Washington University in St. Louis
Washington University Open Scholarship

Olin Business School Electronic Theses and Dissertations | Washington University Open Scholarship

Summer 5-15-2023

Supply Chain Management under Technology Innovation and Pandemic

Xiaoyu Wang
Washington University in St. Louis, xiaoyuwang@wustl.edu

Follow this and additional works at: https://openscholarship.wustl.edu/olin_etds

Part of the Business Administration, Management, and Operations Commons

Recommended Citation

Wang, Xiaoyu, "Supply Chain Management under Technology Innovation and Pandemic" (2023). *Olin Business School Electronic Theses and Dissertations*. 31.
https://openscholarship.wustl.edu/olin_etds/31

This Dissertation is brought to you for free and open access by the Washington University Open Scholarship at Washington University Open Scholarship. It has been accepted for inclusion in Olin Business School Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Olin Business School

Dissertation Examination Committee:
Fuqiang Zhang, Chair
Lingxiu Dong
Xiang Hui
Baojun Jiang
Fasheng Xu

Supply Chain Management under Technology Innovation and Pandemic

by

Xiaoyu Wang

A dissertation presented to
Olin Business School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2023
St. Louis, Missouri

Table of Contents

Page

List of Figures

Figure Page

Figure Page

List of Tables

Acknowledgments

First, I am deeply indebted to my adviser Dr. Fuqiang Zhang for his tremendous support and guidance throughout my Ph.D. study, which have helped me greatly in completing my doctoral study and becoming a qualified scholar. I would never be able to fully repay his effort and time spent on me. I also learn from Fuqiang how to become a person with integrity. I am deeply touched by his kindness, encouragement, and empathy. It is an honor to be his Ph.D. student, and I hope one day to become such a great mentor to others as well.

I would also like to express my deepest gratitude to my coauthor Dr. Xiang Hui and Dr. Fasheng Xu, who have spent innumerable time and effort on our papers and tirelessly supported me at different stages of my Ph.D. study. It is impossible to imagine what I can achieve without their wisdom and most generous help. Both of them have the most beautiful heart in the world, and it is the best memory in my life to work with them.

Moreover, I want to thank Dr. Lingxiu Dong and Dr. Baojun Jiang for serving as my committee members and providing insightful feedback and comments on improving my dissertation.

My heartfelt appreciation to my master's mentor Dr. Bin Hu, who opened the gate of academia and led me to the world of operations management. Without him, my life will be totally different, missing out on the beautiful research world.

Also, I would like to thank our fantastic staff members Kristy Chan, Jessica Hatch, and Shengxin Sun for their extraordinary administrative support through my five years of the doctoral program.

My gratitude also goes to my best WashU friends Dr. Puping Jiang and Dr. Mingyang Cui for all the hardship and joyful moments we had together, and to my best non-WashU friends Runhao Wang, Dr. Cheng Shu, Dr. Jiaxuan Yan, Hao Luo, and Siyu An for always being there to support me and cheer me up. I owe them too much.

Most important, I am extremely grateful to my parents Luyin Wang, Guoying Du, and my girlfriend, Dr. Lei Ye for their unconditional love and support, which give me the courage to overcome all difficulties and chase for my dreams. I could not imagine getting through this journey without their love and support. My forever miss goes to my grandpa Fangzhou Du, who watches over me from heaven with love and pride. Words cannot express my gratitude to my family. To them I dedicate this thesis.

Xiaoyu Wang

Washington University in St. Louis

May 2023

Dedicated to my family.

ABSTRACT OF THE DISSERTATION

Supply Chain Management under Technology Innovation and Pandemic

by

Xiaoyu Wang

Doctor of Philosophy in Business Administration

Washington University in St. Louis, 2023

Professor Fuqiang Zhang, Chair

The main purpose of this dissertation is to study supply chain issues under the challenge of technology innovation and pandemic; and to identify the implications for individuals and businesses.

In Chapter 1, "Consumer Privacy in Online Retail Supply Chains", we study the implications of newly adopted privacy policies such as the GDPR (General Data Protection Regulation) for online retail supply chains consisting of a retailer and a supplier. Exploitation of consumer data allows online retailers to enhance services provided to consumers, but at the risk of causing unintended privacy issues. There has been debate about whether to devise regulation policies to restrict data collection and usage by online retailers. We find that, although the GDPR is designed to protect consumer privacy, it may actually hurt consumer surplus while benefiting the retailer. In fact, the GDPR may even lead to a triple-lose situation for the retailer, supplier, and consumers. We further explore two coordinated supply chain arrangements, i.e., agency selling and vertical integration. We show that the GDPR always enhances the social welfare under these two arrangements, but it may still decrease the consumer surplus. Our results have significant implications for consumers, supply chain firms, and policymakers, and contribute to the literature evaluating the impact of privacy regulation on technology innovation and adoption.

In Chapter 2, "The Value of Smart Contract in Trade Finance", we investigate how smart contract adoption could facilitate trade finance activities and create value for supply chain firms. As the emerging blockchain technology could potentially reshape the trade financ-

ing landscape, understanding the impact of smart contract adoption and its interaction with trade finance activities is practically relevant and of great importance. We develop a two-stage game-theoretic model and adopt supply chain finance theory to characterize the strategic interactions between supply chain firms in the presence of both operational risk (demand uncertainty) and financial risks (credit and liquidity risks). We find that the value of smart contract depends critically on the trade finance structures, including both pre-shipment and post-shipment financing schemes. Under the baseline trade finance model (with purchase order financing as pre-shipment financing and factoring as post-shipment financing), smart contract alleviates the supplier's overpricing behavior caused by commitment frictions and helps restore the supply chain efficiency. When buyer direct financing serves as an alternative pre-shipment financing, smart contract might discourage the retailer from offering buyer direct financing, which significantly hurts the supplier and thus reduces the supply chain profit. When invoice trading serves as the alternative post-shipment financing, the supplier always chooses invoice trading over factoring due to its trading flexibility which, in turn, makes the commitment frictions ubiquitous and unresolvable (namely, *commitment trap*). As a result, invoice trading could unexpectedly lead to a lower supplier's profit. Luckily, such an adoption dilemma can be resolved by smart contract adoption in conjunction with factoring. Our findings provide guidelines for and insights into when smart contract should be adopted and its interactions with different trade finance schemes. In particular, smart contract adoption does not always benefit the supply chain.

In Chapter 3, "Impact of COVID-19 on Online Share of Expenditure and the Mediating Role of Digital Infrastructure: Evidence from a Two-year Consumer Panel", we document changes in consumption behaviors after the COVID-19 outbreak in 2020. Our unique dataset from the largest digital payment platform in China allows us to track online and offline consumption for a given consumer over two years. The identification of the COVID impact on consumption is based on two strategies: a year-on-year comparison and a comparison among cities with different numbers of COVID cases. We find that the pandemic disproportionately reduced online and offline consumption, causing a higher online share of consumption during and one year into the pandemic. This result suggests that the pandemic may have a long-lasting impact on consumption structure. Our second main finding is that consumers who live in cities with better digital infrastructure experienced a smaller reduction in online and offline consumption during the pandemic. This result suggests that digital infrastructure leads to consumption resilience against macroeconomic shocks, and that the impact of digital infrastructure on consumption goes beyond the digital economy. We discuss the policy and managerial implications of these findings.

Chapter 1: Consumer Privacy in Online Retail Supply Chains

1.1 Introduction

[1]Online retail expands fast and reshapes the global economy. A Deloitte report shows that online retail represents 40% of total retail spending in the U.S. in 2020.[2] Online retailers substantially utilize web cookies to access browsers and provide services that alleviate hassles which would otherwise impair the online shopping experience in daily operations.[3] Web cookies enable companies to collect data to identify and remember specific consumers in order to provide website features such as shopping cart, location sharing, and saving preference.[4] We call this *upside data exploitation*, which benefits consumers. For example, using web cookies, Costco's website is able to record the consumer's location information, which automatically helps find the nearest store for pickup services. With such website features, consumers avoid the time and effort spent on store searching, leading to improved convenience in the online shopping experience.

However, there is a major privacy concern about the data collection (cookie utilization). Web cookies allow companies to track consumers and sell their data to third parties like data brokers for additional profits, which hurts consumer rights. We call this *downside data exploitation*. Over two-thirds of consumers think retailers use their data for marketing reasons, and 55% of consumers believe that retailers share data with third parties or sell data to outside buyers.[5] High-profile data breaches and privacy infraction incidents have made consumers more aware that their personal information is at risk. Deloitte's

[1]This chapter is based on the author's early work [1] jointly with Fuqiang Zhang and Fasheng Xu.
[2]https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consumer-business/us-2021-retail-industry-outlook.pdf
[3]https://www.businessnewsdaily.com/10625-businesses-collecting-data.html
[4]https://www.hostpapa.com/blog/marketing/what-internet-cookies-are-and-how-they-can-help-your-business/
[5]https://www2.deloitte.com/us/en/pages/about-deloitte/articles/press-releases/deloitte-consumer-privacy-in-retail-survey-the-next-regulatory-and-competitive-frontier.html

U.S. Consumer Data Privacy Survey shows that one in every three Americans has been exposed to data compromise. Nearly half of consumers (47%) feel they have little control over their personal data, and the majority of consumers (86%) believe they should be able to opt out of the sale of their data.[6]

Consumer data is valuable, but the ownership of such assets is not well defined under traditional data policy. To handle this issue, various measures have been taken to protect consumers' personal data. A notable suggestion is that consumers should be given control over their data, so they can decide what data to share with companies and how companies can exploit the shared data. This gives rise to the European Union's General Data Protection Regulation (GDPR).[7] The GDPR policy has opted for personal data protection rights, including prohibition of data access by companies and restrictions on secondary use of personal data without consumer consent. Figure 1.1 gives an example of a website cookie declaration after the GDPR. In this example, consumers can manage cookie settings and choose whether to allow LEGO to use data for upside exploitation such as necessary functions, or for downside exploitation such as data sharing with third-party partners.

Note that even under traditional data policy, consumers can still decide whether to allow data access by companies. For example, many web browsers enable consumers to block tracking by websites if they do not want to give data access to companies. However, without the GDPR, once data consent is given (web cookies are allowed), consumers are unable to control how companies use the data. Specifically, companies use data for necessary functions like saving the shopping cart (the upside exploitation). Meanwhile, companies do not mention they might sell data to third parties (the downside exploitation), which is not controlled by consumers.[8] As a quick summary, contrary to the "take it or leave it" data-disclosing choice for consumers under traditional data policy, the GDPR identifies

[6] https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consumer-business/us-retail-privacy-survey
-2019.pdf

[7] The GDPR is the cornerstone of European privacy law and is considered the most comprehensive, globally leading privacy regime. It establishes common rules on data processing throughout the EU and is directly binding for companies and residents in the EU and beyond [2].

[8] https://www.businessnewsdaily.com/10625-businesses-collecting-data.html

and categorizes the purpose of data exploitation (i.e., upside and downside exploitation), and enables consumers to decide which type of data exploitation is allowed.

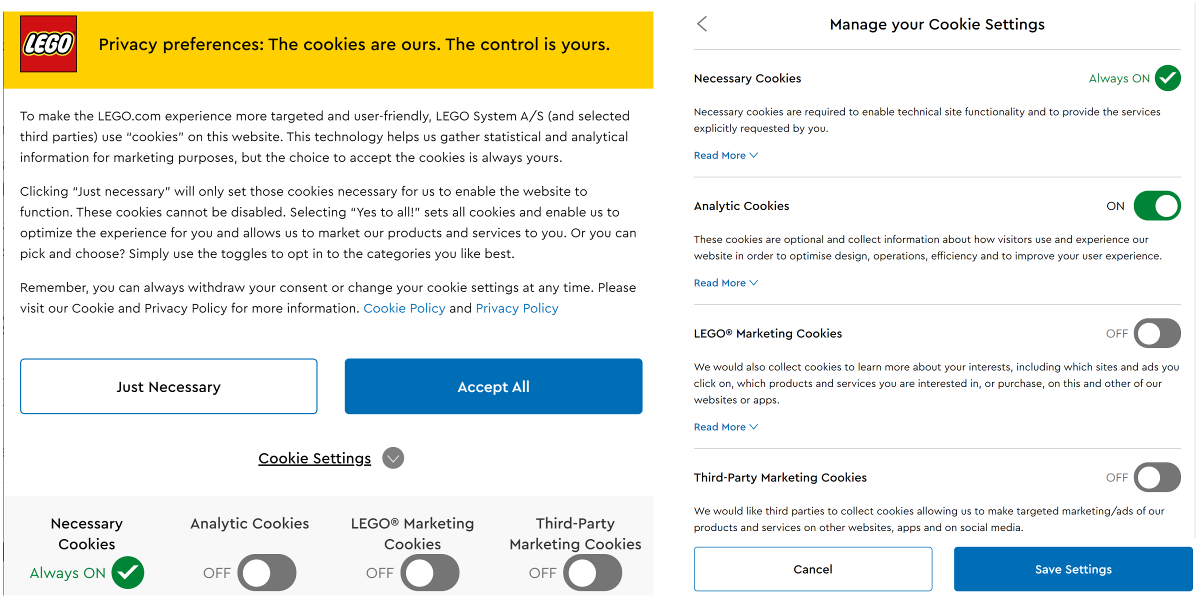


Figure 1.1.: An example of consent notifications after GDPR

The GDPR is widely adopted in the retail sector. Established retailers such as Costco, Dillard's, and Marshall all comply with the GDPR. A survey shows that 78% of U.S. firms are compliant to some extent.[9] Motivated by the above observations, we study the impact of privacy policies like the GDPR on online retail supply chains. We examine the welfare implications of the GDPR policy for each party in the supply chain (in particular, whether the GDPR can protect privacy and benefit consumers), and our analysis yields the following main results.

First, we find that although initiated to protect consumers, the GDPR may hurt the consumer surplus. As previously mentioned, consumers enjoy the upside data exploitation, but dislike the downside data exploitation. The GDPR enables consumers to decide how their data is exploited, and thus they would only allow the upside exploitation but block the downside exploitation. As a result, the consumers who originally block data collection (i.e., privacy-sensitive consumers) may disclose data and benefit from the upside exploitation. The privacy-insensitive consumers, who originally disclose data under the traditional data policy, are able to avoid the downside exploitation cost. Clearly, the

[9] https://legaljobs.io/blog/gdpr-statistics/

GDPR improves the consumer surplus with regard to data exploitation. However, such a positive *privacy effect* from data exploitation may be dominated by a negative *retail price effect*, which can be explained as follows. The GDPR expands market demand by encouraging data disclosure and product purchasing. Thus, the retailer is able to charge a higher retail price to increase the profit margin. When the consumer valuation is low, market demand is sensitive to the retail price, which leaves the retailer little pricing flexibility under the GDPR. Hence, the retail price effect is less significant compared to the privacy effect and the GDPR benefits consumers. However, when the consumer valuation is sufficiently high, the product becomes very "popular" and the market demand is highly insensitive to the price. That is, consumers can tolerate a significant price escalation in exchange for the utility increase generated by the GDPR. Hence, the GDPR enables the retailer to substantially raise the retail price, and thus the pricing effect dominates the privacy effect, leading to an overall decrease in the consumer surplus.

Second, our analysis shows that even though the GDPR imposes more constraints on the retailer (i.e., the retailer cannot decide how consumer data is exploited anymore), the retailer could be better off with the GDPR, especially when the consumer valuation is low. The retailer's profit consists of two parts: the product reselling profit and the downside data exploitation profit. It is clear that the GDPR deprives the retailer of the downside exploitation profit. However, this profit loss may be outweighed by the increased product reselling profit. To explain, we look into the two factors that affect the product reselling profit: (i) Market size: consumers are free of the downside exploitation cost under the GDPR, and thus they are more willing to disclose data and purchase, leading to a market expansion; (ii) Profit margin: under the GDPR, the retail price decision and the supplier's wholesale price decision affect the profit margin in opposite directions. On one hand, the GDPR eliminates the downside exploitation benefit, which demotivates the retailer to reduce the retail price (so as to persuade more consumers to disclose data and gain higher downside exploitation profit). Hence, the GDPR leads to the profit margin increase. On the other hand, without the downside exploitation cost, the product becomes more "popular" and consumers are more price insensitive, which encourages the supplier to "squeeze" the retailer harder by charging a higher wholesale price, leading to the decrease of the profit margin. When the consumer valuation is low (i.e., the market is price sensitive), the upstream supplier cannot squeeze the retailer too hard (so as to maintain the order quantity), and thus the supplier's wholesale price effect is dominated by the retailer's retail price effect, leading to an overall increase of the retailer's profit margin. Therefore, both the market expansion and the margin increase contribute to the increase of the product reselling profit, which may dominate the loss of the data exploitation profit. As a result, the GDPR policy could actually benefit the retailer. When the consumer valuation is sufficiently high, on the contrary, the supplier's wholesale price effect dominates, leading to a decrease of the profit margin. In addition, the market is already highly penetrated due to the high consumer valuation, and the GDPR cannot help further expand the market size. Hence, the product reselling profit is reduced and the GDPR would make the retailer worse off.

Third, we show that the GDPR policy may either benefit or hurt the supplier's profit. Note that the supplier only gains from the wholesale revenue. Given that the retailer collects data, the GDPR helps expand the market size, which allows the supplier to increase the wholesale price and make a higher profit. However, the retailer has less incentive to collect consumer data under the GDPR, which hurts the supplier's profit if the retailer switches her decision from data collection to no collection. This can be explained by comparing the retailer's data-collection decisions with and without the GDPR. Note that without the GDPR, collecting data benefits the retailer in two ways: First, it generates the downside exploitation profit for the retailer; second, it expands the market demand. Meanwhile, collecting data has a negative effect on the retailer when the consumer valuation is sufficiently high, since the market demand becomes less sensitive to price and the upstream supplier can increase the wholesale price to squeeze the retailer's profit. As a result, the retailer's data-collection decision hinges upon the trade-off between the above positive and negative effects. In contrast, in the presence of the GDPR, the downside data exploitation is disabled by consumers, which reduces the benefits of data collection for the retailer. At the same time, market demand becomes even less sensitive to price due to the absence of the downside exploitation cost for consumers, encouraging the upstream supplier to further increase the wholesale price and exacerbating the negative effect. Thus the retailer is less likely to collect consumer data under the GDPR. In particular, When introducing GDPR induces the retailer to forego data collection, the market demand shrinks and the wholesale price decreases. In this case, the supplier would be worse off under the GDPR.

Fourth, we find that the GDPR policy could even lead to a triple-lose situation (i.e., all three parties in our model are worse off). As mentioned earlier, data collection has a two-fold impact on both the retailer profit and the consumer surplus: on one hand, it leads to market expansion and benefits the retailer; on the other hand, the market expansion induces the supplier to increase the wholesale price, yielding a higher cost for the retailer and consumers. The introduction of GDPR affects both driving forces and, in particular, has a stronger impact on the wholesale pricing effect when the consumer valuation is high. Thus the GDPR may hurt the retailer profit and consumer surplus under high consumer valuations. Furthermore, when the consumer valuation is sufficiently high, the GDPR may induce the retailer to switch from collecting data to not collecting data, and thus hurt the supplier as well. Therefore, we show that the GDPR may lead to a triple-lose outcome under sufficiently high consumer valuations.

Finally, we consider two model variants that provide better supply chain coordination: (i) the agency selling model, where the supplier sells to consumers directly and the retailer only charges a pre-determined commission fee;[10] (ii) vertical integration, where the supplier and the retailer are integrated into one firm. Our analysis suggests that, under the agency selling (resp., vertical integration), the retailer (resp., integrated firm) always collects data with the GDPR policy. The GDPR still may hurt the consumer surplus.

The rest of the chapter is organized as follows. Section 1.2 reviews the related literature. Section 1.3 introduces the model setting. In Sections 1.4 and 1.5, we investigate the scenarios with and without the GDPR policy, respectively, and study the value of the GDPR to the supplier, retailer as well as consumers. Section 1.6 further discusses how

[10] In our baseline model, we focus on the reselling model, where the supplier sells to the retailer, who resells to the consumers afterwards.

the GDPR policy performs with supply chain coordination. We conclude the chapter in Section 1.7. All proofs are given in the appendix.

1.2 Literature Review

Our work is related to the literature on the economics of privacy, which focuses on the firm's behavior-based price discrimination and the consumer's endogenous privacy choice [3–7].[11] [9] study the case in which the monopoly seller could ask for ex ante registration before purchasing and showed that this behavior could enhance the seller's profit. In [10] and [11], consumers can choose to disclose their private information to a monopolist in return for reduced search cost due to more accurate product recommendation or targeted advertising. A similar trade-off arises in [12], who study a retailer's choice of disclosing consumers' preferences to advertisers. In [13], consumers can take costly actions to protect their identities and make the profiling technology less effective. [14] studies a setting where consumers may disclose personal information to a firm that may use it both for product recommendations and price discrimination. There are also papers studying consumer data privacy with competing retailers [15, 16]. All the aforementioned papers analyze settings that feature consumer data-based price discrimination. In contrast, our model is designed to study the interplay between the upside data exploitation (i.e., reducing consumer inconvenience cost) and the downside data exploitation (i.e., data monetization via data selling) without allowing for price discrimination.

The second strand of related literature is about (pre-GDPR) data regulation, which focuses on whether the usage of firms' collected data should be regulated [17, 18]. [19] investigates whether firms should be allowed to sell consumer data and the corresponding economic effect. It has been found that data regulation might not be useful if consumers are sophisticated enough. Empirical papers also discuss the welfare implication of privacy regulation. [20] study how privacy regulation would affect the effectiveness of online advertisement. They show that regulation would indeed reduce the efficiency of online advertisement and harm the economic output. [21] find that offering consumers the right

[11] We recommend that readers refer to [8] for a comprehensive literature review.

of voluntarily profiling their data could hurt the consumers than no profiling. [22] study the commonly used consent-based regulation method and investigate the data regulation under the multi-firm competition case. They show that the consent-based approach would benefit the firm but hurt the consumers. The existing literature on data regulation assumes that once accessed by the firm, data can be used for profiting and enhancing the consumer experience at the same time. Our chapter contributes to this literature by proposing an analytical model where we separate using data from accessing data. Under the GDPR, consumers could decide which purpose to allow the firm to use their data for; in other words, having access to consumer data does not mean the firm could use data for any purpose. For example, consumers might allow the firm to use data to enhance the service while forbidding the firm from profiting from consumer data.

More recent work investigates the impact of privacy regulation on revenue management and the operations of platforms or service systems [23–26]. [27] study the effectiveness of privacy regulation (e.g., requiring customers' consent or prohibiting data tracking) in scenarios in which customers interact with firms sequentially and the customers' data can be shared among these firms (i.e., data linkages). [28] find that due to competition, the absence of data tracking may lead to a decrease in consumer surplus, even when consumers are myopic. [29] investigate whether the platform, or the users, should have the right to decide which data the platform commercializes. [30] study how firms' revenue model affects their optimal data strategy regarding collection and protection of users' data. They find a business with a more data-driven (source of revenue: selling data or data-based services to third parties) revenue model will collect more users' data and provide more data protection than a similar business that is more usage-driven (source of revenue: collect users' payments in the form of subscription fees). We contribute to this literature stream by exploring the impact of privacy regulation in a new business environment: online retail supply chains, where the strategic interactions between the supplier and the retailer (in addition to the strategic consumer behaviors) can result in considerably different insights.

We also add to an emerging literature on the economic implications of the GDPR, a new data regulation. Some empirical studies have reported that the GDPR had a negative impact on marketing effectiveness and firm profitability. [31] show that GDPR-induced privacy rights reduce consumer data opt-in and mostly increase consumer surplus, but they have a negative impact on firm profits, especially in competitive markets. [32] show that the GDPR increased the businesses' cost of collecting consumer data, and [33] find negative post-GDPR effects on investment in new and emerging technology firms in Europe. [2] document that although all firms suffer losses, the largest vendor, Google, loses relatively less and significantly increases market share in important markets such as advertising and analytics. However, some studies find a neutral or even positive effect of the GDPR. [34] show that the GDPR roll-out has increased consumers' data opt-in, which led to more effective targeted advertising and an increase in sales. [35] show that despite the increase in opt-outs as a result of the new GDPR requirement, the ability of firms to predict consumer behavior did not significantly worsen, because the average value of the remaining opt-in consumers to advertisers has increased. Our manuscript extends this literature and complements the emerging body of work on GDPR in the following ways. We develop a full-fledged economic model that endogenizes the strategic interactions among three important parties (retailer, supplier, and consumers) under different data policies or privacy regulations. Our model offers an analytical framework to evaluate the policy impact on the operations and profitability of online retailing supply chains, as well as consumer welfare. Notably, we find that the GDPR may lead to a triple-lose situation for the retailer, supplier, and consumers.

1.3 Model Setting

We develop a game-theoretic model for an online retail supply chain consisting of two parties: a supplier (he) and an online retailer (she). The supplier determines the wholesale price w and the retailer purchases goods from the supplier. Then the retailer decides the retail price p and resells to consumers. Without loss of generality, we set the supplier's production cost to zero, normalize the total number of consumers to one, and assume each consumer requires at most one unit of the product.[12] Consumer valuation of the product is v. Online shopping incurs a heterogeneous inconvenience cost θ for each consumer, where θ follows a uniform distribution on $[0,1]$. The inconvenience cost could refer to unsatisfactory user experience due to the lack of website functions (e.g., the function of saving the shopping cart for future visits and the function of location sharing). A consumer with inconvenience cost θ would purchase only if it can generate a non-negative utility: $v - p - \theta \geq 0$.

The online retailer is able to access, collect, and exploit consumer data from those who use the retailer's e-commerce service. The quality of consumer data is denoted δ, where $\delta \in (0,1)$ is exogenously given. The value of δ represents how effectively the collected consumer data can be used for exploitation. Two conditions must be satisfied for the retailer to successfully collect consumer data. First, the retailer is willing to collect data from consumers visiting the website. We use $\ell \in \{0,1\}$ to denote the retailer's binary decision on data collection. Second, some consumers are willing to give consent to data collection if the retailer chooses to collect data (i.e., $\ell = 1$). Let $\alpha \in \{0,1\}$ denote the consumer's data disclosure choice, where $\alpha = 1$ means the consumer gives consent to data collection. In summary, a consumer's data would be successfully collected only if the retailer collects data and the consumer gives consent at the same time (i.e., $\alpha\ell = 1$).

Data exploitation by the retailer can be a double-edged sword for consumers. On one hand, data can be used to reduce the consumer inconvenience cost from θ to $(1-\delta)\theta$ (the *upside* data exploitation). On the other hand, the retailer can monetize consumer data, which benefits the retailer but hurts the consumer surplus (the *downside* data exploitation). We assume the profit from data exploitation is $\lambda r \delta$ for the retailer, where $\lambda \in [0,1]$ is the probability that the consumer data is valuable[13], and $r \geq 0$ represents the retailer's data monetization intensity. The retailer has to incur a cost cr^2 with $c > 0$ to achieve an intensity level r.[14] Data monetization intensity shows the retailer's efforts

[12] The model can be extended to non-zero production costs without affecting the main results.

