021). Klößner and Sekkel (2014) has noted that international spillovers considerably affect other countries' dynamics. Additionally, Colombo (2013) has noted that USA EPU has a significant negative impact on the European market. Considering the strong trade ties with China, we expect that as China's EPU increases, it can have a negative effect on the USA market, thereby leading us to hypothesize that

Hypothesis 1 (b): Increased China's EPU negatively affects the domestic review rate.

1.2.2.1 Effect of EPU on review sharing

Reviews have become an essential tool for customers to gather information about a product (Dellarocas, 2003), and they serve as a proxy for word-of-mouth (Chrysanthos Dellarocas, 2004). Literature has shown that customers contribute to reviews based on the product's popularity (Dellarocas et al., 2010) and brand reputation (Amblee and Bui, 2008), meaning that customers review popular products more. Furthermore, more reviews can equate to more sales (Spool, 2009). The positive effect of reviews on sales has been well documented in economics (Chevalier and Mayzlin, 2006; Liu, 2006; Duan et al., 2008). By sharing reviews, Amazon is increasing the number of reviews. As hypotheses 1(a) and 1(b) mention, the product review rate decreases when USA or China EPU volatility increases. Policy uncertainty increases domestic prices (Handley and Limão, 2017), reduces customer spending (Tamara Charm, 2022; Lai, 2022), and in turn, reduces customer consumption (Gudmundsson, 2012; Romer, 1990). Literature has also found a spillover effect of a foreign country's EPU on another country (Antonakakis et al., 2018). Specifically, China's EPU impacts the economy of other countries (Biljanovska et al., 2021). As the global reviews can be in multiple languages, we split the global reviews into English and non-English reviews. Therefore, we hypothesize that Amazon shares more global reviews to account for the reduction in review rate during highly volatile conditions, improving the number of reviews customers can see. From Amazon's perspective, when the EPU increases, the sales go down even in their global markets. So they will have fewer reviews to share. As we do not know which direction of sharing is stronger, we hypothesize that

Hypothesis 2(a): USA EPU volatility positively affects the number of shared reviews from a non-USA market that Amazon shares in the USA market.

Hypothesis 2(b): USA EPU volatility negatively affects the number of shared reviews from a non-USA market that Amazon shares in the USA market.

Hypothesis 2(c): China EPU volatility positively affects the number of shared reviews from a non-USA market that Amazon shares in the USA market.

Hypothesis 2(d): China EPU volatility negatively affects the number of shared reviews from a non-USA market that Amazon shares in the USA market.

1.2.2.2 Effect of Global Review Sharing

Amazon shares more reviews, showing their customer more reviews than the product got domestically. According to Chevalier and Mayzlin (2006), past reviews positively impact future sales. Additionally, Xu et al. (2021) found a positive correlation between review ratings and demand. At the same time, giving more information is only sometimes beneficial. However, Dellarocas (2006) demonstrated that review manipulation has negative effects. Wu et al. (2019) has shown the customer's perceived risk when fake reviews persist. While Amazon is not faking or manipulating reviews, it artificially increases the number of reviews by selectively sharing reviews from its other markets and displaying them to customers in the USA, which we consider a form of artificial inflation.

At the same time, the literature suggests that customers feel burdened when presented with more information (Eppler and Mengis, 2004). Lee and Lee (2004) has shown that an overload of information to consumers can negatively affect their satisfaction. Information has an inverted U-curve relation with a consumer's decision accuracy Eppler and Mengis (2004); Roetzel (2018), stating that consumers have an optimal information processing capacity, and if the threshold is crossed, the burden of information will confuse the consumers. In product reviews, customers are more satisfied with their decision-making when the information overload is reduced (fen Hu and Krishen, 2019). The abundance of online information is linked to the possibility of experiencing information overload depending on how reviews are shown. Information overload can lead to negative outcomes, such as decreased purchase intention and heightened perceived risk (Soto-Acosta et al., 2014). In this research, we also test if sharing global reviews can benefit customers coming early in the market. When the product has fewer reviews, global reviews can show a positive impact, but not when the product already has enough domestic reviews. So we split the customers into two groups: (i)early customers: one who sees less than or equal to 100 domestic reviews for the product, and

(ii) late customers: who see more than 100 reviews for the product. Specifically, we hypothesize that

Hypothesis 3(a): The shared global reviews on Amazon have a positive impact on the "early" customers for the product.

Hypothesis 3(b): The shared global reviews on Amazon have a negative impact on the "late" customers for the product.

1.3 Data

1.3.1 Data Collection

To test our hypothesis, we collected data from three sources. The first dataset was scraped from Amazon.com, specifically reviews for Nike products sold on the site. We chose Nike as it is a USA-based company with a global market presence, similar to Amazon. Moreover, Nike is the world's most valuable sports brand, according to sportsmedia.com, making it an appropriate setting to study our research question. We used Python language and Selenium WebDriver to scrape the reviews for multiple Nike products from Amazon.com. Given the many products to scrape, the process took ten days, from July 3rd, 2021, to July 12th, 2021. We collected the review date, text, rating for all reviews scraped, and product details from each product page.

The second dataset was secondary data from previous studies by Baker et al. (2016) and Baker et al. (2013). The authors developed an Economic Policy Uncertainty (EPU) index for the USA and China. We chose China's EPU as a proxy for the global EPU, given that China is the leading manufacturing economy globally and Nike has significant manufacturing operations there. The EPU index was computed using three components: newspaper coverage related to policy-related economic uncertainty, reports by the Congressional Budget Office, and the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. The newspaper coverage was used to construct an index for the volume of news articles discussing economic policy uncertainty. The Congressional Budget Office provided lists of temporary federal tax code provisions, which were used to compile annual dollar-weighted numbers for each tax code set to expire in the next ten years, giving a measure of the uncertainty regarding the future of the federal tax code. The Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters was used to create indices of uncertainty about policy-related macroeconomic variables using the dispersion between individual forecasters' predictions of future levels of the Consumer Price Index, Federal Expenditures, and State and Local Expenditures.

The construction of the EPU index for China by Baker et al. (2013) involved several steps. Firstly, monthly counts of articles containing at least one term from three sets - economics, policy, and uncertainty - were obtained. The total number of articles for the same newspaper and month was then scaled to these counts. The sample was split into three periods, and each period's newspaper's monthly series of scaled frequency counts were standardized for unit standardization. The standardized series were averaged over newspapers by month and then normalized to an average of 100 for each period. Detailed methodology for constructing the EPU index can be found at www.policyuncertainty.com. Figure 2.6 displays the dispersion of USA and China EPU values. Notably, China's EPU has been higher than the USA's EPU in recent times, indicating an increase in global EPU values compared to the domestic USA market.

The third dataset is the quarterly e-commerce report developed by U.S. Census Bureau. This dataset comprises data on e-commerce activity in the United States. We use this dataset to obtain information on quarterly e-commerce sales in the United States.

1.3.2 Data Cleaning and Summary

Our data collection from Amazon.com consisted of a sample of 300,878 reviews. However, before conducting any analysis, we followed a data-cleaning process with four steps:

1. The first step was to remove any repeated reviews from the sample. 54,503 reviews were repeated, leaving us with a sample of 246,375 reviews.

2. The second step involved filtering out reviews that were assigned to products without a unique ASIN (Amazon Standard Identification Number). There were 2,910 reviews without an ASIN, attributed to 38 products. Eliminating these reviews resulted in a sample size of 243,465

reviews.

3. The third step was to remove three reviews that did not have a date assigned to them. This step reduced our sample size to 243,462 reviews.

4. The fourth step involved removing reviews that were posted before the product was first available on Amazon.com. We deemed these reviews fake and removed 19,975 such reviews from our sample. This step reduced our sample size to 223,487 reviews.

5. Finally, we removed products with only one review, leading to a final sample size of 174,334 reviews.

Our data consists of reviews from 15 countries. As we scraped Amazon.com, the US website, more than 50% of the reviews come from US customers. As shown in Table 1.1, the US has the highest number of reviews, with 117,729 reviews, Germany with 11,578 reviews, Italy with 7,733 reviews, and the UK with 5,160 reviews. This is consistent with the growth of Amazon, as it was first established in the US and then expanded into European and Asian markets. In our sample, out of the total 175,092 reviews in our final sample size, 117,729 reviews are from the US, and the rest of the 57,363 reviews are global reviews shared on Amazon.com. Figure 1.4 demonstrates the number of shared reviews by language over time. The number of non-English reviews shared is increasingly more than the English language reviews.

This information suggests that most of the products in our sample have relatively recent reviews. Only 41 products have had domestic reviews since 2013, and only nine products have shared global reviews as old as 2013, indicating that most of our data are centered on more recent years of 2020 and 2021. Additionally, 135 products (6%) are yet to have domestic reviews in 2021, suggesting that these products may be less popular or less frequently reviewed than other products in our sample. This may have implications for the representativeness of our data, mainly if these products are disproportionately represented in specific categories or markets.

1.4 Variable Construction

1.4.1 Variables for Hypothesis 1

1.4.1.1 Independent Variables

Under the group of hypothesis 1, three major independent variable groups were built.

- ***EPU variables***: The EPU variables were collected from Baker et al. (2013) and Baker et al. (2016). The USA and China EPU are the dataset's economic policy uncertainty index values. As shown in Table 1.3, the average EPU of China is greater than that of the USA. Using these EPU variables, we built the EPU volatility variable using a six-month window period.

- ***Recent review variables***: These variables are constructed to test the impact of the most recent five reviews on the domestic review rate. These variables are constructed from the US customer reviews on Amazon.com. Prior literature has established that customers weigh recent events more than past events (Gaur and Park (2007), Davis et al. (2021). Similarly, we expect a customer recency bias concerning online reviews too. We constructed the "recent reviews" variable group based on the most recent five reviews for the product. The recent review rate is the average review rate for the most recent five domestic reviews. The recent reviews score is the average rating of the recent five domestic reviews as seen by the customer for the product.

- ***Cumulative variables***: The cumulative variables are intended to measure the impact of historical reviews on the domestic review rate. The "domestic score" variable represents the average rating of all reviews before the most recent five reviews, as observed by customers on Amazon.com. The "domestic reviews" variable represents the total number of domestic reviews a customer can see for the product on Amazon.com. Finally, the "domestic review rate" variable represents the average review rate of all reviews before the most recent five reviews for the product as seen by the customer on Amazon.com. We can test whether past reviews impact the domestic review rate by examining these variables.

1.4.1.2 Dependent Variables

To test the hypotheses under the first group of hypotheses, the primary dependent variable of interest is the review rate for a product. The review rate is calculated as the inverse of the time difference (measured in days) between two consecutive reviews for a product in the system. The longer the time difference, the slower the reviews come in, and vice versa. The idea aligns with the concept of cycle time in Little's law. The review rate is measured as the number of reviews per month, and it provides a quantitative measure of the frequency and pace of reviews for a given product. By analyzing the factors that influence the review rate, we can gain insights into how the online business is affected by the market conditions and how businesses can leverage this information to optimize their marketing and sales strategies.

$$NextReviewRate = 30/(Date_{r+1} - Date_r) \tag{1.4.1}$$

Since multiple reviews could come on the same date making the denominator in the review rate variable zero, we had to adjust the formula. In cases where a day had multiple reviews, we used 1/number of reviews to calculate the denominator. Upon analyzing the data from Table 1.3, we can observe that the second quartile for the next review rate is 30, indicating that more than half of the reviews were posted within a day of their previous review. This suggests that the review rate was relatively high, and customers were actively reviewing the product.

A second dependent variable we used is ΔNext review rate. This variable is calculated as the difference between the next review rate and the cumulative average of the review rate calculated as follows:

$$\Delta NextReviewRate = NextReviewRate - CumulativeReviewRate \tag{1.4.2}$$

This variable captures the difference between the current performance of a product to the past performance of a performance. This accounts for any selection bias in the products we have. Products can be at multiple stages of their life cycle. Using this dependent variable, we capture the product's performance with respect to its past performance. In contrast, the "Next Review Rate" variable compares the product's performance against all other products in the dataset.

1.4.2 Variables for Hypothesis 2

1.4.2.1 Independent Variables

Under the group of hypothesis 2, three major independent variable groups were built.

- ***EPU variables***: The EPU variables were collected from Baker et al. (2013) and Baker et al. (2016). The USA and China EPU are the dataset's economic policy uncertainty index values. As shown in Table 1.3, the average EPU of China is greater than that of the USA.

- ***Recent review variables***: These variables are constructed to test the impact of the most recent five reviews on the domestic review rate. These variables are constructed from the US customer reviews on Amazon.com. Prior literature has established that customers weigh recent events more than past events (Gaur and Park (2007), Davis et al. (2021)). We constructed the "recent reviews" variable group based on the most recent five reviews for the product. The recent review rate is the average review rate for the most recent five domestic reviews. The recent review score is the average rating of the recent five domestic reviews as seen by the customer for the product.

- ***Cumulative variables***: These variables are constructed to test the impact of the history of the reviews on the domestic review rate. These variables are constructed from the US customer reviews on Amazon.com. We excluded the most recent five reviews to avoid the bias introduced by the "recency effect." The cumulative rating is the cumulative average rating of all previous reviews for the product as observed by the customer on Amazon.com. The cumulative count variable tells us how many domestic reviews a customer can see for the product on Amazon.com. The cumulative review rate is the average review rate of all previous reviews for the product, as seen by the customer on Amazon.com.

1.4.2.2 Dependent Variables

There are four primary dependent variables of interest to test our hypotheses 2. We observed that the reviews posted on Amazon.com are in multiple languages. To investigate the effect of EPU volatility on the language of shared reviews, we divided the data into two groups based on language: English and non-English. As shown in Table 1.3, the average number of shared English reviews seen by the customer is 91, while the average number of shared non-English reviews seen by the customer is 414. This suggests that Amazon shares more non-English reviews with its customers.

Furthermore, the average rating of English reviews is 4.38, compared to 4.40 for non-English reviews. This is an exciting finding where, on average, Amazon shares more non-English reviews while the rating is similar to its English counterparts. However, the average rating of shared English and non-English reviews is considerably different from the domestic review's average rating of 4.22. To detect the language of the reviews, we used the "fastText" package in R.

- ***English variables***: We developed the "English variables" based on the shared English reviews from Amazon.com. This group includes two variables: "count" and "score." The count represents the total number of shared English reviews as seen by the customer, while "score" represents the average rating of all shared English reviews as seen by the customer at the time of the review for the product.
- ***Non-English variables***: We developed the "Non-English variables" based on the shared non-English reviews from Amazon.com. This group of variables also includes two variables: "reviews" and "score." The count represents the total number of shared non-English reviews as seen by the customer, while "score" represents the average rating of all shared non-English reviews as seen by the customer at the time of the review for the product.

1.4.3 Variables for Hypothesis 3

In our study, we encountered reviews in languages other than English in our dataset. To account for the potential impact of review language on customer behavior, we utilized the fastText package in R to detect the language of each review. However, upon reviewing a random sample of reviews, we found that the language detection accuracy was unsatisfactory for a significant portion of the reviews, with an approximate error rate of 10%. To address this issue, we implemented a probabilistic categorization approach based on the language probabilities provided by the fastText package. Here are the steps we followed for language categorization:

- We estimated the language tag for each review using the fastText package.
- We collected the language probability associated with each review.

- If the language probability exceeded 0.75, we assigned the review to the corresponding language category determined by the package.

- If the language probability was below 0.75, we categorized the review as part of an unconfident group.

As a result, we classified all the reviews into three groups:

1. English confident: This category comprises reviews confidently detected as being in English with a probability greater than 0.75.

2. Non-English confident: This category includes reviews that were confidently identified as being in a non-English language with a probability greater than 0.75.

3. Unconfident: This category consists of reviews for which the language probability fell below 0.75, indicating a lack of confidence in the language classification.

By employing this categorization approach, we aimed to address the errors associated with natural language processing packages. Notably, a significant portion of the unconfident group reviews may contain similar words between two languages. For example, "excellent" in English is "excelente" in Spanish. When the textual content of a review is closely related to another language, the language detection packages may struggle to determine the review's language confidently. Moreover, when the shared words are easily understandable across languages, the language effect on reviews may be limited.

1.4.3.1 Independent Variables

Under the group of hypothesis 3, four major independent variable groups were built.

- ***EPU variables***: The EPU variables were collected from Baker et al. (2013) and Baker et al. (2016). The USA and China EPU are the datasets of economic policy uncertainty index values. As shown in Table 1.3, the average EPU of China is greater than that of the USA.

- ***Global variables***: We have four variables under the group of "Global variables," which are derived from the residual values obtained from the first stage regression. The procedure for calculating these variables is described in detail in the variable construction section of the paper. The four variables are: curated English count, curated English score, curated non-English count, and curated non-English score. Curated English count represents the number of additional English reviews shared by Amazon beyond what was predicted; curated English score represents the additional score of the English reviews shared by Amazon beyond what was predicted; non-English count represents the number of additional non-English reviews shared by Amazon beyond what was predicted, and non-English score represents the additional score of the non-English reviews shared by Amazon beyond what was predicted. The English and non-English variables are calculated based on the groups categorized as English confident and non-English confident respectively.

- ***Recent review variables***: These variables are constructed to test the impact of the most recent five reviews on the domestic review rate. A detailed description of these variables was given under the variables for Hypothesis 1 section.

- ***Cumulative variables***: These variables are constructed to test the impact of the history of the reviews on the domestic review rate. A detailed description of these variables was given under the variables for Hypothesis 1 section.

1.4.3.2 Dependent Variables

To test our hypothesis 3, the primary dependent variable of interest is the review rate for a product, as discussed in section 4.1.2.

1.4.4 Control Variables

- ***Price variables*** We collected the listed price data for each product from Amazon.com, which included multiple options based on color and size. Each product option was sold at a different price. However, considering all options, we observed the product's minimum and maximum prices on each product page. These two listed price values were used as control variables. Price is an important feature of any product when a customer evaluates it. As customer-level price data was unavailable, we used the static values of the listed price as control variables.

- ***Time trend variables*** Considering the increasing trend in the e-commerce business, we built a time trend variable: the number of months since the first review month in our dataset. These variables account for any time effects present in the data.

1.5 Analysis

In this section, we will develop an econometric model to test our hypotheses. We will build the analysis in three sections, each representing the three groups of hypotheses we have constructed.

1.5.1 Econometric setup for hypothesis 1

The analysis focuses on the domestic review rate as the dependent variable. The independent variables are divided into four groups: Macroeconomic variables, and cumulative review variables, as shown in Table 1.3.

We will test the direct impact of the EPU variables on the next review rate, Δnext review rate, next review score, and Δnext review score by using four econometric equations in combination with other independent variable groups. The independent variables in all the equations are the same, with only the dependent variable being different.

$$\mathbf{RV}_{d,r+1}^{Domestic} = \alpha_0 + \alpha_1^{MV} * MV_r^{Domestic} + \alpha_2^{MV} * MV_r^{China} + \alpha_3^{RV} * \mathbf{RV}_{p,d,1:r-5}^{Domestic} + \alpha_4^{CV} * \mathbf{CV}_r^{Domestic} \tag{1.5.1}$$

where $\mathbf{RV}_{dp,r+1}^{Domestic}$ is the vector of our dependent variables: next review rate, Δnext review rate, next review score, and Δnext review score, which we use one at once to analyze the impact of all the independent variables one at once. $MV_r^{Domestic}$ is the EPU volatility value for the USA, and MV_r^{China} is the EPU volatility value for China, $\mathbf{RV}_{p,d,r-4:r}^{Domestic}$ is a vector of recent five review variables which includes the average review rate and score, $\mathbf{RV}_{p,d,1:r-5}^{Domestic}$ is a vector of cumulative review variables, which includes the average rate, average score and the number of reviews. $\mathbf{CV}_{dp,r}$ is a vector of price control variables, including minimum product price and price range for multiple product options listed on Amazon.com and the time trend variables.

When using the cumulative review rate variables, we divide it by the cumulative reviews to account for the correlation between a product's review rate and the number of reviews a product has. In general, products have a higher review rate when they have more reviews.

1.5.2 Econometric setup for hypothesis 2

To address the second research question of what drives Amazon's global reviews in the USA market, we use only the sample data for products with reviews shared from a different country and for the reviews coming after January 2020.

We will analyze solely global products, which have additional variables compared to domestic products. These products contain "Global cumulative variables," which consist of the cumulative score and count for both English and non-English reviews. However, the curation mechanism behind Amazon's shared global reviews needs to be clarified and can result in endogeneity issues. To address this, we will use a two-stage model to correct the curated global variables and estimate their effect. In the first stage, we will estimate the global variables using the EP volatility, recent review, and cumulative variables. The estimation equations are as follows:

First stage:

$$\begin{aligned}\underbrace{\mathbf{RV}^{Global}_{p,g,1:r}}_{\text{uncorrected}} = \beta_0 + \beta_1^{MV} * MV_r^{Domestic} + \beta_2^{MV} * MV_r^{China} + \beta_3^{RV} * \mathbf{RV}^{Domestic}_{p,d,r-4:r} + \\ \beta_4^{RV} * \mathbf{RV}^{Domestic}_{p,d,1:r-5} + \beta_5^{CV} * \mathbf{CV}_r^{Domestic} + \mathbf{FE}_{gp,r}\end{aligned} \tag{1.5.2}$$

We use this equation to analyze the second set of hypotheses we developed, with four dependent variables, one for each equation. We use English reviews and English score, non-English reviews, and non-English score as dependent variables in four separate regression equations with the same independent variables. Since we are assessing the effect of the global variables, which are a part of the global review-sharing program, we only use the data after 2020 for the analysis.

1.5.3 Econometric setup for hypothesis 3

The expected values for the global variables are calculated from the first stage regression in equation 5.2, and residual values are then calculated for each data point. The residual value is used as an independent variable in the second stage of regression.

We calculate the residuals from equation 5.2 as:

$$\hat{\epsilon}_{p,g,r} = \tilde{\mathbf{RV}}^{Global}_{p,g,1:r} = \hat{\mathbf{RV}}^{Global}_{p,g,1:r} - \mathbf{GR}^{Global}_{p,g,1:r} \tag{1.5.3}$$

where,

$$\hat{\mathbf{RV}}^{Global}_{p,g,1:r} = \hat{\beta}_0 + \hat{\beta}^{MV}_1 * MV^{Domestic}_r + \hat{\beta}^{MV}_2 * MV^{China}_r + \hat{\beta}^{RV}_3 * \mathbf{RV}^{Domestic}_{p,d,r-4:r} + \hat{\beta}^{RV}_4 * \mathbf{RV}^{Domestic}_{p,d,1:r-5} + \hat{\beta}^{CV}_5 * \mathbf{CV}^{Domestic}_r + \epsilon_{p,r} \tag{1.5.4}$$

In the second stage of our analysis, we utilized the residuals obtained from equation 5.5 as independent variables to predict the domestic review. To understand the effect of sharing more or fewer reviews than expected, we employed the difference between the predicted and actual international reviews as a proxy for curation. We predicted the direct effect of the EPU volatility, global cumulative, recent domestic, and cumulative domestic variables using the following equations:

$$RV^{Domestic}_{p,d,r+1} = \delta_0 + \delta^{MV}_1 * MV^{Domestic}_r + \delta^{MV}_2 * MV^{China}_r + \delta^{RV}_3 * \tilde{\mathbf{RV}}^{Domestic}_{p,d,1:r} + \delta^{RV}_4 * \mathbf{RV}^{Domestic}_{p,d,r-4:r} + \delta^{RV}_5 * \mathbf{RV}^{Domestic}_{p,d,1:r-5} + \delta^{CV}_6 * \mathbf{CV}^{Domestic}_r \tag{1.5.5}$$

We estimate the direct effect of the China EPU volatility and USA EPU volatility on the review rate of the products using equation 5.5. We conducted a multicollinearity check (r > 0.75) before running our regression models. Additionally, we examined the dataset for rare variables (i.e., variables with less than 30 observations). We also included interaction terms in all equations.

1.6 Results

1.6.1 EPU Volatility: Results from hypothesis 1

In this section, we discuss how EPU volatility impacts the domestic review rate of a product.

1.6.1.1 USA EPU drives domestic sales

We analyzed whether the USA EPU volatility impacts the review rate of domestic products. In Table 1.5, we presented the estimation results using Equation (5.1). In column (1), we considered the effect of the USA and China EPU volatility and other independent variable groups of cumulative reviews, price, and time trend variables. Each column has a different dependent variable. The dependent variables used are next review rate, Δnext review rate, current review rate, next review score, and Δnext review score and recent review score. Our results showed that the USA

EPU volatility has a negative and statistically significant effect on all the dependent variables. This indicates that USA EPU volatility affects future review rates negatively. The coefficient of the EPU volatility index in column (1) suggested that a 1% increase in USA EPU decreases the review rate by 0.11%. It is important to note that the USA EPU and China EPU volatility values are logged variables.

Furthermore, as USA EPU increase by 1%, the Δnext review rate decreases by 0.25%, indicating that the future review rates depend on the macroeconomic conditions. Interestingly customer scores also had a negative effect because of USA EPU. The coefficient for the USA EPU volatility in column (4) indicates that as USA EPU increase by 1%, the following review score decreases by 0.04%, indicating that future ratings depend on the EPU too. These results are consistent with the literature on policy uncertainty (Handley and Limão, 2017). In summary, table 1.5 supports Hypothesis 1(a).

1.6.1.2 China EPU improves domestic sales

Table 1.5 presents China's EPU volatility estimation results. The coefficient in columns (1) and (2) for China EPU volatility suggests that as China EPU volatility increases by 1%, the review rate increases by 0.095%, and the Δnext review rate increases by 0.318%. This indicates that the review rate increases as the China EPU volatility increases, indicating more sales in the USA market. This can be because of low sales in the Chinese market and a possible overproduction of goods that might help the USA consumers. This is counter-intuitive to hypothesis 1(b). The spillover effect of policy uncertainty is not straightforward and needs more investigation.

Interestingly customer scores also had a positive effect because of China EPU. The coefficient for the China EPU volatility in column (4) indicates that as USA EPU increase by 1%, the next review score increases by 0.06%, indicating that future ratings depend on the EPU too. While we did not find evidence to support Hypothesis 1(b), we did observe a significant positive effect which refutes the literature findings.

1.6.2 Review sharing: Results from hypothesis 2

1.6.2.1 EPU drives review sharing

We analyzed the conditions under which Amazon shares global reviews and presented the results of the first stage estimation in table 1.6 and 1.7. In columns (1) - (3) of table 1.6, the number of shared non-English reviews is the dependent variable. The values indicate that when the USA EPU volatility increases by 1%, the shared non-English reviews decrease by 1.09%. Similarly, when the China EPU volatility increases by 1%, the shared non-English reviews decrease by 0.958%. In columns (4) - (6) of table 1.6, the number of shared English reviews is the dependent variable. The values indicate that when the USA EPU volatility increases by 1%, the shared English reviews decrease by 0.861%. Similarly, when the China EPU volatility increases by 1%, the shared non-English reviews decrease by 0.292%. This shows that when global markets are uncertain, Amazon has fewer reviews to share with the customers, reducing the shared reviews for USA customers. Combining the results from columns (2) and (5), we can conclude that sales are reduced when markets become uncertain, as mentioned in the traditional economics literature, thereby giving Amazon fewer reviews to share. These results have found support for hypotheses 2(b) and 2(d), thereby refuting hypotheses 2(a) and (c).