[13] https://techmonitor.ai/techonology/data/85-of-data-is-useless-to-business-and-is-creating-a-33-trillion-drain-on
-resources-4840818

[14] Data monetization cost involves software license and labor cost, as well as the cost of computing power. See https://www.itsasap.com/blog/data-analytics-cost

to extract profitable information such as contact information, phone number, etc., which increases the downside exploitation profit (data selling or sharing with third parties), but is irrelevant to the upside exploitation (inconvenience cost reduction). As a benchmark, we assume that the consumers' disutility from downside data exploitation is also $\lambda r\delta$. Thus, $\lambda r\delta$ represents the data value that originally belongs to the consumer, but is captured by the retailer. In particular, such a downside data exploitation will neither increase nor decrease the total surplus of the two parties. It is simply a profit/utility transfer from the consumer to the retailer. Let $\mathbb{1}_{\{\mathrm{ue}=1\}}$ (resp., $\mathbb{1}_{\{\mathrm{de}=1\}}$) be an indicator function, where ue $= 1$ (resp., de $= 1$) represents the adoption of upside exploitation (resp., downside exploitation). Then consumer's utility can be written as

$$v - p - \theta + \mathbb{1}_{\{\mathrm{ue}=1\}}\delta\theta - \mathbb{1}_{\{\mathrm{de}=1\}}\lambda r\delta, \tag{1.1}$$

where $v - p - \theta$ represents consumer's base utility, and $\mathbb{1}_{\{\mathrm{ue}=1\}}\delta\theta$ (resp., $\mathbb{1}_{\{\mathrm{de}=1\}}\lambda r\delta$) represents the benefit of upside exploitation (resp., cost of downside exploitation).

Regarding the data control right, this chapter focuses on two different data privacy policies: (i) *retailer-controlled data policy* (referred to as the $\mathbb{R}$ mode), in which the retailer controls the consumer data and decides how to exploit it once data is collected; (ii) *consumer-controlled data policy* (referred to as the $\mathbb{C}$ mode), in which the consumers control their own data and decide how the retailer can exploit the data. In this chapter, the $\mathbb{C}$ mode reflects the newly-adopted GDPR policy, in which consumers, instead of the retailer, make data exploitation decisions (see Appendix A.2 for more detailed discussions of GDPR). We assume that the retailer cannot discriminate between consumers (offer different retail prices) based on their data-disclosing decisions.[15]

The distinction between the $\mathbb{R}$ mode and the $\mathbb{C}$ mode is the adoption of the downside data exploitation. Under the $\mathbb{R}$ mode, consumers cannot observe the data monetization intensity chosen by the retailer.[16] In addition, the retailer is unable to commit to not adopting downside exploitation due to the lack of transparency and supervision. Once

[15]Price discrimination is illegal in the U.S. See https://definitions.uslegal.com/p/price-discrimination/

[16]For example, users are unaware until Facebook is caught sharing consumer private data with third parties. https://www.bbc.com/news/technology-46618582

data is collected, the retailer always adopts the downside data exploitation and determines the optimal data monetization intensity for profit maximization. Therefore, if data is collected under the $\mathbb{R}$ mode, both the upside and downside data exploitation will happen. The consumer utility can be written as $v - p - (1 - \alpha\ell\delta)\theta - \alpha\ell\lambda r\delta$, where $\alpha\ell = 1$ represents data being collected and $\alpha\ell = 0$ represents data not being collected. Under the $\mathbb{C}$ mode, consumers control the data and decide whether the retailer can use data for the upside/downside exploitation. As a result, if data is collected under the $\mathbb{C}$ mode, consumers only allow the upside exploitation while prohibiting the downside data exploitation. Hence, the consumer utility becomes $v - p - (1 - \alpha\ell\delta)\theta$.

The sequence of events is summarized in Figure 1.2. First, the retailer makes the (long-term) data-collection decision $\ell \in \{0, 1\}$. Second, the supplier sets the wholesale price w. Third, the retailer decides the retail price p. Then, consumers make the purchasing and data-disclosure decisions. Under the $\mathbb{R}$ mode, the retailer uses the collected data (if $\alpha\ell = 1$) for both the upside and the downside exploitation, and decides the data monetization intensity r. Under the $\mathbb{C}$ mode, consumers only allow the upside exploitation while blocking the downside exploitation, and thus the retailer can only use the collected data for the upside exploitation if $\alpha\ell = 1$.

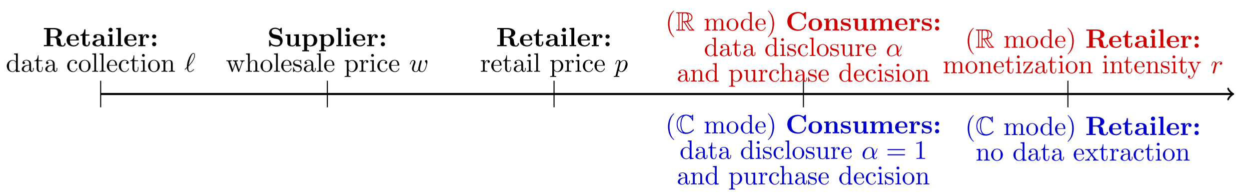


Figure 1.2.: Sequence of events

1.4 Equilibrium Analysis

In this section, we first derive the equilibrium under the retailer-controlled data policy ($\mathbb{R}$ mode), and then analyze the model under the consumer-controlled data policy ($\mathbb{C}$ mode).

1.4.1 Retailer-Controlled Data Policy ($\mathbb{R}$ Mode)

Under the $\mathbb{R}$ mode, we only present the analysis where the retailer is willing to collect consumer data ($\ell = 1$) for concision. The analysis for no data collection ($\ell = 0$) is similar and relegated to the online appendix (proof of Proposition 1.4.1). Following the backward induction approach, we start with the retailer's data monetization intensity decision r, given the wholesale price w, the retail price p, and the number of consumers disclosing data s (i.e., the quantity of collected consumer data). The retailer chooses the monetization intensity r to maximize the downside exploitation profit $\lambda r\delta s - cr^2$. Recall that $\lambda r\delta$ is the downside exploitation value extracted from each consumer who has disclosed data. Thus $\lambda r\delta s$ is the total downside exploitation profit and cr^2 is the corresponding exploitation cost. We derive the retailer's optimal monetization intensity as $r(s) = \frac{\lambda\delta}{2c}s$, which increases in the number of consumers who have disclosed data s.

We then turn to the consumer's data-disclosing and purchasing decisions. Note that in our setting, a consumer's upside exploitation benefit is contingent on the purchasing decision. For example, location sharing saves consumer time finding the nearest store for order pickup, and such a benefit only exists with purchasing. Therefore, consumers choose not to disclose data if they do not purchase, and make the data-disclosing decision α only when purchasing. If the consumer with specific θ purchases, the utility is given by:[17]

$$U_1(\alpha, p, s, \theta) = (v - p) - (1 - \alpha\delta)\theta - \alpha\lambda r(s)\delta.$$

Note that s is the consumer's belief of the data-disclosing quantity and thus $r(s)$ is the belief of the retailer's data monetization intensity. The first term $(v - p)$ represents the gross benefit from purchase. The second term $(1-\alpha\delta)\theta$ represents the inconvenience cost. The third term $\alpha\lambda r(s)\delta$ represents consumers downside exploitation cost. The consumer has three choices: (i) Purchase and disclose data ($\alpha = 1$), which generates the consumer utility $v - p - (1 - \delta)\theta - \lambda r(s)\delta$; (ii) purchase and hide data ($\alpha = 0$), which generates the

[17]Remember here we only present the case where the retailer collects data ($\ell = 1$). If the retailer's data-collection decision is considered, consumer utility becomes $U_1(\alpha, \ell, p, s, \theta) = (v-p)-(1-\alpha\ell\delta)\theta-\alpha\ell\lambda r(s)\delta$.

consumer utility $v - p - \theta$; (iii) no purchase, which leads to zero consumer utility. Each consumer chooses the one with the highest utility.

Table 1.1: Conditions for consumers' purchasing and data-disclosure decisions

(purchase, disclose)	$v - p - (1-\delta)\theta - \lambda r(s)\delta \geq v - p - \theta$	$v - p - (1-\delta)\theta - \lambda r(s)\delta \geq 0$
(purchase, hide)	$v - p - \theta > v - p - (1-\delta)\theta - \lambda r(s)\delta$	$v - p - \theta \geq 0$
(no purchase, hide)	$v - p - (1-\delta)\theta - \lambda r(s)\delta < 0$	$v - p - \theta < 0$

The results are summarized in Table 1.1. Taking the example of the first row, the consumer purchases and discloses data if disclosing data dominates not disclosing data $v-p-(1-\delta)\theta-\lambda r(s)\delta \geq v-p-\theta$, and purchasing generates positive utility $v-p-(1-\delta)\theta-\lambda r(s)\delta \geq 0$. From these two conditions, we derive that consumers with $\theta \in [\lambda r(s), \mathcal{T}(s,p)]$ will purchase and disclose data, where $\mathcal{T}(s,p) = \min\left\{\frac{v-p-\lambda r(s)\delta}{1-\delta}, 1\right\}$. We are also able to show that consumers with $\theta > \mathcal{T}(s,p)$ neither purchase nor disclose data, whereas consumers with $\theta < \lambda r(s)$ purchase the goods but hide their data. Combining the above analysis, Figure 1.3 illustrates consumer's purchasing and data-disclosing decisions with respect to the heterogeneous inconvenience cost θ. From Figure 1.3, it is also clear that $\mathcal{T}(s,p)$ represents the sales quantity, i.e., the number of consumers who would purchase.

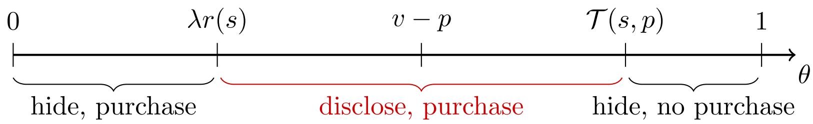


Figure 1.3.: An illustrative example of consumer data-disclosure and purchase decisions ($\mathbb{R}$ Mode)

Given consumers' belief of data-disclosing quantity s, consumers are able to anticipate the optimal intensity $r(s)$ the retailer sets. Recall that under the belief s, only consumers with $\theta \in [\lambda r(s), \mathcal{T}(s,p)]$ will purchase and disclose data. Hence, under belief s, the realized data-disclosing quantity is $\mathcal{T}(s,p) - \lambda r(s)$. A rational expectations equilibrium can be derived [36–38]. Specifically, the consumers form a common belief that the data-disclosing quantity is s, and if each consumer makes the purchasing and data-disclosing decisions based on this belief, then the actually data-disclosing quantity is s. Mathematically,

$\mathcal{T}(s,p) - \lambda r(s) = s$. From this equation, s can be derived, which is a function of p. Throughout the chapter, we denote $\mathcal{Z} = 2c + \lambda^2\delta$ for concision. Replacing s, we can further simplify the data monetization intensity $r(p)$ and the sales quantity $\mathcal{T}(p)$ as follows:

$$r(p) = \frac{\lambda\delta(v-p)}{2c(1-\delta)+\lambda^2\delta}, \qquad \mathcal{T}(p) = \min\left\{\frac{(v-p)\mathcal{Z}}{2c(1-\delta)+\lambda^2\delta}, 1\right\}. \tag{1.2}$$

We characterize the consumer's optimal purchasing and data-disclosing decisions in Lemma 1.

Lemma 1 *Given the retail price p, consumers with $\theta \in [0, \lambda r(p))$ will purchase the product but hide data; consumers with $\theta \in [\lambda r(p), \mathcal{T}(p)]$ will purchase the product and disclose data; consumers with $\theta \in (\mathcal{T}(p), 1]$ neither purchase nor disclose data.*

Next, we consider the retailer's pricing decision p. Given the wholesale price w and the retail price p, the retailer profit is:

$$\pi_r(w,p) = (p-w)\mathcal{T}(p) + \lambda r(p)\delta\left[\mathcal{T}(p) - \lambda r(p)\right] - cr(p)^2. \tag{1.3}$$

The first term $(p-w)\mathcal{T}(p)$ is the retailer's reselling (product selling) profit, where $\mathcal{T}(p)$ represents the aggregate sales quantity and $(p-w)$ represents the profit margin. The second term $\lambda r(p)\delta\left[\mathcal{T}(p) - \lambda r(p)\right]$ is the downside data exploitation profit, in which $\mathcal{T}(p) - \lambda r(p)$ represents the data-disclosing quantity. The last term $cr(p)^2$ is the quadratic data monetization cost with the corresponding optimal intensity $r(p)$. For expositional convenience, we denote $\mathcal{M} = 4c^2(1-\delta) + c\lambda^2\delta(4-3\delta) + \delta^2\lambda^4$ throughout the rest of the chapter. Based on the profit function $\pi_r(w,p)$, we can derive the retailer's optimal pricing decision $p(w) = \frac{v+w}{2} - \frac{c\lambda^2\delta^2(v-w)}{2\mathcal{M}}$, which is a function of the wholesale price w. We can write the upstream supplier's profit function as:

$$\pi_s(w) = w\mathcal{T}(p(w)).$$

Plugging $p(w) = \frac{v+w}{2} - \frac{c\lambda^2\delta^2(v-w)}{2\mathcal{M}}$ into $\mathcal{T}(p)$ defined in equation (1.2), we can derive $\mathcal{T}(p(w)) = \min\{1, \frac{(v-w)\mathcal{Z}^2}{2\mathcal{M}}\}$, which depends on the wholesale price w. Recall that $\mathcal{T}(p(w))$

represents the aggregate sales quantity. Finally, the retailer makes the data-collection decision ℓ to maximize profit. We summarize the equilibrium in the following proposition.

Proposition 1.4.1 *Under the $\mathbb{R}$ mode, the retailer's data monetization intensity r^*, retail price p^*, the supplier's wholesale price w^*, the sales quantity $\mathcal{T}(p^*)$, and the retailer's data-collection decision ℓ^* are given as follows:*

(i) given that the retailer collects consumer data, we have

$$(r^*, p^*, w^*, \mathcal{T}(p^*)) = \begin{cases} \left(\dfrac{v\lambda\delta\mathcal{Z}}{4\mathcal{M}}, 3v/4 - \dfrac{vc\lambda^2\delta^2}{4\mathcal{M}}, v/2, \dfrac{v\mathcal{Z}^2}{4\mathcal{M}}\right) & \text{if } \ v \leq \dfrac{4\mathcal{M}}{\mathcal{Z}^2}, \\ \left(\dfrac{\lambda\delta}{\mathcal{Z}}, v - 1 + \dfrac{2c\delta}{\mathcal{Z}}, v - \dfrac{2\mathcal{M}}{\mathcal{Z}^2}, 1\right) & \text{if } \ v > \dfrac{4\mathcal{M}}{\mathcal{Z}^2}; \end{cases}$$

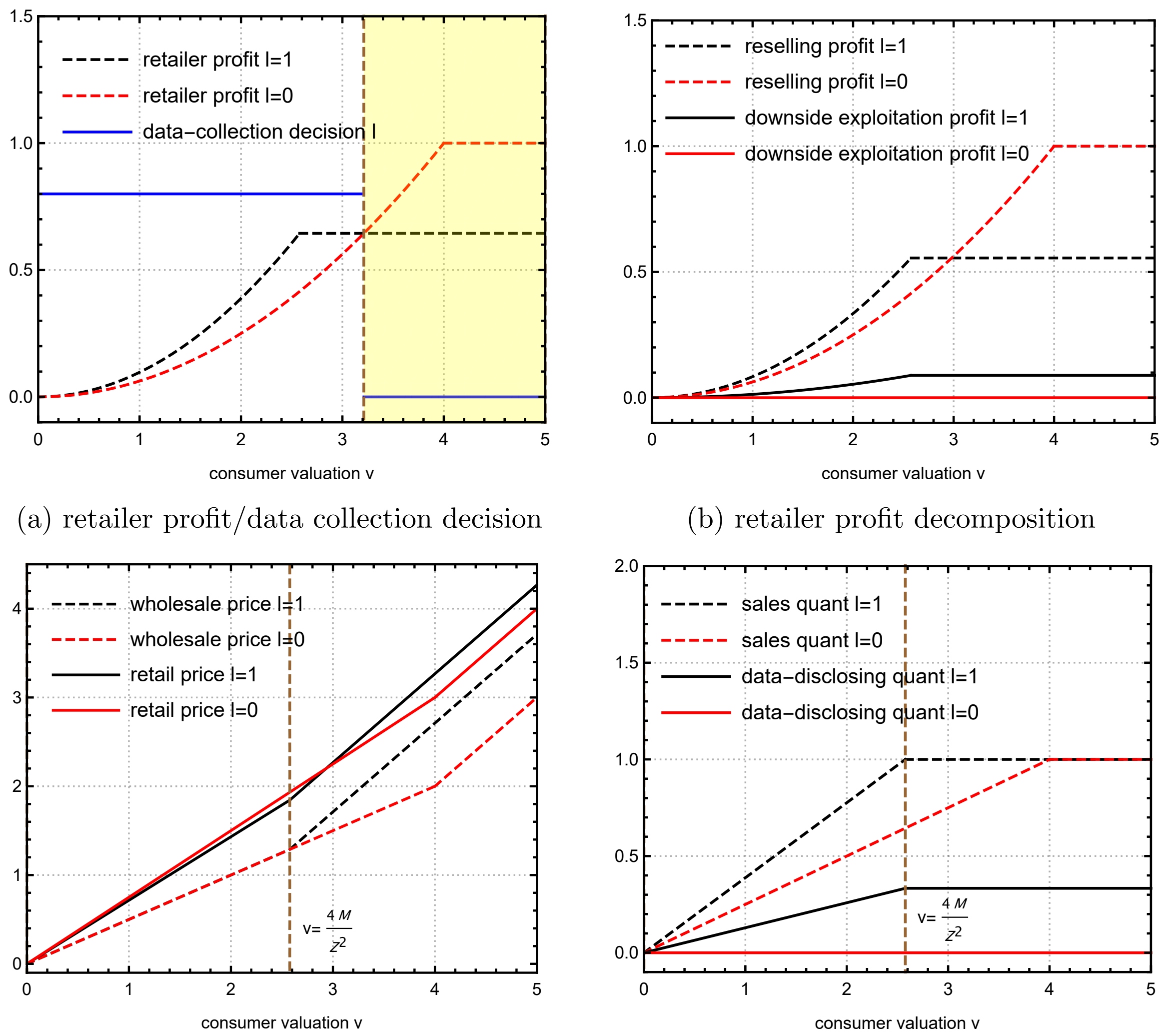
(ii) given that the retailer does not collect consumer data, we have

$$(r^*, p^*, w^*, \mathcal{T}(p^*)) = \begin{cases} (0, 3v/4, v/2, v/4) & \text{if } \ v \leq 4, \\ (0, v-1, v-2, 1) & \text{if } \ v > 4; \end{cases}$$

(iii) the retailer collects consumer data ($\ell^* = 1$) if and only if $v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, where $\frac{4\mathcal{M}}{\mathcal{Z}^2} \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}} \leq 4$.

Proposition 1.4.1 implies that collecting consumer data does not always benefit the retailer, and the retailer would strategically forgo data collection when the consumer valuation v is sufficiently high (i.e., $v > \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$). Recall that the retailer's profit consists of two parts: the product reselling profit $(p-w)\mathcal{T}(p)$ and the downside exploitation profit $\lambda r\delta\,[\mathcal{T}(p) - \lambda r] - cr^2$. Clearly, collecting consumer data allows the retailer to obtain the downside exploitation profit, but the reselling profit can be either increased or decreased (see Figure 1.4b), leading to the overall increase or decrease in the retailer's profit (see Figure 1.4a). To understand the nature of the impact of data collection on the product reselling profit, we decompose the product reselling profit and analyze the retailer's sales quantity $\mathcal{T}(p)$ and the profit margin $(p-w)$, respectively.

We first investigate the sales quantity $\mathcal{T}(p)$. Recall that Lemma 1 divides consumers into the following three types. The *privacy sensitive* consumers with $\theta \in [0, \lambda r)$ purchase the product but hide their data, since the downside exploitation cost (i.e., privacy cost) $\lambda r\delta$ dominates the upside exploitation benefit $\theta\delta$. The *privacy insensitive* consumers with $\theta \in [\lambda r, \mathcal{T}(p)]$ purchase the product and disclose data, since the privacy cost is lower than the upside exploitation benefit, and purchasing the product generates a positive profit as well. Finally, the *inactive* consumers with $\theta \in (\mathcal{T}(p), 1]$ neither purchase the product nor disclose their data. Clearly, data collection and exploitation benefit the privacy insensitive consumers while not affecting the privacy sensitive consumers, leading to the expansion of the market size $\mathcal{T}(p)$ (see Figure 1.4d).



(a) retailer profit/data collection decision

(b) retailer profit decomposition

(c) wholesale/retail price

(d) sales quantity/data-disclosing quantity

Figure 1.4.: (Color Online) Retailer's profit decomposition under the $\mathbb{R}$ mode w.r.t. v ($c = 0.2$, $\delta = 0.8$, $\lambda = 1$)

Next we consider the profit margin, which is determined by the retailer's retail price and the supplier's wholesale price. Data collection enables the retailer to make a profit from the downside data exploitation in addition to the product reselling profit, which motivates the retailer to reduce the retail price to attract more consumers and thus boost the downside exploitation profit. Specifically, given the wholesale price w, collecting data encourages the retailer to reduce the retail price from $\frac{v+w}{2}$ to $\frac{v+w}{2} - \frac{c\lambda^2\delta^2(v-w)}{2\mathcal{M}}$, where $\frac{c\lambda^2\delta^2(v-w)}{2\mathcal{M}}$ reflects the price "discount" the retailer is willing to sacrifice. Meanwhile, the expanded market demand $\mathcal{T}(p)$ drives the supplier to increase the wholesale price as shown in Proposition 1.4.1. As a quick summary, data collection reduces the profit margin, which is caused by the retailer and the supplier's pricing decisions (see Figure 1.4c).

The above discussion reveals the two countervailing forces that shape the value of data collection, i.e., the increased market size and the reduced profit margin. In general, either force could be dominant. Specifically, collecting data hurts the product reselling profit only if the consumer valuation is sufficiently high. In what follows, we provide more detailed discussions using Figure 1.4c and Figure 1.4d. When the consumer valuation v is low (i.e., $v \leq \frac{4\mathcal{M}}{\mathcal{Z}^2}$), data collection substantially increases the sales quantity, while it does not lead to a severe profit margin reduction. Hence, the product reselling profit is enhanced. As v increases across the threshold $v = \frac{4\mathcal{M}}{\mathcal{Z}^2}$, the market is fully penetrated with data collection (i.e., $\mathcal{T}(p) = 1$), meaning that the product becomes extremely "popular" and consumers are price insensitive. As a response, the upstream supplier is able to take advantage of such a price-insensitive market by aggressively increasing the wholesale price, which causes a significant decrease in profit margin compared to the case without data collection (see Figure 1.4c). Meanwhile, the effect of market expansion is diminishing, since the market is already fully penetrated and thus is unable to further grow (see Figure 1.4d). Hence, collecting data cannot increase the product reselling profit anymore. When the consumer valuation is large enough ($v > \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$), collecting data leads to a substantial loss of the product reselling profit, which counteracts and dominates the benefit of the downside data exploitation. As a result, the retailer would deliberately choose not to collect consumer data ($\ell = 0$), which sacrifices some downside exploitation profit, but maintains the more significant reselling profit.

Impact of Key Parameters.

Facilitated by the fast development of technologies, the data marketplace has experienced rapid growth in recent years. Industry reports show that the global data broker market was valued at \$232.634 billion in 2019 and is expected to increase at 5.80% to reach \$345.153 billion in 2026.[18] As the data marketplace keeps growing, understanding how the data is extracted and monetized is of great importance to both the industry and consumers. In this part, we focus on the retailer's data monetization intensity decision r and study how it is affected by the probability of having downside exploitation value (i.e, λ) and the data quality (i.e., δ). To avoid the trivial case of no data collection (where the retailer does not monetize data $r = 0$), here we only focus on $v < \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$. Our result is given in the following proposition:

Proposition 1.4.2 *When the retailer collects consumer data ($\ell = \delta$) under the $\mathbb{R}$ mode, the retailer's data monetization intensity r^* is (i) quasi concave in λ; (ii) increasing in δ.*

Several remarks are in order. First, higher probability λ does not necessarily encourage the retailer to increase the data monetization intensity r. To understand the nature of such non-monotonicity, the left panel of Figure 1.5 plots the retailer's monetization intensity r, the sales quantity $\mathcal{T}(p)$, and the data-disclosing quantity $\mathcal{T}(p) - \lambda r$. When λ is low, the market is fully penetrated (i.e., $\mathcal{T}(p) = 1$), meaning the product is very "popular" and consumers respond insensitively to the increase of either the exploitation cost $\lambda r \delta$ or the purchasing cost p. As λ increases, each consumer's exploitation value $\lambda r \delta$ is sharply increased, while the data-disclosing quantity $\mathcal{T}(p) - \lambda r$ remains sufficiently large. Hence, raising the monetization intensity significantly increases the exploitation revenue $\lambda r \delta \left[\mathcal{T}(p) - \lambda r\right]$, which dominates the increased extraction cost cr^2. As λ increases across $\bar{\lambda}$, the market becomes sensitive. Specifically, the data-disclosing quantity

[18] https://www.knowledge-sourcing.com/report/global-data-broker-market

$\mathcal{T}(p) - \lambda r$ drops sharply in λ, which leads to the decrease of the downside exploitation revenue $\lambda r\delta\,[\mathcal{T}(p) - \lambda r]$, albeit the individual exploitation value $\lambda r\delta$ is slightly increased. When this happens, raising the monetization intensity cannot efficiently enhance the exploitation revenue, while the extraction cost cr^2 escalates. Therefore, the retailer reduces the monetization intensity r.

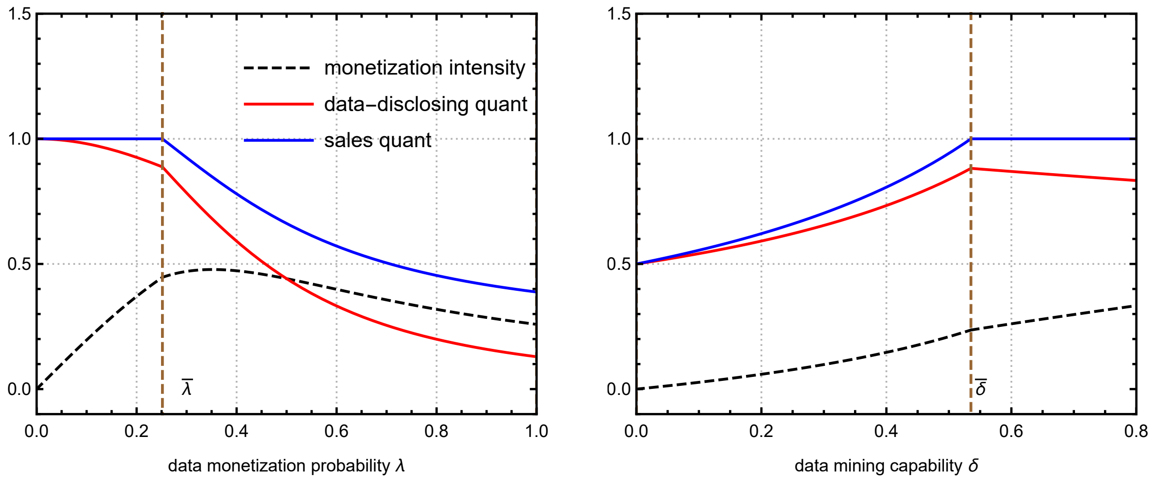


Figure 1.5.: (Color Online) Monetization intensity r^*, sales quantity $\mathcal{T}(p^*)$, and data-disclosing quantity $\mathcal{T}(p^*) - \lambda r$ under the $\mathbb{R}$ mode w.r.t. λ ($c = 0.2$, $\delta = 0.8$, $v = 1$) and δ ($c = 0.5$, $\lambda = 0.5$, $v = 2$)

Second, a higher data quality δ always leads to a higher data monetization intensity r. Recall that Lemma 1 divides the consumers into three types. All else being equal, as δ increases, the privacy sensitive consumers (with $\theta < \lambda r$) keep hiding their data and thus are not affected by the change in the retailer's data quality δ. The privacy insensitive and inactive consumers have sufficiently large θ, where the upside exploitation benefit $\delta\theta$ dominates the downside exploitation cost $\lambda r\delta$. Hence, the increased data quality δ makes product purchasing and data disclosing more appealing to these two types of consumers, leading to the expansion of the market size and the data-disclosing quantity. In addition, the increased δ also enhances the individual data value $\lambda r\delta$. Similar to the above discussion of λ, as δ increases, the increased exploitation revenue motivates the retailer to raise the monetization intensity r. When δ increases across the threshold $\bar{\delta}$, the market is fully penetrated (i.e., $\mathcal{T}(p) = 1$). Hence, as δ increases, the sales quantity and the data-disclosing quantity are unchanged. Meanwhile, the individual data value

$\lambda r\delta$ is still increasing. By similar reasoning, the retailer increases the data monetization intensity r, which does not reduce the sales quantity $\mathcal{T}(p)$, but slightly reduces the data-disclosing quantity $\mathcal{T}(p) - \lambda r$ (see the right panel of Figure 1.5).

1.4.2 Consumer-Controlled Data Policy ($\mathbb{C}$ Mode)

In this subsection, we investigate the GDPR policy. Similar to the $\mathbb{R}$ mode, the discussion of the retailer not collecting data case (i.e., $\ell = 0$) is relegated to the online appendix (proof of Proposition 1.4.3). Recall that under the $\mathbb{C}$ mode, consumers would allow upside data exploitation but block downside data exploitation if they disclose data. Given the retail price p, the consumer utility of purchasing the product is:

$$U_2(\alpha, p, \theta) = (v - p) - (1 - \alpha\delta)\theta.$$

It is clear that all consumers would disclose data ($\alpha = 1$) when purchasing. Thus a consumer chooses to purchase if the corresponding utility $v - p - (1 - \delta)\theta \geq 0$, leading to a threshold rule for the consumer purchasing decision. That is, a consumer would disclose data and purchase the product only if $\theta \leq \frac{v-p}{1-\delta}$ (see Figure 1.6). Similarly, we denote $\mathcal{T}'(p) = \min\{\frac{v-p}{1-\delta}, 1\}$ as the sales quantity under the $\mathbb{C}$ mode. We adopt the backward induction approach and derive the equilibrium, which is summarized in the next proposition.

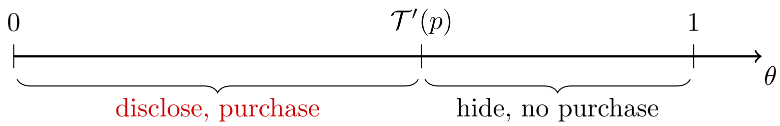


Figure 1.6.: Consumer data-disclosing and purchasing decisions ($\mathbb{C}$ Mode)

Proposition 1.4.3 *Under the $\mathbb{C}$ mode, the retailer's retail price $\tilde{p}^*$, the supplier's wholesale price $\tilde{w}^*$, the sales quantity $\mathcal{T}'(\tilde{p}^*)$, and the retailer's data-collection decision $\tilde{\ell}^*$ are given as follows:*

(i) given that the retailer collects consumer data, we have

$$(\tilde{p}^*, \tilde{w}^*, \mathcal{T}'(\tilde{p}^*)) = \begin{cases} \left(3v/4, v/2, \dfrac{v}{4(1-\delta)}\right) & \text{if } \ v \leq 4(1-\delta), \\ (v-(1-\delta), v-2(1-\delta), 1) & \text{if } \ v > 4(1-\delta); \end{cases}$$

(ii) given that the retailer does not collect consumer data, we have

$$(\tilde{p}^*, \tilde{w}^*, \mathcal{T}'(\tilde{p}^*)) = \begin{cases} (3v/4, v/2, v/4) & \text{if } \ v \leq 4, \\ (v-1, v-2, 1) & \text{if } \ v > 4; \end{cases}$$

(iii) the retailer collects consumer data ($\tilde{\ell}^* = \delta$) if and only if $v \leq 4\sqrt{1-\delta}$, where $4(1-\delta) \leq 4\sqrt{1-\delta} \leq 4$.

Proposition 1.4.3 shows that, similar to the $\mathbb{R}$ mode, collecting consumer data under the $\mathbb{C}$ mode affects the sales quantity and the profit margin in the opposite direction. Specifically, data collection expands the sales quantity $\mathcal{T}'(\tilde{p})$ due to the upside data exploitation benefit to consumers, but reduces the profit margin $\tilde{p} - \tilde{w}$ (see Figure A.1). To better understand the difference between the $\mathbb{R}$ mode and the $\mathbb{C}$ mode, Figure 1.7 plots the sales quantity and the profit margin under these two modes. The left panel shows that as v increases, the $\mathbb{C}$ mode expands the market demand faster than the $\mathbb{R}$ mode, and thus is earlier to achieve full market penetration (i.e., $v \geq 4(1-\delta)$). As previously discussed, such a market expansion is driven by the data exploitation benefit to consumers. In addition to the upside exploitation benefit both modes provide, the $\mathbb{C}$ mode also eliminates the downside exploitation cost, leading to a further expansion of the market demand compared with the $\mathbb{R}$ mode.

The right panel of Figure 1.7 plots the profit margin. The discussion of the $\mathbb{R}$ mode reveals the profit margin is affected by both the supplier's wholesale price and the retailer's retail price. When the market is not fully penetrated (i.e., $v < 4(1-\delta)$), consumers are price sensitive. Hence, the upstream supplier cannot "squeeze" the downstream retailer. The supplier always charges the wholesale price $v/2$ and there exists no wholesale price escalation under both modes. Meanwhile, the retailer offers a discounted retail price under

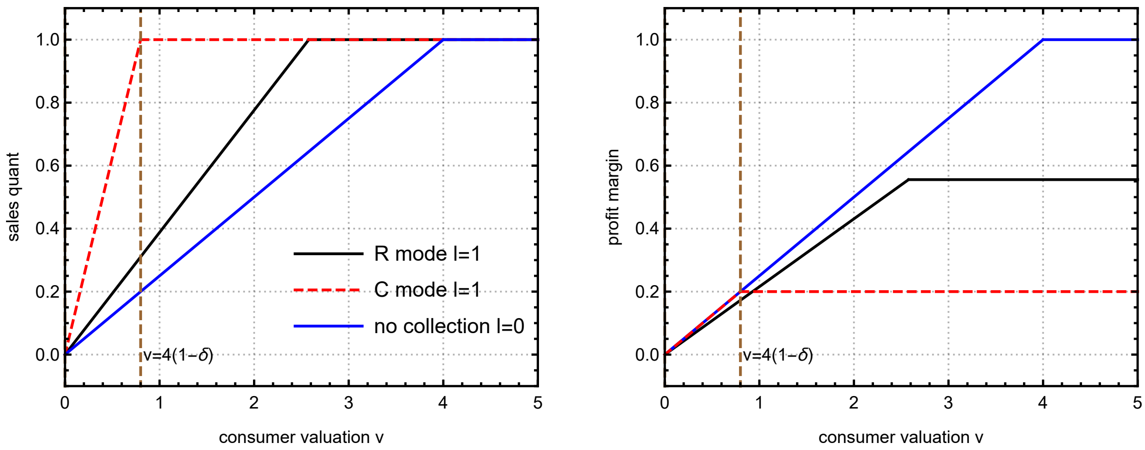


Figure 1.7.: (Color Online) Comparison of the sales quantity and the profit margin w.r.t. v ($c = 0.2$, $\delta = 0.8$, $v = 1$)

the $\mathbb{R}$ mode (to persuade more consumers to disclose data and earn higher downside exploitation profit), but does not have incentives to offer such a discount without the opportunity of downside exploitation under the $\mathbb{C}$ mode. Hence, collecting data reduces the profit margin under the $\mathbb{R}$ mode, but does not affect the margin under the $\mathbb{C}$ mode. As v increases across the threshold $v = 4(1-\delta)$, the market under the $\mathbb{C}$ mode is the first to become price insensitive (i.e., fully penetrated), leading to the sharp increase of the wholesale price. As a result, the profit margin under the $\mathbb{C}$ mode is substantially lower than that under the $\mathbb{R}$ mode.

1.5 Welfare Comparison and Policy Implications

A fundamental question studied in this chapter is how the GDPR policy affects the consumer surplus, the retailer's profit, the supplier's profit, and the overall social welfare. In this section, we compare the previously studied data policies (i.e., the $\mathbb{R}$ mode and the $\mathbb{C}$ mode) and show how the GDPR adoption impacts each party's profit.

1.5.1 Data-Collection Decision

We first examine the retailer's optimal data-collection decision, ℓ. As previously discussed, consumers become price insensitive when the consumer valuation is sufficiently high. As such, the benefit of data collection to consumers enables the upstream supplier to substantially raise the wholesale price, leading to the decrease of the retailer's profit. Hence, when the consumer valuation v is high enough, the retailer strategically chooses not to collect data under both modes. We further notice that the retailer is less likely to collect consumer data under the GDPR policy (i.e., $\mathbb{C}$ mode), because the parameter region of v for data collection under the $\mathbb{C}$ mode (i.e., $v \leq 4\sqrt{1-\delta}$) is a subset of that under the $\mathbb{R}$ mode (i.e., $v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$). This is also depicted in Figure 1.8. The intuition is as follows. The retailer's profit consists of both the product reselling profit and the downside exploitation profit. On one hand, the $\mathbb{C}$ mode disables the downside exploitation profit. On the other hand, data collection also generates lower product reselling profit under the $\mathbb{C}$ mode. As Figure 1.7 shows, when the consumer valuation v is sufficiently high, data collection under the $\mathbb{C}$ mode cannot expand the market size (since the market is fully penetrated with the sales quantity $\mathcal{T}'(\tilde{p}) = 1$), but significantly reduces the profit margin, leading to the sharp decrease of the product reselling profit. In summary, collecting data under the $\mathbb{C}$ mode loses the downside exploitation profit, and reduces the product reselling profit. Hence, data collection under the $\mathbb{C}$ mode is less appealing to the retailer. This result suggests that the GDPR policy indeed helps decrease the collection/usage of consumer data.

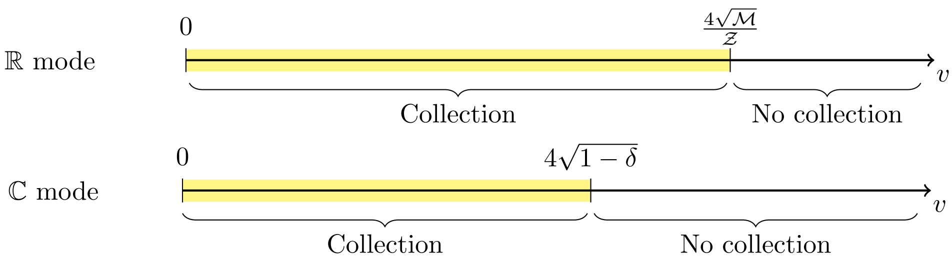


Figure 1.8.: Comparison of data collection under two policies

1.5.2 Consumer Surplus

The GDPR policy is aimed at protecting the consumer data and enhancing the consumer surplus. However, it only emphasizes one side of the picture: the direct *privacy effect*.

That is, consumers can block the downside data exploitation to save the privacy cost $\lambda r\delta$, while still enjoying the upside exploitation benefit by reducing the inconvenience cost $\theta\delta$. The GDPR ignores the other side of the picture: the indirect *pricing effect*. Proposition 1.4.1 and Proposition 1.4.3 imply the retailer charges a higher retail price under the $\mathbb{C}$ mode, due to the block of the downside exploitation. Clearly, the indirect pricing effect hurts the consumer surplus. For expositional clarity, we further denote $\mathcal{F} = 4c^2(1-\delta) + 4c\lambda^2\delta(1-\delta) + \delta^2\lambda^4$. The trade-off between the direct privacy effect and the indirect pricing effect determines the value of the GDPR policy to consumers, which we summarize in the next proposition.

Proposition 1.5.1 *(i) If $v < \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$ or $\frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}} < v < \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the GDPR policy strictly increases the consumer surplus;*

(ii) If $\frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}} < v < \frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}}$, the GDPR policy strictly reduces the consumer surplus;

(iii) Otherwise, the GDPR policy has no effect on the consumer surplus.

Proposition 1.5.1 asserts that the GDPR policy can hurt the consumer surplus when the consumer valuation is relatively large (i.e., $v > \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$).[19] We may explain the intuition using Figure 1.9. As previously mentioned, when the market is price sensitive (i.e., $v \leq 4(1-\delta)$), the supplier charges the same wholesale price $v/2$ regardless of the data-collection decision and the data policy. In other words, the upstream supplier cannot raise the wholesale price to take advantage of the benefit generated by the data collection. Hence, the difference of the retail price (i.e., the indirect pricing effect) is only driven by the retailer's price decision, but not affected by the pricing pressure from the upstream supplier. Recall the retailer offers a price discount $\frac{c\lambda^2\delta^2(v-w)}{2\mathcal{M}}$ under the $\mathbb{R}$ mode, into which we plug $w = v/2$ and obtain the indirect pricing effect $\frac{vc\lambda^2\delta^2}{4\mathcal{M}}$. The left panel of Figure 1.9 also depicts the indirect pricing effect, which is insignificant for a price sensitive market (i.e., v is small). Hence, the direct privacy effect is dominant, leading to the overall improvement of the consumer surplus by the GDPR. As v increases across $4(1-\delta)$, the market becomes price insensitive, allowing the upstream supplier to make

[19] We will discuss $\frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}} < v < \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$ case later.

good use of the direct privacy benefit of the GDPR and tremendously raise the wholesale price. As a result, the indirect pricing effect dominates the direct privacy effect and the GDPR hurts the consumer surplus.

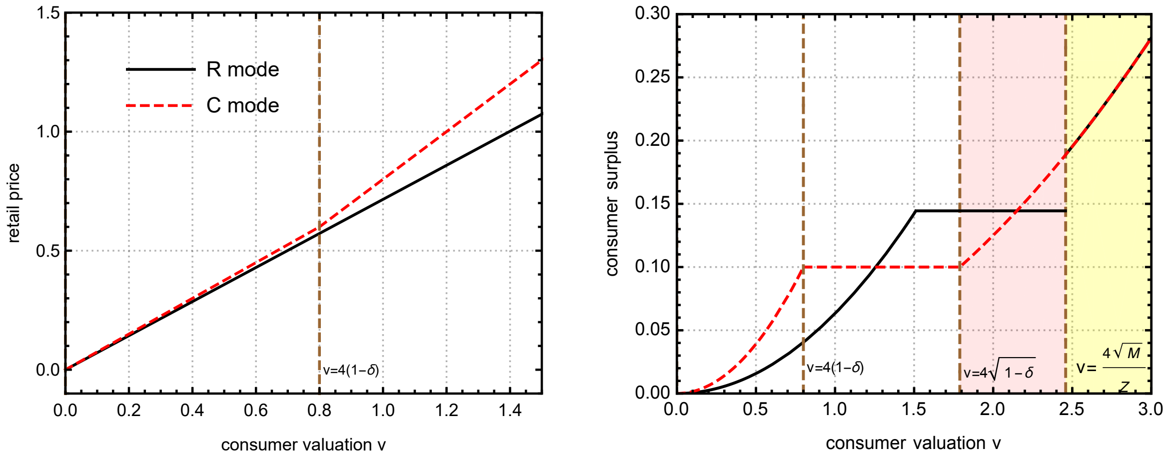


Figure 1.9.: (Color Online) Comparison of the retail price and the consumer surplus w.r.t. v ($c = 0.2$, $\delta = 0.8$, $v = 1$)

To better understand the nature of such a "two-way" result, we also explain it in the mathematical way. Recall there are three types of consumers under the $\mathbb{R}$ mode. (i) Compared to the $\mathbb{R}$ mode, the GDPR policy eliminates the downside exploitation cost $\lambda r^*\delta$ for the privacy insensitive consumers, but exerts the indirect pricing effect $\tilde{p}^* - p^*$. Note that $\mathcal{T}(p^*) \leq \mathcal{T}'(\tilde{p}^*)$ always holds (see Figure 1.7), which is equivalent to $v - p^* - \lambda r^*\delta \leq v - \tilde{p}^*$. Clearly, the direct privacy effect $\lambda r^*\delta$ dominates the indirect pricing effect $\tilde{p}^* - p^*$, and thus the GDPR policy always benefits the privacy insensitive consumers. (ii) Note that $\mathcal{T}'(\tilde{p}^*) - \mathcal{T}(p^*)$ represents the number of inactive consumers that switch to purchasing under the GDPR, and thus the GDPR also benefits the inactive consumers. (iii) The GDPR policy persuades the privacy sensitive consumers to disclose data, providing the upside exploitation benefit $\theta\delta$. Meanwhile, the indirect pricing effect $\tilde{p}^* - p^*$ hurts the consumers. Hence, only consumers with $\theta \geq \bar{\theta}$ would benefit from the GDPR policy, where $\bar{\theta} = \frac{\tilde{p}^* - p^*}{\delta}$. Combining the above analysis, the GDPR policy hurts consumers if and only if their inconvenience parameter $\theta < \bar{\theta}$. It is clear that $\bar{\theta}$ is increasing in v (see the left panel of Figure 1.9), indicating that the GDPR benefits fewer consumers as v increases. When v is sufficiently large, the GDPR hurts the overall consumer surplus.

Next, we discuss the region with $\frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}} < v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, where the GDPR enhances the consumer surplus as v increases to a sufficiently high level, driven by the changing relative impacts of data exploitation on the consumer surplus and on the retailer's profit. We start with the $\mathbb{C}$ mode. Under the GDPR, data collection has the same impacts on the consumer surplus and on the retailer's profit, resulted from the upside exploitation. Specifically, the upside exploitation benefits consumers and causes market expansion, allowing the retailer to share part of the increased surplus (we call this the positive effect P of data collection). Meanwhile, the market expansion leads to wholesale price escalation (we call this the negative effect N of data collection), which hurts both the consumers and the retailer. When $v = 4\sqrt{1-\delta}$, the upside exploitation effect equals the price escalation effect (i.e., $P = N$), and both the retailer's profit and the consumer surplus are indifferent between collecting and not collecting data (see Figure A.4a and the right panel of Figure 1.9). Hence, the impacts of data exploitation on the two downstream parties are equivalent under the $\mathbb{C}$ mode. Under the $\mathbb{R}$ mode, however, there exists a downside exploitation effect in addition to the upside exploitation effect. The downside exploitation causes a value transfer from the consumers to the retailer (we call this the transfer effect T of data collection). Such a value transfer causes unbalanced impacts of data collection on the two downstream parties. Specifically, when the consumer surplus is indifferent between collecting and not collecting data at $v = \frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}}$ (i.e., $P - T = N$), the retailer still prefers collecting data under the $\mathbb{R}$ mode ($P + T > N$).[20] As v increases, the price escalation effect becomes more severe (i.e., N is increasing). As a consequence, the consumer surplus would decrease due to data collection ($P - T < N$), while the retailer still chooses to collect data ($P + T > N$). Proposition 1.4.3 shows that the retailer does not collect data under the $\mathbb{C}$ mode when $v > 4\sqrt{1-\delta}$. Hence, consumer surplus under the $\mathbb{C}$ mode is equivalent to that of not collecting data when $v \in [\frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}}, \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}]$, whereas

[20]Recall that the retailer switches to not collecting data at $v = \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, where $4\sqrt{1-\delta} \leq \frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}} \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$.

the retailer still collects data under the $\mathbb{R}$ mode. Hence, the GDPR policy benefits the consumers compared to the $\mathbb{R}$ mode when $\frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}} < v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$.

1.5.3 Retailer Profit

We find that the GDPR policy can either increase or decrease the retailer's profit. On one hand, the GDPR blocks the downside data exploitation, which hurts the retailer's profit. On the other hand, the product reselling profit can be either enhanced or reduced by the GDPR, driven by the trade-off between the sales quantity and the profit margin. As discussed in §1.4.2, compared to the $\mathbb{R}$ mode, the GDPR saves consumers' downside exploitation cost, and always leads to further market expansion. Meanwhile, the profit margin can be increased or decreased. Specifically, when the market is price sensitive (i.e., $v \leq 4(1-\delta)$), there exists no wholesale price escalation. Due to the elimination of the downside exploitation profit, the retailer charges a higher retail price under the $\mathbb{C}$ mode, which increases the profit margin. When the market is price insensitive, the benefit provided by the GDPR encourages the upstream supplier to substantially increase the wholesale price, leading to the profit margin reduction, albeit the slight retail price increase (see the right panel of Figure 1.7). As a result, when v is relatively low, the increased product reselling profit under the GDPR counteracts and dominates the loss of the downside exploitation profit, leading to an overall increase of the retailer's profit. As v keeps increasing, the product reselling profit sharply decreases, caused by the upstream supplier's wholesale price escalation. Eventually, the GDPR reduces the retailer's profit. We formally summarize this result in Proposition A.0.2, and the comparison of the retailer's profits under the two modes is also depicted in Figure A.4a.

1.5.4 Supplier Profit

We now turn to the analysis of the supplier's profit. One might intuit that the GDPR policy always benefits the supplier. On one hand, the supplier makes no data-related profit, and thus the elimination of the downside exploitation by the GDPR does not directly hurt the supplier's profit. On the other hand, the GDPR benefits consumers with regard to the data exploitation, leading to the market expansion that increases the supplier's profit. Our analysis demonstrates that, however, the GDPR may hurt the supplier's profit under some conditions. If the retailer always collects the consumer data, it is true the GDPR always benefits the supplier. As the left panel of Figure 1.7 shows, compared to the sales quantity $\mathcal{T}(p^*)$ under the $\mathbb{R}$ mode, the $\mathbb{C}$ mode generates a larger sales quantity $\mathcal{T}'(\tilde{p}^*)$, leading the increase of the supplier's profit. Note that when $v > \frac{4\mathcal{M}}{\mathcal{Z}^2}$, the market is fully penetrated under both modes (i.e., $\mathcal{T}(p^*) = \mathcal{T}'(\tilde{p}^*) = 1$), disallowing the GDPR to expand the market size anymore. Nevertheless, the improved consumer surplus enables the supplier to increase the wholesale price from $w^* = v - \frac{2\mathcal{M}}{\mathcal{Z}^2}$ to $\tilde{w}^* = v - 2(1-\delta)$, which enhances the supplier's profit as well. However, the result is reversed considering the retailer's data-collection decision. As discussed in §1.5.1, the retailer is less "likely" to collect consumer data under the $\mathbb{C}$ mode. Specifically, the GDPR switches the retailer to not collecting consumer data when $4\sqrt{1-\delta} < v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, leading to the decrease of the market size without the data exploitation benefit. As a consequence, the GDPR hurts the supplier's profit (see Figure A.4b). The result is summarized in Proposition A.0.3.

1.5.5 Social Welfare

Next we investigate the social welfare. We find that the GDPR can either increase or decrease the social welfare, depending on the retailer's data-collection decision (see Figure A.4c). Data collection has both the upside and downside data exploitation effects on the social welfare. The upside exploitation reduces the consumers' inconvenience cost within the whole system, which increases the social welfare. The downside exploitation simply transfers the value $\lambda r^* \delta$ from the consumers to the retailer, without any welfare creation. Hence, the downside exploitation does not directly affect the social welfare, except the data monetization cost incurred by the retailer. However, the elimination of downside exploitation indirectly affects the social welfare, leading to the increase of the social welfare under the $\mathbb{C}$ mode. Specifically, the $\mathbb{C}$ mode allows consumers to block the downside exploitation, which motivates more data disclosing by consumers (see Figure A.3). As a result, the benefit of upside exploitation is more likely to happen under the $\mathbb{C}$ mode, implying that the GDPR increases the social welfare when data is collected under both modes (i.e., $v \le 4\sqrt{1-\delta}$). However, when $4\sqrt{1-\delta} < v \le \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the $\mathbb{C}$ mode discourages the retailer from collecting data, which, on the contrary, deprives the benefit of data collection and reduces the social welfare. When $v > \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the retailer never collects consumer data regardless of the data policy, and thus the $\mathbb{R}$ mode and the $\mathbb{C}$ mode are equivalent. In this case, the GDPR policy has no impact on the social welfare. Not only could the GDPR reduce the social welfare, it may also hurt the supplier's profit, the retailer's profit, and the consumer surplus at the same time (i.e., triple-lose situation). Figure 1.10 plots the welfare comparison for each party under the two modes, where the shaded area (i.e., $\mathbb{C} \le \mathbb{R}$) represents the situation where the GDPR hurts the corresponding party. In particular, when $4\sqrt{1-\delta} < v < \frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}}$, the GDPR policy leads to the triple-lose result. We summarize our findings in Proposition 1.5.2.

Proposition 1.5.2 *(i) If $4\sqrt{1-\delta} < v < \frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}}$, all three parties (consumers, retailer, and supplier) are strictly worse off when the GDPR policy is adopted;*

(ii) If $v < \frac{4\sqrt{(1-\delta)\mathcal{M}}}{\mathcal{Z}}$, the adoption of GDPR policy leads to a Pareto improvement and all three parties are strictly better off.