1.6.2.2 EPU drives review sharing scores

We analyzed what drives the ratings of shared reviews. In columns (1) - (3) of table 1.7, the score of shared non-English reviews is the dependent variable. The values indicate that when the USA EPU volatility increases by 1%, the shared non-English reviews decrease by 1.07%. Similarly, when the China EPU volatility increases by 1%, the shared non-English reviews decrease by 0.31%. In columns (4) - (6) of table 1.7, the score of shared English reviews is the dependent variable. The values indicate that when the USA EPU volatility increases by 1%, the score of shared English reviews decreases by 0.89%.

Similarly, when the China EPU volatility increases by 1%, the score of shared non-English reviews decreases by 0.24%. This shows that when global markets are uncertain, Amazon receives lower product scores, leaving them with low-score reviews to share with the customers. Combining the results from table 1.6 and 1.7, we can conclude that when markets become uncertain, the Amazon sharing program also takes a hit by the reduced number of reviews and by sharing low-rated reviews.

1.6.3 Global Products: Results from Hypothesis 3

1.6.3.1 Effect of Global Reviews

We analyzed whether the review rate of domestic products is driven by global review variables based on whether the customers are early or late into the market. The estimation results using equation (5.5) are reported in Table 1.8. Each column has a different dependent variable: next review rate, Δnext review rate, recent review rate, next review score, and Δnext review score. We use cumulative, price, and time variables in all models as controls. The results show that early customers perceive the language of shared reviews differently. Interestingly, there is a negative effect because of the English reviews, while a positive effect is because of non-English reviews. This is an interesting result to investigate more. However, the results show that the next review rate of early customers is impacted by the non-English score but does not impact the late customers. This implies that early customers look at the content of the review. However, the late customers are only impacted by the review volume indicating an information overload effect on the late customers (Eppler and Mengis, 2004; Lee and Lee, 2004). These results partly support hypothesis 3(a) and 3(b)

1.7 Counterfactual Analysis

Although, in the analysis of global reviews, we assessed the effect of the global review sharing program, the question of what would have the performance of the products if the global reviews were absent needs to be answered. We answer this question using the counterfactual analysis models. Counterfactual models (Heckman, 2001; Cartwright, 2007) are used to answer the what-if questions. In this part of the analysis, we study what would happen if global reviews were absent. We followed the steps listed below to employ the counterfactual techniques:

- Select the reviews which USA customers post.
- Select the reviews posted only after the review sharing date of 2020.

- Split the data into two parts (i) the USA customers who see at least one English review from a different country, (ii) the USA customers who see at least one non-English review from a different country.
- Estimate the global count and score variables individually using the cumulative variables and price variables
- Predict the global expected reviews for all the customers in the USA using the estimation in the previous step.
- Estimate the curated values for all global variables
- Estimate the review rate on the entire sample using the estimated curated values. This estimation is our "Ex-ante" prediction of the review rate.
- Convert all the global variables to zero for products with global reviews, and for products without global reviews, add the global variables using their predicted values.
- Re-estimate the review rate when the global variables are changed. This estimation is our "counterfactual" prediction of the review rate.
- Compare the counterfactual and Ex-ante prediction.

We present the findings from the counterfactual analysis in table **??**, which explores the impact of global review sharing (GRS) on customer behavior. The analysis is conducted using multiple review groups, categorized based on the year and the popularity of the products.

The results consistently show a positive effect of global reviews for products with less than or equal to 100 reviews, which are considered early-stage products. This group's "Percentage" column demonstrates a significant positive impact, suggesting that global reviews are beneficial on average, regardless of the year or specific product. This finding implies that additional global reviews can enhance customer perception and decision-making when a product is relatively new to the market.

On the other hand, for products with more than 100 reviews, classified as popular products, the results demonstrate a consistent negative effect of global reviews. The "Percentage" column in this group reveals a significant negative impact, indicating that global review sharing has a detrimental effect on average, irrespective of the year or specific product's popularity. This finding aligns with the concept of information overload in the literature (Eppler and Mengis, 2004; Lee and Lee, 2004). When a product already has a substantial number of reviews, introducing additional global reviews overwhelms customers and negatively affects their decision-making process.

In summary, the analysis reveals that global review sharing positively impacts customer perception and decision-making for early-stage products. However, introducing more global reviews has a negative effect on popular products with an adequate number of reviews. The counterfactual analysis suggests the importance of considering the stage of a product's lifecycle and the existing review volume when implementing global review-sharing strategies. By understanding the dynamics between product popularity, review volume, and the impact of global reviews, businesses can effectively leverage review sharing to enhance customer experiences and drive positive outcomes.

1.8 Conclusions

We contribute to the existing literature on online reviews by examining the impact of authentic reviews shared from different markets and the language of reviews. We provide valuable insights into how review-sharing programs affect e-commerce businesses and shed light on the influence of review language on customer behavior. Through empirical analysis of a rich dataset from Amazon.com, we demonstrate that review sharing has a differential effect on the domestic review rate based on the timing of the customer into the market. Our findings indicate that adding a shared English review increases the domestic review rate by approximately 0.02% for a late customer. In contrast, an additional non-English review increases it by approximately 0.08% for an early customer. We also find that the number of reviews Amazon shares is influenced by the product's domestic performance and economic policy uncertainty (EPU) level. In periods of low EPU, Amazon shares lesser global reviews irrespective of the language, suggesting sales, and in turn, reviews are reduced during economic certainty. We argue that Amazon's strategy of sharing genuine reviews from different markets is effective only under certain conditions.

Review sharing is only sometimes viewed positively by customers. This result has implications for sellers, consumers, and policymakers interested in cross-border trade on digital platforms. The findings from the counterfactual analysis have critical managerial implications and contribute to the existing literature on global review sharing. The results highlight the need for managers to consider the stage of a product's lifecycle and the volume of existing reviews when implementing global review-sharing strategies. Early-stage products can benefit from including global reviews, as they enhance customer perception and decision-making. However, introducing additional global reviews may lead to information overload for popular products with many reviews and negatively impact customer behavior. These insights guide managers in designing effective global review-sharing initiatives and contribute to understanding the complex dynamics between product popularity, review volume, and customer response in online platforms.

However, our study has certain limitations. Firstly, we need price data for the products, limiting our ability to analyze the effects of pricing on review rates. Future research should incorporate price data to understand its influence comprehensively. Additionally, our study focuses on a single brand, Nike, within the fashion category on Amazon, which may restrict the generalizability of our findings. Expanding the scope to include other product categories would provide a more comprehensive understanding of the impact of review sharing across different industries. Furthermore, analyzing individual customer responses to global reviews by collecting data at a customer level would deepen insights into consumer behavior.

In summary, our study highlights the importance of sharing authentic reviews to enhance the reputation of global brands on e-commerce platforms. However, caution should be exercised regarding the language of shared reviews, particularly during periods of economic certainty. Overall, our study provides new insights into the impact of review-sharing programs on e-commerce businesses and contributes to the existing literature on online reviews. Considering the limitations identified, future research should explore the interplay between economic policy uncertainty, market competition, product characteristics, and customer demographics to gain a more comprehensive understanding of online reviews. Broadening the scope of the study to include different product categories and brands would enhance the generalizability of findings. Qualitative research methods can be employed to explore the underlying mechanisms that drive the relationships between economic policy uncertainty, global reviews, and consumer decision-making. Comparative studies on review-sharing practices across various e-commerce platforms would deepen understanding of the role of online reviews in shaping consumer behavior. Additionally, investigating the impact of online reviews in conjunction with other types of user-generated content and promotional strategies and exploring the implications of emerging technologies on review systems would further advance the field. By addressing these research gaps, scholars can enhance their understanding of the multi-faceted nature of online reviews, the influence of macroeconomic factors, and their implications for businesses operating in the global e-commerce landscape.

1.9 Exhibits

Table 1.1: Nike review collection on Amazon.com by country (in July 2021)

	Products	Reviews	Score
USA	2,185	117,516	4.22
Germany	730	17,615	4.41
Italy	556	9,788	4.55
UK	557	9,501	4.45
Spain	462	7,180	4.41
France	443	4,192	4.34
Japan	282	4,042	4.37
Mexico	384	1,855	4.47
India	217	1,558	3.59
Canada	237	783	4.14
Brazil	37	86	4.62
Australia	12	38	4.55
Netherlands	13	15	3.27
Sweden	2	2	4.00
Singapore	1	1	5.00
Total	2,197	174,172	4.28

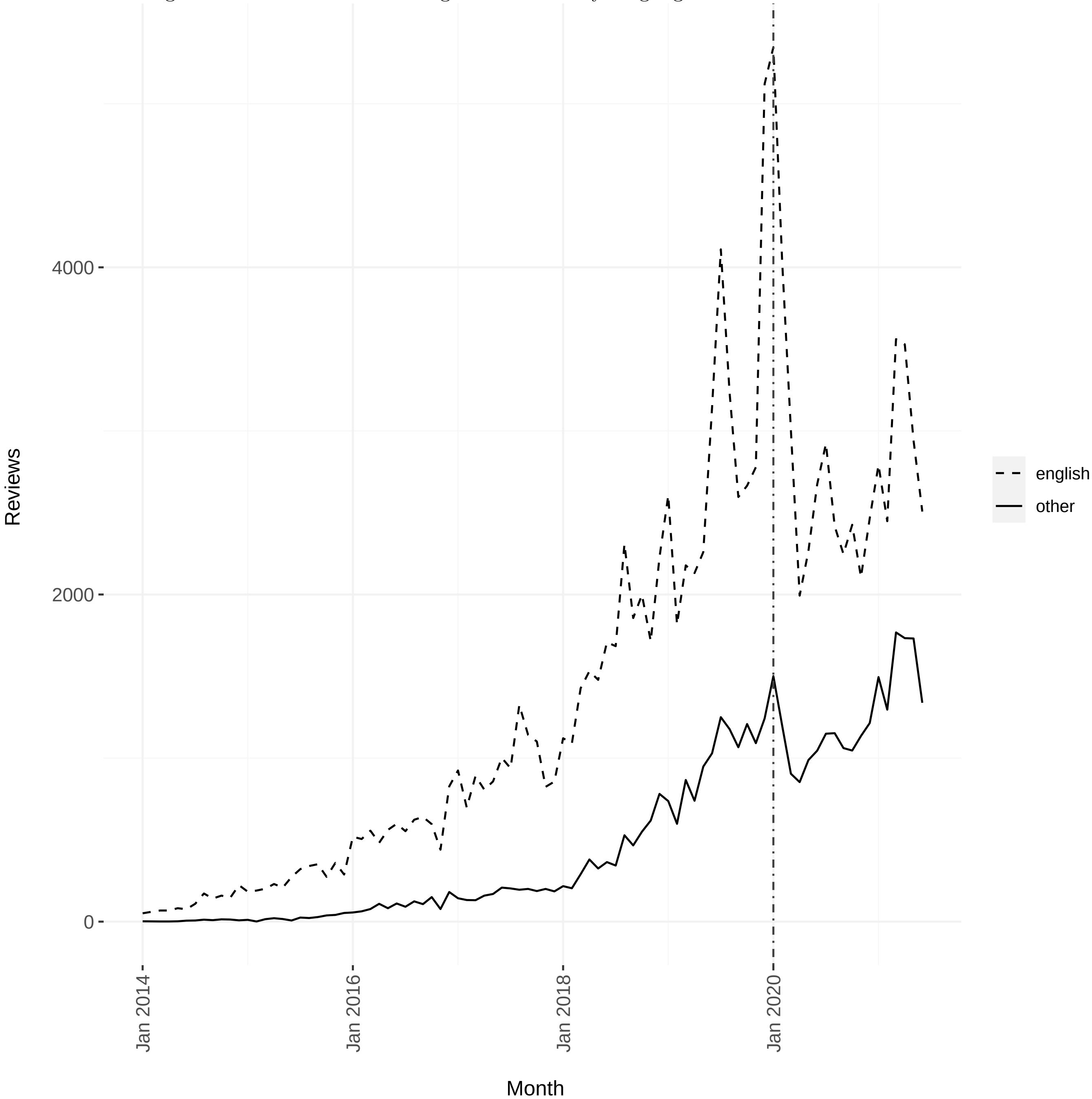
The table summarizes the number of unique products, the number of unique reviews, and the average ratings for the reviews from each country are mentioned in the table. The countries are sorted (from higher to lower) according to the number of reviews from each country in our dataset.

Table 1.2: Review data overview by customer location and year

	Products with Reviews		Products without Reviews		Total Reviews		Average Review Score	
	Domestic	Global	Domestic	Global	Domestic	Global	Domestic	Global
2013	41	9	4	36	733	22	4.22	4.83
2014	76	22	8	62	1,297	133	4.30	4.28
2015	135	51	14	98	2,968	528	4.36	4.22
2016	268	101	11	178	6,247	1,881	4.34	4.40
2017	388	190	37	235	10,179	3,294	4.27	4.19
2018	631	316	55	370	18,359	6,850	4.22	4.33
2019	1,149	618	109	640	32,430	14,159	4.26	4.26
2020	1,442	848	171	765	28,834	18,317	4.10	4.24
2021	1,250	636	135	749	16,469	11,472	4.00	4.08

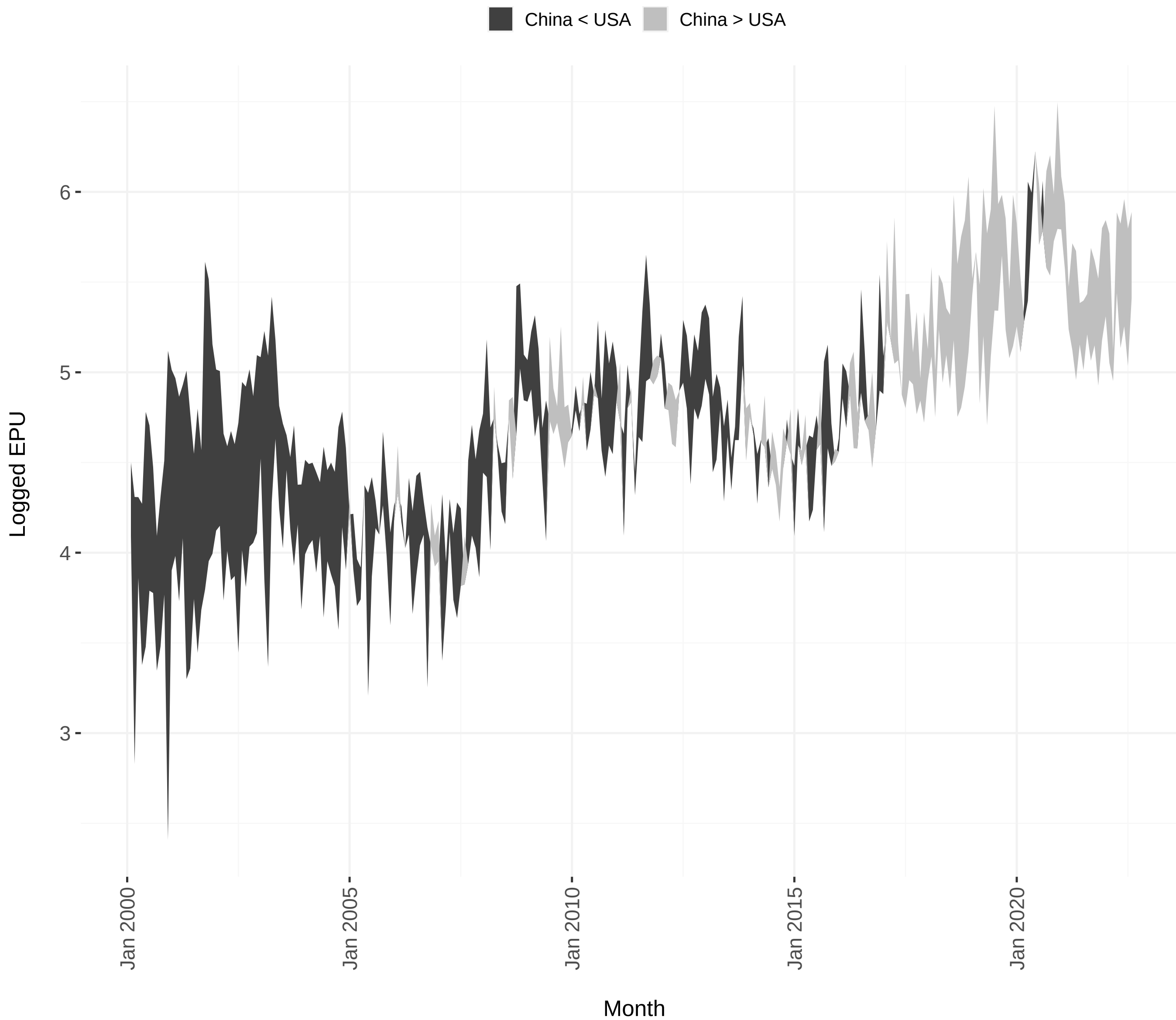
The table summarizes the number of products in each year without domestic and international reviews, the percentage of international reviews, and the difference between the mean international rating and domestic rating.

Figure 1.1: Timeline of shared global reviews by language and month



The plot displays the comparison between the number of reviews in each month by the language of the review. We used the fastText package in R to detect the language of the review.

Figure 1.2: Comparison of USA and China EPU indices



The figure compares the economic uncertainty of the USA and China over the years. The black-shaded region is when USA EPU is higher than the China EPU, and the grey-shaded region is when the China EPU is greater than the USA EPU. In the most recent times, China's EPU is predominantly greater than USA's EPU.

Table 1.3: Descriptive statistics of review and product variables

			Min	Q1	Q2	Q3	Max	Mean	SD	Description
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	USA EPU volatility	2.08	16.25	24.87	30.54	197.01	33.70	37.06	Volatility of economic uncertainty index of USA for the last 6 months
	$\mathbf{MV}_r^{China}$	China EPU volatility	16.88	68.16	76.41	96.62	176.62	86.42	34.97	Volatility of economic uncertainty index of China for the last 6 month
	$\mathbf{MV}_r^{Sales}$	USA e-commerce sales	96.17	122.59	147.59	207.72	256.60	163.99	45.31	E-commerce sales in USA in the quarter measured in billions
Reviews	$\mathbf{RV}_{p,d,r+1}^{Domestic}$	Next review rate	0.02	10.00	30.00	60.00	1260.00	60.46	87.56	The number of reviews per month, calculated as shown in Equation (4
		Next review Δrate	-5.38	-2.95	-1.54	3.07	6.95	-0.22	3.18	Difference between the next review rate to the cumualtive review rate
		Next review score	1.00	4.00	5.00	5.00	5.00	4.22	1.34	Review rating given by the customer
		Next review Δscore	-4.00	-0.41	0.51	0.70	4.00	-0.09	1.33	Difference between the next review's rating to the cumulative rating
	$\mathbf{RV}_{p,d,r-4:r}^{Domestic}$	Recent rate	0.12	18.20	37.33	73.00	1200.00	60.48	77.60	Average review rate of recent five domestic reviews
		Recent score	1.00	3.80	4.40	5.00	5.00	4.23	0.77	Average score of the recent five domestic reviews
	$\mathbf{RV}_{p,d,1:r-5}^{Domestic}$	Cumulative rate	0.12	18.23	34.28	64.13	221.88	55.98	57.41	Ratio of log average cumulative review rate to the log cumulative cou
		Cumulative count	11.00	78.00	274.00	939.00	9502.00	943.10	1656.20	Ratio of log average cumulative review rate to the log cumulative cou
		Cumulative relative rate	0.04	0.56	0.63	0.71	1.86	0.65	0.15	Cumulative number of domestic reviews as seen by the customer
		Cumulative score	1.00	4.17	4.35	4.49	5.00	4.31	0.30	Cumulative average score of the domestic reviews as seen by the custo
	$\mathbf{RV}_{p,g,1:r}^{Global}$	Cumulative English count	0.00	0.00	3.00	24.00	2386.00	90.24	308.19	Cumualtive number of shared global English reviews for the product a
		Cumulative English score	1.00	4.00	4.50	4.75	5.00	4.28	0.76	Cumulative average score of shared global English reviews as seen by t
		Cumulative English Δscore	-0.76	0.00	0.00	0.01	0.84	-0.00	0.04	Value by which the total rating increased by shared English reviews
		Cumulative Non-English count	0.00	0.00	6.00	72.00	7396.00	291.45	998.49	Cumulative number of shared global non-English reviews for the produ
		Cumulative Non-English score	1.00	4.22	4.50	4.71	5.00	4.41	0.56	Cumulative average score of the global non-English reviews as seen by
		Cumulative Non-English Δscore	-0.89	-0.00	0.00	0.01	1.79	0.01	0.06	Value by which the total rating increased by shared non-English revie
Prices	$\mathbf{PV}_{p,d,r}^{Domestic}$	Product's minimum price	8.99	30.01	47.78	84.99	812.50	60.04	42.45	Minimum listed price for product options of Amazon.com in July 2021
		Product's price range	0.00	0.00	25.44	113.17	560.04	61.86	75.63	Listed price range for product options on Amazon.com in July 2021
Controls	$\mathbf{CV}_r^{Domestic}$	Monthly time trend	26.00	97.00	110.00	120.00	133.00	106.32	18.66	Number of months since the first review in the dataset

This table summarizes the dependent and independent variables we use in our model. All the values are calculated for products after they receive at least ten reviews.

Table 1.4: Effect of product's past performance on future

			Next Review Rate (1)	Next Review Rate (2)	ΔNext Review Rate (3)	Recent Review Rate (4)	Next Review Score (5)	Next Review Score (6)	ΔNext Review Score (7)	Recent Review Score (9)
		Intercept	1.098*	−2.655*	−3.187	−0.571*	4.212*	0.395	−2.646*	−0.661†
			(0.014)	(0.433)	(2.209)	(0.071)	(0.016)	(0.571)	(0.959)	(0.314)
Reviews	$\mathbf{RV}^{Domestic}_{p,d,r-4:r}$	Recent rate	1.471*				−1.133*			
			(0.015)				(0.018)			
		Recent score	−0.496*	−0.570*	−2.573*		−0.034*	0.257*	0.242*	
			(0.005)	(0.008)	(0.023)		(0.006)	(0.010)	(0.010)	
		Recent rate * score	0.337*	0.389*	1.678*		0.292*	0.065*	0.070*	
			(0.004)	(0.004)	(0.013)		(0.005)	(0.005)	(0.005)	
	$\mathbf{RV}^{Domestic}_{p,d,1:r-5}$	Cumulative count		0.156*	−0.213*	0.082*		−0.383*	−0.549*	−0.534*
				(0.020)	(0.079)	(0.003)		(0.027)	(0.032)	(0.015)
		Cumulative rate * count		0.074*	−0.105*	0.011*		−0.007*	−0.008*	−0.007*
				(0.001)	(0.004)	(0.000)		(0.002)	(0.002)	(0.001)
		Cumulative relative rate		0.488*	−4.294*	0.560*		0.003	−0.028	0.209*
				(0.027)	(0.081)	(0.004)		(0.036)	(0.036)	(0.019)
		Cumulative count * score		−0.066*	−0.025	−0.007*		0.091*	0.128*	0.131*
				(0.005)	(0.019)	(0.001)		(0.007)	(0.008)	(0.004)
		Cumulative score		0.270*	0.270*	0.013*		−0.030	−1.163*	0.042*
				(0.020)	(0.078)	(0.003)		(0.026)	(0.030)	(0.014)
Prices	$\mathbf{PV}^{Domestic}_{p,d,r}$	Product's minimum price		−0.042*	0.153*	−0.012*		0.025*	0.020*	0.029*
				(0.005)	(0.015)	(0.001)		(0.007)	(0.007)	(0.004)
		Product's price range		0.006*	−0.015*	0.002*		0.002	0.001	0.003*
				(0.002)	(0.005)	(0.000)		(0.002)	(0.002)	(0.001)
Controls	$\mathbf{CV}^{Domestic}_{r}$	Monthly time trend		1.911*	3.997*	0.612*		1.315*	2.953*	2.266*
				(0.192)	(0.976)	(0.031)		(0.253)	(0.426)	(0.139)
		Monthly time trend^2		−0.238*	−0.471*	−0.079*		−0.169*	−0.349*	−0.284*
				(0.022)	(0.110)	(0.004)		(0.029)	(0.048)	(0.016)
Statistics		R-squared	0.467	0.560	0.151	0.871	0.055	0.062	0.065	0.651
		Adjusted R-squared	0.467	0.560	0.151	0.871	0.055	0.062	0.065	0.651
		N	115331	115329	107052	115329	115331	115329	108050	115329

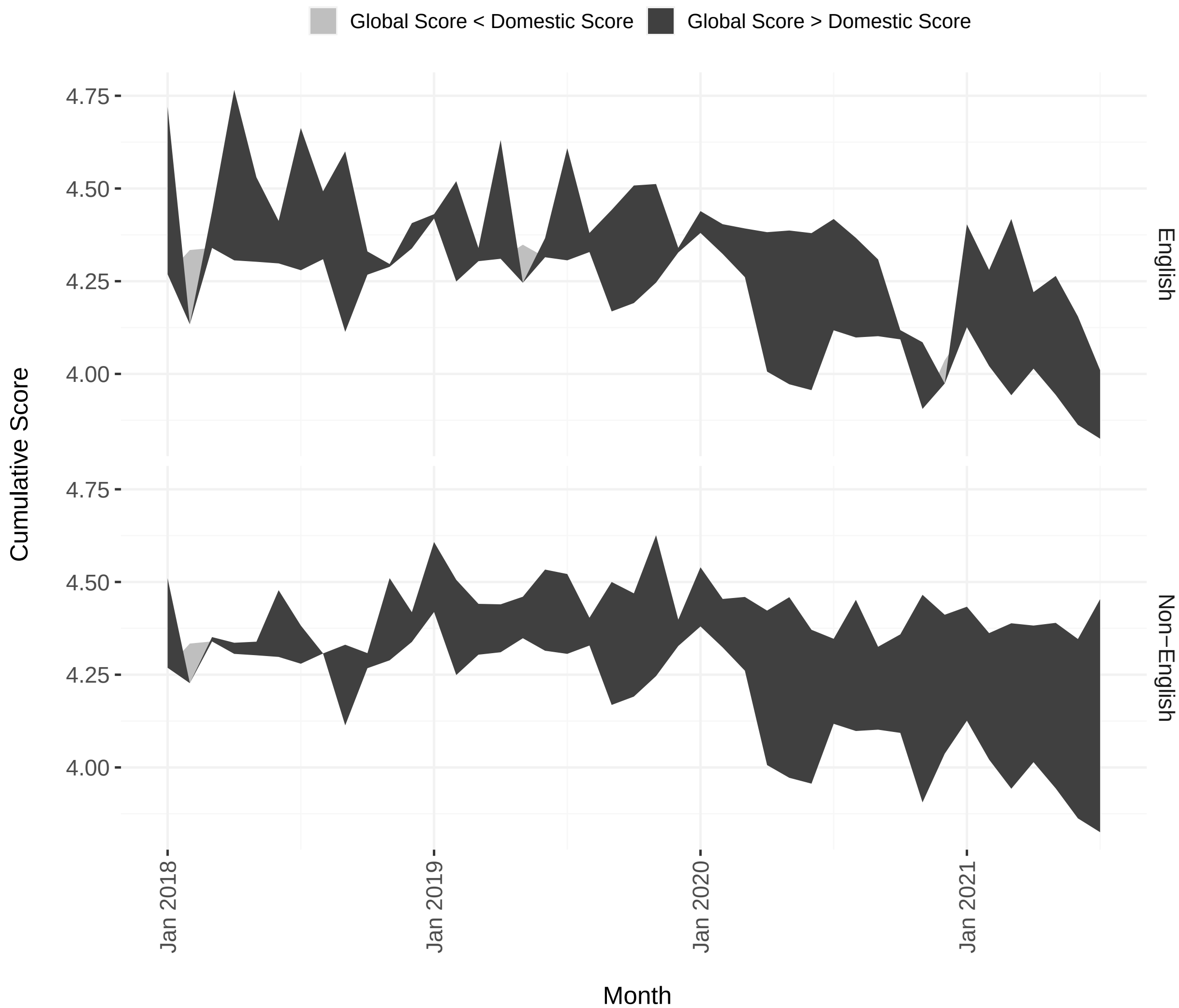
This table reports the results for domestic products. The headers are the dependent variables. $\mathbf{RV}^{Domestic}_{p,d,r-4:r}$ is a vector of domestic last five review variables, which includes the domestic review rate and the average score for the recent five reviews, $\mathbf{RV}^{Domestic}_{p,d,1:r-5}$ is a vector of domestic cumulative review variables, which includes the cumulative domestic reviews, and the average score for all but the recent five reviews, $\mathbf{PV}^{Domestic}_{p,d,r}$ is a vector of price variables, $\mathbf{CV}^{Domestic}_{r}$ is a vector of price control variables including the minimum price and price range. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 1.5: Impact of EPU volatility on domestic review rate of domestic products

			Next Review Rate (1)	ΔNext Review Rate (2)	Recent Review Rate (3)	Next Review Score (4)	ΔNext Review Score (5)	Recent Review Score (6)
		Intercept	-2.491^*	-18.793^*	0.057	0.161	-4.389^*	-4.650^*
			(0.471)	(2.557)	(0.074)	(0.595)	(1.049)	(0.364)
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	Adjusted USA EPU volatility	-0.110^*	-0.250^*	-0.017^*	-0.040^*	-0.034^*	-0.027^*
			(0.005)	(0.017)	(0.001)	(0.007)	(0.007)	(0.004)
	$\mathbf{MV}_r^{China}$	Adjusted China EPU volatility	0.095^*	0.318^*	0.023^*	0.061^*	0.057^*	0.058^*
			(0.008)	(0.023)	(0.001)	(0.009)	(0.010)	(0.006)
Controls	$\mathbf{RV}_{p,d,1:r-5}^{Domestic}$	Cumulative variables	Yes	Yes	Yes	Yes	Yes	Yes
	$\mathbf{PV}_{p,d,r}^{Domestic}$	Price variables	Yes	Yes	Yes	Yes	Yes	Yes
	$\mathbf{CV}_r^{Domestic}$	Time trend variables	Yes	Yes	Yes	Yes	Yes	Yes
Statistics		R-squared	0.500	0.013	0.864	0.025	0.027	0.552
		Adjusted R-squared	0.500	0.013	0.864	0.025	0.027	0.552
		N	115329	107052	115329	115329	108050	115329

This table reports the results for domestic products. The headers are the dependent variables. $\mathrm{MV}_r^{domestic}$ is the volatility of China EPU, and $\mathrm{MV}_r^{Domestic}$ is the volatility of USA EPU in the past six months. To check the coefficients of other variables in this analysis, refer to table 8
Significance at $^*p < 0.05$, $^\dagger p < 0.01$

Figure 1.3: Average rating of shared global reviews and domestic reviews



The figure compares the domestic and shared global ratings of a product by language. The black-shaded region shows that the global ratings (both English and non-English) are higher than the domestic rating, and the grey-shaded region shows when the domestic rating is higher than the shared global ratings (both English and non-English). The graph shows that the shared global ratings are usually higher than the domestic ones.