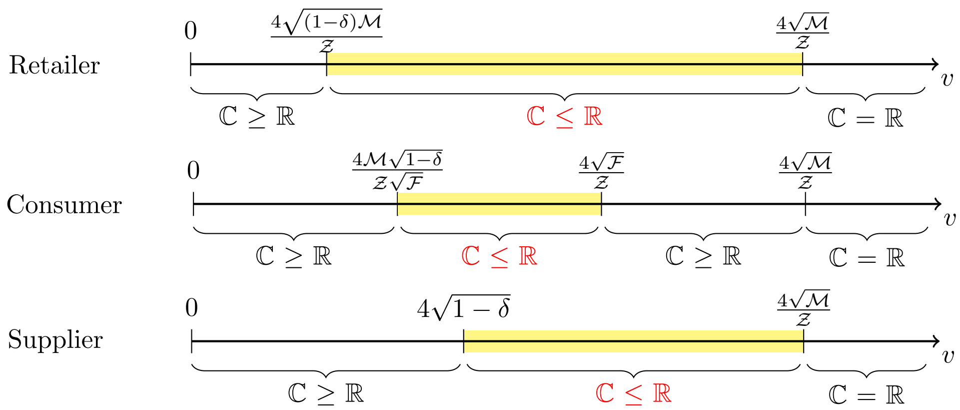


Figure 1.10.: Comparison of retailer profit, consumer surplus and supplier profit under two policies

There are two points worth attention. First, when the consumer surplus is indifferent between the $\mathbb{C}$ mode and the $\mathbb{R}$ mode at $v = \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$, the retailer is already hurt by the $\mathbb{C}$ mode ($v > \frac{4\sqrt{(1-\delta)\mathcal{M}}}{\mathcal{Z}}$). Recall the discussion in §1.5.2 that the upside exploitation exerts both the positive effect P from the market expansion and negative effect N from the wholesale price escalation on the retailer as well as the consumers. As v increases, the market is easier to get fully penetrated under the $\mathbb{C}$ mode, where the negative effect N dominates. In other words, the $\mathbb{C}$ mode hurts the retailer and consumers through the upside exploitation when v is sufficiently large. For illustration, we use PN to denote such a value loss of the $\mathbb{C}$ mode. Besides, data monetization has a value transferring effect, T, from the consumers to the retailer. When the consumer surplus is identical under both modes at $v = \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$, the value loss PN under the $\mathbb{C}$ mode equals the transfer value loss T under the $\mathbb{R}$ mode. However, under the $\mathbb{R}$ mode, the retailer avoids the value loss PN (same as the consumer surplus), and benefits from the transfer value T (contrary to the consumer surplus). As a result, when the consumer surplus is identical under both modes (i.e., $v = \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$), the retailer strictly prefers the $\mathbb{R}$ mode, which saves the value loss PN from the upside exploitation, and adds a transfer value T from the downside exploitation, resulting in a smaller indifference threshold for the retailer's profit (see Figure 1.10). Second, when the retailer switches to not collecting data under the $\mathbb{C}$ mode (i.e., $v = 4\sqrt{1-\delta}$), it already hurts the consumer surplus (and the retailer's profit by previous discussion) compared with the $\mathbb{R}$ mode. The intuition is explained as follows. When the market is price sensitive, the upstream supplier charges the same wholesale price $v/2$ regardless of the data-collection decision and the data policy, meaning the wholesale price escalation due to the data collection does not exist. Compared to the no data collection case, the $\mathbb{R}$ mode with data collection benefits consumers via the upside exploitation, and the $\mathbb{C}$ mode with data collection further benefits consumers by saving the downside exploitation cost. Hence, when the market is price sensitive (data being collected under both modes), the $\mathbb{C}$ mode generates the largest consumer surplus, followed by the $\mathbb{R}$ mode and the no data collection. As v increases, the product becomes more popular and eventually the market is price insensitive, under which even a small improvement of the consumer surplus would induce the upstream supplier to raise the wholesale price. As a result, the $\mathbb{C}$ mode suffers the most from the wholesale price escalation, which sharply reduces the consumer surplus. As v increases, the consumer surplus under the $\mathbb{C}$ mode first becomes worse off than the $\mathbb{R}$ mode at $v = \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$, and then gets worse off than not collecting data at $v = 4\sqrt{1-\delta}$, where the retailer switches to no data collection. Hence, when the GDPR causes no data collection, it already hurts the consumer surplus, and the retailer's profit is decreased as well. Moreover, switching to no data collection also reduces the supplier's profit. As a consequence, the triple-lose situation arises.

1.5.6 Data Collection Ban

Finally, we briefly discuss an extreme data regulation policy: the data collection ban, where the retailer is prohibited from collecting any data. We find that the data collection ban is dominated by the GDPR policy, meaning the data collection ban leads to a lower consumer surplus, retailer's profit, and supplier's profit. The results are formalized in the following proposition:

Proposition 1.5.3 *The data collection ban is dominated by the GDPR policy. When $v < 4\sqrt{1-\delta}$, the data collection ban strictly hurts all players.*

Recall when $v > 4\sqrt{1-\delta}$, the retailer does not collect consumer data under the $\mathbb{C}$ mode, which is identical to the data collection ban. When $v < 4\sqrt{1-\delta}$, compared with the data collection ban, the $\mathbb{C}$ mode enables the upside data exploitation, which benefits consumers and leads to the market expansion. Meanwhile, such expansion allows the retailer and the supplier to share part of the increased surplus. As a result, the GDPR dominates the data collection ban.

1.6 Supply Chain Coordination

In this section, we examine two cases with better supply chain coordination: (i) agency selling, in which the supplier sells directly to consumers using the retailer's online platform; (ii) vertical integration, in which the upstream supplier and the downstream retailer are integrated into one firm.

1.6.1 Agency Selling

Now consider the agency selling model. For notational clarity, we denote $\mathcal{M}' = 4c^2(1-\delta) + c\lambda^2\delta\,(4-5\delta/2) + \delta^2\lambda^4$ and $\mathcal{F}' = (2c+\delta\lambda^2)[2c(1-\delta)+\delta\lambda^2]$. As the first mover, the retailer decides the commission fee u to charge the supplier at the beginning of the game, and then the supplier makes the retail price decision p. We derive the equilibrium of the $\mathbb{R}$ mode and the $\mathbb{C}$ mode under the agency selling model (see Proposition A.0.5 and Proposition A.0.6), compare the two modes, and summarize our findings in the next proposition.

Proposition 1.6.1 *In the agency selling model, data collection always happens. The GDPR policy always benefits the retailer, but only increases the consumer surplus (resp., the supplier profit) if* $v < \frac{4\mathcal{M}'\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$ *(resp.,* $v < \frac{4\mathcal{M}'\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}'}}$*).*

We first look into the retailer's data collection decision and her profit. Recall that under the baseline reselling model, the retailer may not collect data to avoid the upstream supplier's wholesale price escalation pressure. However, under the agency selling model, as the first mover who sets the fixed fee u before the supplier's pricing decision p, the retailer no longer has such pressure, so the retailer always collects the consumer data. Besides, under the agency selling, the retailer always gets better off with the $\mathbb{C}$ mode, in contrast to the reselling model where the retailer only gets better off with the $\mathbb{C}$ mode when v is sufficiently small (i.e., $v < \frac{4\sqrt{(1-\delta)\mathcal{M}}}{\mathcal{Z}}$). Note that when v is large enough, the product becomes very "popular", which causes the wholesale price escalation under the baseline reselling model, leading to the reduction of the retailer's profit under the $\mathbb{C}$ mode. However, under the agency selling model, as the first mover, the retailer no longer faces such upstream pressure, and thus her profit is always enhanced by the $\mathbb{C}$ mode. We next investigate the supplier's profit. Under the agency selling model, the supplier chooses the retail price p after the retailer's commission fee decision. Hence, the supplier plays the same role as the retailer under the baseline reselling model, meaning that the supplier also faces the upstream pricing pressure. Specifically, following the similar logic, when v is sufficiently large, the product becomes "popular" and the market becomes price insensitive, which allows the retailer to sharply increases the fixed fee u to seize more profit. Recall from the baseline model that such pricing pressure is more severe under the $\mathbb{C}$ mode (we do not repeat the discussion here). Such a high pricing pressure hurts the supplier, and thus he can become worse off with the GDPR policy. Finally, we analyze the consumer surplus. As in the baseline model, the consumer surplus is still determined by the trade-off between the direct privacy effect and the indirect pricing effect. Hence, the consumer surplus can be either increased or decreased under the GDPR policy.

1.6.2 Vertical Integration

We next investigate the centralized system in which the supplier and retailer are integrated into one firm. Our analysis shows that under vertical integration, the GDPR policy always increases the total supply chain profit, but it might increase or decrease the consumer surplus.

Proposition 1.6.2 *Under the centralized system, the integrated firm always collects consumer data; the GDPR policy always leads to higher supply chain profit and social welfare, but hurts the consumer surplus iff* $v > \frac{2\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$.

Similar to the agency selling model, vertical integration eliminates the upstream pricing pressure. Hence, the integrated firm always collects the consumer data, which benefits the whole supply chain and the social welfare following the similar logic in the previous subsection. Under the centralized system, the trade-off between the indirect pricing effect and direct privacy effect still exists. As previously discussed, when v is large enough, data collection becomes less appealing to consumers due to the increased privacy cost. Hence, vertical integration hurts the consumer surplus when the retailer does not collect data in the decentralized system.

1.7 Conclusion

Consumer data exploitation has emerged as a prominent trend in recent years due to the rapid advancement in data analytics. As consumer data becomes increasingly valuable, it also causes substantial privacy concerns. This gives rise to the General Data Protection Regulation (GDPR), which aims to protect consumer privacy and give consumers more data control power. This chapter develops an analytical model that captures several important features: (i) The retailer is able to use the consumer data for the upside data exploitation as well as the downside exploitation; (ii) The data acquirement depends on both the retailer's data-collection decision and the consumer's data-disclosing decision; (iii) Consumers have heterogeneous preferences of the retailer service (upside exploitation), and therefore make the data-disclosing decision based on their own benefits. Our research focuses on the question of how the transfer of data control rights (from the retailer to the consumers) by the GDPR policy would affect the welfare of the supplier, the retailer, and the consumers.

There are several main findings. First, the GDPR policy benefits the consumers by protecting their privacy, which leads to market demand expansion. Meanwhile, such expansion causes price escalation when consumers are price insensitive, which may reduce the overall consumer surplus under the GDPR. Second, although the GDPR policy imposes some constraints on the retailer, the retailer's profit may actually be increased by the GDPR. On one hand, the GDPR deprives the retailer's downside exploitation profit. On the other hand, the product reselling profit can be increased or decreased, depending on the upstream supplier's pricing pressure. When the market is price sensitive, the upstream pricing pressure is mild (the market shrinks rapidly if the supplier increases the wholesale price), and thus the increased product reselling profit counteracts the loss of the downside exploitation profit, leading to an increase of the retailer's profit under the GDPR. Third, the GDPR policy can hurt the supplier's profit since the GDPR discourages the retailer from data collection, which reduces the market size and thus hurts the supplier's profit. Furthermore, the GDPR policy could even lead to a triple-lose outcome. As discussed, the GDPR discourages the data collection due to the elimination of the downside exploitation profit. When the GDPR leads to no data collection, all parties lose the data exploitation benefit and become worse off. Finally, we find that under the agency selling model and vertical integration, the upstream pressure is resolved, and thus data is always collected. The consumer surplus still can be enhanced or reduced by the GDPR policy due to the trade-off between the direct privacy effect and the indirect pricing effect.

We conclude by pointing out some caveats and potential directions for future research. First, our current work has been focused on the monopoly retailer case. In many business environments, there could be multiple retailers who are competitors in the same market. It would be promising to investigate the impact of the privacy policy in the presence of the retailers' competition. Second, we have assumed that disallowing the data collection would incur zero cost for consumers. Although we believe our main results still hold with a positive data blocking cost, it would be interesting to study how consumers' data blocking cost affects each party's decisions and in what direction the welfare would be affected.

Chapter 2: The Value of Smart Contract in Trade Finance

2.1 Introduction

[1]The supplier's delivery of an order to a buyer does not always coincide with the payment by the buyer. In many cases, the buyer makes a delayed payment to the supplier, which is called trade credit (or open account). Trade credit is a widely observed form of inter-firm credit, accounting for about 15% of the assets of US manufacturing firms [40]. When trade credit is used, the supplier undertakes financing pressures and has trade finance needs both before the shipment of products and after the shipment of products. Before the shipment, the capital-constrained supplier needs to raise funds for production. For instance, the supplier can rely on *purchase order financing* (i.e., borrowing from the bank with the back of the buyer's purchase order) to cover the upfront cost.[2] In other cases, the buyer, if in good position of cash flow, can make a prepayment to the supplier (a fraction of the purchase order) to cover the supplier's production cost and help alleviate the supplier's financial distress. This is referred to as *buyer direct financing*. Both purchase order financing and buyer direct financing fall into the pre-shipment financing category as these financing schemes are adopted before the orders are delivered.

Trade finance also involves various financing activities after the order delivery but before the buyer's payment date, i.e., the post-shipment financing. After the product shipment and delivery, due to the common practice of trade credit, only accounts receivable is recorded on the balance sheet of the supplier as short-term claims. For capital-constrained suppliers, they are looking to improve their working capital and free up cash that can be used to improve their short-term liquidity, pay expenses for daily business operations, or make new investments to grow their businesses. One type of post-shipment financing is

[1]This chapter is based on the author's work [39] jointly with Fasheng Xu.

[2]POF in our chapter is equivalent to bank financing. However, we call it POF to highlight its pre-shipment nature and to be in line with related literature [41, 42].

factoring, where the supplier independently sells accounts receivable to the factor against a premium and receives immediate cash for working capital needs. A similar financing option is on-demand *invoice trading*, which is a fast and flexible way for suppliers to meet working capital needs by selling their outstanding accounts receivable to individual or institutional investors via FinTech platforms.[3] On-demand invoice trading is particularly flexible and provides quick access to liquidity to meet cash flow needs, whereas traditional factoring takes a much longer time to complete the transaction with the factor due to labor-intensive verification and time-consuming paperwork. Thus, with invoice trading, working capital financing is available at any point throughout the entire payment term when the supplier needs it.

The increasing prevalence of blockchain-related applications has brought into sharp focus the potential value of smart contracts, with frequent claims that they will likely transform trade finance landscape by supporting more efficient processes, facilitating real-time data recording and sharing, and thus improving supply chain efficiency in various industries. Smart contracts are digital contracts allowing terms contingent on decentralized consensus that are tamper-proof and typically self-enforcing through automated execution [43, 44]. The key characteristics of smart contract are two-fold. The first is the decentralized consensus, which is a description of the state of the world (for example, whether the supplier has sold the accounts receivable to the factor during the post-shipment stage) universally accepted and acted on by all agents in the blockchain system (for example, the retailer and the factor can be record-keepers to provide verifications of the supplier's factoring action). The second is their automated algorithmic execution based on a mapping from certain detectable states of the world (e.g., successful order delivery) to corresponding actions (e.g., invoice insurance).[4]

Tech companies, global banks, and logistic providers around the world are engaged in various proof of concept operations and pilot projects of smart contracts for trade finance.

[3] For example, MarketFinance is UK's first and largest invoice trading platform. https://marketfinance.com/

[4] It is important to note that smart contracts are neither merely digital contracts (many of which rely on trusted authority for agreement and execution, e.g., Letter of Credit) nor are they entailing artificial/business intelligence (on the contrary they are rather robotic) [43].

For example, Trusple (AntChain's blockchain trade platform) generates a smart contract when a buyer and a seller upload a trading order onto the platform.[5] Smart contract automatically updates crucial information (such as order placements, logistics, payment agreement, and financing options) when executing an order. Using AntChain, the involved banks can automatically process the payment settlements via the smart contract. Another example is Skuchain's EC3 platform which has been adopted by banks to obtain independent visibility into the operational process and financial activities of the supply chain participants.[6] Once the financial terms are agreed to, they are executed via smart contracts. As a result, smart contract can help the bank set a more accurate interest rate contingent on the borrower's (supplier's) financial and operational activities.

In our trade finance setting, it is costly, if not impossible, for the bank to closely monitor the borrower's business activities after the loan issuance. Smart contract has the potential to replace established processes that rely on verification procedures by correspondent banks, which involves manual and time-consuming documentary evidence and coordination efforts. In addition, our model of smart contract shares the same spirit as [43] in that smart contract can resolve contract incompleteness via decentralized consensus on the blockchain. In their setting, smart contract enables sellers to offer prices contingent on the success of delivering the goods. Similarly, in our trade finance setting, smart contract enables the bank to offer interest rates contingent on the supplier's post-shipment financing choice (i.e., whether to adopt factoring or not).

In light of the above, smart contract is expected and claimed to hold the promises of reshaping the trade finance market. We explore how the adoption of smart contract can affect the incentives of supply chain members and whether and how its anticipated benefits may be realized under different trade finance structures. To do so, we propose a supply chain finance model where the capital-constrained supplier is in need of both pre-shipment financing and post-shipment financing, and examine how the adoption of smart contract affects the supply chain operations and profits. Our analysis provides three sets of insights, which we summarize below.

[5] https://www.trusple.com/
[6] https://www.skuchain.com/banking-financial-services/

Results.

First, we find that there exists commitment frictions between the supplier and the bank under the baseline model, where the supplier relies on bank financing as pre-shipment financing and traditional factoring as post-shipment financing. More specifically, when the bank determines the interest rate upon the pre-shipment loan request, the supplier is unable to commit the post-shipment factoring adoption decision, which in turn affects the bank's risk exposure. Therefore, the bank would anticipate the supplier's factoring decision based upon the purchase order information (i.e., wholesale price and order quantity). As a consequence, the supplier might deliberately "overprice" in wholesale price offering to convince the bank of post-shipment factoring adoption and get access to a cheaper loan rate. Such overpricing leads to a sharp reduction of the order quantity and tremendously hurts the retailer's profit. Smart contract adoption makes it possible for the supplier to credibly commit to adopting factoring, which alleviates the overpricing issue caused by commitment frictions, and thus restores the supply chain efficiency.

Second, when the retailer can offer BDF as alternative pre-shipment financing, smart contract might reduce the supplier's profit (and even the supply chain profit) if the cost of capital is within a certain medium range. The key driving force of such negative impact is that smart contract discourages the retailer from offering BDF, whose decision boils down to the trade-off between the capital cost (associated with BDF) and the wholesale cost (due to the supplier's overpricing behavior without BDF). Specifically, when BDF is in place, the financing cost of production is reallocated to the retailer side (capital cost). Reciprocally, the supplier would be able to retain the relatively low optimal wholesale price in the absence of interactions (and thus commitment frictions) with the bank, leading to a cheaper wholesale cost for the retailer. Furthermore, when smart contract is adopted, the commitment frictions is fully resolved and the supplier's overpricing behavior is remedied. As a result, the wholesale cost reduction benefit of BDF is wiped out, which discourages the retailer from offering BDF, leading to a sharp decrease in the supplier's profit due to the loss of an interest-free financing option.

Third, when the supplier has access to invoice trading as the post-shipment financing in addition to factoring, the supplier always chooses invoice trading over factoring due to its trading flexibility: allowing the supplier to hold accounts receivable on hand if liquidity shock does not occur so as to save the unnecessary post-shipment financing cost. However, we find, unexpectedly, such a preferable option might turn out to make the supplier worse-off. The intuition can be explained as follows. Along with the trading flexibility benefit, invoice trading, by its nature, also makes the commitment frictions ubiquitous and completely unresolvable whereas it could be partially resolved with over-pricing in factoring. We refer to this phenomenon as the *commitment trap*. Further, the value of invoice trading (relative to factoring) can be decomposed into two components: (i) the *explicit benefit*, which is the combined effect of trading flexibility and premium difference; (ii) the *implicit cost*, which is driven by the commitment trap introduced by invoice trading. When the invoice trading premium is relatively high, the implicit cost of commitment trap can dominates the explicit benefit, leading to the adoption dilemma of invoice trading. Luckily, in such a case, smart contract (in conjunction with factoring) can bring positive value by fixing both commitment trap and commitment frictions.

Contributions.

The chapter makes three main contributions. The first is to provide a general supply chain finance modeling framework to examine trade finance activities at both pre-shipment and post-shipment stages, whereas the extant supply chain finance literature has so far focused almost exclusively on pre-shipment financing. As such, we are able to generate novel insights regarding the interaction between pre-shipment and post-shipment financing schemes. The second is to show that the decision to adopt smart contract should depend on specific trade finance situations. Smart contract is not a “cure” for all business settings. For example, when the retailer’s direct financing is available, smart contract adoption might reduce the supply chain profit. The third is to quantify the value of invoice trading that enables speed-up transaction process and real-time financing offering.

Despite its importance (and widespread adoption), invoice trading has received little attention in the literature.

The rest of the chapter is organized as follows. Section 2.2 reviews the related literature. Section 2.3 lays out the model framework, key assumptions, and notations. In section 2.4, we investigate the baseline model, where POF is adopted as pre-shipment financing and factoring is used as post-shipment financing. We study the value of smart contract under the baseline trade finance model. Section 2.5 discusses the adoption of BDF as an alternative pre-shipment financing scheme and the value of smart contract under this structure. Section 2.6 studies on-demand invoice trading as an alternative post-shipment financing scheme and how smart contract could add value to the supply chain. Table 2.1 provides an overview of different trade finance activities covered in our chapter. Section 2.7 provides further discussions about our model extension of fire sale of accounts receivable and the value of FinTech-driven digitalization. We conclude in Section 2.8 with key managerial insights and future research directions. All proofs and supplemental materials are given in the appendix.

Table 2.1: Overview of various trade finance activities

	Pre-Shipment Financing	Post-Shipment Financing
Section 4	Purchase Order Financing (POF)	Traditional Factoring
Section 5	*Buyer Direct Financing* (BDF)	Traditional Factoring
Section 6	Purchase Order Financing (POF)	*Invoice Trading*

2.2 Literature Review

There are two primary streams of research related to our work: supply chain finance and economics of FinTech (e.g., smart contract and blockchain technology).

Our research is related to the emerging literature on supply chain finance [45, 45–54]. Our model of pre-shipment financing is closely related to the literature on purchase order financing and buyer direct financing. [42] study purchase order financing under market frictions, and show how the financial characteristics of the supplier influence the operational decisions and profits of the supply chain participants. In a pull supply chain setting, [55] compare buyer direct financing with bank financing in a supply chain with one assembler and multiple heterogeneous capital-constrained component suppliers. [56] find that in a consignment selling environment with debt seniority choice, buyer direct financing weakly improves the expected profits of both the retailer and the supplier. Closely related to our work, [41] attempt to understand the relative efficiency of purchase order financing and buyer direct financing under the supplier's endogenous effort and the manufacturer's private information. However, we focus on the interactions between these two pre-shipment financing schemes and the two post-shipment financing schemes (i.e., factoring and invoice trading), and how smart contract could facilitate such trade finance activities and create value for supply chain firms.

Our model of post-shipment financing builds upon the literature on factoring and reverse factoring, which help suppliers (sellers) liquidate accounts receivable for working capital. An early work by [57] examines the optimal factoring contract and finds that the preference of recourse over non-recourse factoring depends on the credit quality of accounts receivable and the seller's solvency. [58] is among the first to provide an econometric analysis of the benefit of factoring and reverse factoring for financing SMEs. [59] find that a reverse factoring program can help reduce financial frictions and provide suppliers with inexpensive financing. [60] empirically confirm the substantial benefits of reverse factoring for both retailers and suppliers. Closely related to our work, [61] develop a supply chain theory of factoring and reverse factoring showing when these post-shipment financing schemes should be adopted and who benefits from the adoption.

Moreover, our chapter contributes to the emerging literature on the FinTech [38,62–65].[7] Using data from an online invoice trading platform, [69] show that the pricing mechanism (auction vs. fixed prices) properly reflects the default risk. [70] study the impact of FinTech competition in payment services where banks rely on consumers' payment data to obtain information about their credit quality. [71] study the value of dynamic trade

[7] [66] provide a comprehensive discussion and review of FinTech innovations for supply chains by focusing on supply chain finance and risk management. [67] and [68] provide recent reviews of finance literature in the emerging area of FinTech.

finance under which the bank dynamically adjusts loan interest rate as an order passes through different steps in the trade process, with a focus on the information frictions related to process uncertainties and its interaction with FinTech.

As an emerging topic within the FinTech literature, blockchain and smart contract have raised much attention from academia recently [72–78]. [79] and [80] study the value of blockchain-driven traceability in supply chains. [81] and [82] study entrepreneurial financing problems in initial coin offerings (ICOs) where crypto-tokens are issued on existing blockchain platforms. Closely related to smart contract, [83] study the value of smart contract which enables automated trade in the fresh produce industry and [84] show that Internet of Things (IoT) and smart contract have different impacts on contracting outcome and efficiency. Motivated by the application of trade finance, [43] show that smart contracts can mitigate contracting incompleteness and improve welfare and consumer surplus through enhanced entry and competition. Our model of smart contract shares the same spirit as [43] in that smart contract resolves contract incompleteness via decentralized consensus on the blockchain. Different from the above papers, we study the commitment value of smart contract in supply chain finance settings and its interaction with various types of pre-shipment and post-shipment financing schemes (e.g., BDF, invoice trading).

2.3 Modeling Framework

Building upon the classic selling-to-the-newsvendor model, we develop a game-theoretic model of a supply chain, where a single type of good is produced by the supplier (he) and sold by the retailer (she) to the end market. The sequence of basic events can be described as follows (illustrated in Figure 2.1). Note that t_1 represents the production period, t_2 is the payment period, and $t_c = t_1 + t_2$ is the total length of the business cycle. At time 0, the upstream supplier first determines the wholesale price w, and then the downstream retailer decides the order quantity q. At time t_1 (end of the production period), the supplier finishes production and delivery and gets the accounts receivable from the retailer with face value wq. At time t_c (end of the payment period), the demand is realized and the retailer collects revenue and pays the receivable outstanding to the supplier.

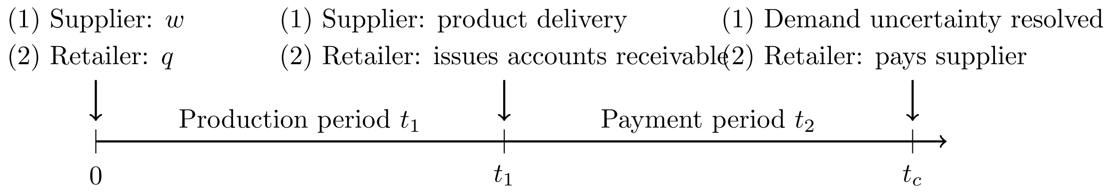


Figure 2.1.: Timing of basic supply chain events

We assume that the supplier is capital-constrained with zero initial cash position. Given the wholesale contract (w, q), the supplier needs to raise capital at time 0 to cover the production cost cq, where c is the unit production cost. The above mentioned financing is before the production and shipment of goods, so we call it the *pre-shipment financing*. In this chapter, we consider two types of pre-shipment financing schemes: (i) *purchase order financing* (POF), in which the supplier borrows cq from the bank with the support of the retailer's purchase order wq, and pays off the bank loan at the end of the business cycle t_c; (ii) *buyer direct financing* (BDF), in which the downstream retailer prepays the supplier cq to cover his production cost, and pays off the remaining accounts receivable $(w - c)q$ at time t_c with the collected sales revenue.

Note that, due to the retailer's payment delay, the supplier has zero cash on hand during the payment period. During this time period, the supplier may face a liquidity shock, which leads to bankruptcy if there are insufficient funds to meet the financial obligations [61, 85]. We assume that the liquidity shock happens with probability ρ_s. Our model of liquidity risk is motivated by the practical observations where firms may have enough value in total assets, but they will default and could eventually enter bankruptcy if there isn't enough cash to meet the short-term obligations (namely, liquidity crisis). To avoid such a potential liquidity crisis, firms can maintain a sufficient self-financed reserve on hand.[8] In our supply chain setting, the supplier could adopt *post-shipment financing* to avoid such liquidity risks, where accounts receivable is sold to increase cash position in order to meet the working capital needs of daily business operations [61]. In

[8] https://www.investopedia.com/terms/l/liquidity-crisis.asp

§2.7.1, we extend our model to consider fire sale of accounts receivable, where, instead of going bankrupt upon unprepared liquidity shock, the supplier can liquidate his accounts receivable via fire sale, but at heavily discounted prices.

In this chapter, we study two types of post-shipment financing schemes: (i) factoring, in which the supplier's accounts receivable is sold to a financial institution (e.g., factor or bank) for immediate cash; (ii) invoice trading, in which accounts receivable is sold on an invoice trading platform. Note that traditional factoring requires tedious verification and paperwork, which takes time (e.g., up to a week) to complete the transaction.[9] To prepare for the potential liquidity shock, the supplier has to make the factoring decision early in the payment period. Without loss of generality, we assume the supplier's factoring decision is made at time t_1 immediately after the accounts receivable is issued.[10] However, invoice trading has the advantage of fast transaction speed (e.g., within 24 hours) thanks to the online FinTech platforms (e.g., MarketFinance).[11] Therefore, the supplier is still able to sell accounts receivable to avoid bankruptcy immediately after observing the liquidity shock.

The retailer has adequate cash to pay off the outstanding accounts receivable and cover the potential profit loss due to demand uncertainty. Since the retailer has adequate cash on hand, she faces no liquidity risk. However, the retailer is exposed to credit risk, leading to breach of contract and default to pay off the accounts receivable at time t_c with probability ρ_r.[12] When the retailer offers direct financing (i.e., BDF) to the supplier, an opportunity cost of capital r_v is incurred, where the retailer could otherwise invest in other businesses to make profits.

We normalize the exogenous retail price to $p = 1$ and the risk-free interest rate to zero ($r_f = 0$). We adopt continuous compounding for all interest-related calculations. We assume the financial market is fully competitive. All interest rates are fairly priced, with

[9] https://fundbox.com/resources/guides/invoice-factoring/

[10] Note that the cash advance from factoring cannot be used to early repay the bank loan at time t_1 because otherwise the supplier does not have sufficient funds to fully hedge against the liquidity risk over the payment period.

[11] https://marketfinance.com/solutions/selective-invoice-finance

[12] Credit risk is a default risk where one of the two counterparts breaks the business contract and causes loss for the other party [61]. Incorporating the supplier's credit risk into our model does not affect our main analysis and results, so we assume away the supplier's credit risk for expositional clarity.

different risk premiums charged by different financial institutions or financing providers. The bank's premium is denoted as $\eta_{\mathcal{B}}$; factoring requires a premium $\eta_{\mathcal{F}}$; invoice trading has a premium $\eta_{\mathcal{I}}$. We assume the unit production cost c is not too large, i.e., $c \leq (1-\rho_s)(1-\rho_r)e^{-\eta_{\mathcal{B}} t_c}$. Otherwise, the supply chain transaction is not profitable and the supplier never produces. The factoring premium $\eta_{\mathcal{F}}$ is not too large to avoid the trivial case where the supplier never adopts factoring: $(1-\rho_s)e^{\eta_{\mathcal{F}} t_2} \leq 1$. The market demand D is a non-negative random variable with p.d.f $f(\cdot)$, c.d.f. $F(\cdot)$, and complementary c.d.f. $\bar{F}(\cdot) = 1 - F(\cdot)$. We make the following assumptions about the demand distribution F: (i) it has a continuous p.d.f., with $f(\xi) > 0$ in $[0, \mathbb{Z}]$ ($\mathbb{Z} \leq +\infty$); (ii) the generalized failure rate $\frac{\xi f(\xi)}{1-F(\xi)}$ is strictly increasing in ξ, i.e., IGFR. We denote the failure rate function as $z(\xi) = f(\xi)/\bar{F}(\xi)$.

2.4 The Baseline Trade Finance Model

In this section, our analysis focuses on the baseline trade finance model, where purchase order financing (POF) is provided as pre-shipment financing to cover the production cost, factoring as post-shipment financing to liquidate accounts receivable to mitigate liquidity risk over the payment period. We start with the analysis of the case without smart contract, followed by the one with smart contract, where a comparison between the profits of the two cases is provided and we characterize the value of smart contract.

2.4.1 Equilibrium Analysis under Baseline Model

Following the backward induction approach, we first analyze the supplier's factoring decision at time t_1 given all the time-0 decisions (i.e., the loan rate r, the wholesale price w, and the retailer's order quantity q). The supplier's expected profit of adopting factoring can be written as

$$\pi_{\mathcal{B}}^{t}(w,q,r) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{rt_c}. \tag{2.1}$$

The first term $(1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq$ is the value of accounts receivable fairly priced by factoring (i.e., cash amount received by the supplier at time t_1), in which the term $(1-\rho_r)$ is the retailer's probability to survive the credit risk, and $e^{-\eta_{\mathcal{F}}t_2}$ is the premium discounting over the payment period t_2. Note that factor does not receive the payment if the retailer defaults due to credit risk, and thus the accounts receivable wq is discounted by $(1-\rho_r)$. The second term cqe^{rt_c} is the bank loan principal plus accrued interest to be paid at time t_c. On the one hand, the supplier's expected profit of not adopting factoring is

$$\pi_{\mathcal{B}}^{n}(w,q,r) = (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c}), \tag{2.2}$$

where $(1-\rho_s)(1-\rho_r)$ is the probability of surviving the supplier's own liquidity risk and the retailer's credit risk, as either of the two risks would lead to the supplier's bankruptcy. $(wq - cqe^{rt_c})$ is the supplier's profit in the absence of those two risks, in which the bank loan cqe^{rt_c} is deducted from the collected wholesale revenue wq. To summarize, given decisions (w,q,r), the supplier adopts factoring if and only if $\pi_{\mathcal{B}}^{t}(w,q,r) \geq \pi_{\mathcal{B}}^{n}(w,q,r)$. Hereafter we use the superscript t (resp., n) to denote the case that accounts receivable is traded (resp., not traded).

Next, we consider the bank's interest rate decision at time 0. Note that the supplier's factoring decision at time t_1 affects the default risk throughout the payment period, and thus has an impact on the bank loan repayment. For the convenience of illustration, let us temporarily assume the bank knows the supplier's time-t_1 factoring decision when making the interest rate decision at time 0. If factoring is adopted, the supplier is guaranteed to repay the bank loan cqe^{rt_c} with the liquidated wholesale revenue $(1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq$. As a response, the bank simply charges $r_t = \eta_{\mathcal{B}}$ for such a risk-free loan. If factoring is not adopted (i.e., holding accounts receivable on hand), the bank's interest rate r_n is determined by the competitive lending equation $e^{\eta_{\mathcal{B}}t_c} = (1-\rho_s)(1-\rho_r)e^{r_n t_c}$, where the bank is indifferent between using cash cq to earn a total capital return $cqe^{\eta_{\mathcal{B}}t_c}$ and issuing the loan with repayment $cqe^{r_n t_c}$ if the supplier survives the liquidity risk and the retailer's credit risk.

However, the supplier is unable to commit the time-t_1 factoring decision when borrowing from the bank at time 0. As a consequence, upon loan request, the bank anticipates the supplier's factoring decision by evaluating the wholesale contract (w, q). For notational convenience, we denote

$$\bar{w} = \frac{[\rho_s + (1-\rho_s)\rho_r]ce^{\eta_{\mathcal{B}}t_c+\eta_{\mathcal{F}}t_2}}{[1-(1-\rho_s)e^{\eta_{\mathcal{F}}t_2}](1-\rho_r)}. \tag{2.3}$$

which, as shown in Lemma 2, represents the wholesale price threshold above which the bank is convinced that the supplier would adopt factoring. Given the wholesale contract (w, q), the bank's loan rate and the supplier's factoring decision are summarized as follows.

Lemma 2 *If $w \leq \bar{w}$, the bank offers the loan rate r_n and the supplier does not adopt factoring; otherwise, the bank offers the loan rate r_t and the supplier adopts factoring.*

Lemma 2 reveals an important observation that the bank loan rate and the supplier's factoring decision critically depend on the wholesale price w. From Equation (2.1) and (2.2), factoring adoption helps the supplier mitigate the potential liquidity shock, though at some expense of factoring discount $e^{-\eta_{\mathcal{F}}t_2}$, which overall increases the supplier's expected wholesale revenue from $(1-\rho_s)(1-\rho_r)wq$ to $(1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq$. However, on the cost side, factoring increases the supplier's survival probability from $(1-\rho_s)(1-\rho_r)$ to 1 (which means guaranteed loan repayment), and thus also increases the expected bank loan cost (i.e., effective production cost) from $(1-\rho_s)(1-\rho_r)cqe^{rt_c}$ to cqe^{rt_c}. The intuition of factoring decision boils down to the supplier's trade-off between a higher wholesale revenue and a lower bank loan cost. Therefore, there exists an indifference level of the wholesale price (referred to as the *safe loan threshold* $\bar{w}$) above which the supplier chooses to adopt factoring. Anticipating that, the bank offers the risk-free loan rate $r_t = \eta_{\mathcal{B}}$. From the safe loan threshold $\bar{w}$ formula in (2.3), we see the impact of both the supplier's liquidity risk ρ_s and the retailer's credit risk ρ_r. These two risks are countervailing forces in play. On the one hand, $\bar{w}$ is decreasing in ρ_s. The intuition is that as ρ_s increases, the supplier's expected profit of not adopting factoring $\pi_{\mathcal{B}}^n(w, q)$ is decreasing while the profit of adopting factoring $\pi_{\mathcal{B}}^t(w, q)$ is unchanged. Hence, the bank anticipates that the supplier is more likely to adopt factoring, and a lower safe loan threshold $\bar{w}$ is required.

On the other hand, $\bar{w}$ is increasing in ρ_r. The reason is that a higher credit risk ρ_r reduces the expected accounts receivable value $(1-\rho_r)wq$, which makes the wholesale revenue less weighted relative to the bank loan cost. As a result, the supplier is less likely to adopt factoring and the bank requires a higher safe loan threshold $\bar{w}$.

Next, we turn to the retailer's order quantity decision. Given the wholesale price w, the retailer's optimal order quantity is derived from the expected profit $\Pi_{\mathcal{B}}(w,q) = (1-\rho_r)[S(q)-wq]$, where $(1-\rho_r)$ is the probability of surviving the credit risk, and $S(q) = \mathbb{E}[\min(D,q)]$ is the expected sales revenue given the inventory position q. Hence, it is straightforward to deduce that the retailer's optimal order quantity is $q^*(w) = \bar{F}^{-1}(w)$. It is worth noticing the retailer's profit is not affected by the supplier's liquidity risk, and the retailer has to make a payment to whoever owns the accounts receivable, regardless of whether the supplier goes bankrupt or not. Besides, our stylized model of liquidity shock-driven bankruptcy only serves as a proxy of the supplier's cost due to liquidity shock. For example, expensive fire sales of the accounts receivable can be adopted to avoid bankruptcy (see §2.7.1 for more details). Therefore, the retailer's liability of the accounts receivable is unaffected.

Finally, we derive the supplier's optimal wholesale price. From Lemma 2, we know the supplier's wholesale price w at time 0 signals to the bank his factoring decision at time t_1, and the bank determines the corresponding loan rate. With this in mind, we can further derive the supplier's wholesale price decision by optimizing it over the two regions with respect to w, divided by the safe loan threshold $\bar{w}$ (see Figure B.2 for a numerical illustration of the supplier's profit function). We first analyze the following two scenarios independently and then we derive the equilibrium based on the analysis: (i) adopting factoring is exogenously given and the bank offers the corresponding loan rate r_t; (ii) not adopting factoring is exogenously given and the bank offers loan rate r_n. For expositional convenience, we define $c^t_{\mathcal{B}} = \frac{ce^{\eta_{\mathcal{B}}t_c+\eta_{\mathcal{F}}t_2}}{1-\rho_r}$ (resp., $c^n_{\mathcal{B}} = \frac{ce^{\eta_{\mathcal{B}}t_c}}{(1-\rho_s)(1-\rho_r)}$) as the supplier's effective unit production cost when factoring is (resp., not) adopted. Under scenario (i), plugging loan rate r_t into Equation (2.1), we derive the supplier's profit of adopting factoring:

$$\pi_{\mathcal{B}}^{t}(w,q) = (1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq - cqe^{\eta_{\mathcal{B}}t_c}. \tag{2.4}$$

Combining with the retailer's order quantity $q^*(w) = \bar{F}^{-1}(w)$, we can derive the optimal wholesale price and order quantity $(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$ from $\bar{F}(q_{\mathcal{B}}^{t*})[1-q_{\mathcal{B}}^{t*}z(q_{\mathcal{B}}^{t*})] = c_{\mathcal{B}}^{t}$ and $w_{\mathcal{B}}^{t*} = \bar{F}(q_{\mathcal{B}}^{t*})$ (see the proof of Proposition 2.4.1 for technical details). Under scenario (ii), plug bank loan rate r_n into Equation (2.2), and we derive the supplier's profit of not adopting factoring:

$$\pi_{\mathcal{B}}^{n}(w,q) = (1-\rho_s)(1-\rho_r)wq - cqe^{\eta_{\mathcal{B}}t_c}. \tag{2.5}$$

Similar as the previous analysis, we are able to derive the optimal decisions $(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$ from equation $\bar{F}(q_{\mathcal{B}}^{n*})[1-q_{\mathcal{B}}^{n*}z(q_{\mathcal{B}}^{n*})] = c_{\mathcal{B}}^{n}$ and $w_{\mathcal{B}}^{n*} = \bar{F}(q_{\mathcal{B}}^{n*})$. To simplify the exposition of the next proposition, we denote $\rho_s = \beta_2$ as the solution to $\bar{w} = w_{\mathcal{B}}^{t*}$. So $\rho_s = \beta_2$ represents the situation where the supplier is just able to charge the optimal wholesale price $w_{\mathcal{B}}^{t*}$ without the need to "overprice" the wholesale price (i.e., $\bar{w} \leq w_{\mathcal{B}}^{t*}$) in order to convince the bank of adopting factoring. However, if $\bar{w} > w_{\mathcal{B}}^{t*}$, to convince the bank of adopting factoring, the supplier needs to overcharge the wholesale price to $\bar{w}$. We denote $\rho_s = \beta_1$ as the solution to $\pi_{\mathcal{B}}^{n}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*}) = \pi_{\mathcal{B}}^{t}(\bar{w}, q^*(\bar{w}))$, in which $q^*(\bar{w}) = \bar{F}^{-1}(\bar{w})$. So $\rho_s = \beta_1$ represents the situation where the profit of adopting factoring with the least distorted price $\bar{w}$ just equals the supplier's optimal profit of not adopting factoring $\pi_{\mathcal{B}}^{n}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$. In other words, if $\bar{w}$ becomes higher, the price distortion is too severe. The supplier finds adopting factoring not profitable anymore and would switch to not adopting factoring. For notational convenience, we further denote $\tilde{w} = \bar{w}(\beta_1)$ as the above mentioned largest safe loan threshold such that adopting factoring is still profitable for the supplier. We summarize the equilibrium of the baseline model in the following proposition.

Proposition 2.4.1 *Under the baseline trade finance model, the equilibrium is*

$$(r_{\mathcal{B}}^{o*}, w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*}) = \left\{ \begin{array}{ll} (r_n, w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*}) & \textit{if} \;\; \rho_s < \beta_1 \\ \left(r_t, \bar{w}, \bar{F}^{-1}(\bar{w})\right) & \textit{if} \;\; \beta_1 \leq \rho_s < \beta_2 \\ \left(r_t, w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*}\right) & \textit{if} \;\; \rho_s \geq \beta_2 \end{array} \right\}.$$

The supplier adopts factoring if and only if $\rho_s \geq \beta_1$.

Proposition 2.4.1 shows that the equilibrium outcome can be divided into three cases based on the value of the supplier's liquidity risk ρ_s. When liquidity risk is high (i.e., $\rho_s \geq \beta_2$), the supplier has a strong incentive to adopt factoring (so as to mitigate highly possible liquidity shock). As such, the supplier is able to credibly convince the bank of his factoring adoption by charging the optimal wholesale price $w_{\mathcal{B}}^{t*}$, because in this case, the safe loan threshold $\bar{w}$ is even lower (i.e., $w_{\mathcal{B}}^{t*} \geq \bar{w}$). When liquidity risk is medium (i.e., $\beta_1 \leq \rho_s < \beta_2$), factoring is adopted with an upward wholesale price distortion. That is, the supplier offers the wholesale price $\bar{w}$, which exceeds the optimal wholesale price of factoring $w_{\mathcal{B}}^{t*}$. Recall that $\bar{w}$ decreases in ρ_s. As ρ_s decreases from high to medium level, $\bar{w}$ exceeds $w_{\mathcal{B}}^{t*}$, where the overpricing of wholesale price $\bar{w}$ is required for factoring adoption. When liquidity risk is low (i.e., $\rho_s < \beta_1$), the safe loan threshold $\bar{w}$ is substantially high. To convince the bank of factoring adoption, a heavily overpriced wholesale price is required, leading to a sharp decrease in the supplier's profit. Meanwhile, the supplier is exposed to a relatively low liquidity risk if factoring is not adopted. As a result, the supplier chooses not to adopt factoring. As we can see from the above discussion, the supplier's profit is reduced when $\rho_s < \beta_2$, caused by the high safe loan threshold. Note that the safe loan threshold exists because the supplier cannot commit the post-shipment financing decision upon the pre-shipment loan request. Here and in the sequel, we refer to such inefficiency as the *commitment frictions*, and Figure 2.2 offers a succinct summary of the key ideas discussed above.

There are two points worth attention. First, the supplier's self-interest causes the overpricing in wholesale price offering, but such price distortion can substantially reduce the retailer's profit. As we can see, from the left panel of Figure 2.3, the retailer's profit has a

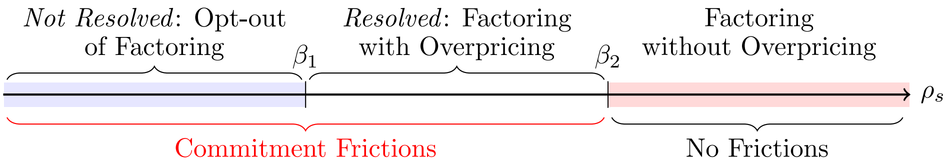


Figure 2.2.: An illustration of commitment frictions and factoring adoption in different liquidity risk regions

discontinuity at β_1, specifically, a downward jump. Recall that this is the point at which the supplier is indifferent between adopting and not adopting factoring. To understand the nature of this discontinuity, the right panel in Figure 2.3 plots the supplier's wholesale price and the retailer's order quantity. As ρ_s increases across this threshold β_1, the supplier switches to adopting factoring, which is achieved by an upward jump in wholesale price decision (so as to convince the bank about such factoring adoption decision). As a response, the retailer's order quantity is sharply reduced, leading to the downward jump of retailer's profit at β_1. Second, contrary to the conventional wisdom, a higher liquidity risk might benefit both the supplier and the retailer within a certain medium range. That is, when $\beta_1 \leq \rho_s < \beta_2$, the supplier's and the retailer's profits increase in ρ_s (see Figure 2.3). The intuition can be explained as follows. As factoring is adopted in this region, the liquidity risk can be completely eliminated, and thus the liquidity risk increase has no direct impact on the supplier's profit. Meanwhile, a larger liquidity risk makes it less costly for the supplier to credibly convince the bank of the factoring adoption decision (as $\bar{w}$ decreases). In other words, the wholesale price distortion narrows thanks to the liquidity risk increase. As a result, the supply chain interaction becomes more efficient, adding value to both the supplier and the retailer.

We conclude this subsection with a brief discussion of the value of factoring for the supply chain. Factoring serves as a financing option that allows the supplier to cash out from the illiquid accounts receivable so as to prepare against potential liquidity shock. Factoring adoption has the following two benefits: (i) It helps the supplier successfully eliminate the potential liquidity risk; (ii) It can help smooth out the supplier's revenue volatility by selling the retailer's credit risk (embedded in the accounts receivable) to the factor. Therefore, if factoring is adopted, the supplier is guaranteed to be able

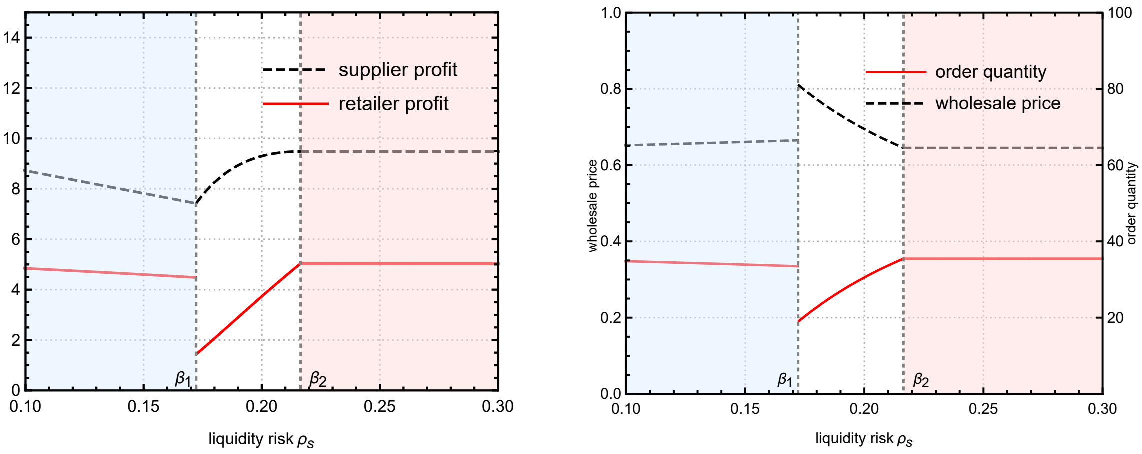


Figure 2.3.: Optimal profits and decisions under the baseline model
($c = 0.2$, $\rho_r = 0.2$, $\eta_{\mathcal{F}} = 0.3$, $\eta_{\mathcal{B}} = 0.3$, $t_1 = 0.1$ and $t_2 = 0.2$)

to repay the bank loan, which leads to a risk-free loan. As expected, such liquidity risk mitigation benefits the supplier, and the benefit is significant when the supplier's liquidity risk is relatively high. Meanwhile, as discussed above, there exists a region of the liquidity risk within which factoring would cause the supplier's wholesale price distortion. Such overpricing would increase the retailer's ordering cost and thus decreases the order quantity. Hence, the retailer's profit might be hurt due to the supplier's adoption of factoring (see Proposition B.0.4 for a formal statement and Figure B.3 for an illustration).

2.4.2 The Value of Smart Contract

Smart contract contains different sets of pre-defined conditions and the corresponding executions. When one specific condition is met, the corresponding decision will be automatically executed [86]. More specifically, the supplier's factoring decision clearly affects his cash flow and associated liquidity risk, and therefore, his capability of full repayment of the bank loan. However, the supplier cannot commit the future factoring decision (at time t_1) when borrowing from the bank (at time 0), which causes the commitment frictions. Traditional banks fail to make credible contracts to resolve such a commitment issue due to contract incompleteness.[13] By adopting smart contract, such frictions can

[13]For a good reference on the costs of writing and enforcing complete contracts, please refer to [87] and [88].

be resolved: the bank loan interest rate is contingent on the supplier's factoring decision. More specifically, in our setting, the bank is able to set different loan rates at time 0 (corresponding executions) based on the supplier's future factoring adoption decision at time t_1 (pre-defined conditions). Therefore, smart contract can help resolve contract incompleteness and mitigate the commitment frictions as we highlighted above.

We briefly discuss the game structure before presenting the equilibrium outcome under smart contract adoption. Following backward induction, we start with the supplier's factoring decision. In §2.4.1, we have shown that if the supplier can commit to adopting factoring (i.e., scenario (i) as discussed), the bank loan rate is $r_t = \eta_{\mathcal{B}}$ and the supplier's profit is $\pi^t_{\mathcal{B}}(w, q) = (1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq - cqe^{\eta_{\mathcal{B}}t_c}$. If the supplier can commit to not adopting factoring (i.e., scenario (ii) as discussed), the bank loan rate is r_n and the supplier's profit is $\pi^n_{\mathcal{B}}(w, q) = (1-\rho_s)(1-\rho_r)wq - cqe^{\eta_{\mathcal{B}}t_c}$. It is clear that $\pi^n_{\mathcal{B}}(w, q) \leq \pi^t_{\mathcal{B}}(w, q)$ for $\forall(w, q)$. That is, factoring helps avoid the supplier's liquidity risk (though at some mild factoring cost) and the overall effect is an increased expected wholesale revenue. Compared with the baseline model, the negative effect of factoring adoption (i.e., wholesale price distortion) is avoided since the commitment frictions are resolved due to smart contract implementation. As a result, factoring is always adopted by the supplier in the presence of smart contract. Next, as for the retailer's order quantity decision, we have shown in §2.4.1 that it is independent of the supplier's factoring adoption decision, and thus is not affected by the smart contract adoption. In light of the above, we can readily characterize the equilibrium outcome and the value of smart contract in the next proposition.

Proposition 2.4.2 *(i) With smart contract, we have the unique equilibrium* $(q^{t*}_{\mathcal{B}}, w^{t*}_{\mathcal{B}})$*, where the supplier always adopts factoring and the bank offers loan rate* $r_t = \eta_{\mathcal{B}}$*.*

(ii) If $\rho_s < \beta_2$*, smart contract adoption strictly reduces the wholesale price and increases the order quantity, which strictly benefits both the supplier and the retailer.*

Recall that without smart contract, the supplier does not have the commitment power and might intentionally deviate from the optimal wholesale price so as to convince the bank of his adoption of factoring, or simply opt out of factoring (as illustrated in Figure

2.2). Driven by such commitment frictions, the retailer's order quantity is significantly reduced, resulting in strong shrinkage in both the supplier's and the retailer's profits. Smart contract adoption helps eliminate such commitment frictions and thus benefits the supply chain. The result is depicted in Figure 2.4. When $\rho_s \geq \beta_2$, high liquidity risk makes factoring appealing to the supplier, and the bank can be easily convinced without price distortion (i.e., $\bar{w} \leq w_{\mathcal{B}}^{t*}$). Hence, the supplier always charges the optimal wholesale price $w_{\mathcal{B}}^{t*}$ with and without smart contract. As there is no commitment issue involved, smart contract does not add value to the supply chain. When $\rho_s < \beta_2$, without smart contract, the supplier either overprices the wholesale price or even forgoes the adoption of factoring. With smart contract, the supplier is able to commit to adopting factoring, which enables the supplier to get a cheaper loan rate from the bank without any wholesale price distortion. As a result, smart contract adoption strictly benefits both the supplier and the retailer when $\rho_s < \beta_2$.