Table 1.6: Impact of EPU on shared review count

			Non-English Count	Non-English Count	Non-English Count	English Count	English Count	English Count
		Intercept	-4.493^{*}	1.239^{*}	-5.211^{*}	-3.683^{*}	1.688^{*}	-4.437^{*}
			(0.359)	(0.372)	(0.466)	(0.324)	(0.336)	(0.421)
Macroeconomic	$\mathbf{MV}_{r}^{Domestic}$	Adjusted USA EPU volatility		-0.958^{*}			-0.861^{*}	
				(0.015)			(0.013)	
	$\mathbf{MV}_{r}^{China}$	Adjusted China EPU volatility		-0.290^{*}			-0.307^{*}	
				(0.028)			(0.026)	
Reviews	$\mathbf{RV}_{p,d,r-4:r}^{Domestic}$	Recent rate			0.780^{*}			0.786^{*}
					(0.207)			(0.187)
		Recent score			0.032			0.062
					(0.075)			(0.067)
		Recent rate * score			$-0.105^{\dagger}$			-0.111^{*}
					(0.048)			(0.044)
	$\mathbf{RV}_{p,d,1:r-5}^{Domestic}$	Cumulative count	1.165^{*}	0.964^{*}	1.152^{*}	0.955^{*}	0.771^{*}	0.940^{*}
			(0.079)	(0.075)	(0.084)	(0.072)	(0.068)	(0.076)
		Relative rate	-0.369^{*}	-0.358^{*}	-0.585^{*}	-0.508^{*}	-0.494^{*}	-0.715^{*}
			(0.076)	(0.072)	(0.089)	(0.069)	(0.065)	(0.080)
		Cumulative count * score	-0.158^{*}	-0.113^{*}	-0.165^{*}	-0.115^{*}	-0.074^{*}	-0.121^{*}
			(0.018)	(0.017)	(0.019)	(0.016)	(0.015)	(0.017)
		Cumulative score	0.084	0.009	$0.182^{\dagger}$	0.032	−0.036	0.113
			(0.078)	(0.074)	(0.081)	(0.070)	(0.067)	(0.073)
Controls	$\mathbf{CV}_{r}^{Domestic}$	Price variables	Yes	Yes	Yes	Yes	Yes	Yes
Statistics		R-squared	0.139	0.232	0.142	0.150	0.241	0.153
		Adjusted R-squared	0.139	0.231	0.142	0.150	0.241	0.153
		N	35352	35352	35352	35352	35352	35352

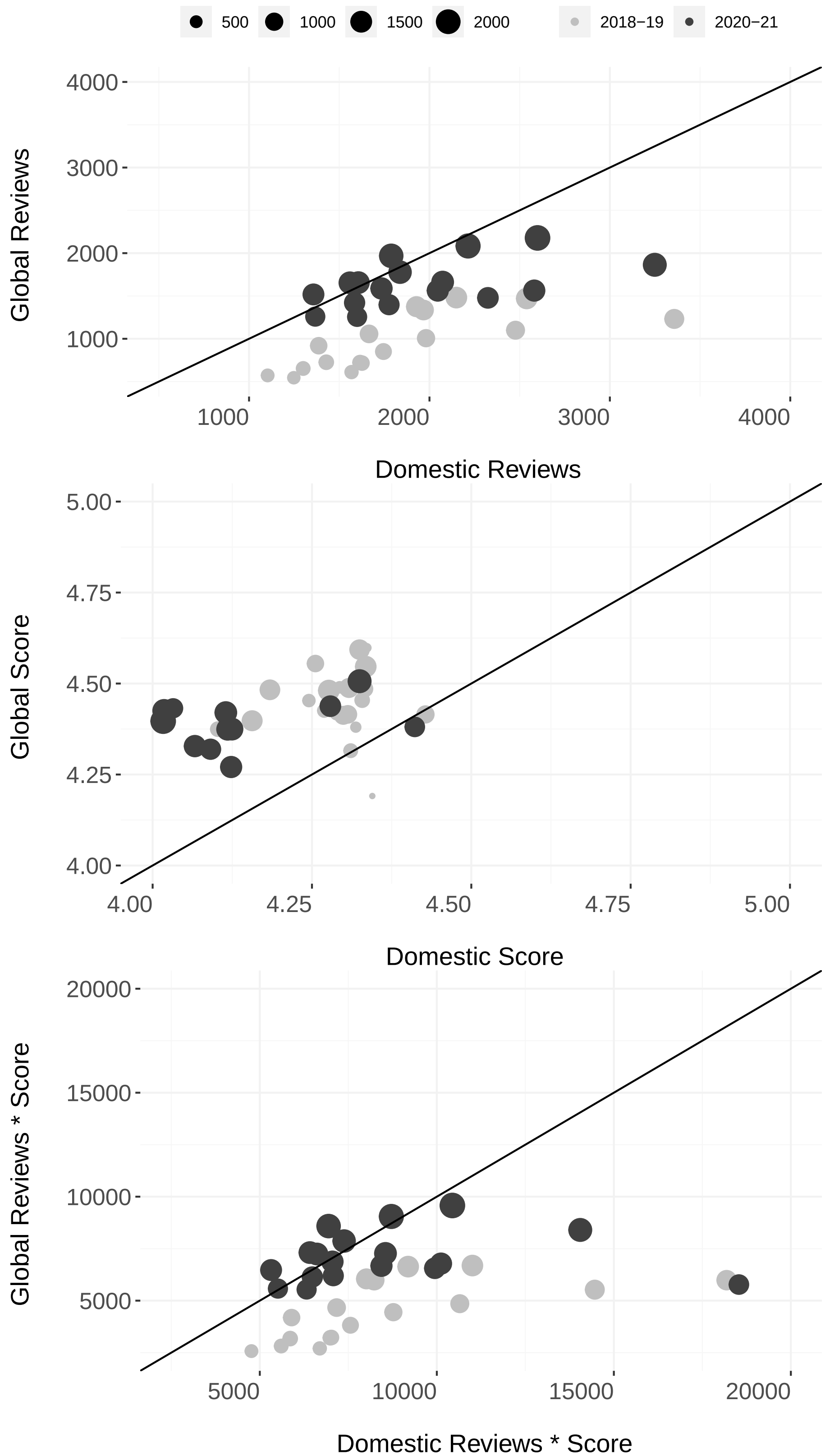
This table reports the results for domestic products. $\mathrm{EPU}_{m}^{Global}$ is the volatility of China EPU, and $\mathrm{EPU}_{m}^{Domestic}$ is the volatility of USA EPU in the past six months. $\mathrm{RV}_{gp,r+1}^{Domestic}$ is the dependent variable, next-review rate of the customer, $\mathbf{RV}_{gp,r-4:r}^{Domestic}$ is a vector of domestic last five review variables, $\mathbf{RV}_{gp,1:r-5}^{Domestic}$ is a vector of domestic cumulative review variables, $\mathbf{CV}_{gp,r}^{Domestic}$ is a vector of price control variables, and $\mathbf{FE}_{gp,r}$ is a vector of fixed effect variables. Significance at $^{*}p < 0.05$; $^{\dagger}p < 0.01$

Table 1.7: Impact of EPU on shared review score

			Non-English Score	Non-English Score	Non-English Score	English Score	English Score	English Score
		Intercept	-3.359^{*}	2.917^{*}	-4.890^{*}	-4.090^{*}	1.770^{*}	-5.800^{*}
			(0.342)	(0.347)	(0.444)	(0.325)	(0.332)	(0.423)
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	Adjusted USA EPU volatility		-1.057^{*}			-0.974^{*}	
				(0.014)			(0.013)	
	$\mathbf{MV}_r^{China}$	Adjusted China EPU volatility		-0.310^{*}			-0.301^{*}	
				(0.027)			(0.025)	
Reviews	$\mathbf{RV}_{p,d,r-4:r}^{Domestic}$	Recent rate			1.288^{*}			1.266^{*}
					(0.197)			(0.187)
		Recent score			0.207^{*}			0.230^{*}
					(0.071)			(0.068)
		Recent rate * score			-0.232^{*}			-0.229^{*}
					(0.046)			(0.044)
	$\mathbf{RV}_{p,d,1:r-5}^{Domestic}$	Cumulative count	1.756^{*}	1.535^{*}	1.770^{*}	1.687^{*}	1.483^{*}	1.745^{*}
			(0.076)	(0.070)	(0.080)	(0.072)	(0.067)	(0.076)
		Relative rate	-0.587^{*}	-0.576^{*}	-0.791^{*}	-0.635^{*}	-0.624^{*}	-0.838^{*}
			(0.073)	(0.067)	(0.085)	(0.069)	(0.064)	(0.080)
		Cumulative count * score	-0.365^{*}	-0.315^{*}	-0.377^{*}	-0.342^{*}	-0.296^{*}	-0.364^{*}
			(0.017)	(0.016)	(0.018)	(0.016)	(0.015)	(0.017)
		Cumulative score	0.562^{*}	0.480^{*}	0.682^{*}	0.674^{*}	0.598^{*}	0.815^{*}
			(0.074)	(0.069)	(0.077)	(0.071)	(0.066)	(0.074)
Controls	$\mathbf{CV}_r^{Domestic}$	Price variables	Yes	Yes	Yes	Yes	Yes	Yes
Statistics		R-squared	0.059	0.195	0.065	0.067	0.194	0.072
		Adjusted R-squared	0.059	0.194	0.064	0.067	0.193	0.072
		N	35352	35352	35352	35352	35352	35352

This table reports the results for domestic products. EPU_m^{Global} is the volatility of China EPU, and $\mathrm{EPU}_m^{Domestic}$ is the volatility of USA EPU in the past six months. $\mathrm{RV}_{gp,r+1}^{Domestic}$ is the dependent variable, next-review rate of the customer, $\mathbf{RV}_{gp,r-4:r}^{Domestic}$ is a vector of domestic last five review variables, $\mathbf{RV}_{gp,1:r-5}^{Domestic}$ is a vector of domestic cumulative review variables, $\mathbf{CV}_{gp,r}^{Domestic}$ is a vector of price control variables, and $\mathbf{FE}_{gp,r}$ is a vector of fixed effect variables. Significance at $^{*}p < 0.05$; $^{\dagger}p < 0.01$

Figure 1.4: Relation between domestic product performance and shared global reviews



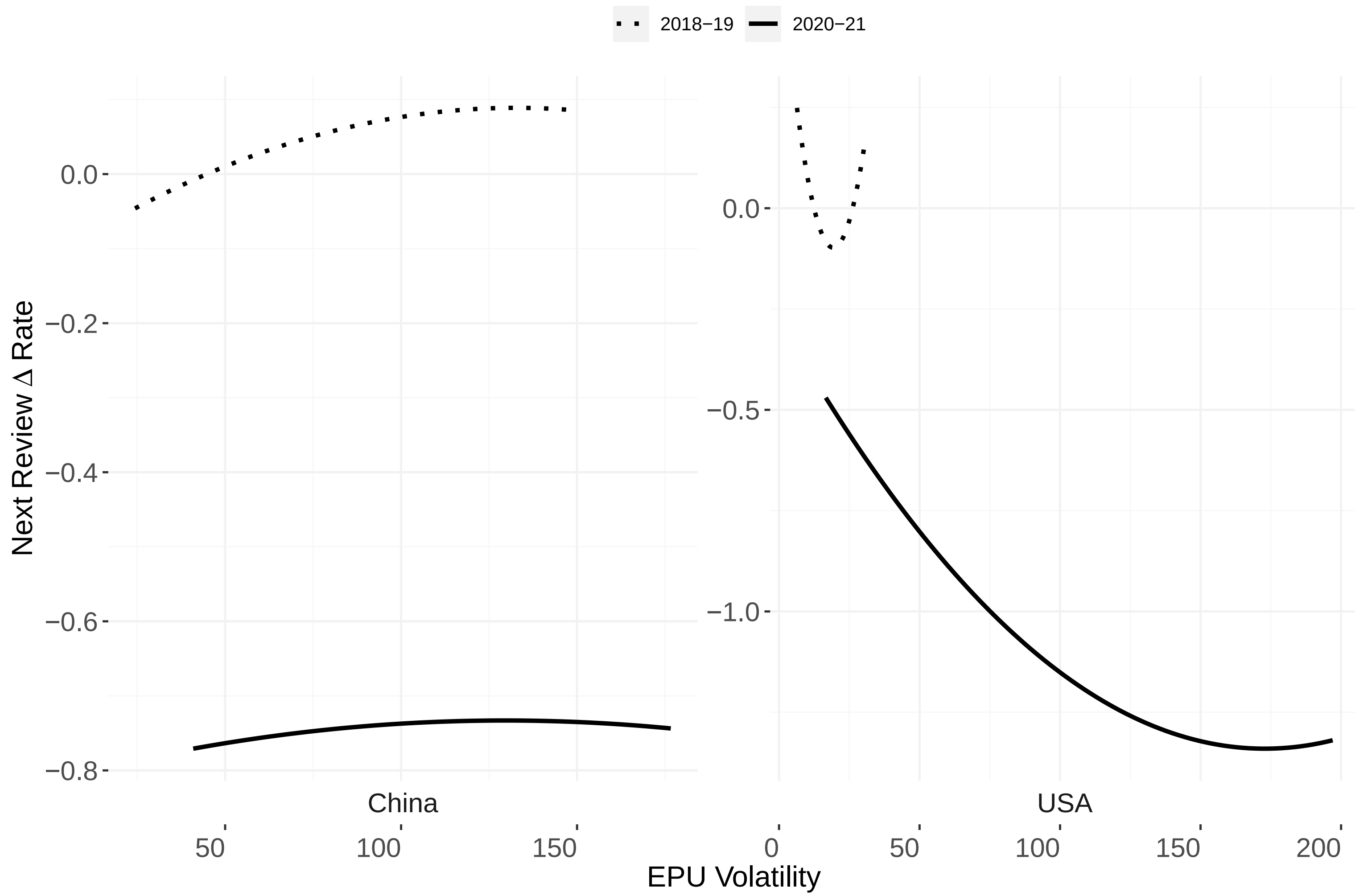
The figure depicts three different relationships. Firstly, a product's number of domestic reviews in the past month is plotted against the number of global reviews Amazon shared in the same month. Secondly, the average rating of domestic reviews in the past month is plotted against the average rating of the shared global reviews the product has received in the same month. Finally, the interaction between the number of domestic reviews and the average rating of those reviews in the past month is compared with the interaction between the number of global reviews Amazon shared and the average rating of those reviews in the same month for the product.

Table 1.8: Effects of curated global reviews on domestic reviews

			Next Review Rate		ΔNext Review Rate		Next Review Score		ΔNext Review Score	
			Early	Late	Early	Late	Early	Late	Early	Late
		Intercept	−0.225	−22.851*	3.216	−72.916*	−0.017	−14.122*	−0.017	−14.122*
			(1.219)	(1.665)	(2.633)	(5.863)	(1.426)	(2.308)	(1.426)	(2.308)
Reviews	$\mathbf{RV}^{Global}_{p,g,1:r}$	Curated English count	−0.039	−0.017	0.068	−0.107†	−0.009	0.022	−0.009	0.022
			(0.029)	(0.013)	(0.062)	(0.046)	(0.034)	(0.018)	(0.034)	(0.018)
		Curated English score	0.011	0.026*	−0.020	0.005	0.052*	−0.034*	0.052*	−0.034*
			(0.013)	(0.009)	(0.029)	(0.033)	(0.015)	(0.013)	(0.015)	(0.013)
		Curated non-English count	−0.018	0.012	−0.004	0.120*	−0.056*	−0.003	−0.056*	−0.003
			(0.020)	(0.012)	(0.044)	(0.041)	(0.024)	(0.016)	(0.024)	(0.016)
		Curated non-English score	0.031*	0.001	0.017	0.019	−0.009	0.001	−0.009	0.001
			(0.010)	(0.009)	(0.021)	(0.032)	(0.011)	(0.013)	(0.011)	(0.013)
Controls	$\mathbf{RV}^{Domestic}_{p,d,1:r-5}$	Cumulative variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	$\mathbf{PV}^{Domestic}_{p,d,r}$	Price variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	$\mathbf{CV}^{Domestic}_{r}$	Time variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Statistics		R-squared	0.216	0.418	0.016	0.014	0.028	0.031	0.061	0.006
		Adjusted R-squared	0.215	0.418	0.016	0.013	0.027	0.031	0.060	0.006
		N	28164	78888	28164	78888	28164	78888	28164	78888

This table reports the results for global products. $\mathrm{RV}^{Domestic}_{gp,r+1}$ is the dependent variable, next-review rate of the customer, $\tilde{\mathbf{RV}}^{Global}_{gp,r}$ is the group of variables constructed using the shared global reviews, $\mathrm{EPU}^{Domestic}_{g,r}$ is the EPU value for USA and $\mathrm{EPU}^{Global}_{g,r}$ is the EPU value for China, $\mathbf{RV}^{Domestic}_{gp,r-4:r}$ is a vector of domestic last five review variables, $\mathbf{RV}^{Domestic}_{gp,1:r-5}$ is a vector of domestic cumulative review variables, $\mathbf{CV}^{Domestic}_{gp,r}$ is a vector of price control variables, and $\mathbf{FE}_{gp,r}$ is a vector of fixed effect variables. "Early" and "Late" denote early customers for the product. To check the coefficients of other variables in this analysis, refer to table 9
Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$.

Figure 1.5: Impact of Cumulative and Recent reviews on the next review rate under the influence of USA EPU



The figure shows the relationship between the log of cumulative domestic reviews and the cumulative score to the future review rate and the product of the log of recent domestic review rate and the recent score to the future review rate. The lines are the fit regression lines for the variables mentioned on the axis of the plot. The solid line represents the fit line when the EPU in the USA is high, and the dashed line represents when the EPU in the USA is low.

Table 1.9: State Table - Counterfactual Results

			Ex-post Global Reviews	Ex-ante Global Reviews	Predicted Global Reviews	Overcuration Global Reviews	Counterfactual Global Reviews	Counterfactual Curated Reviews
English cumulative count	Before 2020	Products without GRS	0.00	0.00	80.01	−80.01	80.01	0.00
		Products with future GRS	56.37	0.00	70.20	−70.20	70.20	0.00
	After 2020	Products without GRS	161.38	161.38	72.20	89.18	0.00	−72.20
		Products active GRS	0.00	0.00	83.60	−83.60	83.60	0.00
Non-English cumulative count	Before 2020	Products without GRS	0.00	0.00	203.29	−203.29	203.29	0.00
		Products with future GRS	187.70	0.00	182.73	−182.73	182.73	0.00
	After 2020	Products without GRS	512.56	512.56	186.83	325.73	0.00	−186.83
		Products active GRS	0.00	0.00	226.74	−226.74	226.74	0.00
Relative English score	Before 2020	Products without GRS	0.00	−0.00	−0.00	0.00	−0.00	0.00
		Products with future GRS	0.01	0.00	−0.02	0.02	−0.02	0.00
	After 2020	Products without GRS	−0.01	−0.01	−0.02	0.01	0.00	0.02
		Products active GRS	0.00	0.00	−0.01	0.01	−0.01	0.00
Relative Non-English score	Before 2020	Products without GRS	0.00	−0.00	0.02	−0.02	0.02	0.00
		Products with future GRS	−0.00	0.00	0.02	−0.02	0.02	0.00
	After 2020	Products without GRS	0.02	0.02	0.03	−0.01	0.00	−0.03
		Products active GRS	0.00	0.00	0.03	−0.03	0.03	0.00

GRS represents "Global Review Sharing." This table reports the counterfactual results of global sharing. We split the results using three subsamples: (i) products without GRS, (ii) products with future GRS, and (iii) products with active GRS. The results tell how the domestic review rate estimation under multiple conditions changes when compared to a scenario when there is no review sharing at all by the year.

Table 1.10: Quantification of the impact of global reviews

		0 - 100 reviews			100+ reviews		
		Ex-ante	Counterfactual	Percentage	Ex-ante	Counterfactual	Percentage
2018	Products with future GRS	13.90	13.39	3.80	89.15	90.01	−0.96
	Products without GRS	7.93	7.21	9.97	42.96	43.06	−0.23
2019	Products with future GRS	14.74	14.33	2.88	74.54	75.49	−1.26
	Products without GRS	7.45	6.68	11.47	53.94	54.04	−0.20
2020	Products with active GRS	15.87	15.47	2.58	73.75	74.12	−0.50
	Products without GRS	9.11	8.44	7.98	59.41	59.66	−0.42
2021	Products with active GRS	14.42	14.13	2.02	71.51	71.52	−0.02
	Products without GRS	15.50	14.94	3.77	58.52	58.73	−0.36

GRS represents "Global Review Sharing." This table reports the counterfactual results of global sharing. We split the results using two samples: (i) only global products and (ii)global and domestic products. "$<$ expected" represents the set of reviews when the expected shared global reviews were less than expected, and "$>$ expected" represents when the expected shared global reviews were more than expected. The results tell how the domestic review rate estimation under multiple conditions changes when compared to a scenario when there is no review sharing at all by the year.

Table 1.11: Moderating effect of EPU on global review sensitivity

			Non-English Curation IRF (1)	ΔNon-English Curation IRF (2)	English Curation IRF (3)	ΔEnglish Curation IRF (4)
		Intercept	395.993	−2813.606	878.474	−6491.102†
			(2379.196)	(3552.792)	(2103.028)	(3514.241)
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	Adjusted USA EPU volatility	−0.045	0.080	0.080	0.029
			(0.087)	(0.129)	(0.077)	(0.128)
	$\mathbf{MV}_r^{China}$	Adjusted China EPU volatility	0.123†	0.010	−0.021	0.010
			(0.060)	(0.089)	(0.053)	(0.088)
Controls	$\mathbf{RV}_{p,d,1:r-5}^{Domestic}$	Cumulative variables	Yes	Yes	Yes	Yes
	$\mathbf{PV}_{p,d,r}^{Domestic}$	Price variables	Yes	Yes	Yes	Yes
	$\mathbf{CV}_r^{Domestic}$	Time variables	Yes	Yes	Yes	Yes
Statistics		R-squared	0.120	0.014	0.127	0.015
		Adjusted R-squared	0.119	0.013	0.126	0.014
		N	12194	11893	11677	11503

The table reports the sensitivity analysis between the dependent variable of the previous models, the next review rate, and our significant independent variables of non-English count and English count. IRF stands for Impulse response function sensitivity. In this table, each column represents a model with column (1) having a dependent variable as non-English count IRF, meaning the dependent variable is the ratio of the next review rate and the non-English count, and the independent variables are USA and China EPU volatility. In column (2), Δtells us it is the ratio between Δnext review rate and the non-English count. Columns (3) and (4) follow the same strategy as columns (1) and (2) but replace the non-English count with the Enlgish count.

Chapter 2

Understanding the Impact of "Made in China" and "Made in USA" Mentions on Amazon Reviews

2.1 Introduction

The globalization of markets has facilitated the sourcing of products from various countries worldwide, providing businesses greater access to international markets. However, alongside this convenience, there is a prevailing notion that consumers prefer locally produced goods. According to a survey conducted by Nielsen, a significant majority of consumers, approximately 75%, deem it essential to purchase local products (Washington, 2020). This preference has several reasons, including the desire to support local economies and a stronger sense of connection to locally sourced products. Despite the growing emphasis on local sourcing, it remains unanswered whether this consumer preference for local products translates into actual online purchasing behavior, considering the wide availability of products from different geographical locations.

Furthermore, it is noteworthy that imports from China continue to exhibit a rising trend despite the ongoing trade tensions between the United States and China. In the United States, for instance, imports from China have demonstrated consistent growth (Bown, 2023). This increase raises questions about the impact of product origin information on consumer behavior in online businesses. This paper examines the influence of product origin source information on the review rate of products in online businesses. By analyzing the relationship between consumer reviews and product origin, this study sheds light on how product origin affects consumer behavior and purchasing decisions in an online setting.

Online reviews play a crucial role in enabling customers to assess the quality of products (Kaemingk, 2020). These reviews serve as a valuable resource for customers, allowing them to make well-informed purchasing decisions and evaluate the credibility of a brand (Chakraborty and Bhat, 2018). Moreover, reviews provide customers with insights into the country of origin of a product, which has gained increasing importance in recent years. Since 2021, Amazon has mandated that sellers provide information regarding the country of origin (Dawson, 2021; Verma, 2020). However, before this requirement, no reliable source of origin information was available to Amazon's customers. This research focuses on examining the impact of origin information derived from reviews on the business of Amazon.com. Specifically, we aim to address how references to "made-in USA/China" in customer reviews influence Amazon.com's customers. By exploring the effects of such mentions on customer behavior, this study aims to provide insights into the significance of origin information in shaping customer perceptions and decisions within the context of Amazon.com.

The awareness of a product's source location notably impacts customers' perception, satisfaction, and purchasing behavior (Zhang and Merunka, 2015). Positive associations with specific countries can enhance customers' satisfaction and confidence in their purchase decisions, while negative biases towards specific locations may discourage customers from purchasing. Furthermore, the source location of a product plays a significant role in influencing customers' purchasing decisions. Some customers prioritize supporting local economies or seeking products with ethical manufacturing practices, while others prioritize cost considerations. Businesses can capitalize on this knowledge to customize their marketing strategies and effectively meet customer expectations. By understanding the importance of product source location to customers, businesses can align their messaging, branding, and sourcing practices to resonate with customers' preferences. This strategic approach allows businesses to establish stronger connections with their target audience, enhance customer satisfaction, and drive positive purchasing behavior.

Transparency into a product can also build trust among customers (Buell et al., 2021). This can create a positive impression among customers, who are more likely to trust and remain loyal to brands that prioritize transparency. Overall, the information captured in online reviews, including the source country of the product, can greatly impact customer satisfaction and decision-making. Brands that prioritize transparency and ethical sourcing can benefit from positive reviews and increased customer loyalty.

In addition to transparency, the economic conditions of a country significantly impact customer buying behavior (Tamara Charm, 2022). A volatile economy can profoundly influence customer behavior, leading to cautious spending and changes in purchasing priorities. Customers tend to adopt a more conservative approach to their purchases during economic uncertainty (Lai, 2022). They prioritize essential items and may opt for lower-cost alternatives(Eberly, 1994). This cautious behavior stems from concerns about job security and income stability, as customers aim to maintain financial resilience amidst economic fluctuations. Furthermore, customer preferences may undergo a shift during uncertain economic times. Increased price sensitivity becomes evident, with customers placing a greater emphasis on obtaining value for their money. In response to these changes in customer behavior, businesses must adapt their marketing strategies and offerings accordingly. Emphasizing value, affordability, and stability becomes crucial in attracting and retaining customers during volatile economic periods. Understanding economic indicators and trends becomes essential for businesses to anticipate customer behavior shifts and proactively adjust their strategies. By keeping a pulse on economic conditions and aligning their marketing efforts with customer behavior, businesses can navigate the challenges of an uncertain economy. By providing value and stability to customers, businesses can enhance their competitive edge and maintain customer loyalty even in challenging economic circumstances.

The main dataset used in this research was a secondary dataset collected by (Ni et al., 2019), which consisted of product reviews from Amazon.com, scraped in 2018. This dataset provided information on review attributes such as rating, reviewer name, review text, and the Amazon Standard Identification Number (ASIN). Additionally, secondary data on the Economic Policy Uncertainty (EPU) index for the USA and China, developed by , was utilized to capture macroeconomic variables.

Furthermore, macroeconomic variables were incorporated into the analysis using secondary data from the Economic Policy Uncertainty (EPU) index for the USA and China. The EPU index, developed by Baker et al. (2013) and Baker et al. (2016), provided insights into the macroeconomic conditions and uncertainty levels prevailing in the respective countries during the study period. We considered these macroeconomic variables to understand the broader economic context and its potential influence on customer behavior. By leveraging this combined dataset, encompassing product reviews and macroeconomic indicators, this research aimed to understand the relationship between product origin information, customer reviews, and macroeconomic factors.

We formulated multiple sets of variables by considering the historical and recent performance of the products to investigate the research question. The initial set of variables concentrated on mentions of "made in" and specifically examined references to "made in USA" and "made in China." These variables were derived from the review texts available in the dataset, enabling us to analyze the effect of product origin on customer behavior. We also calculated additional variables such as review rate, review score, and their corresponding changes (Δ) to assess the impact of "made in" mentions on customer review rates and scores. This comprehensive approach allowed us to examine the influence of product source information on various aspects of customer reviews and ratings.

We conducted data analysis using Ordinary Least Squares (OLS) methods to examine the data. Using OLS methods allowed us to evaluate the direct effects of the "made in" mention variables on the changes in review rates and review scores (Δreview rate and Δreview score). By incorporating controls for relevant factors such as the monthly time trend and product category, we aimed to isolate the impact of product source on customer behavior. This approach gave us valuable insights into the relationship between "made in" mentions and customer review rates and scores. By investigating the effects of product sources on customer behavior, this study contributes to our understanding of how consumers perceive and evaluate products based on their origin. The results of our analysis hold significant implications for businesses, as they provide insights that can inform the development of targeted marketing strategies. By effectively leveraging the influence of product sources, businesses can drive customer satisfaction and enhance their competitive advantage. Overall, this analysis offers valuable findings that shed light on the relationship between "made in" mentions and customer behavior. By examining the impact of product sources on customer perception and evaluation, this study provides valuable insights for businesses seeking to optimize their marketing efforts and capitalize on the influence of product origin.

2.2 Literature

2.2.1 Local Made

The literature extensively recognizes the significance of the country of origin in consumer decision-making. With the advent of online platforms and the proliferation of social media, consumers have increasingly relied on online reviews as a vital source of information to guide their purchasing decisions (Podium, 2017). Within this context, consumers tend to attribute importance to the origin of a product and perceive products manufactured in specific countries as possessing higher quality (Moriuchi, 2021). In the United States, customers often positively perceive products made within the country (Nagashima, 1970; Elliott and Cameron, 1994). This favorable perception can influence their review behavior. When customers perceive a product as being made in the USA, it can shape their evaluation and subsequently impact the nature and tone of their online reviews. Understanding these underlying consumer perceptions is crucial for businesses seeking to effectively manage their online reputation and capitalize on the positive associations with a specific country of origin.

When customers encounter mentions of products made in the USA, these mentions are observed to evoke a positive association with the product's origin (Nagashima, 1970). This association influences customers' perception of the product's quality and subsequently impacts their review behavior. Customers attribute higher quality to products made in the USA, prompting them to leave positive reviews (Moriuchi, 2021). Furthermore, actively mentioning the USA as the country of origin elicits feelings of patriotism or national pride among USA customers (Yu et al., 2022; Maronick, 1995). These vibrant connections to the country influence customers' review behavior, potentially resulting in more positive reviews. The positive perception and emotional resonance associated with products made in the USA actively influence customers' review rates and the tone of their reviews. Businesses aiming to leverage the impact of product origin on customer behavior and reviews must actively understand these dynamics.

Numerous studies have demonstrated that customers' knowledge of a product's source location actively influences their perception, satisfaction, and purchasing behavior (Zhang and Merunka, 2015; Chen and Zhong, 2023). In the case of products manufactured in the USA, customers often have preconceived notions regarding the quality and reliability of such products, increasing their inclination to provide positive reviews. This positive association between the USA as the country of origin and customer perception fosters a sense of trust and confidence, further reinforcing their propensity for positive review behavior (Nagashima, 1970). It is worth noting that the extent to which the USA mentions affect the review rate may vary across different product categories or industries. To summarize, based on the extensive literature, an augmented frequency of USA mentions will positively influence the review rate. This hypothesis is grounded in consumers' perception of products made in the USA as being of higher quality and the positive associations and feelings of national pride that such mentions evoke. Specifically, we propose the following hypothesis:

Hypothesis 1 (a): The number of USA mentions will influence the review rate positively.

The influence of country of origin on consumer behavior extends beyond products made in the USA. Extensive research suggests consumers hold diverse perceptions of products from different countries, including China. Historically, products made in China are associated with lower quality (Schniederjans et al., 2004, 2011). However, when customers encounter mentions of products made in China, it can shift their perception.

Recent research by Yu et al. (2022) suggests that the rise of nationalism has positively affected consumers' perception of a product's country of origin. As consumers become more aware of China's manufacturing capabilities and the quality of its products, their perception of products made in China may become more negative. This change in perception can be attributed to the evolving awareness and changing attitudes toward China's manufacturing industry and the associated quality standards. It is essential to acknowledge that consumers' perceptions of products made in China are subject to various factors, including their individual experiences, cultural influences, and exposure to information. While historical associations may have shaped the initial perception of products made in China, the evolving narrative driven by increased awareness and changing perceptions should be considered when examining the impact of product origin on consumer behavior.

Furthermore, the impact of China's mentions on the review rate can be contingent upon the context. On online platforms like Amazon.com, the price often emerges as the most significant factor for customers (Mohsin, 2022). While positive perceptions of products made in China may exist, customers' review behavior can be influenced by additional factors such as price (Drozdenko and Jensen, 2009). As a result, the mention of products made in China can positively influence the review rate if customers associate these products with affordability. However, the review rate may be negatively affected if customers associate products made in China with nationalism or other negative connotations. Hence, it is essential to consider the various factors that shape customers' perceptions and preferences when examining the impact of product origin on the review rate. Factors such as price and customer value perceptions can interact with the perception of products made in China, leading to different outcomes in review behavior.

Specifically, based on the considerations above, we propose the following hypothesis:

Hypothesis 1 (b): The number of China mentions will influence the review rate positively.

Hypothesis 1 (c): The number of China mentions will influence the review rate negatively.

2.2.2 Economic Policy Uncertainty

Previous research has extensively examined the impact of Economic Policy Uncertainty (EPU) on economic variables and consumer behavior. During economic uncertainty, businesses adopt more cautious strategies (Çolak et al., 2017). These strategies often involve delaying investments (Jens, 2017) and reducing capital spending (Gulen and Ion, 2015) to mitigate risks and maintain financial stability. However, such decisions can adversely affect sales and revenue generation for businesses. Moreover, policy uncertainty has increased domestic prices and reduced consumers' income (**??**). These factors, in turn, can lead to changes in consumer spending patterns. Corporate reports have also highlighted the negative impact of uncertainty on customer spending (Tamara Charm, 2022; Lai, 2022), which may decrease overall consumption (Romer, 1990; Gudmundsson, 2012). The findings from previous research highlight the intricate relationship between Economic Policy Uncertainty, economic variables, and consumer behavior. Understanding the impact of policy uncertainty on businesses and consumers is vital for policymakers and industry practitioners alike, as it allows for better decision-making and strategic planning in uncertain economic environments.

Furthermore, uncertainty notably impacts household consumption (Eberly, 1994). Research has demonstrated that higher levels of uncertainty are associated with reduced consumption and a delay in non-essential purchases (Bernanke, 1983). During times of uncertainty, households tend to prioritize essential goods and services over non-essential items. Consequently, businesses may experience decreased product demand, resulting in declining sales. Moreover, previous studies have observed the volatility of Economic Policy Uncertainty (EPU), as indicated by the clustering of volatility around the EPU index (Yu et al., 2017; Phan et al., 2018), suggesting that EPU volatility carries more economic significance than the EPU itself (Goodell et al., 2021). The volatility in EPU can have broader implications for economic variables and consumer behavior, affecting business operations and market dynamics. Understanding the relationship between uncertainty, household consumption, and economic variables is crucial for policymakers and businesses to navigate uncertain economic conditions effectively. By considering the impact of uncertainty on consumer behavior, businesses can adapt their strategies to align with changing consumption patterns and mitigate the effects of economic volatility.

Based on the existing literature, an increase in the USA's Economic Policy Uncertainty (EPU) negatively affects mentions of products made in the USA. This hypothesis is grounded in the potential impact of uncertainty on business sales and consumer preferences. As EPU increases, businesses may adopt more conservative strategies, which can result in decreased sales. The uncertainty surrounding economic policies and conditions can also influence consumer behavior, leading to changes in spending patterns and potentially affecting preferences for products made in the USA.

Furthermore, fluctuations in domestic prices and income levels, driven by EPU, can further contribute to consumer behavior and preferences alterations. It is worth noting that products made in the USA often carry a perceived higher price than those made in other countries (Ha-Brookshire and Yoon, 2012). This perception of higher prices may interact with consumer responses to EPU, potentially impacting mentions of products made in the USA.

Based on these considerations, we propose the following hypothesis:

Hypothesis 2 (a): Increased USA's EPU negatively affects the mentions of products made in the USA.

The spillover effects of economic uncertainty from one country to another have received significant attention in prior research. Economic conditions and trade relations between countries can contribute to the transmission of uncertainty across borders. Specifically, an increase in Economic Policy Uncertainty (EPU) in one country can potentially adversely affect the economy of another country (Kelly et al., 2016). Recognizing and comprehending these spillover effects is vital, particularly when investigating the relationship between China's EPU and the mentions of products made in China. The interconnectedness of global economies and the interdependence of countries in the global market highlight the significance of understanding how economic uncertainty in one country can reverberate and impact other countries. The transmission of economic uncertainty can influence various economic indicators, consumer behavior, trade patterns, and overall market dynamics. Therefore, when examining the mentions of products made in China, it is essential to consider the potential spillover effects of China's EPU on consumer perception, market conditions, and the overall economic environment both domestically and internationally.

Numerous studies have underscored the significant impact of international spillovers on the dynamics of other countries (Klößner and Sekkel, 2014). One country's Economic Policy Uncertainty (EPU) has been observed to exert substantial adverse effects on the market of another country (Ghirelli et al., 2021). These findings emphasize the interconnectedness of economies and the potential transmission of uncertainty across borders. Given the strong trade ties between the USA and China, it is reasonable to anticipate that an increase in China's EPU can adversely affect the USA market; potentially influencing consumer perceptions and preferences for products made in China. The transmission of uncertainty between the two countries can impact various economic factors, trade relations, and consumer behavior. Understanding these spillover effects is essential for comprehending the broader context in which consumer preferences and market dynamics operate, particularly when examining the mentions of products made in China in the USA market.

In conclusion, drawing from the existing literature, an increase in the USA's Economic Policy Uncertainty (EPU) negatively affects the mentions of products made in the USA. Similarly, an increase in China's EPU negatively affects the mentions of products made in China. These hypotheses are grounded in research that explores the influence of economic uncertainty on a range of economic variables, consumer behavior, and the potential spillover effects across countries. The impact of EPU on consumer behavior, market dynamics, and international spillovers has been extensively studied. The interconnectedness of economies, trade relations, and the transmission of uncertainty highlights the potential effects on consumer perceptions and preferences for products made in the USA and China. By examining the implications of EPU on the mentions of products made in each country, we can gain valuable insights into the relationship between economic uncertainty and consumer behavior. However, it is essential to note that further empirical analysis is required to validate these hypotheses and understand the subtle effects of EPU on consumer behavior and product mentions. Future research can delve deeper into these relationships, considering additional factors such as market conditions, industry-specific dynamics, and the evolving nature of consumer preferences.

Hypothesis 2 (b): Increased China's EPU negatively affects the mentions of products made in China.

2.3 Data

We collected data from four sources to investigate our research question. The first dataset is a secondary dataset obtained from Ni et al. (2019), who scraped and updated the dataset in 2018. This dataset includes all product reviews from Amazon.com in 2018, providing information on various review attributes such as rating, reviewer name, and review text. We extracted data on review rating, review date, Amazon Standard Identification Number (ASIN), and review text from this dataset.

The second dataset is also from Ni et al. (2019), who scraped the product description data and price for the related products from the first dataset mentioned above. We gathered a textual description of the products and the Amazon listed price from this dataset.

The third dataset is secondary data from previous studies by Baker et al. 2016 and Baker et al. 2013. The authors developed an Economic Policy Uncertainty (EPU) index for the USA and China. We chose China's EPU as a proxy for the global EPU, given that China is the leading manufacturing economy globally, and Nike has significant manufacturing operations there. The EPU index was computed using three components: newspaper coverage related to policy-related economic uncertainty, reports by the Congressional Budget Office, and the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. The newspaper coverage was used to construct an index for the volume of news articles discussing economic policy uncertainty. The Congressional Budget Office provided lists of temporary federal tax code provisions, which were used to compile annual dollar-weighted numbers for each tax code set to expire in the next ten years, giving a measure of the uncertainty regarding the future of the federal tax code. The Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters was used to create indices of uncertainty about policy-related macroeconomic variables using the dispersion between individual forecasters' predictions of future levels of the Consumer Price Index, Federal Expenditures, and State and Local Expenditures.

Constructing the EPU index for China by Baker et al. 2013 involved several steps. Firstly, monthly counts of articles containing at least one term from three sets - Economics, Policy, and Uncertainty - were obtained. The total number of articles for the same newspaper and month was then scaled to these counts. The sample was divided into three periods, and each period's newspaper's monthly series of scaled frequency counts was standardized to have a unit standardization. The standardized series were averaged over newspapers by month and then normalized to an average of 100 for each period. Detailed methodology for constructing the EPU index can be found at www.policyuncertainty.com. Figure 2 displays the dispersion of USA and China EPU values.

Notably, China's EPU has been higher than the USA's EPU in recent times, indicating an increase in global EPU values compared to the domestic USA market.

The fourth dataset is the quarterly e-commerce report developed by U.S. Census Bureau. This dataset comprises data on e-commerce activity in the United States. We use this dataset to obtain information on quarterly e-commerce sales in the United States.

2.3.1 Data Cleaning and Summary

Our study utilized a comprehensive dataset consisting of 27 product categories. However, we excluded specific categories such as gift cards, Kindle Store, music, and magazines from the analysis due to their content-dependent nature, which makes their reviews less influenced by the country of manufacture. Our dataset comprises over 178 million reviews, ensuring a robust and representative sample for our analysis. Among these reviews, the largest dataset pertained to clothing and jewelry, encompassing approximately 19 million reviews, followed closely by the electronic products dataset. It is worth noting that most product mentions in our dataset were related to products made in China. This observation underscores the significant role of Chinese manufacturing in the context of our study.

An Oberlo report on Amazon.com in 2022 (Mohsin 2022) revealed that electronics, apparel, and home and kitchen goods were the most popular shopping categories on the platform. Interestingly, as depicted in Table 3.2, our dataset predominantly consists of data from these categories. Even though the data was scraped in 2018, this alignment between our dataset and the popular shopping categories observed in 2022 adds relevance and applicability to our findings. It suggests that our dataset can is representative of the current trends on Amazon.com. Furthermore, it is essential to highlight that our dataset covers a period of 252 months, as indicated by the monthly time trend variable presented in Table 3.3. This extensive time range allows us to capture the dynamics and changes in consumer behavior over an extended period, enhancing the depth and breadth of our analysis.

The initial data size across all product categories was 148,830,623. However, while extracting information from the dataset, some reviews needed values for rating, year, month, and date variables, essential for building our group of independent variables. After removing these missing values, the data size was reduced to 136,607,712. Furthermore, we excluded reviews from categories such as gift cards, Kindle Store, magazines, music, and software, further reducing the sample size to 131,152,449. Additionally, as the dependent variable in the subsequent analysis is the review rate, which cannot be calculated for the first review of a product, a portion of the dataset was eliminated, further condensing the dataset to 118,967,739 observations. The final dataset included a total of 2,921,293 unique products. However, the description information could only be extracted and matched for 524,196 products, possibly due to data storage issues. Despite this, the final dataset remains rich and substantial, allowing us to effectively study the impact of "made in" mentions on customer review rates. In summary, the final dataset consists of 118,967,739 observations, providing a robust and extensive dataset for analyzing the relationship between "made in" mentions and customer review rates.

2.3.2 Variable Construction

2.3.2.1 Dependent and Independent Variables

To study the research questions, we have built variables under multiple categories.

- ***Macroeconomic Variables***: The macroeconomic variables related to Economic Policy Uncertainty (EPU) were obtained from the studies conducted by Baker et al. (2013) and Baker et al. (2016). To measure the volatility of EPU, we calculated the standard deviation of the EPU values over the past six months.
  - *USA EPU volatility* was determined by finding the standard deviation of the USA EPU values for the six months preceding a given month. For instance, to calculate the USA EPU volatility for July 2020, we considered the monthly USA EPU values from January 2020 to June 2020 and computed their standard deviation.
  - *China EPU volatility* was determined by finding the standard deviation of the China EPU values for the six months preceding a given month. As detailed in Table 3.3, it can be observed that the volatility of China's EPU is greater than that of the USA on average in this dataset.

- *USA e-commerce sales* is the quarterly e-commerce sales value in the USA in billions of dollars

- ***Next Review variables***: The next review variables serve as our primary dependent variables in the model. We define and calculate these variables as follows:

  - *Review Rate*: This variable measures the rate at which reviews are generated for a product. It is calculated using the equation:

$$ReviewRate = 30/(Date_{r+1} - Date_r) \tag{2.3.1}$$

    Here, $Date_{r+1}$ represents the date when the current review is submitted into the system, and $Date_r$ represents the date of the most recent review before the current one. The review rate is analogous to the cycle time, indicating the time between two successive units in a system. We compute the review rate as the number of monthly reviews, including 30 in the numerator. A higher value for this variable indicates a faster influx of reviews for a product, while a lower value implies a slower rate.

  - *Relative Review Rate*: This variable captures the difference between the next review rate and the cumulative average of all past review rates for a particular product. It compares the product's current review rate to its historical performance, reflecting how much the current rate deviates from its average past performance.

  - *Score*: The score variable represents the rating assigned by the current customer for the product. As reported in Table 3.3, it can be seen that more than 50% of the reviews have a five-star rating.

  - *Relative Score*: The relative score is computed as the difference between the current customer's rating and the product's average rating from prior reviews.

- ***Description variables***: These are the variables built from the product description text dataset.

  - *Made in China*: This is a binary variable, which indicates if there is a "made in China" mention in the product description.

- *Made in USA*: This is a binary variable, which indicates if there is a "made in USA" mention in the product description.

- **Review variables**
  - ***Recent review variables***: According to a report by Invesp, most customers read 4-6 reviews before making a purchase. Using this idea, we constructed variables to test the impact of the recency bias by building variables based on the five most recent reviews.
    * *Recent rate* is the average review rate of the recent five reviews. It captures the average rate at which reviews are generated for a product based on the most recent customer feedback.
    * *Recent USA mentions* is the number of "Made in USA" mentions the current customer can see in the five most recent reviews. It reflects the frequency of references to products manufactured in the USA, as customers perceive based on recent feedback. From table 3.3, we can see that most customers do not see any made-in-USA mentions in the recent five variables.
    * *Recent China mentions* is the number of "Made in China" mentions the current customer can see in the five most recent reviews. It represents the occurrence of references to products manufactured in China, as observed by customers based on recent feedback.
    * *Recent score* is the average rating customers give in the five most recent reviews for the product. It provides insights into the average customer satisfaction level expressed through ratings in recent feedback.
  - ***Cumulative variables***: We created these variables to investigate the influence of the review history on product review rate. We excluded the five most recent reviews when computing the cumulative variables to avoid potential confounding effects with the recent variables.
    * *Cumulative rate* is the cumulative average review rate of all previous reviews for the product, as perceived by the customer. It captures the overall review generation rate over time.
    * *Cumulative reviews* is the number of reviews the product has received, as seen by the customer. It reflects the total volume of feedback available for the product.

- *Cumulative USA mentions* is the count of the number of mentions of the product being made in the USA from all of its reviews.
- *Cumulative China mentions* is the count of the number of mentions of the product being made in China from all of its reviews.
- *Cumulative score* is the average rating derived from all past reviews for the product.
- *Cumulative made in USA score* is the average rating derived from all reviews made in USA mentioned in their text.
- *Cumulative made in China score* is the average rating derived from all reviews which have been made in China mentions in their text.

2.3.2.2 Control Variables

Monthly time trend: The number of months since the first review in the dataset. It captures the temporal effects and trends in sales over time.

Political Events: These are a collection of multiple binary variables. We gathered multiple political events that could have impacted the trade relations between USA and China, which could have impacted customer response on Amazon.com. We shortlisted the events from the Council of Foreign Relations website, www.cfr.org. The events we considered for the analysis are shortlisted in table 2.2. The significant events we considered are the election times in both countries as they can impact the leader's strategy about foreign trade policies. We used the great recession as it impacted most of the economies. Other events like tariff wars between USA and China have also been used. Each variable tells if a review occurred six months after the event was initiated. Except for the great recession event, which lasted for 19 months, we recorded if the review came within 19 months of the recession.

2.3.2.3 Fixed Effects

Product Category: We have a total of 23 product categories, as mentioned earlier. This variable allows us to account for any unobservable effects of the review rate specific to each product category. Our model does not incorporate time-fixed effects, as the monthly time trend variable already captures the temporal changes.

2.4 Analysis

This section will introduce an econometric model to analyze the relationship between EPU volatility, "Made in" mentions, and review rates. Our analysis will be structured into three main sections to examine these variables' associations thoroughly.

2.4.1 Econometric setup to test the affect of made in mentions

In the first section of our analysis, we will examine the impact of the "made in" mention variables on the changes in review rate (Δreview rate) and changes in review score (Δreview score). To do so, we will employ two econometric equations to test the direct effects of these variables. The equations are as follows:

$$RV_{p,r+1}^{Rate} = \alpha_{D0} + \alpha_{D1}^{RV} * \mathbf{RV}_{p,r-4:r} + \alpha_{D2}^{RV} * \mathbf{RV}_{p,1:r-5} + \alpha_{D3}^{PD} * \mathbf{PD}_r + \alpha_{D4}^{CV} * \mathbf{CV}_r + FE_p \quad (2.4.1)$$

$$RV_{p,r+1}^{Score} = \beta_{D0} + \beta_{D1}^{RV} * \mathbf{RV}_{p,r-4:r} + \beta_{D2}^{RV} * \mathbf{RV}_{p,1:r-5} + \beta_{D3}^{PD} * \mathbf{PD}_r + \beta_{D4}^{CV} * \mathbf{CV}_r + FE_p \quad (2.4.2)$$

where in the first equation, $RV_{p,r+1}^{Rate}$ represents the next-review rate of the customer. The variables $\mathbf{RV}_{p,r-4:r}$ and $\mathbf{RV}_{p,1:r-5}$ are vectors that capture the recent and cumulative "Made in" mentions variables, respectively. These variables reflect the frequency of "made in" mentions in customer reviews over the total past and the most recent five reviews, respectively.