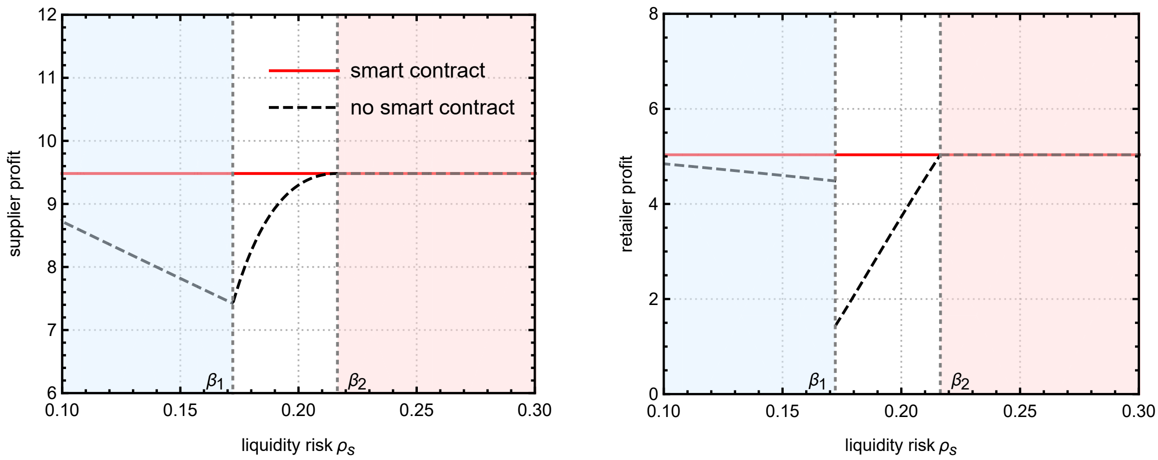


Figure 2.4.: Supplier and retailer profits with and without smart contract under the baseline model
($c = 0.2$, $\rho_r = 0.2$, $\eta_{\mathcal{F}} = 0.3$, $\eta_{\mathcal{B}} = 0.3$, $t_1 = 0.1$ and $t_2 = 0.2$)

2.5 Alternative Pre-Shipment Financing: Buyer Direct Financing

In this section, we investigate an alternative pre-shipment financing scheme: buyer direct financing (BDF). Capital adequate retailers like Walmart offer buyer direct financing, which serves as an alternative financing option to the supplier in addition to POF (bank financing). Note that under BDF, the retailer does not charge any interest for the financing offered to the supplier. Thus, our model of BDF is essentially equivalent to cash-in-advance (also called advance payment or prepayment) and such zero-interest BDF is widely adopted in practice [40, 89]. For example, [90] finds that financially constrained suppliers are partly paid in advance by their customers. Survey evidence in [91] indicates that cash-in-advance accounts for 22% of global trade finance. However, the retailer would incur an opportunity cost of capital r_v when offering BDF to the supplier. As [92] points out, banks evaluate SMEs' credits by critically relying on soft information such as past interactions with suppliers, customers, competitors, and other businesses. Therefore, when the supplier is financed by BDF to cover production cost, the post-shipment factoring premium (denoted as $\eta_{\mathcal{E}}$) will be lower compared to the POF case. That is, we assume $\eta_{\mathcal{E}} \leq \eta_{\mathcal{F}}$ throughout this section. In this section, we first explore the case where BDF is exogenously adopted (§2.5.1), then we endogenize the retailer's BDF offering decision (§2.5.2). We attempt to answer the following two research questions. First, how much value does BDF bring to the supply chain? Second, what's the value of smart contract when BDF is available as an alternative pre-shipment financing scheme?

2.5.1 Equilibrium Analysis under BDF

We start with the supplier's factoring decision. BDF covers the supplier's production cost cq at time 0, and the remaining wholesale revenue $wq - cq$ is due at time t_c. If the supplier adopts factoring at time t_1, the profit is $(1-\rho_r)e^{-\eta_{\mathcal{E}} t_2}(wq - cq)$; otherwise, the profit is $(1-\rho_s)(1-\rho_r)(wq - cq)$. Since $(1-\rho_s)e^{\eta_{\mathcal{F}} t_2} \leq 1$ and $\eta_{\mathcal{E}} \leq \eta_{\mathcal{F}}$, it is straightforward to see that factoring benefits the supplier, and thus is always adopted under BDF. Compared with the baseline model, BDF poses the following two positive impacts on factoring: (i) There is not commitment frictions between the bank and the supplier like in the baseline model; (ii) The factoring cost is also lower as $\eta_{\mathcal{E}} \leq \eta_{\mathcal{F}}$. Next, we derive the retailer's optimal order quantity. The retailer's profit becomes $\Pi_{\mathcal{E}}(w, q) = (1-\rho_r)[S(q) - (wq - cq) - cqe^{r_v t_c}]$, where $S(q)$ is the expected sales revenue, $wq - cq$ represents the outstanding accounts receivable due at time t_c, and $cqe^{r_v t_c}$ represents the BDF amount cq plus the associated cost of capital throughout the whole business cycle. Given the wholesale price w, we can derive the retailer's optimal order quantity as $\tilde{q}^*(w) = \bar{F}^{-1}[w + cq(e^{r_v t_c} - 1)]$. Given the above discussion, we can readily derive the equilibrium given exogenously adopted BDF (see Lemma 4).

Bypassing the technical details of the equilibrium outcome, it is important to emphasize the following point regarding the financing cost reallocation with BDF adoption. Note that for a given wholesale price w, the retailer orders $\tilde{q}^*(w) = \bar{F}^{-1}[w + cq(e^{r_v t_c} - 1)]$ under the BDF, which is smaller than the order quantity $q^*(w) = \bar{F}^{-1}(w)$ under the POF. The reason is that when offering BDF, the retailer bears extra financing cost $cq(e^{r_v t_c} - 1)$ in addition to the wholesale procurement cost wq. On the other hand, BDF saves the supplier's financing cost (no bank loan is needed), which motivates the supplier to offer a lower wholesale price. A natural and interesting question to ask is: How would such financing cost reallocation impact the supply chain profitability, in particular, when the two financing schemes are equally costly? We say the BDF shares the same financing cost as POF when $r_v = \eta_{\mathcal{B}}$.

It is interesting to observe that BDF has an overall better performance with increased order quantity and supply chain profit. Mathematically speaking, under BDF, the order quantity is determined by $\bar{F}(q_{\mathcal{E}}^*)[1 - q_{\mathcal{E}}^* z(q_{\mathcal{E}}^*)] = ce^{r_v t_c}$; under POF, the order quantity is determined by $\bar{F}(q_{\mathcal{B}}^{t*})[1 - q_{\mathcal{B}}^{t*} z(q_{\mathcal{B}}^{t*})] = c_{\mathcal{B}}^t = ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}/(1 - \rho_r)$. When $r_v = \eta_{\mathcal{B}}$, it is straightforward to see $c_{\mathcal{B}}^t = ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}/(1 - \rho_r) \geq ce^{\eta_{\mathcal{B}} t_c} = ce^{r_v t_c}$. Hence, BDF increases the order quantity, and the impact on profit immediately follows. This is driven by the fact that the financing cost saving for the supplier is more valuable relative to the financing cost saving for the retailer if not offering BDF. Summarising, this finding sheds light on the value of financing cost reallocation, or more specifically, downstream capital injection, for the whole supply chain system.

2.5.2 The Value of BDF

To fairly quantify the value of BDF, we further consider the BDF adoption incentives for both the retailer and the supplier, i.e., endogenizing the retailer's and the supplier's adoption decisions. Long term BDF decision is made by the retailer first, then the supplier determines the wholesale price, followed by the retailer's order quantity decision. We denote $r_v = \bar{r}_v^o$ as the solution to $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$, which represents the threshold value of cost of capital at which the retailer is indifferent between offering and not offering BDF. The equilibrium is summarized in the following proposition.

Proposition 2.5.1 *(i) The retailer offers BDF if and only if $r_v \leq \bar{r}_v^o$, and the supplier always accepts BDF if it is offered by the retailer.*

(ii) The threshold $\bar{r}_v^o$ increases in the retailer's credit risk ρ_r, and it first increases, then decreases, and finally remains unchanged as the supplier's liquidity risk ρ_s increases.

Proposition 2.5.1(i) asserts that the retailer would only offer BDF when the cost of capital is below a threshold, i.e., the point at which her profit is equal to the profit she will make when the supplier borrows from the bank (i.e., POF). As we have discussed above, BDF essentially reallocates the financing cost from the supplier to the retailer. As the cost of capital r_v increases, offering BDF becomes more expensive for the retailer. Therefore, the retailer would offer BDF if and only if $r_v \leq \bar{r}_v^o$. While this can be anticipated, what is more interesting is how this BDF adoption threshold $\bar{r}_v^o$ would be affected by the retailer's own credit risk ρ_r and the supplier's liquidity risk ρ_s. Proposition 2.5.1(ii) summarizes the result and the intuition boils down to the retailer's trade-off between the capital cost (associated with BDF offering) and the wholesale cost (due to the supplier's overpricing behavior without BDF offering).

First, BDF offering decision does not directly affect the retailer's attitude towards her own credit risk since this risk is inevitable regardless of whether offering BDF or not. However, if BDF is not adopted, the retailer's credit risk increase would cut down the supplier's wholesale margin, induce the supplier to charge a higher wholesale price, and consequently increase the wholesale cost for the retailer. As a result, the retailer is more willing to sacrifice the capital cost by offering BDF, under which the supplier's wholesale price decision can be successfully corrected. This explains why the adoption threshold $\bar{r}_v^o$ increases as the retailer's credit risk increases. Second, if BDF is offered/adopted, the supplier always adopts factoring and thus his liquidity risk has no impact on the supply chain interaction and efficiency. Hence, the supplier's liquidity risk increase only affects the wholesale cost for the retailer given BDF is not offered. The non-monotonic impact of the supplier's liquidity risk can be readily explained using the same line of reasoning given the non-monotonic impact on the retailer's profit illustrated earlier in Figure 2.3. In summary, when the retailer's cost of capital is not too high, BDF benefits the supply chain in the following two ways (the value of BDF is quantified in Proposition B.0.5). First, BDF resolves the commitment issue. As previously discussed, without BDF, the supplier might need overpricing in the wholesale price offering in order to get a cheaper loan rate from the bank, which heavily cuts down the supply chain efficiency. With BDF, the supplier would be able to retain the optimal wholesale price, in the absence of interactions with the bank. Second, BDF also helps reallocate the financing cost from the capital-constrained supplier to the capital-adequate retailer, which could be a win-win solution if the retailer's cost of capital is relatively low. Taken together, these benefits also provide a novel economic microfoundation to the adoption of BDF in supply chains.

2.5.3 The Value of Smart Contract under BDF

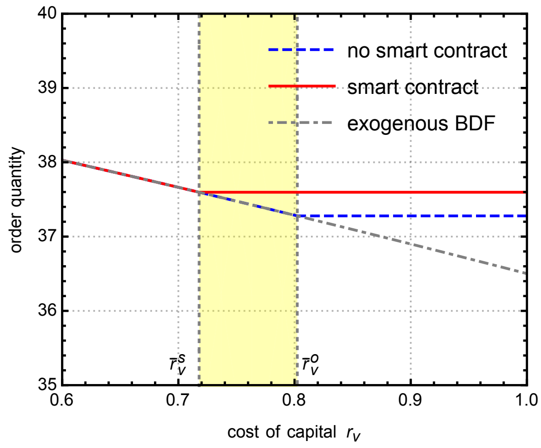
We have shown that smart contract enhances the supply chain value under the baseline model. A natural question to ask is whether smart contract could still benefit the supply chain under BDF. Interestingly, our analysis suggests that smart contract adoption could hurt the supply chain profit when the cost of capital r_v is within a certain medium range. Similar to the definition of $\bar{r}_v^o$, we denote $r_v = \bar{r}_v^s$ as the solution to $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$, to present the cost of capital at which the retailer is indifferent between offering and not offering BDF with smart contract adoption. We also denote $r_v = \bar{r}_v^k$ as the solution to $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) + \pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*}) + \pi_{\mathcal{B}}^t(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$, where the supply chain profits are the same with and without BDF. The equilibrium adoption of BDF and the value of smart contract under BDF are summarized in Proposition 2.5.2.

Proposition 2.5.2 *(i) With smart contract adoption, the retailer offers BDF if and only if $r_v \leq \bar{r}_v^s$, and the supplier always accepts BDF if offered by the retailer.*

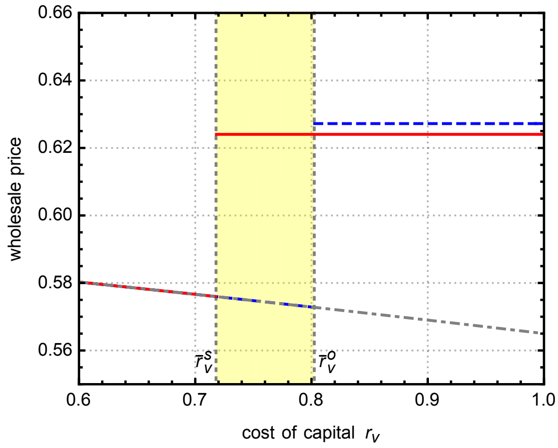
(*ii*) *Under BDF, smart contract always benefits the retailer, but hurts the supplier's profit when* $\bar{r}_v^s < r_v \leq \bar{r}_v^o$*, and even reduces the supply chain profit when* $\bar{r}_v^s < r_v < \min(\bar{r}_v^k, \bar{r}_v^o)$.

Proposition 2.5.2(i) implies that the equilibrium BDF adoption bears a similar structure as in the baseline model, with a different adoption threshold of the retailer's capital cost $\bar{r}_v^s$. It is worth noticing that this adoption threshold is smaller than the one without smart contract, $\bar{r}_v^s \leq \bar{r}_v^o$, which indicates that smart contract discourages the retailer from offering BDF to the supplier. Given the preceding discussion about the retailer's trade-off in determining the BDF offering, this finding is somewhat in line with our expectation. Recall that the retailer's trade-off is between the capital cost (with BDF offering) and the wholesale cost (without BDF offering), where the latter cost is largely cut down due to the adoption of smart contract (i.e., successful mitigation of commitment frictions). In a nutshell, the retailer is less willing to offer BDF when smart contract is in place.

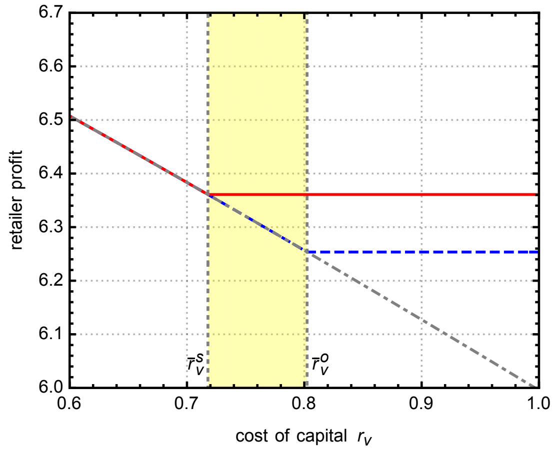
Proposition 2.5.2(ii) highlights the value of smart contract from the perspectives of different parties. It is not surprising to see that smart contract always benefits the retailer as it enhances the value of the no-BDF-offering option. However, the realization of such benefit may significantly hurt the supplier. In what follows, we provide more detailed discussions with the help of Figure 2.5, where the two thresholds $\bar{r}_v^s$ and $\bar{r}_v^o$ partition the range of r_v into three regions. When the cost of capital is low (i.e., $r_v \leq \bar{r}_v^s$), BDF enjoys an evident benefit of financing cost reallocation, and thus is adopted by the retailer regardless of whether smart contract is place or not. When the cost of capital is medium (i.e., $\bar{r}_v^s < r_v \leq \bar{r}_v^o$), the retailer only offers BDF when smart contract is not adopted. As previously discussed, when smart contract is not available, the high wholesale cost driven by the supplier's overpricing behavior forces the retailer to offers BDF, albeit at some mild capital cost. With smart contract, the retailer switches to not offering BDF in this region, which increases her own profit (Figure 2.5(c)), but sharply decreases the supplier's profit (Figure 2.5(d)) due to the losses from forgoing the financing cost reallocation benefit of BDF. When the cost of capital is high (i.e., $r_v > \bar{r}_v^s$), BDF becomes



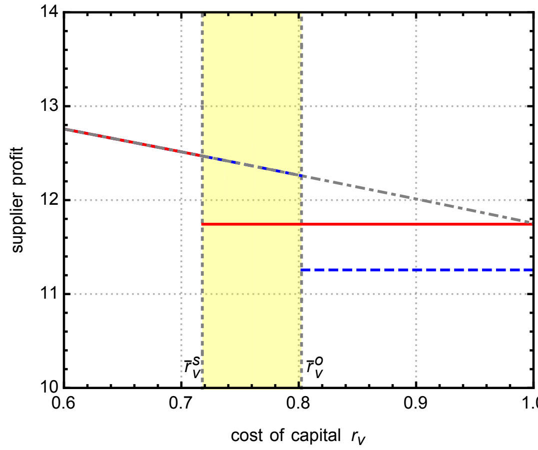
(a) Order quantity



(b) Wholesale price



(c) Retailer profit



(d) Supplier profit

Figure 2.5.: Equilibrium outcomes when BDF is available
($\rho_s = 0.1$, $\rho_r = 0.1$, $c = 0.2$, $\eta_{\mathcal{F}} = 0.4$, $\eta_{\mathcal{B}} = 0.1$, $\eta_{\mathcal{E}} = 0.1$, $t_1 = 0.1$ and $t_2 = 0.2$)

a nonviable option, and equilibrium outcome reduces to the one in the baseline model, where smart contract benefits the whole supply chain.

2.6 Alternative Post-Shipment Financing: Invoice Trading

In this section, we further study an emerging type of post-shipment financing scheme: on-demand invoice trading, as an alternative to traditional factoring. In recent years, the market for online invoice trading has grown substantially. [93] reported that the market volume of invoice trading more than tripled between 2013 and 2015 in the UK. From a global perspective, online invoice trading is likely to continue to grow further with the increasing popularity of various FinTech platforms (e.g., MarktetFinance, Timelio). As introduced earlier in the modeling framework section, invoice trading has faster transaction speed compared with traditional factoring. Facilitated with on-demand invoice trading platforms, the invoice financing requests typically get approved within 24 hours, whereas traditional factoring has to go through tedious inspections and paperwork (e.g., up to a week). As a result, the supplier is able to sell accounts receivable upon liquidity shock. Without loss of generality, we assume that the supplier's liquidity shock occurs at time t_1' during the payment period ($t_1' \geq t_1$), where there is t_2' time remaining towards the end of the payment period t_c (i.e., $t_1' + t_2' = t_c$).

Moreover, extensive anecdotal evidence suggests that businesses favor invoice trading over traditional factoring because of its main advantages in providing flexible and on-demand financing access via easy-to-use, efficient, and fast online application processes and funding platforms. However, the interest rates charged by the investors on invoice trading platforms vary over a wide range. For example, the data from MarketFinance shows that the annualized interest rate for funding the invoices ranges between 4.03% and 48.16% [69]. But the average rate is 12.28%, which is still relatively low compared to the factoring rates (24% to 54%). In this regard, we assume that invoice trading premium is not too high, i.e., $\eta_{\mathcal{I}} \leq \bar{\eta}_{\mathcal{I}}$ to be in line with piratical observations. It is necessary to point out that this assumption is less restrictive than $\eta_{\mathcal{I}} \leq \eta_{\mathcal{F}}$ (i.e., a lower premium than factoring). We relegate the detailed definition of $\bar{\eta}_{\mathcal{I}}$ into the proof of Proposition B.0.6. Given the generic preference of invoice trading, one would naturally expect the existence of the invoice trading option shall always benefit the supplier. However, our analysis finds counterintuitively that it might not be true in certain circumstances, which we will explain later.

2.6.1 Equilibrium Analysis under Invoice Trading

Note that when invoice trading is available, the supplier always adopts this post-shipment financing option (as preferable relative to factoring given its premium is not too high), where the trading decision (i.e., selling accounts receivable on the invoice trading platform) is made at time t_1' immediately after the liquidity shock is realized. If liquidity shock occurs, accounts receivable is sold via invoice trading with the discounted value $(1-\rho_r)e^{-\eta_{\mathcal{I}} t_2'}wq$, which is then used to pay off the bank loan. Hence, the supplier's profit of selling accounts receivable via invoice trading is $(1-\rho_r)e^{-\eta_{\mathcal{I}} t_2'}wq - cqe^{rt_c}$. If liquidity shock does not occur, holding accounts receivable on hand generates a higher profit of $(1-\rho_r)(wq-cqe^{rt_c})$. In brief, the supplier adopts invoice trading and liquidates accounts receivable if and only if liquidity shock occurs. Hence, the supplier's profit function can be written as $\pi_{\mathcal{I}}(w,q) = \rho_s[(1-\rho_r)e^{-\eta_{\mathcal{I}} t_2'}wq - cqe^{rt_c}] + (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$. Next, we solve for the bank's loan rate r_i. The bank anticipates the supplier's invoice trading decision as described above. With probability ρ_s, the supplier sells accounts receivable upon liquidity shock and the bank loan is repaid for sure. With probability $1-\rho_s$, the supplier holds accounts receivable on hand. The bank loan is paid off only if the retailer survives the credit risk ρ_r. Thus the bank loan rate r_i is determined by the following competitive pricing equation: $e^{\eta_{\mathcal{B}} t_c} = [\rho_s + (1-\rho_s)(1-\rho_r)]e^{r_i t_c}$. Further, we derive the retailer's optimal order quantity. The retailer's profit is still $\Pi_{\mathcal{I}}(w,q) = (1-\rho_r)[S(q)-wq]$, which gives the optimal order quantity as $q^*(w) = \bar{F}^{-1}(w)$. Finally, we solve for the supplier's wholesale price decision. Plugging r_i and $w = \bar{F}(q)$ into the supplier's profit $\pi_{\mathcal{I}}(w,q)$, we can rewrite it as a function of the order quantity q:

$$\pi_{\mathcal{I}}(q) = \left[1-\rho_s+\rho_s e^{-\eta_{\mathcal{I}} t_2'}\right](1-\rho_r)\bar{F}(q)q - cqe^{\eta_{\mathcal{B}} t_c}.$$

Similar as before, the equilibrium can be derived by solving a similar system of equations as in the baseline model, but with a different effective unit production cost $c_{\mathcal{I}}$ (details in Proposition B.0.6).

2.6.2 The Value of Invoice Trading

We have shown that BDF (as an alternative pre-shipment financing) benefits the supplier. A natural question to ask is whether invoice trading (as an alternative post-shipment financing) would benefit the supplier as well. Similar to BDF, invoice trading also corrects the supplier's wholesale price distortion (i.e., no overpricing). However, the mechanism is different. BDF resolves the price distortion since the supplier does not borrow from the bank. Therefore, the supplier does not need to distort the wholesale price to signal to the bank of his factoring decision anymore. With invoice trading, the supplier still borrows from the bank (at time 0), but the post-shipment financing decision (at time t_1) is now independent of the wholesale price decision (at time 0). As the bank knows that the supplier sells accounts receivable via invoice trading if and only if the liquidity shock occurs, the supplier has no incentive to go the extra mile to convince the bank about his future invoice trading decision through wholesale price distortion, like in the baseline model.

In light of the above discussion, one would expect the access to the invoice trading option should always benefit the supplier. Surprisingly, under mild conditions, we find such a preferable option might turn out to make the supplier worse-off, which is summarized in Proposition 2.6.1.

Proposition 2.6.1 *Compared with the baseline model,*

(i) *if $\eta_{\mathcal{I}} \leq \tilde{\eta}_{\mathcal{I}}$, invoice trading always benefits the supplier;*

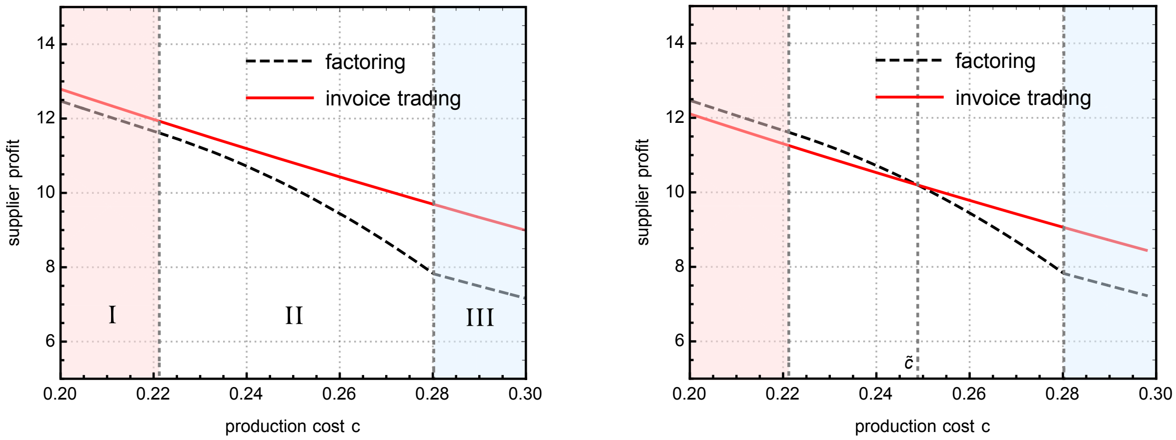
(ii) *if $\eta_{\mathcal{I}} > \tilde{\eta}_{\mathcal{I}}$, there exists a threshold $\tilde{c}$ of the unit production cost such that invoice trading hurts the supplier profit if and only if $c < \tilde{c}$.*

Recall that the supplier always chooses invoice trading over factoring at time t_1 given any time-0 decisions (w, q, r). However, Proposition 2.6.1(ii) identifies the condition under which invoice trading unexpectedly leads to a lower supplier's profit. Such an adoption dilemma of invoice trading is also driven by the lack of creditable commitment between the supplier and the bank. Invoice trading provides the supplier with trading flexibility, allowing the supplier to hold accounts receivable on hand if liquidity shock does not occur (so as to save the unnecessary post-shipment financing cost). However, from the bank's standpoint, such trading flexibility indirectly increases the failure risk of bank loan repayment. Specifically, in the absence of liquidity shock, the supplier would directly collect the wholesale revenue from the retailer. But, the retailer is subject to credit risk and thus may default, which inevitably leads to the supplier's default on the bank loan. Anticipating such increased risk exposure, the bank would charge a higher loan rate when invoice trading is adopted. Recall that if the supplier adopts factoring, the bank loan has a risk-free rate. Along with the trading flexibility benefit, the existence of invoice trading also makes the commitment frictions ubiquitous and completely unresolvable whereas it is partially existent and could be partially resolved under traditional factoring (see Figure 2.2). We refer to this phenomenon as the *commitment trap*. The relative cost of such a commitment trap depends on how much commitment frictions can be resolved via factoring adoption.

To better understand the nature of this adoption dilemma of invoice trading and get more insights into its implications on the actions of the supplier and the bank, let us decompose the value of invoice trading (relative to factoring) into two components: (i) the *explicit benefit*, which is the combined effect of trading flexibility and premium difference; (ii) the *implicit cost*, which is driven by the commitment trap of invoice trading. Since we focus on the case where the invoice trading premium is not too high (i.e., $\eta_{\mathcal{I}} \leq \bar{\eta}_{\mathcal{I}}$), the explicit benefit is always positive. Notice that even though the premium of invoice trading can be higher than factoring, the combined effect with trading flexibility is still positive.