The second equation focuses on the Δreview score as the dependent variable. The variables $\mathbf{RV}_{p,r-4:r}$ and $\mathbf{RV}_{p,1:r-5}$ are again included to account for the recent and cumulative "made in" mentions variables.

The vector $\mathbf{CV}_r$ represents the control variables, such as the recent review rate, recent score, cumulative review rate, cumulative score, and time trend variables. These control variables help account for other factors influencing the review rate and score.

The product category fixed effects denoted as FE_p in the model to capture specific product category effects. These fixed effects control for any inherent differences among product categories that could impact the review rate and review score.By estimating these econometric equations, we aim to assess the direct effects of the "Made in" mentions variables on the review rate and review score changes while accounting for relevant control variables and product category effects.

In our analysis, we incorporated the ratio of these mentions to the total number of reviews when considering the cumulative USA and China mentions variables. This approach was adopted to address the potential bias introduced by products with more reviews having a higher likelihood of receiving more "made in" mentions. We aimed to normalize the impact of the "made in" mentions relative to the overall review activity by using the ratio of mentions to total reviews.

Furthermore, we multiplied the made-in-China mention rate by 100 and the made-in-USA mention rate by 1000 to enhance the interpretability of the results. This scaling was implemented because the frequency of "made in" mentions is generally much lower than the total number of reviews. The scaling allows for a more meaningful comparison and interpretation of the coefficients associated with the "made in" mention variables throughout our analysis. By applying these adjustments consistently in our analysis, we aimed to address potential biases and ensure that the "made in" mention variables are appropriately considered in relation to the total review activity and the overall scale of the dataset.

The second equation tests the impact of volatility on change in review and mention rates. To test this, we use four econometric equations as below:

$$RV_{p,r+1}^{Rate} = \gamma_{D0} + \gamma_{D1}^{USA} * MV_r^{Domestic} + \gamma_{D2}^{MV} * MV_r^{China} + \gamma_{D3}^{CV} * \mathbf{CV}_{p,r}^{Domestic} + FE_p \tag{2.4.3}$$

$$RV_{p,r+1}^{Score} = \delta_{D0} + \delta_{D1}^{MV} * MV_r^{Domestic} + \delta_{D2}^{MV} * \mathbf{MV}_r^{China} + \delta_{D3}^{CV} * \mathbf{CV}_r + FE_p \tag{2.4.4}$$

$$RV^{USA}_{p,r+1} = \zeta_{D0} + \zeta^{MV}_{D1} * MV^{Domestic}_{r} + \zeta^{MV}_{D2} * \mathbf{MV}^{China}_{r} + \zeta^{CV}_{D3} * \mathbf{CV}_{r} + FE_{p} \tag{2.4.5}$$

$$RV^{China}_{p,r+1} = \kappa_{D0} + \kappa^{MV}_{D1} * RV^{USA}_{r} + \kappa^{MV}_{D2} * RV^{China}_{r} + \kappa^{CV}_{D3} * \mathbf{CV}_{r} + FE_{p} \tag{2.4.6}$$

where $RV^{USA}_{p,r+1}$ represents the change in USA mentions, which is the difference between the recent five "made in USA" mentions and the cumulative "Made in USA" mentions. Similarly, $RV^{China}_{p,r+1}$ represents the change in China mentions, calculated as the difference between the recent five "Made in China" mentions and the cumulative "Made in China" mentions.

To account for the potential impact of economic conditions, we introduce the concept of adjusted volatility. This adjusted volatility is obtained by dividing the volatility variables by the USA quarterly e-commerce sales. This adjustment helps normalize the volatility measures relative to the scale of the USA e-commerce market, allowing for a more meaningful interpretation of the results. By incorporating the change in "Made in USA" and "Made in China" mentions and the adjusted volatility measures, we aim to capture the dynamic nature of customer reviews and their association with product origin information.

In our analysis, when utilizing the variables of EPU volatility, we apply a scaling adjustment by dividing the EPU volatility values by the scaled e-commerce sales in the USA. This adjustment accounts for the increasing trend observed in both EPU volatility values and e-commerce sales in the USA. We followed the same strategy throughout our analysis when using the EPU variables.

2.5 Results

2.5.1 Made in mentions drive sales: Results from hypothesis 1

In this section, we present the results of our study, focusing on the impact of ”Made in USA” and ”Made in China” mentions on the domestic review rate of products. Our analysis aimed to determine whether there was a significant relationship between these mentions and consumers' review behavior. We also examined the influence of other factors, including cumulative review variables, recent review variables, and time trend variables, on the review rate and score. The results of our analysis are summarized in Table 2.4, and we discuss the implications of these findings below.

In column (1) of Table 2.4, we examined the impact of ”made in USA” and ”made in China” mentions, as well as other independent variable groups, on the Δreview rate and Δreview score. Our findings reveal that the cumulative ”made in USA” mentions do not significantly affect the review rates or scores. However, we observed a positive and statistically significant impact of the cumulative ”Made in China” mentions on the Δnext review rate and next review score.

The coefficient for cumulative ”Made in China” mentions indicates that for every 1% increase in mentions, the Δnext review rate is expected to increase by 0.26%. This result aligns with our hypothesis 1(b), which proposed that an increased number of ”Made in China” mentions would positively influence the review rate. This result suggests that consumers' knowledge of products made in China and their perception of the associated quality and price can impact their review behavior.

Our findings align with the report by (Mohsin, 2022), which emphasizes the importance of price in Amazon customers' purchasing decisions. While positive perceptions of products made in China may exist, customers' review behavior is influenced by price and perceived value for money. Within the context of Amazon, where the price is a significant factor, customers may be more inclined to leave positive reviews for products made in China due to their affordability and perceived value. This result further supports the notion that price considerations impact customers' review behavior and their evaluation of products.

Contrary to our hypothesis 1(a), our findings reveal that the cumulative "made in USA" mentions do not significantly affect the review rate. This result implies that customers' positive perception of products made in the USA and any associated feelings of patriotism or national pride do not substantially impact their review behavior. These findings suggest that other factors, such as price, quality, or overall customer satisfaction, may play a more prominent role in influencing customers' review behavior rather than the country of origin.

In conclusion, our analysis encompassed various independent variable groups, including cumulative review, recent review, and time trend variables. The results of our study indicate that the cumulative mentions of "Made in China" have a positive and statistically significant impact on the Δnext review rate of products. This finding supports our hypothesis that more "Made in China" mentions would positively influence the review rate. However, we did not find a significant effect for "Made in USA" mentions, contradicting our hypothesis that these mentions would positively impact the review rate. These results suggest that in online platforms like Amazon, factors such as price and perceived value for money may be more influential in customers' review behavior than in the country of origin.

2.5.1.1 EPU does not drive mentions: Results from hypothesis 2

In this section, we present the findings of our study on the impact of Economic Policy Uncertainty (EPU) on consumer behavior and the mentions of products made in the USA and China. We aimed to investigate the relationship between EPU volatility and consumer preferences for products from these countries and explore the potential spillover effects of EPU between the USA and China. As shown in Table 2.6, our results shed light on these dynamics and provide valuable insights. We will now discuss the implications of our findings.

In columns (1) to (4) of Table 2.6, we estimated the impact of EPU volatility on the review rate, review score, mentions of products made in the USA, and mentions of products made in China. Our findings reveal that both USA EPU volatility and China EPU volatility have a negative and statistically significant effect on the review rate and score. Specifically, a 1% increase in USA EPU volatility is associated with a decrease of 0.025% in the Δnext review rate. In comparison, a 1% increase in China's EPU volatility is associated with a 0.1% decrease in the Δnext review rate. These results suggest that higher levels of EPU volatility in both countries are associated with decreased review rates and scores, indicating potential changes in consumer behavior during periods of economic uncertainty.

However, the EPU volatility of both the USA and China does not significantly impact the change in mentions of products made in either country. This finding contradicts our hypotheses 2(a) and 2(b), which suggested that increased EPU volatility in each respective country would negatively influence the mentions of products made in that country. The non-significant relationship between EPU volatility and mentions of products made in the USA or China suggests that the economic uncertainty levels may not significantly influence consumer preferences for products from these countries in those countries alone. It is important to note that our analysis focused on the impact of EPU volatility on consumer behavior and mentions of products. It did not explore the specific factors driving consumer preferences or the underlying mechanisms of their decision-making processes. Further research is necessary to delve deeper into these factors and better understand the dynamics between economic uncertainty and consumer preferences for products made in different countries.

Our findings also shed light on the spillover effects of EPU between the USA and China. Interestingly, our results suggest that EPU volatility in one country does not significantly impact consumer preferences for products made in the other country. This result indicates that factors other than economic uncertainty levels, such as product quality, brand reputation, or cultural associations, may influence consumer perceptions and preferences for products made in the USA or China. Our study provides valuable insights into the impact of EPU volatility on consumer behavior and preferences for products made in the USA and China. The results indicate that higher levels of EPU volatility in both countries are associated with decreased review rates and scores, reflecting potential changes in consumer behavior during uncertain economic times. However, it is important to note that EPU volatility does not significantly affect consumer mentions of products made in the USA or China. These findings highlight the complex nature of consumer preferences and suggest that multiple factors beyond economic uncertainty contribute to consumer decision-making processes. Further research is needed to explore these factors in more detail and better understand the underlying mechanisms of consumer preferences in the context of international trade and economic uncertainty.

2.6 Conclusions

In conclusion, this study sheds light on the influence of product origin information and economic policy uncertainty on customer reviews in the context of online businesses. The findings provide valuable insights into consumer behavior and preferences, particularly regarding "made-in USA/China" mentions on Amazon.com. The results regarding the impact of "made in" mentions reveal that cumulative mentions of products made in China significantly and positively affect the next review rate. This result suggests that increased mentions positively influence customer review behavior, possibly due to consumers' awareness of products made in China and their perceptions of quality and price. On the other hand, cumulative mentions of products made in the USA do not show a significant impact on review rates or scores. This result implies that customers' positive perception of products made in the USA, and any associated feelings of patriotism or national pride, may not substantially influence their review behavior.

Moreover, the study highlights the influence of economic policy uncertainty on consumer behavior. Both USA and China's economic policy uncertainty volatility negatively and significantly affect review rates and scores. Higher levels of economic policy uncertainty in both countries lead to reduced review rates and scores, indicating potential changes in consumer behavior during uncertain economic times. However, volatility of economic policy uncertainty does not significantly affect the mentions of products made in the USA or China. This result suggests that consumer preferences for products from these countries are influenced by factors other than economic uncertainty, such as product quality, brand reputation, or cultural associations.

The findings contribute to a deeper understanding of consumer decision-making processes and the factors that shape customer review behavior. They have important implications for businesses operating on online platforms like Amazon. Businesses can devise targeted marketing strategies and enhance sales by understanding the impact of product origin mentions on customer reviews.

Additionally, recognizing the influence of economic policy uncertainty on consumer behavior can help businesses adapt their strategies during uncertain economic times.

However, it is essential to note that this study has limitations. First, the data cannot identify the sales price of the product. Although the listed price is available, identifying the sales price can significantly improve the results. Also, the description data could not gather all possible product IDs, which would have made the dataset even richer. Furthermore, the study focuses solely on the influence of "made-in USA/China" mentions. Studying other countries' mentions can provide more insights into how the made-in mentions shape consumer behavior based on ethnicity, language, etc. Finally, this research focuses only on the "made in" text; there is a need to incorporate forms of "made in" versions like manufactured in, assembled in, produced in, etc. Future research should consider these factors and incorporate data from websites other than Amazon.com to provide a comprehensive understanding of consumer decision-making in online businesses.

In conclusion, this study contributes to the existing literature on consumer behavior and product origin by examining the impact of "made-in USA/China" mentions on customer reviews. It also explores the influence of economic policy uncertainty on consumer behavior. The findings provide valuable insights for businesses seeking to understand and leverage customer reviews in the online marketplace. By considering the impact of product origin information and economic policy uncertainty, businesses can better tailor their strategies and enhance customer satisfaction, ultimately gaining a competitive advantage in the dynamic online business environment.

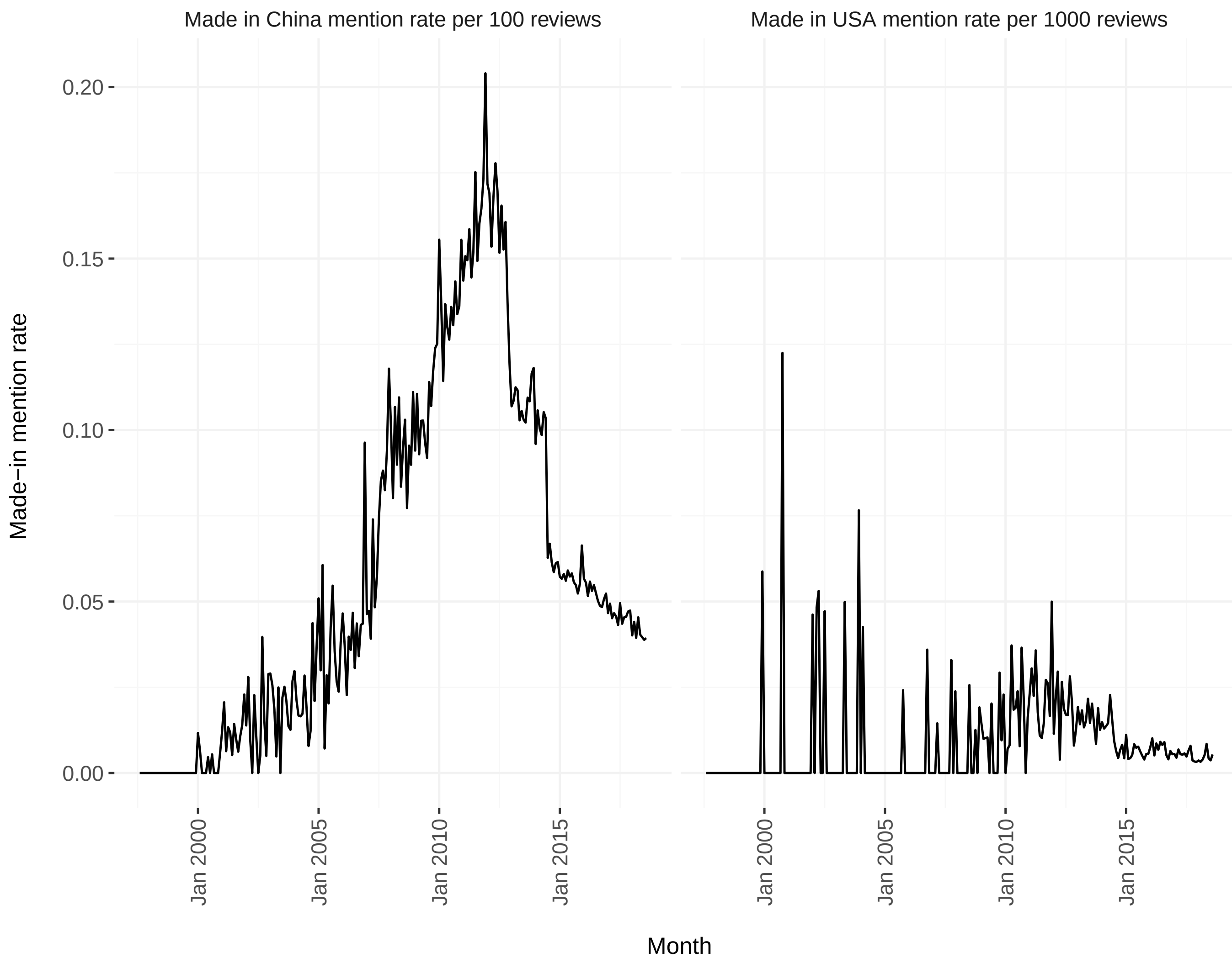
Table 2.1: Data summary

	Total Products	Total Reviews	Mean Rating	Made in USA Mentions	Made in China Mentions
Clothing and Jewelry	1,127,172	24,665,669	4.14	205	17,393
Home and Kitchen	549,241	16,958,390	4.13	221	17,299
Electronics	402,627	16,265,245	3.98	35	6,623
Sports and Outdoors	418,845	10,051,262	4.20	111	9,132
Cell phone and Accessories	257,504	7,706,972	3.84	9	2,001
Tools and Home Improvement	253,982	7,079,173	4.18	87	8,432
Toys and Games	290,887	6,423,963	4.20	21	3,525
Movies and TV	114,306	6,035,245	4.16	31	96
Automotive	357,120	6,010,165	4.24	47	4,736
Pet Supplies	106,820	5,088,673	4.06	123	4,812
Office Products	140,116	4,359,157	4.12	23	1,706
Patio Lawn and Garden	141,691	4,195,488	4.08	100	3,775
Grocery and Gourmet Food	150,748	3,985,562	4.27	33	808
CDs and Vinyl	219,898	3,437,645	4.47	2	10
Arts Crafts and Sewing	114,742	2,129,153	4.29	15	707
Video Games	46,869	2,054,206	3.97	2	288
Industrial and Scientific	60,800	1,325,182	4.26	14	964
Musical Instruments	53,667	1,201,703	4.24	5	1,235
Fashion	54,273	565,674	3.91	3	367
Appliances	15,060	489,232	4.24	3	394
Luxury Beauty	9,554	473,856	4.18	4	49
Prime Pantry	9,286	397,262	4.31	1	21
Beauty	13,657	276,573	4.06	4	105

This table summarizes the data for each product category, sorted by the total number of reviews present in each category. The "Total products" column represents the number of products in each category. The "Total reviews" column displays the total reviews received by the product category. The "Mean rating" column indicates the average rating based on all the review ratings for the product category. The "Made in USA mentions" column shows the total number of reviews that mention "made in USA." The "Made in China mentions" column displays the total number of reviews that mention "made in China" within the product category.

2.7 Exhibits

Figure 2.1: Comparison of USA and China mentions per review



This figure illustrates the monthly comparison of the number of "made in China" and "made in USA" mentions per review. To ensure comparability, the mention rates for "made in China" and "made in USA" are multiplied by 100 and 1000, respectively.

Table 2.2: Major Political events between USA and China during 1998-2018

	Event	Description
2001 January	Presidential Election in USA	G.W. Bush elected as the President of USA
2001 September	September 11 attacks	Terrorist attack on USA. China and USA worked together in the war on terror
2007 December - June 2009	Great Recession	The great recession which lasted for 19 months in USA
2008 September	U.S foreign creditor	China became the largest foreign creditor to USA
2009 January	Presidential Election in USA	Barack Obama elected as the President of USA
2011 November	Pivot to East Asia	USA invested heavily in East Asian countries including China
2012 February	Rising trade tensions	USA and its allies contend China's quota violates international trade norms
2012 November	Presidential Election in China	Xi Jinping elected as the President of People's Republic of China (PRC)
2013 June	Sunnylands Summit	Presidents Obama and Jinping meet in California to ease tensions
2017 January	Presidential Election in USA	Donald Trump elected as the President of USA
2018 March	Tariff war	Trump administration announces tariffs on Chinese imports

This table summarizes major political events that could have influenced trade relations between the USA and China. The first column displays the year and month in which each event occurred. The second column presents the event's name, while the third column briefly describes the event. The timeline of these events is sourced from the Council on Foreign Relations website (www.cfr.org).

Table 2.3: Descriptive statistics of mentions and review variables

			Min	Q1	Q2	Q3	Max	Mean	SD	Description
Macroeconomic Variables	$\mathbf{MV}_r^{Domestic}$	USA EPU volatility	0.00	6.85	10.57	25.87	86.21	15.64	11.79	Volatility of economic uncertainty index of USA for the last 6 months
	$\mathbf{MV}_r^{China}$	China EPU volatility	0.00	17.91	24.50	32.08	103.72	31.57	23.18	Volatility of economic uncertainty index of China for the last 6 months
	$\mathbf{MV}_r^{Sales}$	USA e-commerce sales	1.83	4.32	4.50	4.66	5.05	4.45	0.34	E-commerce sales in USA in the quarter measured in billions
Description	$\mathbf{PD}_r$	Made in China	0.00	0.00	0.00	0.00	1.00	0.00	0.02	Made in China mentions in the product description
		Made in USA	0.00	0.00	0.00	0.00	1.00	0.00	0.06	Made in USA mentions in the product description
Next Review Variables	$\mathbf{NR}_r^{Domestic}$	Rate	0.00	1.43	6.00	30.00	150.00	15.49	21.41	Number of reviews per month, calculated as 30/Interdays
		Relative rate	−149.61	−6.26	0.00	1.18	149.97	0.49	15.76	Difference between the next review rate and the product's cumulative r
		Score	1.00	4.00	5.00	5.00	5.00	4.12	1.33	Rating given by the customer
		Relative score	−4.00	−0.59	0.34	0.75	4.00	−0.07	1.29	Difference between the next review score and the product's cumulative
Review Variables	$\mathbf{RV}_{p,d,r-4:r}^{Domestic}$	Recent rate	0.00	0.96	7.56	20.75	126.00	13.44	16.49	Average review rate of the recent five reviews
		Recent USA mentions	0.00	0.00	0.00	0.00	2.00	0.00	0.01	Number of "made in USA" mentions in the last five reviews as seen by
		Recent China mentions	0.00	0.00	0.00	0.00	5.00	0.00	0.06	Number of "made in China" mentions in the last five reviews as seen b
		Recent score	0.00	3.60	4.20	4.80	5.00	4.12	0.73	Average score of the recent five reviews
	$\mathbf{RV}_{p,1:r-5}^{Domestic}$	Cumulative rate	0.00	0.95	8.67	19.71	150.00	12.90	14.42	Cumulative average review rate of all reviews of the product
		Cumulative reviews	1.00	8.00	42.00	173.00	6453.00	175.05	354.44	Cumulative number of reviews for the product as seen by the customer
		Cumulative USA mentions	0.00	0.00	0.00	0.00	9.00	0.00	0.06	Cumulative number of mentions of "made in USA" as seen by the custo
		Cumulative China mentions	0.00	0.00	0.00	0.00	114.00	0.13	0.89	Cumulative number of mentions of "made in China" as seen by the cus
		Cumulative score	0.00	3.27	4.10	4.48	5.00	3.34	1.74	Cumulative average score of the reviews as seen by the customer
		Cumulative made in USA score	0.00	0.00	0.00	0.00	5.00	0.01	0.19	Cumulative rating of the made in USA mentioned reviews
		Cumulative made in China score	0.00	0.00	0.00	0.00	5.00	0.22	0.90	Cumulative rating of the made in China mentioned reviews
Controls	$\mathbf{CV}_r$	Monthly time trend	0.00	205.00	221.00	233.00	254.00	215.17	29.10	Number of months since the first review in the dataset

This table provides an overview of the dependent and independent variables used in our model. It lists the variables and briefly describes each variable.

Table 2.4: Impact of "made in" mentions on product review rate

			Dependent Variable	
			ΔNext Review Rate	ΔNext Review Score
Review Variables	$\mathbf{RV}_{p,r-4:r}$	Recent rate	1.170^{*}	-0.036^{*}
			(0.081)	(0.009)
		Recent USA mentions	0.108	-0.001
			(0.078)	(0.010)
		Recent China mentions	0.060^{*}	0.001
			(0.013)	(0.002)
		Recent score	0.143^{*}	0.999^{*}
			(0.023)	(0.005)
		Cumulative USA mentions	-0.016	0.003
			(0.036)	(0.002)
		Cumulative China mentions	0.026^{*}	0.004^{*}
			(0.011)	(0.001)
		Made in China score	0.026^{*}	-0.001^{*}
			(0.003)	(0.000)
		Made in USA score	0.018	-0.001^{*}
			(0.014)	(0.000)
Statistics		R-square	0.09	0.27
		Adj.R-square	0.09	0.27
		N (in millions)	118.97	118.97

The column header indicates the dependent variable in the model, which is either the Δnext review rate or the Δnext review score. The model uses the recent mention variables, cumulative review variables, cumulative review variables, time trend variables, and product category fixed effects as independent variables and control variables.
Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 2.5: Effect of Political Events on future reviews

			Dependent variable	
			ΔNext Review Rate	ΔNext Review Score
Political	$\mathbf{PV}_r$	2000 USA Presidential election	-0.103^{*}	-0.135^{*}
			(0.033)	(0.017)
		2001 September 11 attacks	-0.121	-0.133^{*}
			(0.079)	(0.019)
		2007 Great Recession	-0.172^{*}	-0.039^{*}
			(0.058)	(0.013)
		2008 China top foreign creditor	0.155^{*}	0.004
			(0.041)	(0.004)
		2009 USA Presidential election	-0.168^{*}	-0.031^{*}
			(0.031)	(0.006)
		2011 Pivot to East Asia	0.090^{*}	-0.038^{*}
			(0.034)	(0.005)
		2012 Rising trade tensions	-0.171^{*}	-0.060^{*}
			(0.042)	(0.004)
		2012 China Presidential election	0.540^{*}	0.075^{*}
			(0.019)	(0.006)
		2013 Sunnylands summit	0.195^{*}	-0.004
			(0.026)	(0.008)
		2017 Presidential election	-0.276^{*}	-0.028^{*}
			(0.019)	(0.004)
		2018 tariff war	-0.151^{*}	-0.063^{*}
			(0.023)	(0.007)
Statistics		R-square	0.012	0.001
		Adj.R-square	0.012	0.001
		N (in millions)	118.97	118.97

The column header indicates the dependent variable in the model, which is either the Δnext review rate or the Δnext review score. The dummy variables for the political events are used as independent variables in the model. The product category fixed effects are also included as another variable in the model. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 2.6: Effect of EPU Volatility on "Made In" mentions

			Dependent variable			
			ΔNext Review Rate	ΔNext Review Score	ΔMade-in USA Mentions	ΔMade-in China Mentions
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	Adjusted USA volatility	-0.025^{*}	-0.002^{*}	0.000	0.000
			(0.005)	(0.000)	(0.000)	(0.000)
Variables	$\mathbf{MV}_r^{China}$	Adjusted China volatility	-0.101^{*}	-0.004^{*}	0.000	$0.000^{\dagger}$
			(0.012)	(0.001)	(0.000)	(0.000)
Statistics		R-square	0.089	0.266	0.037	0.040
		Adjusted R-square	0.089	0.266	0.037	0.040
		N	118.97	118.97	118.97	118.97

The column header indicates the dependent variable in the model, which is either the Δnext review rate or the Δnext review score. The EPU volatility variables are used as independent variables in the model. The cumulative review variables, time trend variables, and product category fixed effects are included as controls. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 2.7: Model Fit by Product Category

	Recent Review Variables	Cumulative Review Variables	Made in Mention Variables	Macroeconomic Variables	All Variables
Appliances	0.43	0.44	0.07	0.01	0.48
Arts Crafts and Sewing	0.38	0.39	0.02	0.00	0.44
Automotive	0.41	0.43	0.04	0.00	0.47
Beauty	0.44	0.41	0.09	0.00	0.48
CDs and Vinyl	0.51	0.38	0.00	0.01	0.53
Cell phone and Accessories	0.48	0.38	0.02	0.00	0.50
Clothing and Jewelry	0.43	0.38	0.05	0.01	0.47
Electronics	0.49	0.43	0.03	0.01	0.52
Fashion	0.40	0.35	0.02	0.00	0.44
Grocery and Gourmet Food	0.44	0.44	0.02	0.01	0.49
Industrial and Scientific	0.44	0.45	0.06	0.00	0.49
Luxury Beauty	0.40	0.39	0.01	0.01	0.45
Movies and TV	0.57	0.47	0.00	0.04	0.60
Musical Instruments	0.43	0.44	0.04	0.01	0.49
Office Products	0.45	0.44	0.03	0.01	0.50
Patio Lawn and Garden	0.41	0.39	0.04	0.01	0.45
Pet Supplies	0.47	0.47	0.04	0.02	0.52
Prime Pantry	0.31	0.32	0.02	0.02	0.37
Sports and Outdoors	0.42	0.41	0.04	0.01	0.47
Tools and Home Improvement	0.44	0.44	0.05	0.01	0.49
Toys and Games	0.39	0.33	0.03	0.01	0.43
Video Games	0.52	0.40	0.02	0.02	0.54

The table provides the R-squared values for models predicting the next review rate using different independent variable groups. The values represent the goodness of fit of each model when considering specific sets of variables. Only recent review variables are included as independent variables in the first column. In the second column, only cumulative review variables are used. The third column includes only "made in" variables, while the fourth includes only EPU and time trend variables. Lastly, the fifth column includes all variables combined. The next review rate serves as the dependent variable in all models. The R-squared values indicate the proportion of variance in the next review rate explained by the independent variables in each model.

Table 2.8: Moderating effect of EPU on Made in mentions

			Dependent Variable			
			Recent Made-in USA Mentions IRF	Recent Made-in China Mentions IRF	Cumulative Made-in USA Mentions IRF	Cumulative Made-in China Mentions IRF
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	Adjusted USA volatility	-0.328^{*}	-0.327^{*}	-0.331^{*}	-0.301^{*}
			(0.002)	(0.002)	(0.002)	(0.002)
	$\mathbf{MV}_r^{China}$	Adjusted China volatility	-1.140^{*}	-1.139^{*}	-1.160^{*}	-1.063^{*}
			(0.003)	(0.003)	(0.003)	(0.003)
	Controls	Cumulative review variables	Yes	Yes	Yes	Yes
		Category fixed effects	Yes	Yes	Yes	Yes
		Time trends	Yes	Yes	Yes	Yes
	Statistics	R-squaretype	0.013	0.013	0.014	0.014
		Adjusted R-square	0.013	0.013	0.014	0.014
		N	118.968	118.968	116.046	116.046

The table presents the sensitivity analysis between the dependent variable, the next review rate, and the significant independent variables of recent and cumulative mentions of products made in the USA and China. The sensitivity is measured using Impulse Response Function (IRF). Each column represents a model with a different dependent variable. The dependent variable in column (1) is the IRF of recent made-in-USA mentions. In column (2), the dependent variable is the IRF of recent made-in-China mentions, which is calculated as the ratio of the next review rate to the number of made-in-China mentions in the last five reviews. In column (3), the dependent variable is the IRF of cumulative made-in USA mentions, which is calculated as the ratio of the next review rate to the cumulative number of made-in USA mentions. In column (4), the dependent variable is the IRF of cumulative made-in China mentions, which is calculated as the ratio of the next review rate to the cumulative number of made-in China mentions. The independent variables in all models are the adjusted EPU volatility for the USA and China. The sensitivity analysis explores the relationship between these variables and the next review rate, providing insights into the impact of recent and cumulative mentions on consumer behavior.

Chapter 3

Managing Cross-Lingual Inclusiveness in Online Platforms: Evidence from Amazon Reviews and Translation Technology

3.1 Introduction

The rise of e-commerce platforms has transformed the way consumers shop for products and services, with Amazon.com emerging as a dominant player in the United States. With its vast array of products and a large customer base, Amazon.com offers businesses a unique opportunity to reach a wide audience and drive sales. In this context, understanding the factors that influence the customers on Amazon.com is pivotal for it's success. One way Amazon communicates with reviews is by the information Amazon can share to customers through reviews. The language the customer receives a service in plays a pivotal role in shaping consumers' perceptions and purchase decisions (Kelly, 2012). According to Kelly (2012) 72.4% of consumers are more likely to buy a product with information in their own language. Customers prefer to interact in their first language (Holmqvist, 2011). But online reviews are posted in multiple languages. It is not yet studied in a e-commerce setting on how consumers react to reviews from a language other than their own. In this research we focus on how review language can influence the business of Amazon.

According to eDesk, online product reviews, are a way of communicating with the e-commerce customer. Online reviews have gained considerable prominence as a vital source of information for potential buyers. They not only provide insights into product features and functionality but also reflect the opinions and experiences of customers. Consequently, businesses have increasingly recognized the significance of reviews and their impact on sales performance. Understanding how the language used in reviews affects businesses on Amazon.com is therefore a critical area of research. The aim of this study is to explore the relationship between the language employed in reviews on Amazon.com and its influence on businesses' future review rates. Specifically, we will investigate the effects of English reviews count and Spanish reviews count on the review rate for a product on Amazon.com. By analyzing a comprehensive dataset obtained from Ni et al. (2019), which encompasses reviews across 28 distinct product categories, we aim to provide valuable insights into the language dynamics within online reviews and its implications for businesses operating on Amazon.com.

Research in the field of online consumer behavior has established that online reviews significantly impact purchase decisions. Previous studies have investigated various aspects of reviews, including their volume, sentiment, and linguistic characteristics, and their effects on sales performance. Notably, (Chevalier and Mayzlin, 2006) demonstrate the positive influence of reviews on book sales, analyzing Amazon.com and Barnes & Noble. (Dellarocas, 2003) highlights the significance of user feedback mechanisms, such as user forums, in shaping customer behavior. Moreover, Spool (2009) estimates that a single review on Amazon.com can result in 1300 sales, underscoring the positive impact of reviews on customer behavior and businesses. At the same time (Ajjuguttu et al., 2023) have shown that the language of a shared review by Amazon.com can have a significant impact on customers. (Ajjuguttu et al., 2023) have shown that customer respond differently to different review languages under multiple conditions.

While positive effects of English reviews and negative effects of non-English reviews have been studied by (Ajjuguttu et al., 2023) in a global review sharing setting, the influence of domestic peer customer reviews is not studied yet. This represents a significant research gap, considering that Amazon.com caters to a diverse customer base comprising individuals with different linguistic backgrounds. Understanding how domestic English and Spanish reviews affect sales performance is vital for businesses operating on Amazon.com to effectively engage with this diverse customer segment.

Existing research provides some indications regarding the impact of reviews Dellarocas (2003); Xu et al. (2021). Studies in the field of cultural analysis have noted non-English reviews exhibit different expectations than their English counterparts (Schuckert et al., 2015). Antonio et al. (2018) identifies that non-English customers gave lower rating compared to English speaking customers. This suggests that reviews written in languages other than English may convey less positive sentiment, potentially leading to reduced sales. However, further investigation is needed to provide a comprehensive understanding of the specific influence of non-English reviews on businesses' sales performance on Amazon.com.

In light of these research gaps and the significance of language in shaping consumer perceptions and purchase decisions, this study seeks to contribute to the existing body of knowledge by examining the relationship between language in reviews and businesses' sales performance on Amazon.com. By investigating the effects of English reviews count and Spanish reviews count, we aim to provide insights that can guide businesses in optimizing their review strategies and leveraging the language dynamics in online reviews to enhance their sales performance on this influential e-commerce platform. We focus only on Spanish reviews as that is the most widely spoken language in USA after English.

In the era of globalization and digital connectivity, online platforms have become a central hub for product reviews, enabling customers from diverse cultural backgrounds and linguistic preferences to share their experiences and insights. While the influence of reviews on consumer decision-making is widely recognized, the role of review language diversity in motivating customers to write their first reviews remains an intriguing and understudied area. Review language diversity refers to the presence of reviews written in various languages for a given product. In an increasingly interconnected world, products are often accessible to a global customer base, transcending language barriers (Ajjuguttu et al., 2023). When customers encounter a product with multiple language reviews, it suggests widespread usage and acceptance among diverse communities, creating a sense of inclusivity. This research explores whether this inclusivity can motivate customers to contribute to their reviews.

One key aspect of multiple language reviews is the concept of social proof. According to social proof theory, individuals are inclined to look to others for guidance on how to behave, especially in situations with high uncertainty (Cialdini, 2006). When customers see a product with reviews in different languages, it provides a powerful form of social proof, signaling that the product has garnered a broad user base and enjoys global popularity. This collective approval and acceptance of the product may influence first-time customers to feel compelled to participate in the reviewing process, as they perceive their contribution as valuable in shaping the overall reputation and perception of the product. Additionally, review language diversity can foster a sense of community and belonging among customers. When individuals encounter reviews in their native or familiar language, they feel a greater connection to the product and the reviewing community (Holmqvist, 2011). This linguistic affinity creates a shared identity and a desire to engage in the discourse, as customers are more likely to trust and resonate with reviews written in their preferred language. Consequently, customers may be motivated to write their first reviews to actively participate in this linguistic community, express their opinions, and contribute to the overall diversity and richness of the reviewing ecosystem. This research aims to shed light on the impact of multiple language reviews on customers' motivation to engage in the review process for the first time.

3.2 Literature and hypothesis development

3.2.1 Reviews

Customer reviews have become integral in shaping consumer perceptions and influencing purchase decisions, providing valuable information for individuals seeking insights about products or services (Dellarocas, 2003). While previous research has extensively examined the impact of reviews on product popularity, sales, and brand reputation (Dellarocas et al., 2010; Amblee and Bui, 2008; Spool, 2009), the influence of review language diversity on customers' reviewing behavior remains a relatively understudied area. Review language diversity has the potential to foster inclusivity and diversity within the reviewing community, as customers encounter reviews written in various languages. This linguistic diversity can be perceived by English-speaking customers as a reflection of a broader user base, indicating the product's popularity and global acceptance (Papanastasiou and Savva, 2017). As customers often delay their purchase decisions until they see reviews from peers, more reviews for a product can positively influence customer behavior (Papanastasiou and Savva, 2017). Therefore, we hypothesize that Amazon, as a leading online marketplace, would actively post Spanish reviews to customers in the USA, aiming to enhance the review rate among English-speaking customers.

We posit that the count of Spanish reviews can positively affect English-speaking customers' review rates. When English-speaking customers encounter a higher count of Spanish reviews, it can create a sense of inclusivity and diversity within the reviewing ecosystem. English-speaking customers may perceive this linguistic diversity as indicating the product's popularity and global acceptance, influencing their motivation to participate actively in the review process and provide feedback. The increased presence of Spanish reviews can serve as social proof, bolstering the credibility and trustworthiness of the product among English-speaking customers and ultimately leading to a higher review rate.

On the other hand, Spanish-speaking customers often find comfort and validation in reviews written in their native language or a language they understand well. The presence of Spanish reviews signifies that the product has gained attention and usage within their linguistic community, enhancing their trust and relatability (Holmqvist, 2011). The sense of connection and belonging fostered by encountering reviews in their preferred language can motivate Spanish-speaking customers to share their experiences and contribute to the review count actively. Consequently, a higher count of Spanish reviews positively influences the review rate among Spanish-speaking customers. Specifically, we hypothesize that

Hypothesis 1(a): The count of Spanish reviews affects the English customer's next review rate positively.

Hypothesis 1(b): The count of Spanish reviews affects the Spanish customer's next review rate positively.

3.2.2 Review Language

Drawing from the literature on language impact on customer behavior, we propose that the language in which customers encounter reviews plays a crucial role in shaping their subsequent reviewing behavior (Holmqvist, 2011; Holmqvist and Vaerenbergh, 2013). Language holds significant emotional and psychological meaning for individuals, influencing perceptions of brand authenticity (Puntoni et al., 2009).

Research by Vaerenbergh and Holmqvist (2014) highlights that customers are less likely to engage in positive word-of-mouth communication when they receive service or encounter information in a second language, which can hinder effective communication and diminish emotional connections. Additionally, Vaerenbergh and Holmqvist (2013) finds that customers who receive service in their preferred language demonstrate a greater willingness to tip more, indicating a stronger positive response when encountering their native language during service encounters.

Building on this literature, we propose Hypothesis 2(a), which suggests that the count of Spanish reviews negatively affects the review rate of English-speaking customers. When English-speaking customers encounter many Spanish reviews, a lack of familiarity and emotional connection to the Spanish language may decrease their motivation to contribute their reviews. This disconnect or disengagement may reduce their review rate. Furthermore, Spanish reviews may create a perception of exclusion or limited relevance for English-speaking customers, as reviews in a language they need help understanding may reduce their perceived importance within the reviewing community.

Conversely, the count of Spanish reviews positively affects the review rate of Spanish-speaking customers. Spanish-speaking customers react positively to a higher number of Spanish reviews. Reviews in their preferred language enhance their familiarity and emotional connection, increasing their motivation to contribute their own reviews. Consequently, the review rate of Spanish-speaking customers positively influences the count of Spanish reviews.

Based on the literature regarding language's impact on customer behavior, we hypothesize that the count of Spanish reviews negatively affects the review rate of English-speaking customers. In contrast, it positively affects the review rate of Spanish-speaking customers. Specifically, we hypothesize that:

Hypothesis 2(a): The count of Spanish reviews affects the English-speaking customer's next review rate negatively.

Hypothesis 2(b): The count of Spanish reviews affects the non-English-speaking customer's next review rate positively.

3.2.3 First Reviewers

The study by Dellarocas and Narayan (2006) observed that individuals post reviews about a movie if the movie already has more reviews. This self-involvement and social benefits perspective suggests that reviewing is driven by a desire to contribute to an existing online community and gain social recognition. As individuals observe more reviews for a product, they may perceive a thriving and active reviewing community, which can stimulate their motivation to participate and contribute to their reviews.

The literature on social influence further supports the relationship between the number of reviews seen and the propensity to post reviews. According to social proof theory, individuals look to others for behavior guidance, especially in highly uncertain situations (Cialdini, 2006). When individuals observe a more significant number of reviews for a product, it serves as social proof, indicating that the product has been extensively reviewed and experienced by others. This social proof can influence first reviewers by providing reassurance and a sense of validation, increasing their likelihood of posting their reviews (Cialdini, 2006).

In addition to social proof, a higher number of reviews can signal product popularity and demand. Previous research has found that the quantity of reviews positively influences consumers' perceptions (Chevalier and Mayzlin, 2006). When first reviewers observe a substantial number of reviews for a product, they may interpret it as an indication of its popularity and value, further motivating them to contribute their reviews. Moreover, the mere exposure effect suggests that repeated exposure to a stimulus increases individuals' liking and positive evaluation of that stimulus. In the context of online reviews, when first-reviewers encounter a product with a higher number of reviews, they are exposed to more information and opinions about the product. This exposure to diverse perspectives and experiences can enhance their understanding and evaluation of the product, prompting them to share their thoughts and contribute to the growing body of reviews (**?**Gregan-Paxton and Moreau, 2003).

This hypothesis builds upon the findings of Dellarocas and Narayan (2006) and aligns with theories of social proof (Cialdini, 2006), product popularity, and the mere exposure effect. As first reviewers observe more reviews for a product, their motivation to participate in the reviewing process increases, driven by the desire to contribute to the existing reviewing community, gain social recognition, perceive product popularity, and benefit from exposure to diverse opinions. Specifically, we hypothesize that

Hypothesis 3(a): The count of English reviews affects the first-review customer's review rate positively.

Hypothesis 3(b): The count of Spanish reviews affects the first-review customer's review rate positively.

3.3 Data and Methodology

We collected our data from four sources to investigate our research question. The first dataset is a secondary dataset obtained from Ni et al. (2019), who scraped and updated the dataset in 2018. This dataset includes all product reviews from Amazon.com in 2018, providing information on various review attributes such as rating, reviewer name, and review text. We specifically extracted data on review rating, review date, Amazon Standard Identification Number (ASIN), and review text from this dataset.

The second dataset is secondary data from previous study by Baker et al. (2016). The authors developed an Economic Policy Uncertainty (EPU) index for the USA. The EPU index was computed using three components: newspaper coverage related to policy-related economic uncertainty, reports by the Congressional Budget Office, and the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. The newspaper coverage was used to construct an index for the volume of news articles discussing economic policy uncertainty. The Congressional Budget Office provided lists of temporary federal tax code provisions, which were used to compile annual dollar-weighted numbers for each tax code set to expire in the next ten years, giving a measure of the level of uncertainty regarding the future of the federal tax code. The Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters was used to create indices of uncertainty about policy-related macroeconomic variables using the dispersion between individual forecasters' predictions of future levels of the Consumer Price Index, Federal Expenditures, and State and Local Expenditures.

The third dataset is the quarterly e-commerce report developed by U.S. Census Bureau. This dataset comprises data on e-commerce activity in the United States. We use this dataset to obtain information on quarterly e-commerce sales in the United States.

The fourth dataset is from trends.google.com. We used the interests in the topic "Make America Great Again" on trends.google.com over time. Trends.google provides this data from 2004 to present at a monthly level.

3.3.1 Data Cleaning and Summary

Our study encompassed a comprehensive dataset consisting of 28 product categories. Table 3.2 provides a detailed list of the 28 product categories included in our dataset. However, we excluded books, kindle store, digital music, magazine data from the analysis from our analysis due to their content-dependent nature rather than the product themselves. The dataset we utilized comprises 178,828,181 reviews, ensuring a robust and representative sample for our analysis. Among these reviews, the largest dataset pertained to clothing and jewelry, encompassing approximately 19 million reviews, followed closely by the electronic products dataset.

We followed three steps in our process of data cleaning.

- We excluded reviews that did not include a rating, as the rating aspect was integral to our analysis. This step removed 44,392 reviews.

- To mitigate the potential impact of reviews for relatively unpopular products, we established a minimum threshold of 10 reviews for inclusion in our analysis. We removed products which have less than 10 reviews in total. This step aimed to ensure the reliability and validity of our findings by focusing on products that had received a sufficient number of reviews. This step removed more than 5 million reviews.

- We also did not consider the reviews that came before a product has reached 10 reviews, essentially we removed the first 10 reviews of a product out of our analysis to eliminate the early stage introduction effects for a product. And consumers are known to be reading at least 10 reviews on average before trusting a business (Zhou, 2023). This step removed around 3 million reviews.

Our final sample size for the analysis is approximately 170,815,387 million reviews.

According to an Oberlo report on Amazon.com in 2022 (Mohsin, 2022), electronics, apparel, and home and kitchen goods emerged as the most popular shopping categories on the platform. Interestingly, our dataset, as reflected in Table 3.2, predominantly comprises data from these very categories. Consequently, despite the fact that the data was scraped in 2018, our sample can be considered a representation of the current trends observed in 2022. This alignment between our dataset and the popular shopping categories on Amazon.com in 2022 enhances the relevance and applicability of our findings. Furthermore, it is important to highlight that our data encompasses a time span of 254 months, as indicated by the monthly time trend variable presented in Table 3.3.

3.3.2 Variable Construction

3.3.2.1 Dependent and Independent Variables

To study the research questions we have built variables under multiple categories.

- ***Macroeconomic Variables***: The macroeconomic variables related to Economic Policy Uncertainty (EPU) were obtained from the study conducted by Baker et al. (2016). To measure the volatility of EPU, we calculated the standard deviation of the EPU values over the past six months.

- *USA EPU volatility* was determined by finding the standard deviation of the USA EPU values for the six months preceding a given month. For instance, to calculate the USA EPU volatility for July 2020, we considered the monthly USA EPU values from January 2020 to June 2020 and computed their standard deviation.
- *USA e-commerce sales* is the quarterly e-commerce sales value in USA in billions of dollars
- *MAGA* is the interest on the topic "Make America Great Again" in USA. The variable spans over a time period of 2004 to 2023. The peak interest in this time period is considered as value 100 and all other values calculated based on the peak time period ranging from 0-100.
- *Monthly non-English count* is the number of non-English reviews in a month in our dataset. This variable accounts for the trend of non-English review on Amazon.com.

• ***Next Review variables***: The next review variables serve as our primary dependent variables in the model. We define and calculate these variables as follows:

- *Review Rate*: This variable measures the rate at which reviews are generated for a product. It is calculated using the equation:

$$ReviewRate = 30/(Date_{r+1} - Date_r) \tag{3.3.1}$$

  Here, $Date_{r+1}$ represents the date when the current review is submitted into the system, and $Date_r$ represents the date of the most recent review prior to the current one. The review rate is analogous to the cycle time, indicating the time between two successive units in a system. We compute the review rate as the number of reviews per month, hence the inclusion of 30 in the numerator. A higher value for this variable indicates a faster influx of reviews for a product, while a lower value implies a slower rate.

- *Relative Review Rate*: This variable captures the difference between the next review rate and the cumulative average of all past review rates for a particular product. It compares the product's current review rate to its own historical performance, reflecting how much the product's current rate deviates from its average past performance.

- *Score*: The score variable represents the rating assigned by the current customer for the product. As reported in Table 3.3, it can be seen that more than 50% of the reviews have a five star rating.
- *Relative Score*: The relative score is computed as the difference between the rating assigned by the current customer and the average rating of the product from prior reviews.
- *Next English Rate*: This variable captures the next review rate of the English review.
- *Next English Rate*: This variable captures the difference between the next review rate of the English review compared to the cumulative average of all past review rates for a particular product. It compares the product's current English review rate to its own historical performance, reflecting how much the product's current rate deviates from its average past performance.
- *Next Spanish Rate*: This variable captures the difference between the next review rate of the Spanish review compared to the cumulative average of all past review rates for a particular product. It compares the product's current Spanish review rate to its own historical performance, reflecting how much the product's current rate deviates from its average past performance.

In our study, we encountered reviews in languages other than English in our dataset. To account for the potential impact of review language on customer behavior, we utilized the fastText package in R to detect the language of each review. However, upon reviewing a random sample of reviews, we found that the language detection accuracy was unsatisfactory for a significant portion of the reviews, with an approximate error rate of 10%. To address this issue, we implemented a probabilistic categorization approach based on the language probabilities provided by the fastText package. Here are the steps we followed for language categorization:

- We estimated the language tag for each review using the fastText package.
- We collected the language probability associated with each review.
- If the language probability exceeded 0.75, we assigned the review to the corresponding language category determined by the package.
- If the language probability was below 0.75, we categorized the review as part of an unconfident group.

As a result, we classified all the reviews into four groups:

1. English confident: This category comprises reviews confidently detected as being in English with a probability greater than 0.75.

2. Spanish confident: This category includes reviews that were confidently identified as being in a Spanish language with a probability greater than 0.75.

3. Other confident: This category includes reviews that were confidently identified as being in a non-English and non-Spanish language with a probability greater than 0.75.

4. Unconfident: This category consists of reviews for which the language probability fell below 0.75, indicating a lack of confidence in the language classification.

By employing this categorization approach, we aimed to address the errors associated with natural language processing packages. Notably, a significant portion of the unconfident group reviews may contain similar words between two languages. For example, "excellent" in English is "excelente" in Spanish. When the textual content of a review is closely related to another language, the language detection packages may struggle to determine the review's language confidently. Moreover, when the shared words are easily understandable across languages, the language effect on reviews may be limited. The review variables are built based on only the English confident and Spanish confident groups.

- **Review variables**

  - ***Recent review variables***: According to a report by Invesp, most customers read 4-6 reviews before making a purchase. Using this idea, we constructed variables to test the impact of the recency bias by building variables based on the five most recent reviews.

    * *Recent rate* is the average review rate of the recent five reviews. It captures the average rate at which reviews are generated for a product based on the most recent customer feedback.

    * *Recent score* is the average rating given by customers in the five most recent reviews for the product. It provides insights into the average customer satisfaction level expressed through ratings in recent feedback.

* *Recent English count* is the number of English reviews that the customer can see in the five most recent reviews.
* *Recent English rate* is the average review rate of English reviews that the customer can see in the five most recent reviews.
* *Recent English score* is the average rating of English reviews that the customer can see in the five most recent reviews.
* *Recent Spanish count* is the number of Spanish reviews that the customer can see in the five most recent reviews.
* *Recent Spanish rate* is the average review rate of Spanish reviews that the customer can see in the five most recent reviews.
* *Recent Spanish score* is the average rating of Spanish reviews that the customer can see in the five most recent reviews.

– ***Cumulative variables***: We created these variables to investigate the influence of the review history on product review rate. To avoid any potential confounding effects with the recent variables, we excluded the five most recent reviews when computing the cumulative variables.

* *Cumulative count* is the total number of reviews a product has received in the past.
* *Cumulative rate* is the average review rate of all the reviews in the past.
* *Cumulative score* is the average rating given by customers for the product.
* *Cumulative English count* is the number of English reviews that the customer can see for the product.
* *Cumulative English rate* is the average review rate of English reviews that the customer can see for the product.
* *Cumulative English score* is the average rating of English reviews that the customer can see in the for the product.
* *Cumulative Spanish count* is the number of Spanish reviews that the customer can see for the product.
* *Cumulative Spanish rate* is the average review rate of Spanish reviews that the customer can see for the product.

- * *Cumulative Spanish score* is the average rating of Spanish reviews that the customer can see for the product.

- **Reviewer variables**
  - *Reviewer count* is the count of the number of reviews a customer a review posted.

3.3.2.2 Control Variables

Monthly time trend is the number of months since the first review in the dataset. It captures the temporal effects and trends in sales over time.

3.3.2.3 Fixed Effects

Product Category We have a total of 25 product categories, as mentioned earlier. This allows us to account for any unobservable effects of review rate specific to each product category. We do not incorporate time fixed effects in our model, as the monthly time trend variable already captures the temporal changes.

3.4 Analysis

In this section, we will present an econometric model to examine the impact non-English reviews on review rates. Our analysis will be divided into three main sections to provide a comprehensive understanding of the relationships between these variables.

3.4.1 Econometric setup to test the affect of Spanish reviews

In the first section of the analysis, we focus on the Δreview rate, Δreview score as the dependent variable. We test the direct affect of the language variables on the Δnext review rate and Δreview score of the variables by using two econometric equations as follows:

$$\mathbf{RV}^{Rate}_{p,r+1} = \alpha_0 + \alpha_1 * \mathbf{RV}_{p,r-4:r} + \alpha_2 * \mathbf{RV}_{p,1:r-5} + \alpha_3 * \mathbf{CV}_{p,r} + FE_p \tag{3.4.1}$$

where $RV^{Rate}_{p,r+1}$ is vector of multiple dependent variables. the next-review rate, next review score, Δnext review rate, Δnext review score, $\mathbf{RV}_{p,r-4:r}$ is the vector of recent review variables,

$\mathbf{RV}_{p,1:r-5}$ is the vector of cumulative review variables. $\mathbf{CV}_{p,r}$ is the vector of all control variables like recent review rate, recent score, and time trend variables. FE_p is the product category fixed effects used in the model.

3.4.2 Econometric setup to study the impact on different language customers

In the second part of the analysis we split the data into two parts as English and Spanish reviewing customers. In the first group we have only the customer who reviewed in English. We measured the review rate of this subset of customers. This value tells us how fast English reviewing customers are coming into the system. This allows us to measure the rate at which English reviews come into the system. We repeat the same exercise for Spanish reviewing customers too.