In region III of Figure 2.6, the commitment frictions cannot be resolved even with factoring adoption, and thus the gap between the two profit lines represents the (constant) explicit benefit of invoice trading, or we can say the implicit cost of commitment trap is zero. As the unit production cost decreases to region II, the commitment frictions can be resolved with factoring adoption, though at the cost of overpricing. Thus, in this region, invoice trading bears a positive implicit cost due to the commitment trap. As the unit production cost decreases within region II, the cost of mitigating commitment frictions in factoring decreases, and thus the implicit cost component increases. When reaching region I, factoring can help fully resolve the commitment frictions without additional overpricing cost, and thus the implicit cost of invoice trading is maximum. Overall, as the unit production cost decreases, the implicit cost of invoice trading increases while the explicit benefit remains constant. Therefore, there might exist a threshold $\tilde{c}$ (see

Figure 2.6b), below which the supplier's profit under factoring is higher. Lastly, when the premium of invoice trading is not too high (see Figure 2.6a), the explicit benefit of trading flexibility is dominating the implicit cost of the commitment trap, and invoice trading benefits the supply chain.



(a) low premium ($\eta_{\mathcal{I}} = 0.3$) (b) high premium ($\eta_{\mathcal{I}} = 3$)

Figure 2.6.: The value of invoice trading for the supplier relative to factoring ($\rho_s = 0.1$, $\rho_r = 0.1$, $\eta_{\mathcal{F}} = 0.1$, $\eta_{\mathcal{B}} = 0.3$, $t_1 = 0.1$, $t_2 = 0.2$ and $t_2' = 0.15$)

2.6.3 The Value of Smart Contract under Invoice Trading

Finally, we discuss the supply chain value of smart contract under the invoice trading scheme (see Proposition B.0.7 for a formal statement). We have unveiled that the trading flexibility nature of invoice trading introduces the commitment trap, where the commitment frictions becomes ubiquitous and completely unresolvable which otherwise could be partially resolved with overpricing in factoring. When the invoice trading premium is relatively high, the implicit cost of the commitment trap can dominate the explicit benefit of invoice trading, leading to the adoption dilemma of invoice trading. Luckily, in such a case, smart contract (in conjunction with factoring) can bring positive value by fixing both commitment trap and commitment frictions.

More specifically, in addition to the elimination of commitment frictions in factoring, smart contract brings a new dimension of commitment benefit under invoice trading by eliminating the commitment trap. Such commitment is valuable only when the smart contract-equipped factoring could yield a higher profit than invoice trading, which is likely to happen when the invoice trading premium is relatively high (e.g., Figure 2.6b). Otherwise, invoice trading continues to dominate factoring regardless of the commitment value provide by smart contract (e.g., Figure 2.6a). Hence, it is still preferable to adopt invoice trading to take advantage of the trading flexibility while sacrificing the implicit cost of the commitment trap. In this case, smart contract has no value for the supply chain.

2.7 Discussions

In this section, we first extend the main model to consider the fire sale of accounts receivable, which allows the supplier to quickly liquidate the asset (accounts receivable) to meet cash needs upon liquidity shock. Further, we discuss the value of digitalization (which helps decrease the bank premium) and highlight the conditions under which digitalization may unexpectedly hurt the supply chain.

2.7.1 Model Extension: Fire Sale of Accounts Receivable

We have so far assumed that liquidity shock directly leads to the supplier's bankruptcy if no post-shipment financing is adopted. In this subsection, we extend our model and consider fire sale as another option to offset liquidity shock in addition to factoring or invoice trading. Fire sale allows the supplier to liquidate his assets upon any emergent cash needs, though the accounts receivable is heavily discounted by the fire sale premium $\eta_{\mathcal{S}}$. The analysis shows that our main results continue to hold. Recall that under the baseline trade finance model, the supplier simply goes bankrupt and has no value left if fire sale is not available. Therefore, fire sale increases the supplier's profit of not adopting factoring (i.e., $(1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'} wq)$, and thus makes factoring less appealing to the supplier. Nevertheless, the trade-off between ensuring wholesale revenue and saving bank loan cost still exists. Therefore, the bank still uses the threshold rule to infer the supplier's factoring decision, and the supplier might overcharge the wholesale price to convince the bank of his factoring decision. Commitment frictions still exist and smart contract, therefore, brings value to the supply chain. When BDF is offered by the retailer, the supplier always adopts factoring and never uses fire sale, which is more expensive. Therefore, the driving force still exists that smart contract discourages the retailer from offering BDF. As a result, the supplier's profit and even the supply chain profit can be hurt with smart contract. Under the invoice trading scheme, the supplier prefers invoice trading over fire sale upon liquidity shock. Hence, the fire sale does not change the game structure and all the results continue to hold. The detailed analysis is relegated into Appendix A.1.

2.7.2 The Value of Digitalization

In order to implement smart contracts, the supply chain processes need to be digitalized, which brings additional value of digitalization. There exist several main obstacles for financial institutes that keep the financing cost high: (i) traditional financial institutes usually have very high hardware maintenance and labor costs (e.g., paying for their deposit-gathering branch network and ATMs); (ii) financing activities involve a substantial amount of physical paperwork, leading to high inspection and verification cost as well as the cost of being more heavily regulated. With digitalization, such cost reduction can be well reflected by the bank premium $\eta_{\mathcal{B}}$ reduction in our model. We examine the value of digitalization (i.e., decreased bank premium $\eta_{\mathcal{B}}$) under different trade finance models. Interestingly, we find digitalization may actually hurt the supply chain profit under the baseline trade finance model as well as the BDF scheme. POF becomes more appealing with the decreased $\eta_{\mathcal{B}}$, as the wholesale revenue becomes more weighted than the bank loan cost. As a result, digitalization may switches the supplier from not adopting factoring (when $\eta_{\mathcal{B}}$ is significantly high) to expensive overpricing $\bar{w}$ (see Figure B.1a), which hurts the supply chain profit. Under the BDF scheme, with digitalization, the more attractive POF discourages the retailer from offering direct financing, and leads to the supply chain profit reduction as well (see Figure B.1b). Note that smart contract eliminates the commitment frictions, and thus always enables the supply chain to benefit from digitalization under the baseline trade finance model. However, under the BDF scheme, digitalization may still demotivate the retailer to offer direct financing, leading to a decrease in the supply chain profit. We leave the detailed analysis in Appendix A.2.

2.8 Conclusion

The increased attention to blockchain-related technologies and emerging application scenarios (e.g., DeFi: decentralized finance, NFT: non-fungible token, Web 3.0, etc.) has brought smart contract into sharper focus. A central premise of smart contracts (and more broadly blockchain technology) has been that the automated algorithmic execution based on mapping states of the world to corresponding contractual actions, which can create creditable commitments that previously would be expensive or impossible to enforce. In our analysis, we investigate to what extend smart contract indeed allows for this premise to be realized in the trade finance domain; we find that while smart contracts can indeed mitigate the commitment frictions, in many cases the ability to add value to the supply chain critically depends on the underlining trade finance structures.

We have several main findings. First, under the baseline trade finance model, the commitment frictions between the bank and the supplier could lead to either the supplier's opt-out of factoring or adoption with overpricing. Smart contract resolves such issues, reduces the supplier's financing cost, and thus benefits both supply chain members. Second, when BDF is available as an alternative pre-shipment financing scheme, we find that smart contract might hurt the supplier and even the supply chain system. This is because smart contract discourages the self-interest retailer from offering BDF, which transfers the financing burden to the more cost-sensitive supplier's shoulder. As a result, the supplier can only charge a rather high wholesale price, leading to a fairly low order quantity and thus a sharp decline in overall supply chain profit. Third, when invoice trading is available to the supplier as an additional post-shipment financing choice, we find, unexpectedly, that such financing option could make the supplier worse-off due to the adoption dilemma. That is, the supplier always prefers invoice trading over factoring due to its trading flexibility which, in turn, leads to a *commitment trap* (i.e., ubiquitous and unresolvable commitment frictions). Luckily, such an adoption dilemma can be resolved by adopting smart contract in conjunction with factoring.

Our main contribution is to propose a general supply chain finance model framework that enables us to quantify the value of smart contract when different types of trade finance instruments are in place. Our theory can help "debias" the value of smart contract (as might be incorrectly promoted by many FinTech firms) and promote more legitimate implementations of this technology in the trade finance domain (or more broadly, the supply chain area). Caution is warranted if one merely considers the *direct* positive impact of smart contract (e.g., allow creditable commitments that previously would be expensive to enforce), but largely overlooks its *indirect* effect on supply chain firms' incentives as well as their strategic choice of different trade financing schemes.

We conclude by pointing out some model limitations and proposing several potential future directions. First, we have only studied the supply chain structure with one supplier and one retailer. In many business environments, however, large retailers such as Walmart have multiple suppliers and large suppliers such as Intel have multiple downstream buyers. We expect studying such supply chains with retailer or supplier competition might bring in new findings and insights. Second, we only consider a simplified version of FinTech-driven digitalization. In practice, the benefit of digitalization can go beyond lowering the bank premium and may provide other benefits such as improving data-driven decision making and better information sharing among different supply chain members. These could be interesting directions to further explore the value of FinTech-driven digitalization.

Chapter 3: Impact of COVID-19 on Online Share of Expenditure and the Mediating Role of Digital Infrastructure: Evidence from a Two-year Consumer Panel

3.1 Introduction

The novel coronavirus, known as COVID-19, has caused a public health crisis on a global scale.[1] In addition, the pandemic outbreak has led to a reduction in global consumption both directly, through consumers' fear of infection [94], and indirectly, through government containment policies to reduce the spread of the virus [95]. This adverse impact has been documented for offline [96] and online [97] consumption.

While studying online and offline consumption separately is important for understanding COVID's impact on the economy, we lack empirical evidence on the *within-consumer* change in the online share of consumption in the mid- to long-run, which is important to document for two reasons. First, from a macroeconomic perspective, this measure is informative of the relative scale of the digital economy, which has increasingly contributed to the overall consumer welfare [98]. Therefore, documenting the potential shift in this measure has regulatory implications, such as subsidizing the right types of infrastructures (e.g., logistic capability) to accommodate potential changes in consumption structure. Second, from a microeconomic perspective, this measure shows whether consumer behavior in choosing consumption channels has shifted in the mid- to long-term. To the extent that this change reflects changes in consumption habit, managers need to strategize accordingly, for example, by optimizing distribution channels and consumer outreach.

[1] COVID-19 has led to 6,181,850 deaths globally by 2022-04-12, according to World Health Organization.

In this chapter, we document how within-consumer online share of consumption changed post-COVID using a large-scale individual-level monthly panel dataset from China's largest digital payment platform, Alipay, which covers both online consumption and purchases made offline using a digital QR code over a two-year period. The dataset and empirical context are ideal for our research question for a few reasons. First, as revealed in the IPO Prospectus of the Ant Group, Alipay had more than 1 billion active domestic users in 2020, making it the largest digital payment platform in China. Additionally, as more than 80% of daily consumption in China is completed via mobile payment,[2] a large proportion of the payment is through Alipay.[3] The breadth of coverage of consumer spending is rare and allows for more accurate measurement of consumer behavior compared to other payment methods.[4] Second, the dataset allows us to link online and offline consumption for a given individual. This feature in turn allows us to document changes in the online share of consumption without aggregating from potentially different customer pools from different datasets, which can bias the result.[5] Third, the dataset spans 24 months, allowing us to test whether the different consumption structure stabilizes towards a new level in the mid- to long-run. Lastly, the empirical context in China is ideal because there were sharp changes in the pandemic situation in 2020: at the end of January, Wuhan and other cities used strict containment policies to control the spread of the virus. By the end of April, all containment restrictions had been lifted and China reported essentially zero new cases, for the following year. These sharp changes allow us to study the change in consumption structure for well-defined periods of before, during, and after the pandemic, based on the two aforementioned cutoffs.

To get closer to a causal estimate of the COVID impact, we leverage two strategies. First, we compare the temporal change in outcomes in 2020 before and after the pandemic

[2] See http://www.news.cn/politics/2022-01/27/c_1128304402.htm.

[3] According to a recent survey report from iResearch, a well-known third-party consulting firm in China, Alipay has a market share of 55.6% in the second quarter of 2020; see https://report.iresearch.cn/report_pdf.aspx?id=3660 for more details.

[4] As a reference point, the universe of Visa credit and debit card transactions, which is used in [99], covers about 22% of consumption in the United States.

[5] One exception is [100], who study how online share of consumption was affected by COVID using French transaction data. However, their dataset ends in April 2020, roughly one month after the beginning of the COVID pandemic in France.

against the temporal change in the same calendar months in 2019. Second, we compare the temporal change in outcomes in 2020 across cities with different numbers of COVID cases. Towards an interpretation, the first estimate represents the overall impact of COVID on outcomes, and the second estimate represents the differential impacts as a function of the treatment intensity. Although the two strategies should conceptually give us consistent results,[6] the empirical estimation of the two approaches leverages two complementary sources of variation: the variation across calendar years, which allows us to compare the same cities, and the contemporary variation across different cities, which allows us to control for common time trends. Each of these approaches has its limitations; therefore, testing for robustness across specifications is important. Throughout the rest of the chapter we will refer to the estimates as "COVID impact" for conciseness, although we discuss the limitation of interpreting these estimates causally in the concluding remarks. Using the above specifications, we find a large and negative COVID impact on both online consumption (-20.9%) and offline consumption (-55.6%) during the pandemic outbreak, between February and April 2020. After the containment of the pandemic, between May and October 2020, while online consumption bounces back to the pre-pandemic level, the COVID impact on offline consumption persists after the pandemic (at around -15%). Importantly, we find an around 4 percentage points increase in the online share of consumption during the pandemic, and this increase persisted at 1 percentage point one year into the pandemic, between August and October 2020.

Next, we explore whether digital infrastructure moderates the COVID impact. To measure the strength of digital infrastructure, we use the pre-pandemic digitization index constructed by [101].[7] We then study the heterogeneity in treatment effects along this dimension. In doing so, we use the logged GDP per capita in a city to control for unob-

[6]Specifically, if the marginal effect of changing the treatment variable from 0 to 1 is positive, then the marginal effect of changing the treatment variable by a positive fraction should also be weakly positive, and vice versa, under some monotonicity assumption.

[7]The intended goal of this index is to measure the development and inclusion of digital finance in different cities. The index is an aggregate of eight sub-indices, and one of them is the level of digitization in a city, which is our mediating variable of interest. Empirically, all of these sub-indices that capture different dimensions of the digital economy in [101] are highly correlated, with a correlation coefficient of 0.91. The sub-indices are about the coverage breadth of digital finance, the use depth of digital finance, and the digitization level of financial inclusion.

served factors that are correlated with the digitization index and can contribute to the outcome. We find that during the pandemic outbreak, a stronger digital infrastructure mitigates the negative COVID impact on online consumption by 21.6% and on offline consumption by 8.9% for cities in the bottom four quintiles of GDP per capita, i.e., cities excluding the richest ones.[8]

Lastly, we explore treatment effect heterogeneity along other dimensions. Focusing on the online consumption share, we find that the COVID impact is larger for younger consumers, male consumers, and consumers with a low pre-COVID online consumption share. The heterogeneous treatment effects are broadly consistent with consumers' habit formation due to forced experimentation by the pandemic.

Our results have policy and managerial implications. First, we provide suggestive evidence that the online share of consumption may have been permanently altered by COVID. Governments should keep this in mind when designing policies related to the digital economy, as consumers' reliance on it seems to have been accelerated by the pandemic. Managers should also be aware of this change when designing various business strategies. Second, our result suggests that stronger digital infrastructure can increase the resilience of consumption spending to macroeconomic shocks such as COVID. Interestingly, this resilience is found for both online and offline consumption. This result highlights that digital infrastructure is general-purposed and its impact goes beyond the digital world, extending into the physical world. Governments should take this additional benefit into account when designing subsidies and taxes on building and maintaining digital infrastructure.

3.1.1 Related Literature

Our chapter is most related to the literature that estimates the impact of COVID on online or offline consumption. Starting with the papers on online consumption, [97] use sales data from Taobao, the largest e-commerce platform in China, to construct a city-day level panel across three years. Their identification strategy on the COVID impact is similar to

[8]Results using all cities are qualitatively similar but less statistically significant.

ours, leveraging both year-on-year comparisons and across-city comparisons. They find that the negative impact of COVID on online consumption quickly recovered after the initial drop, showing that e-commerce sales are resilient to the pandemic outbreak. This finding is consistent with our result that online consumption reverted to the pre-pandemic (i.e., 2019) level after the lift of containment policies between May and October 2022. In the online grocery shopping setting, [102] use a panel of 5,000 Dutch households over the period 2015-2020 to show that the vast majority of online shoppers abandon the channel relatively quickly. For music consumption, [103] show that music streaming decreased in many countries during COVID, because of the complementarity of music consumption to other activities. They argue that this change is likely to be transient rather than irreversible.

Moving on to papers on offline consumption, [104] compare consumption patterns in 2020 vs. 2019 and find that the percentage drop in consumption is higher for the high-income group, and this gap grew even larger over time. They also find that consumption that requires physical interaction is more likely to be replaced with online shopping. In the offline grocery shopping setting, [105] use loyalty program data from a supermarket chain in the greater St. Louis area and find that trip frequency decreased, expenditure per trip increased, and product variety increased during the pandemic. Lastly, there are a handful of papers that document the effect of government interventions on COVID transmission, either the direct effect [95,106–112] or the geographic spillover effect [113].[9] Towards optimal policy making, [119] argue that policy makers should think about the importance–risk trade-off in different types of public spaces. Our chapter contributes to the two strands of literature by leveraging a unique consumer panel to show how the pandemic changed the share of online consumption. To our knowledge, the only other paper that uses a consumer panel which captures both online and offline consumption is [100], who study how online share of consumption was affected by COVID in France. However, their dataset ends in April 2020, roughly one month into the COVID pandemic

[9] Besides consumption, previous papers have also documented COVID's impact in other contexts, such as the performance of small- and medium-sized enterprises [114, 115], online labor supply ([116]), and media persuasion around COVID-related topics ([117], [118]).

in France. By contrast, we have data one year into the outbreak of the pandemic. Our finding that the change in consumption structure persisted after the pandemic suggests that the pandemic may have caused a long-lasting change in consumption structure, perhaps because of consumers' habit formation from forced experimentation.[10]

Our chapter is also very related to a growing literature that documents the fact that digitization increases the resilience of the economy. In the healthcare sector, [122] show that counties with greater hospital IT capabilities had fewer COVID-19 deaths, a result that is driven by the learning effect by the hospitals. In terms of firm performance, [123] show that the practice of work from home (WFH) induced by digital technology made firms more resilient to COVID shocks. Their identification strategy of comparing firm performance with a high vs. low pre-pandemic WFH index is similar to the first identification strategy in this chapter. For small- and medium-sized enterprises (SME), [124] use administrative universal firm registration data and surveys of small businesses in China. They find that digitization increased SMEs' resilience in terms of sales and spurred digital technology adoption, and the effect persisted one year after. Their identification is based on comparing outcomes in 2020 vs. 2019, which is similar to the second identification strategy in this chapter. Our chapter contributes to this literature by showing the mediating effect of digitization on consumption, which is an important part of the GDP and social welfare.

3.2 Data and Empirical Approach

Consumption: To measure consumption at the individual level, we use data provided by Alipay, which is China's largest digital payment platform. The data we use consists of a random sample of 190,330 consumers from 225 major cities in China, each with an urban district population above 1 million. We sample consumers from cities proportionally based on their population. The sample is representative of 94.5% of China's urban population as of 2019. The data consists of individual consumers' (i.e., Alipay users') monthly online

[10] [120] estimate the benefits of forced experimentation from a sudden disruption in the London underground network. [121] argue that the practice of working from home will likely persist after COVID as companies and employees have become used to this work mode because of the pandemic.

and offline transaction history from November 2018 to October 2020. Specifically, the online consumption is represented by individuals' transactions at Taobao, which reaches a penetration rate of 52.5% of the shopping apps in China as of June 2019.[11] The offline consumption is represented by purchases made through mobile payment by scanning a digital QR code. A report shows that in 2020, around 85% of Alipay users paid by scanning QR codes in China.[12] Hence, transaction data of using a QR code can be used to represent offline consumption. Each individual's monthly total consumption is calculated by adding online and offline consumption, and the online consumption ratio is represented by online consumption divided by total consumption. We plot the aggregated online and offline consumption of our data in Figure C.1. It shows that there exists a spike around November, and consumption drops to the bottom around January to February. Clearly, our data exhibits a pattern of seasonality, which is caused by the Chinese Lunar Spring Festival. In 2019 and 2020, this holiday was in late January and early February, respectively. Consumption typically increases before the Spring Festival and experiences a sharp decrease during the holiday, which suggests that our estimation needs to take into account such seasonality.

Measuring COVID intensity: We focus on two strategies to obtain a causal estimate of the COVID impact. For our baseline analysis, we compare the overall difference between 2019 and 2020. As a robustness check, we also exploit regional variations in COVID intensity, and compare the temporal change in outcomes in 2020 across cities with different numbers of infected cases. For that measure, we use city (prefecture) level cumulative cases from January 2020 to April 2020.[13] Since the number of cumulative cases across different cities remains stable after the end of April, because of the containment of COVID in China, using the number of cumulative cases from January 2020 (the beginning of the pandemic) to April 2020 is representative of the COVID intensity at city level. We divide cities into three intensity groups based on the number of cumulative cases. That is,

[11] https://www.statista.com/statistics/1059076/china-leading-shopping-apps-penetration-rate/
[12] https://global.chinadaily.com.cn/a/202102/02/WS6018c23ea31024ad0baa6b4c.html
[13] The city level cases data is collected from https://www.tianditu.gov.cn/coronavirusmap/.

we classify a city into the low-COVID group (less than 10 cumulative cases), the medium-COVID group (10 to 100 cases), or the high-COVID group (more than 100 cases).

Digital infrastructure before COVID: Our measure for the pre-pandemic local level of digitization comes from the Peking University Digital Financial Inclusion Index of China (PKU-DFIIC) [101], which is constructed to describe and represent how accustomed a city is to digitization. In other words, it defines the level at which a city's population is used to digital transactions and payment instead of non-digital tools. For example, the index incorporates the number of Alipay accounts owned by 10,000 people (breadth of coverage), the number of users with an Internet loan for consumption per 10,000 adult Alipay users (depth of usage), and the proportion of mobile payments (digital inclusion). We use the index at the city level in 2018 as a proxy for the extent to which local consumers can use digital payment as well as other digital services in their life. For our analysis, we use the aggregate index to capture the overall level of digitization. Choosing other sub-indexes such as the depth of digital payment delivers very similar results.[14] Based on the index, we rank the selected 225 major cities from high to low, and define the top 50% as high-digitization cities, and the rest as low-digitization cities.

Consumer profile: The data contains consumers' demographic information such as age and gender, which we use to explore treatment effect heterogeneity. In order to explore the heterogeneity with respect to age, we classify consumers into three groups: young (up to 23 years old), middle aged (23 to 50 years old), and old (50 years old and older). To investigate the heterogeneity effect in consumers' pre-COVID purchasing habit, we calculate individual consumers' average share of online consumption before the pandemic, rank the ratio from high to low, and classify consumers with this ratio in the top 50% in the high (online consumption) group, and those with this ratio in the bottom 50% in the low group.[15]

[14] The aggregated index and sub-index are high correlated. Using the index in 2019 instead of 2018 does not substantially change our results, either.

[15] We define the period for pre-COVID consumption habit as November 2018 to October 2019. Consumption patterns in this period reflect consumers' habits in the baseline year of our empirical specification, minimizing the concern of mean reversion.

Data access: Our study is remotely conducted on the Ant Open Research Laboratory in an Ant Group Environment. The data is sampled and desensitized by the Ant Group Research Institute and stored on the Ant Open Research Laboratory. To ensure Alipay users' privacy, the laboratory is a sandbox environment where we can remotely conduct empirical analysis while individual observations are not visible.

Empirical approach: We give an overview of our empirical approach, and leave the details for the corresponding sections. For our baseline model, we perform various difference-in-differences (DiD) estimations to study the effect of COVID on consumers' consumption behaviors. We then investigate the effect of digitization in mitigating the COVID impact and build a triple DiD model for estimation. Last, we study the difference in COVID impact by various demographics and estimate the corresponding triple DiD models as well.

3.3 COVID Impact on Consumer Behavior

We first evaluate the COVID impact on consumption behavior during and after the pandemic. In order to control for the seasonal differences in purchasing, we compare the same calendar months between 2020 (the COVID year) and 2019. We define the period between November 2018 and October 2019 as year 2019, and the period between November 2019 and October 2020 as year 2020, so that we are able to have a well-defined pre-COVID period as our benchmark for comparison. Specifically, the pre-COVID period should not overlap with the pandemic period, and needs to span a sufficiently large time period within the whole year for valid comparison. Hence, by our design of year 2019 and 2020, we can define November, December, and January as the pre-COVID months and compare the pre-COVID period in the COVID year (November 2018 to October 2019) to the same calendar months in the previous year (November 2019 to October 2020). Note that in China, the COVID outbreak started in late January and was contained by the end of April 2020. Therefore, February to April 2020 can be regarded as the during-COVID period. Specifically, we compare the period between February and April 2020, which represents the COVID period in year 2020, with the same calendar months in the previous year (i.e., February to April 2019). Following the same logic, we also define May to July 2020 as the first post-COVID period, and August to October 2020 as the second post-COVID period. In this way, we evenly partition 12 months into 4 periods, which helps us better observe and understand the COVID impact during and after the pandemic.

Figure C.1 plots monthly aggregate online consumption, offline consumption, and online ratio in year 2019 and year 2020. Compared with the pre-COVID period (the benchmark), offline consumption for year 2020 significantly dropped relative to year 2019 in the during-COVID period. Although the discrepancy shrinks in the post-COVID period, it is still negative compared with the benchmark. Similarly, we find that the online consumption share increased during the pandemic and persisted after the pandemic. Lastly, the pattern of online consumption is less clear. Note that we should be cautious in interpreting the graphical patterns as the causal effects of interest, because the graph reflects not only within-consumer changes in consumption but also changes in the composition of consumers.

To control for the composition effect and identify the impacts of COVID on consumption behaviors, we estimate the following model:

$$\begin{aligned} y_{itm} =& \beta_1 \mathit{During}_m \times \mathit{Year20}_t + \beta_2 \mathit{Post1}_m \times \mathit{Year20}_t + \beta_3 \mathit{Post2}_m \times \mathit{Year20}_t \\ &+ \beta_4 \mathit{Year20}_t + \lambda_m + \mu_i + \varepsilon_{itm}. \end{aligned} \tag{3.1}$$

The dependent variables are different consumption outcomes (i.e., online consumption, offline consumption, and online consumption ratio, respectively) of individual consumer i in month m of year t; $\mathit{Year20}_t$ is a dummy for year 2020, which takes the value of 1 for November 2019 to October 2020, and 0 for November 2018 to October 2019, as previously defined; m denotes calendar month, which takes the value 1 to 12 for January to December. We define During_m as a dummy for months from February to April, which are during the COVID period, as defined. Similarly, dummy variable $\mathit{Post1}_m$ represents the first post-COVID period, and dummy $\mathit{Post2}_m$ represents the second post-COVID period, as described. In addition, λ_m indicates monthly fixed effects, which control the seasonal difference; μ_i represents individual fixed effects, which control for time invariant individual traits. Error term ε_{itm} is clustered at the city level in all of our analyses. The coefficients of interests are the interactions between $Year20_t$ and the within-year temporal dummies for various COVID periods: $During_m$, $Post1_m$, and $Post2_m$. This baseline model identifies the effects of COVID on consumption behaviors by exploiting the variations between two years and across months. It allows us to quantify the COVID impact on consumption not only during the pandemic (short-term) but also after the pandemic (mid- to long-term).