The econometric equation for the English reviewing customers is:

$$\mathbf{RV}_{p,r+1}^{EnglishRate} = \beta_0 + \beta_1 * \mathbf{RV}_{p,r-4:r} + \beta_2 * \mathbf{RV}_{p,1:r-5} + \beta_3 * \mathbf{CV}_{p,r} + FE_p \tag{3.4.2}$$

The econometric equation for the Spanish reviewing customers is:

$$\mathbf{RV}_{p,r+1}^{SpanishRate} = \beta_0 + \beta_1 * \mathbf{RV}_{p,r-4:r} + \beta_2 * \mathbf{RV}_{p,1:r-5} + \beta_3 * \mathbf{CV}_{p,r} + FE_p \tag{3.4.3}$$

For English customers, we use the ΔNext English review rate as a new dependent variable as shown in equation 4.2. And for Spanish customers, we use the ΔNext Spanish review rate as a new dependent variable which tells how fast the next Spanish review is coming into the system as shown in equation 4.3.

3.4.3 Econometric setup to test the affect of first reviewers

In the third section of the analysis, we focus on the Δreview score variable as the dependent variable. We test the affect of the review variables on the Δnext review score of the variables on only the first-review customers by using the econometric equation as follows:

$$\mathbf{RV}_{p,r+1}^{Rate} = \alpha_0 + \alpha_1 * \mathbf{RV}_{p,r-4:r} + \alpha_2 * \mathbf{RV}_{p,1:r-5} + \alpha_3 * \mathbf{CV}_{p,r} + FE_p \tag{3.4.4}$$

where $RV_{p,r+1}^{Rate}$ is vector of dependent variable the next review score, Δnext review rate, Δnext review score, $\mathbf{RV}_{p,r-4:r}$ is the vector of recent review variables, $\mathbf{RV}_{p,1:r-5}$ is the vector of cumulative review variables. $\mathbf{CV}_{p,r}$ is the vector of all control variables like recent review rate, recent score, and time trend variables. FE_p is the product category fixed effects used in the model.

3.5 Results

3.5.1 Effect of Spanish reviews

In this section, we present the estimation results from our analysis, examining the impact of Spanish review variables on Amazon.com customers. Table 3.4 displays the estimation results using equation (4.1), providing insights into the relationship between the Spanish review variables and the customers' reviewing behavior.

Several noteworthy patterns emerge from the results presented in Table 3.4. Firstly, we find that the count of Spanish reviews significantly positively affects the customers' next review score. This result implies that increased Spanish reviews lead to higher future review scores, indicating that Spanish reviews influence customers' perceptions of the product. Furthermore, the count of Spanish reviews also positively impacts the next review rate of Spanish-speaking customers. This result suggests that Spanish-speaking customers are more inclined to contribute their reviews when they observe more Spanish reviews for a product. Spanish reviews serve as social proof and contribute to community engagement, motivating Spanish-speaking customers to participate in the reviewing process.

On the other hand, the count of English reviews demonstrates a positive effect on the next review rate for both English-speaking and Spanish-speaking customers. As the number of English reviews increases, customers are likely to post them. English reviews serve as a social cue, indicating the popularity and engagement surrounding the product, which influences customers to participate and share their opinions actively.

Interestingly, the cumulative English review rate positively impacts the future review rate.

As a measure of the product's past performance, a higher English review rate increases the likelihood of future reviews. This finding suggests that English review rates have a lasting impact on customers' reviewing behavior, as customers are motivated by the positive experiences and perceptions of English-speaking reviewers. However, it is noteworthy that the English score, representing the average rating of English reviews, has a counterintuitive effect on the future review rate. While it positively impacts the next review rate, indicating that higher average ratings encourage customers to post their reviews, it negatively influences future review scores. This result suggests that higher expectations from positive English reviews may lead to a relative decrease in future review scores. This finding emphasizes the importance of managing customer expectations and understanding the complex interplay between review ratings and customers' perceptions.

Turning to the impact of cumulative Spanish count, we find no significant effect on the review rate. This result implies that Spanish reviews do not influence the number of reviewers. However, the next review score and the change in the next review score (ΔNext review score) increase, indicating that Spanish reviews shape customers' future ratings and perceptions of the product. This finding suggests that while Spanish reviews may not directly impact sales or the overall review rate, they influence customers' perceptions and evaluations of the product.

Our results demonstrate that English reviews positively impact the review rate, while Spanish reviews primarily influence the review score and customers' perceptions. These findings highlight the nuanced effects of different language reviews on customers' reviewing behavior and emphasize the importance of considering the diverse linguistic backgrounds of customers in understanding their engagement with and perceptions of products.

3.5.2 Effect of review language on customers from different backgrounds

In this section, we present the estimation results from our analysis, examining the impact of Spanish review variables on Amazon.com customers. Table 3.5 and 3.7 displays the estimation results using equations (4.2) and (4.3), providing insights into the relationship between the customer's review language and the cumulative review variables by language.

In this section, we present the results of our analysis, focusing on the impact of Spanish and Spanish reviews on the reviewing behavior of English-speaking and Spanish-speaking customers. We examine the relationship between various review variables and the customers' next review rate, next review score, and Δnext review score.

Literature suggested two confounding effects on reviews and language. The service customer receives expected to be in the customer's first language to make the customers more satisfied (Holmqvist, 2011; Holmqvist and Vaerenbergh, 2013; Holmqvist et al., 2019; Puntoni et al., 2009; Vaerenbergh and Holmqvist, 2013). In this research, we treat customers receiving reviews in a different language than they prefer as a different service. Reviews enable customers with more information. Providing more information to customers helps them make more informed decisions (Gallino and Moreno, 2018). This information effect is more evident in our research. When designing a new service for the customers, focusing more on providing more information to the consumers is more important than the service's language.

Table 3.5 provides the estimation results for English customers, while Table 3.7 presents the results for Spanish customers. The results indicate that the count of Spanish reviews does not significantly influence English customers' next review score or Δnext review score. However, the count of Spanish reviews positively impacts the next review rate of English customers. This result suggests that when English customers observe more Spanish reviews for a product, they are likelier to contribute their reviews.

The cumulative English count has a minimal impact on the next review rate, but it has a negative effect on Δnext review score. This result suggests that as the number of English reviews increases, the relative change in the current rating compared to the total rating decreases for English customers. Additionally, the cumulative English score has a negative impact on both the next review score and Δnext review score. This result implies that higher prior ratings from English reviews lead to lower future scores for English customers.

On the other hand, for Spanish customers, the results reveal that the English review score has a negative impact on both the next review score and Δnext review score. This result suggests that Spanish customers have different expectations than English customers, and higher English review scores may not align with their preferences.

Regarding the impact of Spanish reviews on Spanish customers, most variables do not have significant effects. This result could be due to the dataset's relatively low sample size of Spanish reviews. However, the English review score has a negative impact on both the next review score and Δnext review score for Spanish customers. This finding indicates that Spanish customers have distinct expectations from English-speaking customers, and higher English review scores may not align with their preferences.

The literature suggests that the language of service has a significant impact on customer satisfaction (Holmqvist, 2011; Holmqvist and Vaerenbergh, 2013; Holmqvist et al., 2019; Puntoni et al., 2009; Vaerenbergh and Holmqvist, 2013). However, our research treats customers receiving reviews in a different language than their preference as a different service. Our results indicate that the service's language may not substantially impact the next review rate, next review score, or Δnext review score. Instead, providing more information to customers through reviews plays a crucial role in enabling them to make more informed decisions (Gallino and Moreno, 2018).

In summary, the count of Spanish reviews positively influences the next review rate of English customers. However, the count of Spanish reviews does not significantly impact the next review score or Δnext review score. Furthermore, the cumulative English count and score have mixed effects on English customers' reviewing behavior. The English review score negatively affects Spanish customers' next and Δnext review scores. However, most of the variables related to Spanish reviews do not show significant effects, potentially due to the low sample size of Spanish reviews. These findings emphasize the need to consider customer expectations and the informational value of reviews in understanding the impact of language diversity on reviewing behavior.

3.5.2.1 First Reviewer

In this section, we present the results of our analysis, focusing on the impact of English and Spanish reviews on the review rate of first-review customers. We examine the relationship between the count of English reviews, the count of Spanish reviews, and the review rate of customers posting their first reviews. The Chevalier and Mayzlin (2006) study found that individuals are more likely to post reviews about a movie if the movie already has more reviews. This finding aligns with social proof theory, which suggests that individuals look to others for guidance on how to behave, particularly in situations characterized by high uncertainty (Cialdini, 2006).

Our results indicate that the count of English reviews positively affects the review rate of first-review customers. Specifically, as the count of English reviews increases by 1%, the Δnext review score, representing the relative change in the current rating compared to the total rating, increases by 0.04 units. This positive effect suggests that first-review customers perceive a more positive experience or have a more favorable perception of the product when they observe more English reviews. The positive impact of English reviews on the review rate of first-review customers aligns with the findings of Chevalier and Mayzlin (2006) and the social proof theory. More English reviews signal popularity, engagement, and a thriving reviewing community. This social proof effect influences first-review customers, encouraging them to participate and contribute to their reviews, thereby increasing the review rate.

In summary, our results support Hypothesis 3(a), indicating that the count of English reviews positively impacts the review rate of first-review customers. This finding aligns with the study by (Chevalier and Mayzlin, 2006). It provides empirical evidence for the influence of social proof and the presence of reviews on the reviewing behavior of first-review customers.

3.6 Conclusions

This study aimed to explore the influence of language dynamics within online reviews on Amazon.com and its implications for businesses' sales performance. Specifically, we investigated the effects of English and Spanish reviews on businesses' future review rates and scores. Our analysis was based on a comprehensive dataset of 128 million reviews across 25 product categories obtained from various sources. In this research, we examined the impact of review language diversity on the reviewing behavior of customers. Specifically, we investigated how the count of English and Spanish reviews influences the review rate of different customer segments, including English-speaking customers, Spanish-speaking customers, and first-review customers. Through our analysis and the presentation of results, we gained valuable insights that can contribute to understanding review diversity, inform managerial decisions, and identify avenues for future research.

Our findings reveal several vital outcomes. Firstly, we found that the count of Spanish reviews positively influences the scores of Spanish-speaking customers. This result suggests that Spanish-speaking customers are more inclined to contribute their reviews when they observe more in their preferred language. This finding emphasizes the importance of catering to the language preferences of diverse customer segments to encourage active participation in the reviewing process. Additionally, we observed that the count of English reviews positively impacts the review score of first-review customers. This result highlights the social proof effect and the role of English reviews in signaling product popularity, engendering a sense of community engagement, and motivating customers to contribute their reviews. Managers can leverage this insight by promoting English reviews and encouraging customer engagement through strategic communication and marketing efforts.

Furthermore, our results demonstrated that the language of encounter, particularly receiving reviews in a different language than preferred, may not substantially impact future reviews. Instead, providing more information to customers through reviews played a crucial role in enabling them to make more informed decisions. This result suggests that managers should prioritize the quality and comprehensiveness of reviews, regardless of the language, to assist customers in their decision-making processes.

The implications of our research extend beyond the academic sphere and provide practical guidance for businesses. By understanding the impact of review language diversity on customer behavior, managers can develop targeted strategies to improve review diversity and engagement. For example, managers can encourage customers from different linguistic backgrounds to share their experiences by providing incentives or implementing user-friendly platforms that support multiple languages. These strategies foster a diverse and inclusive reviewing community, enhancing the overall quality and credibility of the reviews.

However, it is essential to acknowledge the limitations of our study. First, our research focused on a specific online platform and a particular set of products, which may limit the generalizability of our findings to other contexts. Future research should explore different platforms to obtain a more comprehensive understanding of the impact of reviewing language diversity. Additionally, our analysis primarily focused on the count of reviews and did not delve into the content and sentiment of the reviews. Future studies could examine the content-related factors of reviews, such as the use of specific keywords or language styles, to further uncover the influence of review language diversity on customer behavior.

In conclusion, our research sheds light on the significance of review language diversity in shaping customer behavior. The findings highlight the positive impact of Spanish reviews on Spanish-speaking customers and the influence of English reviews on English-speaking, Spanish-speaking, and first-review customers. These insights provide valuable guidance for managers in their decision-making processes, emphasizing the importance of encouraging diverse review contributions and ensuring comprehensive information to customers. By understanding the complexities of review language diversity, businesses can enhance customer satisfaction, foster inclusive communities, and make informed strategic decisions to improve their products and services.

Table 3.1: Language Statistics by Year

	Before 2014		2014-2016		2016-2018	
	Reviews	Percent	Reviews	Percent	Reviews	Percent
English Confident	27097652	0.97	82681032	0.85	39484387	0.85
Spanish Confident	61767	0.00	36388	0.00	31987	0.00
Other Confident	14021	0.00	125497	0.00	61820	0.00
Unconfident	629104	0.02	13907275	0.14	6678795	0.14

This table summarizes the data by each language detected in our data. The columns represent the year groups before 2014, between 2014-2016 and 2016-2018. Percent shows the percentage of the reviews in that language. Reviews columns shows that total number of reviews in the language in the corresponding year category.

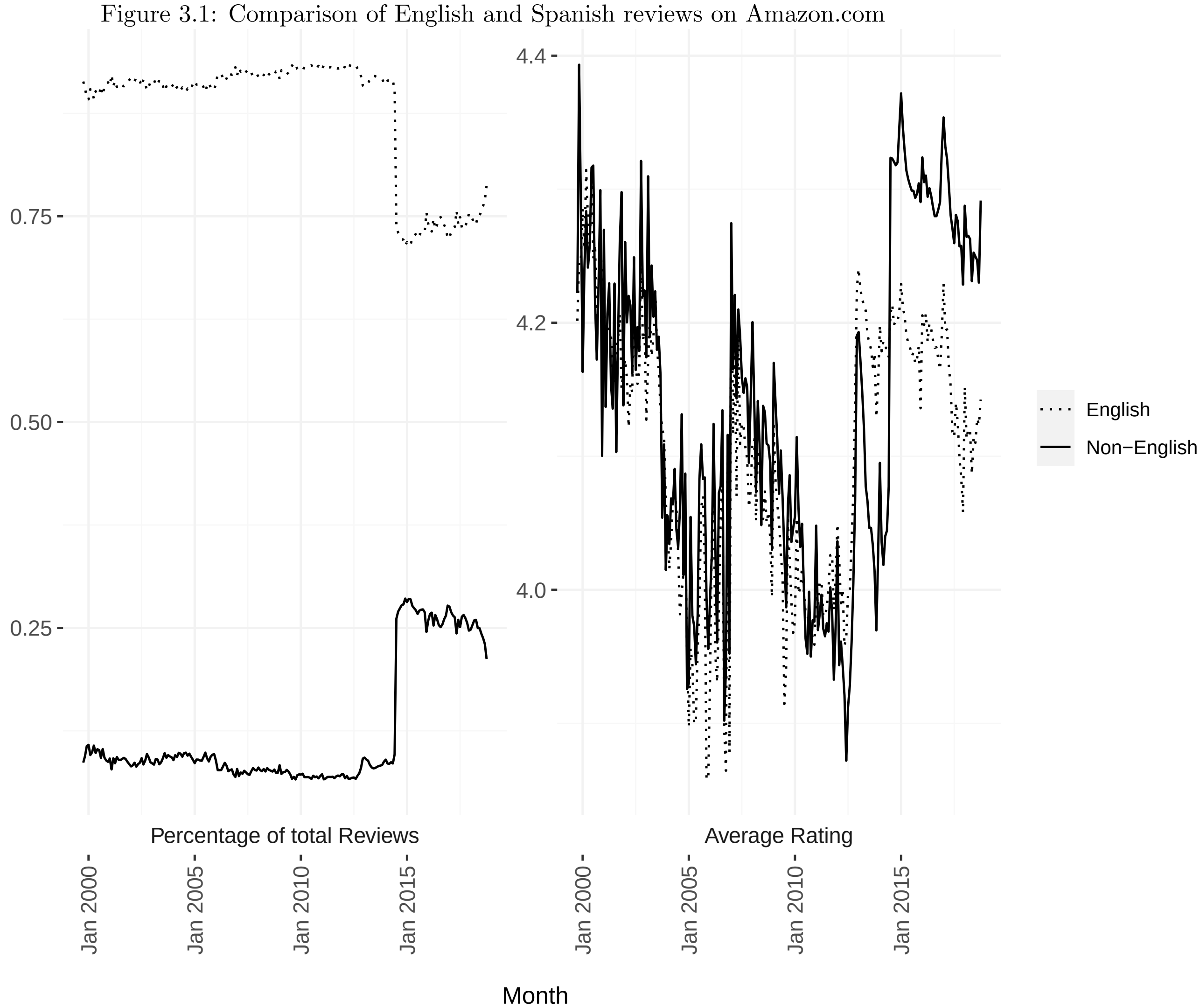
Table 3.2: Data summary

	Total Products with		Total Reviews with		Mean Rating	
	English	Spanish	English	Spanish	English	Spanish
Clothing and Jewelry	2,664,884	45,801	31,872,206	66,656	4.19	4.40
Electronics	753,310	26,238	20,776,033	47,095	4.07	4.45
Cell phone and Accessories	586,352	14,855	9,931,345	27,724	3.93	4.28
Home and Kitchen	1,280,594	14,055	21,729,350	18,644	4.19	4.34
Sports and Outdoors	951,314	13,749	12,823,763	18,225	4.24	4.43
CDs and Vinyl	430,408	10,864	4,474,251	15,132	4.49	4.56
Automotive	919,742	11,427	7,896,584	13,480	4.25	4.40
Movies and TV	181,093	6,725	8,638,321	11,499	4.23	4.12
Toys and Games	621,325	8,896	8,084,740	11,221	4.23	4.46
Tools and Home Improvement	557,420	6,898	8,924,077	8,867	4.22	4.39
Video Games	71,716	3,834	2,526,740	8,709	4.02	4.45
Office Products	305,392	4,897	5,524,830	7,867	4.18	4.53
Grocery and Gourmet Food	281,927	4,296	4,983,510	5,689	4.31	4.48
Arts Crafts and Sewing	301,213	3,122	2,845,697	3,930	4.32	4.47
Pet Supplies	197,742	2,941	6,485,299	3,804	4.15	4.34
Patio Lawn and Garden	275,519	3,115	5,187,816	3,668	4.12	4.24
Musical Instruments	111,633	2,332	1,494,658	3,240	4.26	4.53
Industrial and Scientific	164,835	1,628	1,740,200	2,088	4.29	4.48
Fashion	184,682	1,220	871,409	1,424	3.91	4.04
Luxury Beauty	12,106	503	568,980	876	4.22	4.46
Prime Pantry	10,797	573	461,135	764	4.34	4.56
Beauty	32,386	455	367,172	588	4.11	4.16
Appliances	30,150	407	597,973	504	4.27	4.40

This table summarizes the data by each product category and sorted by the total reviews present in the category. Total products column show the number of products in each category with English and Spanish reviews. Total reviews column show the total number of English and Spanish reviews the product category has. Mean rating column is the average rating of all the English and Spanish review ratings for the product category.

3.7 Exhibits

Figure 3.1: Comparison of English and Spanish reviews on Amazon.com



The figure shows the comparison between the average rating and the number of English and Spanish reviews posted on Amazon.com.

Table 3.3: Descriptive Statistics

			Min	Q1	Q2	Q3	Max	Mean	SD	Description
$\mathbf{MV}_r$	Macroeconomic	USA EPU volatility	1.37	6.85	10.57	18.97	86.21	15.35	11.51	Volatility of economic uncertainty index of USA for the last 6 months
		USA e-commerce sales	5.24	78.02	90.69	106.58	155.06	90.99	23.01	E-commerce sales in USA in the quarter measured in billions
		MAGA	0.00	0.00	0.00	0.00	83.00	3.75	12.30	Interest in Make America Great Again topic on trends.google.com
$\mathbf{RV}_{p,r+1}$	Next Review	Next rate	0.00	3.33	15.00	60.00	62520.00	65.65	525.62	The number of reviews per month, calculated as shown in Equation (4.1)
		Next Δrate	−54486.10	−15.99	−4.23	9.23	61167.18	−10.18	416.96	Difference between the next review rate to the cumulative review rate of th
		Next score	1.00	4.00	5.00	5.00	5.00	4.19	1.30	Review rating given by the customer
		Next Δscore	−4.00	−0.47	0.36	0.69	3.96	−0.05	1.20	Difference between the next review's rating to the cumulative rating
		Next English rate	0.00	3.00	15.00	30.00	61800.00	63.36	511.41	The number of English reviews per month, calculated as shown in Equatio
		Next English Δrate	−53780.15	−16.23	−4.35	8.73	60447.18	−10.84	413.55	Difference between the next English review's rate to the cumulative Englis
		Next Spanish rate	0.00	0.00	0.00	0.00	330.00	0.00	0.51	The number of Spanish reviews per month, calculated as shown in Equatic
		Next Spanish Δrate	−258.47	0.00	0.00	0.00	330.00	−0.17	2.04	Difference between the next Spanish review's rate to the cumulative Spanis
$\mathbf{RV}_{p,r-4:r}$	Review	Recent rate	0.03	7.73	21.20	43.50	62520.00	64.98	520.32	Average review rate of recent five reviews
		Recent score	1.00	3.80	4.40	4.80	5.00	4.19	0.74	Average score of the recent five reviews
		Recent English count	0.00	5.00	5.00	5.00	5.00	4.94	0.25	Number of English reviews in the recent five reviews
		Recent English rate	0.00	7.33	20.77	42.20	61800.00	63.51	509.85	Average review rate of English reviews in the recent five reviews
		Recent English score	0.00	3.80	4.40	4.80	5.00	4.19	0.74	Average score of English reviews in the recent five reviews
		Recent Spanish count	0.00	0.00	0.00	0.00	5.00	0.01	0.09	Number of Spanish reviews in the recent five reviews
		Recent Spanish rate	0.00	0.00	0.00	0.00	330.00	0.00	0.27	Average review rate of Spanish reviews in the recent five reviews
		Recent English score	0.00	0.00	0.00	0.00	5.00	0.03	0.39	Average score of Spanish reviews in the recent five reviews
$\mathbf{RV}_{p,1:r-5}$		Cumulative count	11.00	41.00	130.00	450.00	28539.00	529.42	1267.47	Cumulative number of reviews as seen by the customer
		Cumulative score	1.00	4.00	4.33	4.57	5.00	4.23	0.47	Cumulative average score of the reviews as seen by the customer
		Cumulative rate	0.02	9.64	22.13	46.64	54529.46	75.01	582.21	Cumulative average review rate as seen by the customer
		Cumulative English count	0.00	39.00	125.00	438.00	28387.00	517.70	1246.90	Cumulative number of English reviews as seen by the customer
		Cumulative English score	0.00	4.00	4.33	4.57	5.00	4.24	0.48	Cumulative average of English review score as seen by the customer
		Cumulative English rate	0.00	9.56	21.98	46.36	53823.62	74.20	574.64	Cumulative average of English review rate as seen by the customer
		Cumulative Spanish count	0.00	0.00	0.00	0.00	304.00	0.01	0.57	Cumulative number of Spanish reviews as seen by the customer
		Cumulative Spanish score	−3.74	0.00	0.00	0.00	2.95	0.01	0.12	Cumulative average of Spanish review score as seen by the customer
		Cumulative Spanish rate	0.00	0.00	0.00	0.00	258.47	0.17	2.00	Cumulative average of Spanish review rate as seen by the customer
CV_r	Controls	Monthly time trend	7.00	208.00	223.00	235.00	254.00	218.41	25.96	Number of months since the first review in the dataset

This table summarizes the dependent and independent variables we use in our model. All the values are calculated for product after they receive at least 10 reviews.

Table 3.4: Effect of Spanish Reviews on review rate and scores

			Next Review Rate	Next Review Score	ΔNext Review Score
Cumulative	Cumulative	English count	0.071^{*}	0.001	-0.006^{*}
			(0.015)	(0.002)	(0.001)
	Review	English rate	0.280^{*}	0.003	0.000
			(0.010)	(0.004)	(0.002)
	Variables	English score	0.033^{*}	-0.011^{*}	-0.688^{*}
			(0.003)	(0.003)	(0.015)
		Spanish count	−0.046	0.145^{*}	0.114^{*}
			(0.031)	(0.016)	(0.018)
		Spanish rate	0.030^{*}	-0.002^{*}	0.004^{*}
			(0.009)	(0.001)	(0.001)
		Spanish score	−0.003	0.004^{*}	0.003
			(0.006)	(0.002)	(0.003)
Recent	Recent	English rate	0.607^{*}	$0.003^{\dagger}$	0.004^{*}
			(0.030)	(0.001)	(0.002)
	Review	English score	0.022^{*}	0.989^{*}	0.773^{*}
			(0.006)	(0.001)	(0.010)
	Variables	Spanish count	-0.253^{*}	0.023	−0.024
			(0.061)	(0.027)	(0.023)
		Spanish rate	0.244^{*}	$-0.018^{\dagger}$	−0.013
			(0.045)	(0.009)	(0.011)
		Spanish score	0.002	0.192^{*}	0.149^{*}
			(0.005)	(0.003)	(0.003)
Statistics	Statistics	R-squared	0.638	0.271	0.165
		Adjusted R-squared	0.638	0.271	0.165
		N (in millions)	159.82	170.81	170.81

The table reports the results of the analysis on full dataset. The dependent variables are next review rate and next review score. The major independent variables of interest are the cumulative review variables. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 3.5: Effect of review language on English customers

			Next Review Score	ΔNext Review Score
Cumulative	Review	English rate	0.003	0.000
			(0.008)	(0.003)
	Variables	English score	-0.018^{*}	-0.691^{*}
			(0.005)	(0.016)
		Spanish rate	−0.002	$0.004^{\dagger}$
			(0.005)	(0.002)
		Spanish score	0.002	0.000
			(0.005)	(0.005)
Cumulative1	Cumulative	English count	0.001	-0.007^{*}
			(0.004)	(0.002)
Statistics	Statistics	R-squared	0.261	0.168
		Adjusted R-squared	0.261	0.168
		N (in millions)	40.13	40.13

The table reports the results of the analysis of only English reviews. The dependent variables are next review score, Δnext review score, Δnext review rate. The major independent variables of interest are the cumulative review variables. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 3.6: Effect of review language on Spanish customers

			Next Review Score	ΔNext Review Score
Cumulative	Review	English rate	0.002*	−0.001
			(0.001)	(0.002)
	Variables	English score	−0.005†	−0.642*
			(0.003)	(0.025)
		Spanish count	0.002	−0.010
			(0.004)	(0.007)
		Spanish rate	0.007	0.014†
			(0.005)	(0.008)
		Spanish score	0.006	0.016
			(0.006)	(0.013)
Statistics	Statistics	R-squared	0.789	0.860
		Adjusted R-squared	0.789	0.860
		N (in millions)	0.07	0.07

The table reports the results of the analysis of only Spanish reviews. The dependent variables are next review score, Δnext review score, Δnext review rate. The major independent variables of interest are the cumulative review variables. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 3.7: Effect of review language on first review customers

			ΔNext Review Score	Next Review Score	ΔNext Review Score
Macroeconomic	Variables	M.A.G.A	0.006*	−0.004*	−0.002*
			(0.001)	(0.000)	(0.000)
Recent	Recent	English count			−0.353*
					(0.032)
		English rate			0.636*
					(0.025)
	Review	English score			0.028*
					(0.006)
	Variables	Spanish count			−0.045
					(0.172)
		Spanish rate			0.381*
					(0.056)
		Spanish score			0.013†
					(0.006)
Cumulative	Cumulative	English count		0.073*	0.040*
				(0.015)	(0.010)
	Review	English rate		0.782*	0.299*
				(0.029)	(0.011)
	Variables	English score		0.090*	0.026*
				(0.012)	(0.005)
		Spanish count		0.202*	−0.051
				(0.033)	(0.035)
		Spanish rate		0.090*	0.071*
				(0.031)	(0.019)
		Spanish score		0.039*	0.018*
				(0.017)	(0.007)
Statistics	Statistics	R-squared	0.056	0.575	0.653
		Adjusted R-squared	0.056	0.575	0.653
		N (in millions)	7.88	7.88	7.88

The table reports the results of the analysis of only Spanish reviews. The dependent variables are next review score, Δnext review score, Δnext review rate. The major independent variables of interest are the cumulative review variables. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Appendices

Appendix A Appendix for Chapter 1

Table 8: Impact of EPU volatility on domestic review rate of domestic products

			Next Review Rate (1)	ΔNext Review Rate (2)	Recent Review Rate (3)	Next Review Score (4)	ΔNext Review Score (5)	Recent Review Score (6)
		Intercept	-2.491^{*}	-18.793^{*}	0.057	0.161	-4.389^{*}	-4.650^{*}
			(0.471)	(2.557)	(0.074)	(0.595)	(1.049)	(0.364)
Macroeconomic	$\mathbf{MV}_r^{Domestic}$	Adjusted USA EPU volatility	-0.110^{*}	-0.250^{*}	-0.017^{*}	-0.040^{*}	-0.034^{*}	-0.027^{*}
			(0.005)	(0.017)	(0.001)	(0.007)	(0.007)	(0.004)
	$\mathbf{MV}_r^{China}$	Adjusted China EPU volatility	0.095^{*}	0.318^{*}	0.023^{*}	0.061^{*}	0.057^{*}	0.058^{*}
			(0.008)	(0.023)	(0.001)	(0.009)	(0.010)	(0.006)
Reviews	$\mathbf{RV}_{p,d,1:r-5}^{Domestic}$	Cumulative count	0.174^{*}	0.555^{*}	0.050^{*}	-0.463^{*}	-0.767^{*}	0.482^{*}
			(0.020)	(0.084)	(0.003)	(0.025)	(0.032)	(0.015)
		Cumulative rate	2.164^{*}	-0.585^{*}	0.613^{*}	0.156^{*}	0.126^{*}	0.191^{*}
			(0.018)	(0.061)	(0.003)	(0.023)	(0.023)	(0.014)
		Cumulative score	-0.158^{*}	0.458^{*}	-0.049^{*}	0.079^{*}	-1.172^{*}	0.903^{*}
			(0.019)	(0.080)	(0.003)	(0.024)	(0.029)	(0.015)
Prices	$\mathbf{CV}_r^{Domestic}$	Minimum price	-0.027^{*}	$0.032^{\dagger}$	-0.007^{*}	0.031^{*}	0.023^{*}	0.037^{*}
			(0.005)	(0.016)	(0.001)	(0.007)	(0.007)	(0.004)
		Ending range	-0.035^{*}	0.013^{*}	-0.004^{*}	0.007^{*}	0.006^{*}	0.011^{*}
			(0.002)	(0.005)	(0.000)	(0.002)	(0.002)	(0.001)
Controls		Time trend	1.267^{*}	8.505^{*}	0.331^{*}	1.908^{*}	4.381^{*}	2.511^{*}
			(0.208)	(1.121)	(0.033)	(0.262)	(0.461)	(0.161)
		Time Trend^2	-0.164^{*}	-1.057^{*}	-0.045^{*}	-0.256^{*}	-0.522^{*}	-0.334^{*}
			(0.024)	(0.125)	(0.004)	(0.030)	(0.051)	(0.018)
Statistics		R-squared	0.500	0.013	0.864	0.025	0.027	0.552
		Adjusted R-squared	0.500	0.013	0.864	0.025	0.027	0.552
		N	115329	107052	115329	115329	108050	115329

This table reports the results for domestic products. The headers are the dependent variables. $\mathrm{MV}_r^{domestic}$ is the volatility of China EPU, and $\mathrm{MV}_r^{Domestic}$ is the volatility of USA EPU in the past six months. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Table 9: Effects of curated global reviews on domestic reviews

			Next Review Rate		ΔNext Review Rate		Next Review Score		ΔNext Review Score	
			Early	Late	Early	Late	Early	Late	Early	Late
		Intercept	3.436†	−16.173*	7.088†	−47.748*	2.568	−9.057*	2.568	−9.057*
			(1.692)	(1.107)	(3.260)	(3.656)	(1.948)	(1.491)	(1.948)	(1.491)
Reviews	$\mathbf{RV}^{Domestic}_{p,d,1:r-5}$	Cumulative count	0.247	0.520*	−0.450	1.982*	−1.253*	−0.454*	−1.253*	−0.454*
			(0.201)	(0.045)	(0.387)	(0.148)	(0.231)	(0.060)	(0.231)	(0.060)
		Cumulative rate	1.436*	4.560*	−1.019*	0.502*	0.084*	0.159*	0.084*	0.159*
			(0.030)	(0.032)	(0.058)	(0.105)	(0.035)	(0.043)	(0.035)	(0.043)
		Cumulative count * score	−0.003	0.029*	−0.004	−0.452*	0.294*	0.102*	0.294*	0.102*
			(0.046)	(0.010)	(0.089)	(0.033)	(0.053)	(0.014)	(0.053)	(0.014)
		Cumulative score	0.028	−0.149*	−0.003	1.914*	−0.486*	0.231*	−1.486*	−0.769*
			(0.118)	(0.052)	(0.228)	(0.171)	(0.136)	(0.070)	(0.136)	(0.070)
	$\mathbf{RV}^{Global}_{p,g,1:r}$	Curated English count	−0.125*	0.028*	−0.074	0.062†	0.143*	0.008	0.143*	0.008
			(0.048)	(0.011)	(0.093)	(0.036)	(0.055)	(0.015)	(0.055)	(0.015)
		Curated English score	−0.017	0.030*	−0.060	−0.020	−0.001	0.004	−0.001	0.004
			(0.019)	(0.008)	(0.037)	(0.027)	(0.022)	(0.011)	(0.022)	(0.011)
		Curated non-English count	0.078†	−0.038*	0.149†	−0.038	−0.102†	−0.005	−0.102†	−0.005
			(0.040)	(0.010)	(0.078)	(0.033)	(0.046)	(0.014)	(0.046)	(0.014)
		Curated non-English score	0.037†	0.006	0.006	0.045†	−0.010	−0.020†	−0.010	−0.020†
			(0.017)	(0.007)	(0.032)	(0.025)	(0.019)	(0.010)	(0.019)	(0.010)
Controls	$\mathbf{CV}^{Domestic}_{r}$	Time trend	−1.567†	6.938*	−2.508†	19.324*	1.801†	5.724*	1.801†	5.724*
			(0.750)	(0.478)	(1.446)	(1.579)	(0.864)	(0.644)	(0.864)	(0.644)
		Time Trend^2	0.196†	−0.857*	0.320†	−2.367*	−0.221†	−0.669*	−0.221†	−0.669*
			(0.088)	(0.053)	(0.169)	(0.175)	(0.101)	(0.071)	(0.101)	(0.071)
Statistics		R-squared	0.191	0.450	0.039	0.011	0.023	0.030	0.115	0.007
		Adjusted R-squared	0.190	0.450	0.038	0.010	0.022	0.030	0.114	0.007
		N	12128	94924	12128	94924	12128	94924	12128	94924

This table reports the results for global products. $\mathrm{RV}^{Domestic}_{gp,r+1}$ is the dependent variable, next-review rate of the customer, $\tilde{\mathbf{RV}}^{Global}_{gp,r}$ is the group of variables constructed using the shared global reviews, $\mathrm{EPU}^{Domestic}_{g,r}$ is the EPU value for USA and $\mathrm{EPU}^{Global}_{g,r}$ is the EPU value for China, $\mathbf{RV}^{Domestic}_{gp,r-4:r}$ is a vector of domestic last five review variables, $\mathbf{RV}^{Domestic}_{gp,1:r-5}$ is a vector of domestic cumulative review variables, $\mathbf{CV}^{Domestic}_{gp,r}$ is a vector of price control variables, and $\mathbf{FE}_{gp,r}$ is a vector of fixed effect variables. "Early" and "Late" denote early customers for the product .

Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$.

Appendix B Appendix for Chapter 2

Table 10: Impact of ”made in” mentions on product review rate

			Dependent Variable ΔNext Review Rate	Dependent Variable ΔNext Review Score
Description	$\mathbf{PD}_p$	Made in China	-0.011	0.000
			(0.016)	(0.002)
		Made in USA	-0.037^{*}	$0.002^{\dagger}$
			(0.010)	(0.001)
Review Variables	$\mathbf{RV}_{p,r-4:r}$	Recent rate	1.170^{*}	-0.036^{*}
			(0.081)	(0.009)
		Recent rate * score	-0.079^{*}	0.015^{*}
			(0.015)	(0.002)
		Recent USA mentions	0.108	-0.001
			(0.078)	(0.010)
		Recent China mentions	0.060^{*}	0.001
			(0.013)	(0.002)
		Recent score	0.143^{*}	0.999^{*}
			(0.023)	(0.005)
	$\mathbf{RV}_{p,1:r-5}$	Cumulative rate	-1.228^{*}	0.011^{*}
			(0.039)	(0.004)
		Cumulative rate * score	0.032^{*}	-0.006^{*}
			(0.009)	(0.001)
		Cumulative reviews	$-0.129^{\dagger}$	0.008^{*}
			(0.076)	(0.003)
		Cumulative reviews * score	0.057^{*}	$-0.002^{\dagger}$
			(0.015)	(0.001)
		Cumulative USA mentions	-0.016	0.003
			(0.036)	(0.002)
		Cumulative China mentions * mention score	-0.028^{*}	0.000^{*}
			(0.004)	(0.000)
		Cumulative China mentions	0.026^{*}	0.004^{*}
			(0.011)	(0.001)
		Cumulative USA mentions * mention score	-0.005	0.000^{*}
			(0.007)	(0.000)
		Cumulative score	-0.150^{*}	-1.015^{*}
			(0.033)	(0.002)
		Made in China score	0.026^{*}	-0.001^{*}
			(0.003)	(0.000)
		Made in USA score	0.018	-0.001^{*}
			(0.014)	(0.000)
Political	$\mathbf{PV}_r$	2000 USA Presidential election	0.172	0.007
			(0.144)	(0.005)
		2001 September 11 attacks	0.047	0.007^{*}
			(0.120)	(0.002)
		2007 Great Recession	-0.273^{*}	-0.004^{*}
			(0.056)	(0.001)
		2008 China top foreign creditor	0.122^{*}	0.006^{*}
			(0.030)	(0.002)
		2009 USA Presidential election	-0.165^{*}	-0.013^{*}
			(0.031)	(0.003)
		2011 Pivot to East Asia	$-0.057^{\dagger}$	$-0.004^{\dagger}$
			(0.032)	(0.002)
		2012 Rising trade tensions	-0.220^{*}	-0.008^{*}
			(0.025)	(0.001)
		2012 China Presidential election	0.212^{*}	0.009^{*}
			(0.028)	(0.002)
		2013 Sunnylands summit	$0.031^{\dagger}$	-0.008^{*}
			(0.018)	(0.001)
		2017 Presidential election	-0.141^{*}	-0.005^{*}
			(0.023)	(0.001)
		2018 tariff war	$0.057^{\dagger}$	0.004^{*}
			(0.029)	(0.001)
Controls	$\mathbf{CV}_r$	Time Trend	4.530^{*}	0.191^{*}
			(1.767)	(0.030)
		Time trend^2	-0.475^{*}	-0.019^{*}
			(0.183)	(0.003)
Statistics		R-square	0.09	0.27
		Adj.R-square	0.09	0.27
		N (in millions)	118.97	118.97

The column header indicates the dependent variable in the model, which is either the Δnext review rate or the Δnext review score. The model uses the recent mention variables, cumulative review variables, cumulative review variables, time trend variables, and product category fixed effects as independent variables and control variables. Significance at $^{*}p < 0.05$, $^{\dagger}p < 0.01$

Bibliography

Ajjuguttu, M. R., A. Colak, and L. D.Fredendall (2023): "Do Global Policy Uncertainty Influence Online Reviews? Evidence from Amazon Global Review Sharing," .

Amblee, N., and T. Bui (2008): "Can Brand Reputation Improve the Odds of Being Reviewed On-Line?," *International Journal of Electronic Commerce*, 12(3), 11–28.

Angelovska, N. (2019): "Top 5 Online Retailers: 'Electronics And Media' Is The Star Of E-commerce Worldwide," .

Antonakakis, N., D. Gabauer, R. Gupta, and V. Plakandaras (2018): "Dynamic connectedness of uncertainty across developed economies: A time-varying approach," *Economics Letters*, 166, 63–75.

Antonio, N., A. de Almeida, L. Nunes, F. Batista, and R. Ribeiro (2018): "Hotel online reviews: different languages, different opinions," *Information Technology and Tourism*, 18(1-4), 157–185.

Baker, S., N. Bloom, and S. J. D. an Xiaoxi Wang (2013): "A measure of economic policy uncertainty for China," .

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty∗," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Bernanke, B. S. (1983): "Irreversibility, Uncertainty, and Cyclical Investment," *The Quarterly Journal of Economics*, 98(1), 85.

Biljanovska, N., F. Grigoli, and M. Hengge (2021): "Fear thy neighbor: Spillovers from economic policy uncertainty," *Review of International Economics*, 29(2), 409–438.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

Bloom, N., S. Bond, and J. V. Reenen (2007): "UNCERTAINTY AND INVESTMENT DY-

NAMICS," *The review of economic studies.*

BOJANOWSKI, P., E. GRAVE, A. JOULIN, AND T. MIKOLOV (2017a): "Bag of Tricks for Efficient Text Classification," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pp. 427–431. Association for Computational Linguistics.

——— (2017b): "Enriching Word Vectors with Subword Information," *Transactions of the Association for Computational Linguistics*, 5, 135–146.

BOJANOWSKI, P., E. GRAVE, A. JOULIN, T. MIKOLOV, M. DOUZE, AND H. JEGOU (2016): "FastText.zip: Compressing text classification models," *arXiv preprint arXiv:1612.03651.*

BOWN, C. P. (2023): "US imports from China are both decoupling and reaching new highs. Here's how.," Discussion paper.

BUELL, R. W., E. PORTER, AND M. I. NORTON (2021): "Surfacing the Submerged State: Operational Transparency Increases Trust in and Engagement with Government," *Manufacturing and Service Operations Management*, 23(4), 781–802.

CARTWRIGHT, N. (2007): "'Counterfactuals in economics : a commentary.', in Causation and explanation.," *MIT Press.*

CHAKRABORTY, U., AND S. BHAT (2018): "Credibility of online reviews and its impact on brand image," *Management Research Review*, 41(1), 148–164.

CHEN, N., AND Z. Z. ZHONG (2023): "History and Country-of-Origin Effects," *Marketing Science.*

CHEVALIER, J. A., AND D. MAYZLIN (2006): "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 43(3), 345–354.

CHRYSANTHOS DELLAROCAS, NEVEEN FARAG AWAD, X. M. Z. (2004): "Exploring the Value of Online Reviews to Organizations:Implications for Revenue Forecasting and Planning," .

CIALDINI, R. B. (2006): *Influence: The Psychology of Persuasion, Revised Edition.* Collins Business Essentials.

ÇOLAK, G., A. DURNEV, AND Y. QIAN (2017): "Political Uncertainty and IPO Activity: Evidence from U.S. Gubernatorial Elections," *Journal of Financial and Quantitative Analysis*, 52(6), 2523–2564.

COLOMBO, V. (2013): "Economic policy uncertainty in the US: Does it matter for the Euro area?," *Economics Letters*, 121(1), 39–42.

Davis, A. M., V. Gaur, and D. Kim (2021): “Consumer Learning from Own Experience and Social Information: An Experimental Study,” *Management Science*, 67(5), 2924–2943.

Dawson, C. (2021): “Country of Origin information mandatory on Amazon,” .

Dellarocas, C. (2003): “The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms,” *Management Science*, 49(10), 1407–1424.

——— (2006): “Strategic Manipulation of Internet Opinion Forums: Implications for Consumers and Firms,” *Management Science*, 52(10), 1577–1593.

Dellarocas, C., G. Gao, and R. Narayan (2010): “Are Consumers More Likely to Contribute Online Reviews for Hit or Niche Products?,” *Journal of Management Information Systems*, 27(2), 127–158.

Dellarocas, C., and R. Narayan (2006): “What motivates consumers to review a product online?A study of the product-specific antecedents of online movie reviews,” *Statistical Science*.

Deshpande, V., and P. K. Pendem (2022): “Logistics Performance, Ratings, and Its Impact on Customer Purchasing Behavior and Sales in E-Commerce Platforms,” *Manufacturing & Service Operations Management*.

Drozdenko, R., and M. Jensen (2009): “Translating country-of-origin effects into prices,” *Journal of Product and Brand Management*, 18(5), 371–378.

Duan, W., B. Gu, and A. B. Whinston (2008): “Do online reviews matter? — An empirical investigation of panel data,” *Decision Support Systems*, 45(4), 1007–1016.

Eberly, J. C. (1994): “Adjustment of Consumers’ Durables Stocks: Evidence from Automobile Purchases,” *JOURNAL OF POLITICAL ECONOMY*.

Elliott, G. R., and R. C. Cameron (1994): “Consumer Perception of Product Quality and the Country-of-Origin Effect,” *Journal of International Marketing*.

Eppler, M. J., and J. Mengis (2004): “The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines,” *The Information Society*, 20(5), 325–344.

Facebook, I. (2016): *fastText: Library for fast text representation and classification*.

fen Hu, H., and A. S. Krishen (2019): “When is enough, enough? Investigating product reviews and information overload from a consumer empowerment perspective," *Journal of Business Research*, 100, 27–37.

Gallino, S., and A. Moreno (2018): "The Value of Fit Information in Online Retail: Evidence from a Randomized Field Experiment," *Manufacturing & Service Operations Management*, 20(4), 767–787.

Gaur, V., and Y.-H. Park (2007): "Asymmetric Consumer Learning and Inventory Competition," *Management Science*, 53(2), 227–240.

Ghirelli, C., J. J. Pérez, and A. Urtasun (2021): "The spillover effects of economic policy uncertainty in Latin America on the Spanish economy," *Latin American Journal of Central Banking*, 2(2), 100029.

Ghose, A., P. G. Ipeirotis, and B. Li (2014): "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," *Management Science*, 60(7), 1632–1654.

Goodell, J. W., A. Goyal, and A. Urquhart (2021): "Uncertainty of uncertainty and firm cash holdings," *Journal of Financial Stability*, 56, 100922.

Gregan-Paxton, J., and P. Moreau (2003): "How Do Consumers Transfer Existing Knowledge? A Comparison of Analogy and Categorization Effects," *Journal of Consumer Psychology*.

Gudmundsson, Jørgen; Natvik, G. J. (2012): "That Uncertain Feeling - How Consumption Responds to Economic Uncertainty in Norway," .

Gulen, H., and M. Ion (2015): "Policy Uncertainty and Corporate Investment," *Review of Financial Studies*, p. hhv050.

Ha-Brookshire, J., and S.-H. Yoon (2012): "Country of origin factors influencing US consumers' perceived price for multinational products," *Journal of Consumer Marketing*, 29(6), 445–454.

Handley, K., and N. Limão (2017): "Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States," *American Economic Review*, 107(9), 2731–2783.

Heckman, J. J. (2001): "Econometrics Counterfactuals and Causal Models," Keynote Address, International Statistical Institute.

Holmqvist, J. (2011): "Consumer language preferences in service encounters: a cross-cultural perspective," *Managing Service Quality: An International Journal*, 21(2), 178–191.

Holmqvist, J., and Y. V. Vaerenbergh (2013): "Perceived importance of native language use in service encounters," *The Service Industries Journal*, 33(15-16), 1659–1671.

Holmqvist, J., Y. V. Vaerenbergh, R. Lunardo, and M. Dahlén (2019): "The Language Backfire Effect: How Frontline Employees Decrease Customer Satisfaction through Language Use," *Journal of Retailing*, 95(2), 115–129.

Howarth, J. (2022): "81 Online Review Statistics (New 2022 Data)," .

Hu, G., and S. Liu (2021): "Economic Policy Uncertainty (EPU) and China's Export Fluctuation in the Post-pandemic Era: An Empirical Analysis based on the TVP-SV-VAR Model," *Frontiers in Public Health*, 9.

Jens, C. E. (2017): "Political uncertainty and investment: Causal evidence from U.S. gubernatorial elections," *Journal of Financial Economics*, 124(3), 563–579.

Kaemingk, D. (2020): "Online reviews statistics to know in 2022," Discussion paper.

Kelly, B., L. Pástor, and P. Veronesi (2016): "The Price of Political Uncertainty: Theory and Evidence from the Option Market," *The Journal of Finance*, 71(5), 2417–2480.

Kelly, N. (2012): "Speak to Global Customers in Their Own Language," *Harvard Business Review*.

Klössner, S., and R. Sekkel (2014): "International spillovers of policy uncertainty," *Economics Letters*, 124(3), 508–512.

Ko, D.-G., F. Mai, Z. Shan, and D. Zhang (2019): "Operational efficiency and patient-centered health care: A view from online physician reviews," *Journal of Operations Management*, 65(4), 353–379.

Lai, A. (2022): "Economic Uncertainty: Here's How Consumers Are Responding," *Forrester*.

Lee, B.-K., and W.-N. Lee (2004): "The effect of information overload on consumer choice quality in an on-line environment," *Psychology and Marketing*, 21(3), 159–183.

Lei, Z., D. Yin, S. Mitra, and H. Zhang (2022): "Swayed by the reviews: Disentangling the effects of average ratings and individual reviews in online word-of-mouth," *Production and Operations Management*, 31(6), 2393–2411.

Liu, Y. (2006): "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of marketing*.

Maronick, T. J. (1995): "An empirical investigation of consumer perceptions of "made in USA" claims," *International Marketing Review*, 12(3), 15–30.

Mbabazi, E. (2022): "Amazon to Expand to Africa in 2023," .

Mejia, J., S. Mankad, and A. Gopal (2021): "Service Quality Using Text Mining: Measurement and Consequences," *Manufacturing & Service Operations Management*, 23(6), 1354–1372.

Mohsin, M. (2022): "10 Amazon Statistics You Need to Know in 2022," Discussion paper.

Moreno, A., and C. Terwiesch (2014): "Doing Business with Strangers: Reputation in Online Service Marketplaces," *Information Systems Research*, 25(4), 865–886.

Moriuchi, E. (2021): "The impact of country of origin on consumers' pricing judgments in ecommerce settings," *International Marketing Review*, 38(3), 514–538.

Mouselimis, L. (2022): *fastText: Efficient Learning of Word Representations and Sentence Classification using R*R package version 1.0.3.

Nagashima, A. (1970): "A Comparison of Japanese and U. S. Attitudes toward Foreign Products," *Journal of Marketing*.

Ni, J., J. Li, and J. McAuley (2019): "Justifying Recommendations using Distantly-Labeled Reviews andFine-Grained Aspects," *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*.

Ozdemir, O., W. Han, and M. Dalbor (2021): "Economic policy uncertainty and hotel occupancy: the mediating effect of consumer sentiment," *Journal of Hospitality and Tourism Insights*, 5(2), 253–273.

Papanastasiou, Y., and N. Savva (2017): "Dynamic Pricing in the Presence of Social Learning and Strategic Consumers," *Management Science*, 63(4), 919–939.

Phan, D. H. B., S. S. Sharma, and V. T. Tran (2018): "Can economic policy uncertainty predict stock returns? Global evidence," *Journal of International Financial Markets, Institutions and Money*, 55, 134–150.

Podium (2017): "State of Online Reviews," .

Puntoni, S., B. de Langhe, and S. M. J. van Osselaer (2009): "Bilingualism and the Emotional Intensity of Advertising Language," *Journal of Consumer Research*, 35(6), 1012–1025.

Roetzel, P. G. (2018): "Information overload in the information age: a review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development," *Business Research*, 12(2), 479–522.

Romer, C. D. (1990): "The Great Crash and the Onset of the Great Depression," *The Quarterly Journal of Economics*, 105(3), 597.

Schniederjans, M. J., Q. Cao, and J. R. Olson (2004): "Consumer Perceptions of Product Quality: Made in China," *Quality Management Journal*, 11(3), 8–18.

Schniederjans, M. J., Q. Cao, D. Schniederjans, and V. C. Gu (2011): "Consumer Perceptions of Product Quality Revisited: Made in China," *Quality Management Journal*, 18(3), 52–68.

Schuckert, M., X. Liu, and R. Law (2015): "Hospitality and Tourism Online Reviews: Recent Trends and Future Directions," *Journal of Travel and Tourism Marketing*, 32(5), 608–621.

Soto-Acosta, P., F. J. Molina-Castillo, C. Lopez-Nicolas, and R. Colomo-Palacios (2014): "The effect of information overload and disorganisation on intention to purchase online," *Online Information Review*, 38(4), 543–561.

Spool, M. (2009): "The Magic Behind Amazon's 2.7 Billion Dollar Question," .

Tamara Charm, Jason Rico Saavedra, K. R. T. S. (2022): "The Great Uncertainty: US consumer confidence and behavior during inflationary times," *McKinsey & Company*.

Vaerenbergh, Y. V., and J. Holmqvist (2013): "Speak my language if you want my money," *European Journal of Marketing*, 47(8), 1276–1292.

——— (2014): "Examining the relationship between language divergence and word-of-mouth intentions," *Journal of Business Research*, 67(8), 1601–1608.

Verma, S. (2020): "Amazon makes it compulsory for sellers to display country of origin on product listings," .

Washington, A. (2020): "Your Customers Want Locally Sourced Products. Here's Why.," Discussion paper.

Wu, S., N. Wingate, Z. Wang, and Q. Liu (2019): "The Influence of Fake Reviews on Consumer Perceptions of Risks and Purchase Intentions," *Journal of Marketing Development and Competitiveness*.

Xu, Y., M. Armony, and A. Ghose (2021): "The Interplay Between Online Reviews and Physician Demand: An Empirical Investigation," *Management Science*, 67(12), 7344–7361.

Yu, A., S. Yu, and H. Liu (2022): "How a "China-made" label influences Chinese Youth's product evaluation:The priming effect of patriotic and nationalistic news," *Journal of Retailing and Consumer Services*.

Yu, H., L. Fang, D. Du, and P. Yan (2017): "How EPU drives long-term industry beta," *Finance Research Letters*, 22, 249–258.

Zhang, M., and D. Merunka (2015): "The impact of territory of origin on product authenticity perceptions," *Asia Pacific Journal of Marketing and Logistics*, 27(3), 385–405.

Zhou, L. (2023): "Online Review Statistics: The Definitive List (2023 Data)," luisazhou.com.