Table C.2 reports our baseline estimates exploiting the temporal variation during and after the COVID shock (Equation 3.1). We first evaluate the COVID impact on online and offline consumption. We apply the inverse hyperbolic sine transformation to the consumption values to account for the right-skewness of this variable, and to be able to interpret the estimated coefficient as percentage changes.[16] As shown in column (1) and column (2) of Table C.2, compared to the same calendar months in 2019, online consumption dropped by 21% and offline consumption dropped by 56% during the pandemic in year 2020. Online consumption recovered after the pandemic, consistent with the findings in [97], but the negative impact of COVID on offline consumption was much more persistent, at around -15% six months after the pandemic. We compare our estimates with the figures and find the overall patterns are consistent with our estimates obtained from exploiting within-individual variations. We next focus on the online consumption ratio, which reflects the extent to which consumers' purchasing relies on the online channel. We see a relative shift of consumer consumption from offline to online during and after the pandemic. The estimates in column (3) of Table C.2 show that the ratio of online consumption was boosted by 4.4 percentage points during the pandemic. The impact was persistent: six months after the end of the initial COVID shock, the online consumption ratio was still increased by 1 percentage point, which is also consistent with our findings shown in Figure C.1.

[16] The inverse hyperbolic sine transformation is very similar to the log transformation but with the benefit of allowing for the value of 0. See more details in [125].

Robustness check: In our baseline model, we estimate the overall effects of the COVID shock during and after the pandemic, to control for city-level unobserved heterogeneity and to be able to make a year-on-year comparison of the COVID impact. However, the identification is based on two different years. Besides, we are unable to identify how the COVID impact varies with the level of COVID severity. As an alternative estimation, we compare the temporal change in consumption behaviors across cities with different COVID intensity. This comparison serves as a complementary model that helps us check the robustness of our baseline model. As illustrated in Section 3.2, we divide cities based on their cumulative COVID cases into three groups: high COVID, medium COVID, and low COVID. Since we focus on across-city instead of across-year comparison, we do not need to divide the data into year 2019 and year 2020 and compare the COVID impact within the same calendar month, as in the baseline model. Instead, we simply denote the period from November 2018 to January 2020 as pre-COVID,[17] Februrary to April 2020 as during-COVID, May to July 2020 as the first post-COVID, and August to October 2020 as the second post-COVID period. The DID model is implemented using the following specification:

$$
\begin{aligned}
y_{it} = {} & \theta_1 During_t \times HighCovid_c + \theta_2 Post1_t \times HighCovid_c + \theta_3 Post2_t \times HighCovid_c \\
& + \theta_4 During_t \times MedCovid_c + \theta_5 Post1_t \times MedCovid_c + \theta_6 Post2_t \times MedCovid_c \\
& + \mu_i + \eta_t + \epsilon_{it}.
\end{aligned} \tag{3.2}
$$

The dependent variables are different consumption outcomes (i.e., online consumption, offline consumption, and online consumption ratio, respectively) of individual consumer i at time t (actual data); $HighCovid_c$ and $MedCovid_c$ are dummy variables indicating the city group of COVID intensity, as previously described; $During_t$ is a dummy for the period from February to April 2020, $Post1_t$ for May to July 2020, and $Post2_t$ for August to October 2020. Individual fixed effects μ_i and time fixed effects η_t control for time-invariant individual traits and common shocks that applied to everyone, respectively.

[17] Using November 2019 to January 2020 as the pre-COVID period does not qualitatively change the results.

The coefficients of interest are the coefficients on the interactions between time temporal dummies and $HighCovid_c$ and $MedCovid_c$, which identify the COVID impacts during and after the pandemic, respectively, across intensity levels. Similar specifications have been used in [97] and [124].

The results of this DiD estimation are reported in Table C.3. We first investigate online and offline consumption, which are shown in column (1) and column (2) of Table C.3, respectively. Compared with the benchmark (i.e., low-COVID-intensity regions), during the pandemic, online consumption dropped by 40% and offline consumption dropped by 30% in high-COVID-intensity regions, whereas the COVID impact was relatively mild in regions with medium COVID intensity: online consumption decreased by only 8% and offline consumption by only 13%. Six months after the pandemic, online consumption recovered in both regions, but offline consumption was still decreased by 8% in high-COVID-intensity regions, and by 4% in medium COVID-intensity regions. We next look at the share of online consumption in column (3) of Table C.3. Compared with low-COVID-intensity regions, during the pandemic, the online ratio was increased by 0.8% in high-COVID-intensity regions, and by 0.7% in medium-COVID-intensity regions. Six months after the pandemic, the increase was still at around 0.8% in high-COVID-intensity regions, and only at 0.3% in medium-COVID-intensity regions. Overall, the estimates of COVID impact here are largely consistent with our findings in the baseline model. Specifically, the pandemic negatively affects both online and offline consumption during the pandemic. However, online consumption rebounded after the pandemic, but the negative COVID impact on offline consumption persisted. The online ratio tremendously increased during the pandemic and remained at a significant level even after the pandemic. Besides, the impact became more severe as the COVID intensity increased.

3.4 The Mediating Role of Digital Infrastructure

Digital infrastructure has played an important role in shaping economic activities and improving consumer welfare ([98]. We are interested in studying how better digitization infrastructure potentially mitigates the negative COVID impact on consumption. We first plot the monthly aggregate of online and offline consumption for consumers who live in regions with high vs. low levels of digitization, based on a median split. Figures C.2 and C.3 show patterns consistent with the mediating role of digital infrastructure.[18]

To quantify the heterogeneous COVID impacts by the level of digitization, we further adopt a triple-DiD model that is a modification of equation 3.1. Conceptually, the modification involves two things. First, we incorporate the dummy variable for high-digitization regions and its interactions with other terms in equation 3.1 to essentially compare how the treatment effect estimated based on equation 3.1 differs for consumers who live in cities with high vs. low levels of digitization. Second, we control for the interactions of GDP per capita of each city with the treatment dummies based on temporal variations in equation 1. The reason is that the digitization level of a city is likely correlated with other city characteristics, such as the wealth of the population and medical resources, which can also mediate the COVID impact. Including terms involving GDP per capita mitigates this concern if the measure is positively correlated with the omitted variables (e.g., supply chain capacity of a city) that may create an upward bias of the estimates. Note that we did not need to include the GDP per capita terms when we estimated the overall impact of COVID in the previous section, because the identification using equation 3.1 was based on a year-on-year comparison of the same set of cities, rather than on an across-city comparison. Formally, our triple-DiD model is

[18] For example, in Figure C.2, the gap between the two lines in the second post-period is wider for consumers in high-digitization regions than those in low-digitization regions, relative to the pre-period.

$$
\begin{aligned}
y_{itm} = \quad & \beta_1 During_m \times Year20_t \times HighDigital_c + \beta_2 Post1_m \times Year20_t \times HighDigital_c \\
& + \beta_3 Post2_m \times Year20_t \times HighDigital_c + \beta_4 Year20_t \times HighDigital_c \\
& + \beta_5 During_m \times HighDigital_c + \beta_6 Post1_m \times HighDigital_c \\
& + \beta_7 Post2_m \times HighDigital_c \\
& + \gamma_1 During_m \times Year20_t \times GDP_c + \gamma_2 Post1_m \times Year20_t \times GDP_c \\
& + \gamma_3 Post2_m \times Year20_t \times GDP_c + \gamma_4 Year20_t \times GDP_c \\
& + \gamma_5 During_m \times GDP_c + \gamma_6 Post1_m \times GDP_c \\
& + \gamma_7 Post2_m \times GDP_c \\
& + \delta_1 Year20_t \times HighDigital_c \times GDP_c + \delta_2 During_m \times HighDigital_c \times GDP_c \\
& + \delta_3 Post1_m \times HighDigital_c \times GDP_c + \delta_4 Post2_m \times HighDigital_c \times GDP_c \\
& + \theta_1 Year20_t + \theta_2 During_m \times Year20_t + \theta_3 Post1_m \times Year20_t \\
& + \theta_4 Post2_m \times Year20_t \\
& + \lambda_m + \mu_i + \varepsilon_{itm}
\end{aligned} \tag{3.3}
$$

The dependent variable y_{itm} is the same as in our baseline model; $HighDigital_c$ is a dummy variable indicating the high-digitization group, constructed by the digitization index described in Section 3.2. We use the same definitions of $Year20_t$, $During_m$, $Post1_m$, and $Post2_m$ as in the baseline model. Moreover, λ_M indicates monthly fixed effects, μ_i represents individual fixed effects, and error term ε_{itm} is clustered at the city level. We also control for the city-level GDP per capita in 2018, with the inverse hyperbolic sine transformation to account for the potential right-skewness. For our main results, we rank the 225 cities of our sample by GDP from high to low and remove the top 20% of cities.[19] The coefficients of interest are the parameters on $During_m \times Year20_t \times HighDigital_c$,

[19]This is to make our results more applicable to regular cities that are not the richest mega-cities in China. In the appendix, Table C.8, we also present our estimation with all cities, and the results are qualitatively similar.

$Post1_m \times Year20_t \times HighDigital_c$, and $Post2_m \times Year20_t \times HighDigital_c$, which identify how digitization mediates the COVID impact during and after the pandemic.

The estimation results on the parameters of interest are reported in Table C.4. Column (1) shows that compared to consumers who live in cities with low levels of digitization, those who live in cities with high levels of digitization experienced a 22% smaller drop in online consumption. Additionally, column (2) shows that the digital resilience to the pandemic extended also to offline consumption, where high levels of digitization reduced the negative impact of COVID by almost 9%. Lastly, we see that the share of consumption that takes place online was higher in cities with a high digitization level, suggesting a positive reinforcement between technology adoption and usage, perhaps because of consumers' habit formation.

While we cannot pin down the mechanism through which digitization mitigates the COVID impact, we have some speculations. First, note that the digitization index is based on how common online shopping is in a city before the pandemic. Therefore, it is intuitive that cities in which people were already used to shopping online shopping experienced a smaller reduction in online consumption during the pandemic. The smaller drop in offline consumption could be explained by that people are more comfortable using cashless payment option by scanning the QR code on their phones for offline purchases due to the fear of infection ([126]).

3.5 Heterogeneous Treatment Effects

In this section, we explore more heterogeneous treatment effects by the following consumer demographics: age, gender, and pre-COVID online consumption ratio. These analyses help us understand the impacts of COVID on consumption on different types of consumers, thereby shedding light on the mechanisms of this impact on consumption behaviors. The understanding of the underlying mechanism can then allow for more precise policy targeting different subgroups of consumers.

3.5.1 Heterogeneity by Age

Similar to estimating treatment effect heterogeneity by the level of digitization, we estimate a triple-DiD model to evaluate the heterogeneous effect by consumers' age. The model is given as

$$\begin{aligned}
y_{itm} =& \beta_1 \mathit{During}_m \times \mathit{Year20}_t \times \mathit{Young}_i + \beta_2 \mathit{Post1}_m \times \mathit{Year20}_t \times \mathit{Young}_i \\
&+ \beta_3 \mathit{Post2}_m \times \mathit{Year20}_t \times \mathit{Young}_i + \beta_4 \mathit{Year20}_t \times \mathit{Young}_i \\
&+ \beta_5 \mathit{During}_m \times \mathit{Young}_i + \beta_6 \mathit{Post1}_m \times \mathit{Young}_i \\
&+ \beta_7 \mathit{Post2}_m \times \mathit{Young}_i \\
&+ \gamma_1 \mathit{During}_m \times \mathit{Year20}_t \times \mathit{Med}_i + \gamma_2 \mathit{Post1}_m \times \mathit{Year20}_t \times \mathit{Med}_i \\
&+ \gamma_3 \mathit{Post2}_m \times \mathit{Year20}_t \times \mathit{Med}_i + \gamma_4 \mathit{Year20}_t \times \mathit{Med}_i \\
&+ \gamma_5 \mathit{During}_m \times \mathit{Med}_i + \gamma_6 \mathit{Post1}_m \times \mathit{Med}_i \\
&+ \gamma_7 \mathit{Post2}_m \times \mathit{Med}_i \\
&+ \theta_1 \mathit{Year20}_t + \theta_2 \mathit{During}_m \times \mathit{Year20}_t + \theta_3 \mathit{Post1}_m \times \mathit{Year20}_t \\
&+ \theta_4 \mathit{Post2}_m \times \mathit{Year20}_t \\
&+ \lambda_m + \mu_i + \varepsilon_{itm}
\end{aligned} \tag{3.4}$$

The dependent variable y_{itm} is the same as discussed in Section 3.3, Young_i and Med_i are dummy variables indicating the young group and middle age group, respectively, which are constructed in Section 3.2. The old group is the omitted benchmark group. We use the same definition of $\mathit{Year20}_t$, During_m, $\mathit{Post1}_m$, and $\mathit{Post2}_m$ as in the baseline model. In addition, λ_m indicates monthly fixed effects, μ_i represents individual fixed effects, and error term ε_{itm} is clustered at the city level. The coefficients of interest are $\mathit{During}_M \times \mathit{Year20}_Y \times \mathit{Young}_i$, $\mathit{Post1}_M \times \mathit{Year20}_Y \times \mathit{Young}_i$, and $\mathit{Post2}_M \times \mathit{Year20}_Y \times \mathit{Young}_i$ (resp., $\mathit{During}_M \times \mathit{Year20}_Y \times \mathit{Med}_i$, $\mathit{Post1}_M \times \mathit{Year20}_Y \times \mathit{Med}_i$, and $\mathit{Post2}_M \times \mathit{Year20}_Y \times \mathit{Med}_i$), which identify the heterogeneous effect of age on COVID impact during and after the pandemic, respectively.

The results are presented in Table C.5. Columns (1) and (2) show mostly a decreasing pattern in online and offline consumption for the different age groups, except that online consumption for consumers who are at least 50 years old increased relative to the year before the pandemic. This result is consistent with a forced experimentation story where the pandemic and related containment policies forced older consumers, who are presumably more technologically inept, to try out the online channel. Interestingly, while the online consumption ratio of older consumers experienced a large increase (3.1 percentage points) during the pandemic, the ratio diminished quickly after the end of the containment policy. In comparison, for the two younger groups (i.e., young and middle age), the COVID impact on the online consumption ratio was much larger and more persistent in the post-COVID periods. Compared to the old group, for the young group, in particular, COVID boosted the online consumption ratio by 2.3% even six months after the end of the initial shock.[20]

These findings are consistent with the presumption that learning effects on behavioral changes are larger for younger consumers. While old consumers had to use online purchasing channels during the pandemic, they quickly returned to their familiar traditional offline purchasing approach as soon as their activities were not constrained by COVID. Younger consumers, on the other hand, spent a larger proportion of their money online and continued to do so one year into the pandemic. This behavior suggests a more persistent change in their consumption structure.

3.5.2 Heterogeneity by Gender

Next, we explore the heterogeneous effect of COVID by gender. The triple-DiD model is similar to that in subsection 3.5.1, where the dummy variables *Young* and *Med* are simply replaced with *Female*. The results are reported in Table C.6. Interestingly, we see a persistent COVID effect on online consumption ratio only for males. Specifically, their share of online consumption is increased by 1.79% six months after the pandemic. For females, however, the COVID impact on the online consumption ratio eventually

[20] A similar heterogeneous response by age groups can be seen in Figure C.4 in the appendix.

vanished six months after the pandemic. Specifically, the effect of COVID on the online consumption ratio for *Post2* is $1.7890 - 1.7939 = -0.0047$ percentage points, with an unreported standard error of 0.1239.[21]

Before COVID, the average online consumption of women in our sample was 40.8%, while that of men was only 24.4%; therefore, the online consumption ratio of women was much higher than that of men. The persistent effect of COVID on the online consumption share for men reflects that they caught up with women in terms of online consumption in the pandemic. This is consistent with the interpretation of habit formation: men who used to spend less online were forced to adapt to online consumption during COVID. As a result, they paid the learning cost and learned more about their own (previously unrevealed) preference for online shopping. They continued to purchase more online relative to offline even after the pandemic. Other potential interpretations, such as income effect or changes in relative price between online and offline consumption, seem to be unable to reconcile the contrast in online consumption between men and women.

3.5.3 Heterogeneity by Pre-COVID Online Consumption Ratio

Finally, we investigate the treatment effect heterogeneity by consumers' pre-COVID online consumption ratio to provide a further test for consumers' learning habit formation of online shopping. Similar to the analysis in subsection 3.5.2, we use a triple-DiD model to evaluate the heterogeneous effect of pre-COVID online consumption share, where we replace the dummy *Female* with *High*. The dummy variable *High* represents the group of high pre-COVID online consumption share, defined in section 3.2.

The results are reported in Table C.7. The online consumption ratio increased substantially during COVID for the *Low* group (4.5%) and for the *High* group (4.2%), and the difference between the two groups is statistically non-significant (-0.35% with a standard error of 0.23). After the pandemic shock, the online consumption ratio of the *High* group returned to the pre-COVID level. For the *Low* group, however, the effect of the pandemic on the online consumption ratio was extremely persistent in the first three months

[21]Qualitatively similar results by gender can be seen in Figure C.5 in the appendix.

after the pandemic (4.2%) and still remained increased by 3% increase in the following three months.[22] These patterns are also consistent with the consumer learning story. Consumers who were already familiar with online shopping increased their share of online shopping only temporarily during the pandemic because of the restrictions on offline purchasing. However, they returned to their optimal allocation of consumption choices once the constraint was removed. The other group of consumers, who were previously unfamiliar with online shopping were forced to use online shopping during the pandemic. They either overcame the previous barrier of learning to shop online or learned about the benefits of online shopping previously unknown to them. For these consumers, the increase in the online ratio of consumption persisted even after the end of the pandemic.

3.6 Conclusion

In this chapter, we documented changes in consumption patterns in China after the COVID outbreak in 2020. We found that although the pandemic decreased both online and offline consumption, the effect is more negative for offline consumption, and led to an increase in the share of consumption conducted online. In the six months after the lift of the containment policies (where the number of new cases was essentially zero) in China, we found that online consumption rebounded to the pre-pandemic level, while the negative effect on offline consumption persisted. Additionally, we found that individuals who live in cities with better digital infrastructure were less negatively impacted by the pandemic. Lastly, our heterogeneous treatment effect estimates are broadly consistent with the hypothesis that forced experimentation due to the pandemic induced a behavioral change in consumers' channel choice.

Our results have both policy and managerial implications. Policy-wise, our results suggest that the change in consumption structure due to COVID may be long lasting, and policy makers should consider this when designing post-COVID policies. For example, they may want to allocate more resources that are complementary to e-commerce, instead of traditional brick-and-mortar stores. Also, consumers' larger reliance on online shopping

[22]Similar patterns by pre-COVID online ratio can be seen in Figure C.6 in the appendix.

may accelerate the change in the retail landscape, which can trigger a change in the local job markets over time. Policy makers should keep this in mind when revising unemployment benefits, adult training programs, and related policies. Our second finding that better digital infrastructure mitigates the negative COVID impact is also useful for policy makers in doing cost and benefits analyses on infrastructure planning. Specifically, the result that digitization not only benefits online consumption, but also spills over to the offline consumption is an important point to account for in such analyses. Lastly, for managers, our results provide useful information for adjusting their strategies based on the changing consumption structure. Examples of these strategies include investing in better logistic services that can meet the increasing demand for online shopping and spending the marketing budget more on digital channels than on traditional media.

There are a few limitations of our work. First, both of our identification strategies are imperfect: The year-on-year comparison allows us to control for city-level unobserved heterogeneity, but the identification is based on two different years. The across-city comparison enables a contemporaneous comparison, but the identification is across different cities (although we do control for city fixed effects). Despite the imperfection, the fact that the two approaches, which are based on two different sources of identifying variations, give qualitatively similar results adds confidence to the results, at least qualitatively. The second limitation is that in our analyses on digital resilience, controlling for GDP per capita may not fully resolve potential omitted variable bias in the estimation of the mediating effect of digitization. For example, if a better digital infrastructure may have caused larger GDP growth, then controlling for GDP (hence implicitly treating it as if it were exogeneous) will wrongly attribute the digitization effect to the GDP effect, thereby creating a downward bias on the digitization effect. Lastly, we do not have offline consumption data by category and, therefore, we cannot study the change in consumption structure by types of products. Understanding the changes in consumption structure across different product types can shed further light on the mechanisms behind the changes. We leave this topic for future research.

REFERENCES

[1] Xiaoyu Wang, Fasheng Xu, and Fuqiang Zhang. Consumer privacy in online retail supply chains. Working paper, 2023.

[2] Christian Peukert, Stefan Bechtold, Michail Batikas, and Tobias Kretschmer. Regulatory spillovers and data governance: Evidence from the gdpr. Marketing Science, 2022.

[3] Alessandro Acquisti and Hal R Varian. Conditioning prices on purchase history. Marketing Science, 24(3):367–381, 2005.

[4] Vincent Conitzer, Curtis R Taylor, and Liad Wagman. Hide and seek: Costly consumer privacy in a market with repeat purchases. Marketing Science, 31(2):277–292, 2012.

[5] Daron Acemoglu, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar. Too much data: Prices and inefficiencies in data markets. Technical report, National Bureau of Economic Research, 2019.

[6] Dirk Bergemann, Alessandro Bonatti, and Tan Gan. The economics of social data. The RAND Journal of Economics, 53(2):263–296, 2022.

[7] Alessandro Bonatti and Gonzalo Cisternas. Consumer scores and price discrimination. The Review of Economic Studies, 87(2):750–791, 2020.

[8] Alessandro Acquisti, Curtis Taylor, and Liad Wagman. The economics of privacy. Journal of Economic Literature, 54(2):442–92, 2016.

[9] Florian Morath and Johannes Münster. Online shopping and platform design with ex ante registration requirements. Management Science, 64(1):360–380, 2017.

[10] Byungwan Koh, Srinivasan Raghunathan, and Barrie R Nault. Is voluntary profiling welfare enhancing? MIS Quarterly, 41(1):23–41, 2017.

[11] Luca Braghieri. Targeted advertising and price discrimination in intermediated online markets. Available at SSRN 3072692, 2019.

[12] Alexandre De Corniere and Romain De Nijs. Online advertising and privacy. The RAND Journal of Economics, 47(1):48–72, 2016.

[13] Tommaso Valletti and Jiahua Wu. Consumer profiling with data requirements: Structure and policy implications. Production and Operations Management, 29(2):309–329, 2020.

[14] Shota Ichihashi. Online privacy and information disclosure by consumers. American Economic Review, 110(2):569–95, 2020.

[15] Rodrigo Montes, Wilfried Sand-Zantman, and Tommaso Valletti. The value of personal information in online markets with endogenous privacy. Management Science, 65(3):1342–1362, 2019.

[16] Zhijun Chen, Chongwoo Choe, and Noriaki Matsushima. Competitive personalized pricing. Management Science, 66(9):4003–4023, 2020.

[17] Mary J Culnan and Pamela K Armstrong. Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. Organization science, 10(1):104–115, 1999.

[18] Catherine E Tucker. Social networks, personalized advertising, and privacy controls. Journal of Marketing Research, 51(5):546–562, 2014.

[19] Curtis R Taylor. Consumer privacy and the market for customer information. RAND Journal of Economics, pages 631–650, 2004.

[20] Avi Goldfarb and Catherine E Tucker. Privacy regulation and online advertising. Management science, 57(1):57–71, 2011.

[21] Byungwan Koh, Srinivasan Raghunathan, and Barrie R Nault. Is voluntary profiling welfare enhancing? Management Information Systems Quarterly, Forthcoming, 2015.

[22] James Campbell, Avi Goldfarb, and Catherine Tucker. Privacy regulation and market structure. Journal of Economics & Management Strategy, 24(1):47–73, 2015.

[23] Shota Ichihashi. Dynamic privacy choices. Available at SSRN 3472151, 2019.

[24] Ming Hu, Ruslan Momot, and Jianfu Wang. Privacy management in service systems. Working Paper, 2022.

[25] Yanzhe Murray Lei, Sentao Miao, and Ruslan Momot. Privacy-preserving personalized revenue management. Working paper, 2022.

[26] Can Kucukgul, Ozalp Ozer, and Shouqiang Wang. Recommender systems with privacy concerns. Working paper, 2022.

[27] Rossella Argenziano and Alessandro Bonatti. Data linkages and privacy regulation. Technical report, Working Paper, 2021.

[28] Kostas Bimpikis, Ilan Morgenstern, and Daniela Saban. Data tracking under competition. Available at SSRN 3808228, 2021.

[29] Sarit Markovich and Yaron Yehezkel. Data regulation: who should control our data? Available at SSRN 3801314, 2021.

[30] Itay P Fainmesser, Andrea Galeotti, and Ruslan Momot. Digital privacy. Available at SSRN, 2022.

[31] T Tony Ke and K Sudhir. Privacy rights and data security: Gdpr and personal data driven markets. Available at SSRN 3643979, 2020.

[32] Samuel Goldberg, Garrett Johnson, and Scott Shriver. Regulating privacy online: An economic evaluation of the gdpr. Available at SSRN 3421731, 2021.

[33] Jian Jia, Ginger Zhe Jin, and Liad Wagman. The short-run effects of the general data protection regulation on technology venture investment. Marketing Science, 40(4):661–684, 2021.

[34] Miguel Godinho de Matos and Idris Adjerid. Consumer consent and firm targeting after gdpr: The case of a large telecom provider. Management Science, 2021.

[35] Guy Aridor, Yeon-Koo Che, and Tobias Salz. The economic consequences of data privacy regulation: Empirical evidence from gdpr. Technical report, National Bureau of Economic Research, 2020.

[36] W. Amaldoss and S. Jain. Conspicuous consumption and sophisticated thinking. Management science, 51(10):1449–1466, 2005.

[37] X. Su and F. Zhang. Strategic customer behavior, commitment, and supply chain performance. Management Science, 54(10):1759–1773, 2008.

[38] F. Xu and F. Zhang. Crowdfunding under social learning and network externalities. Working paper, 2018.

[39] Xiaoyu Wang and Fasheng Xu. The value of smart contract in trade finance. Manufacturing & Service Operations Management, 2022.

[40] A. Daripa and J. Nilsen. Ensuring sales: A theory of inter-firm credit. American Economic Journal: Microeconomics, 3(1):245–79, 2011.

[41] C.S. Tang, S.A. Yang, and J. Wu. Sourcing from suppliers with financial constraints and performance risk. Manufacturing & Service Operations Management, 20(1):70–84, 2017.

[42] M. Reindorp, F. Tanrisever, and A. Lange. Purchase order financing: Credit, commitment, and supply chain consequences. Operations Research, 66(5):1287–1303, 2018.

[43] L.W. Cong and Z. He. Blockchain disruption and smart contracts. The Review of Financial Studies, 32(5):1754–1797, 2019.

[44] L. Chen, L.W. Cong, and Y. Xiao. A brief introduction to blockchain economics. In Information for Efficient Decision Making: Big Data, Blockchain and Relevance, pages 1–40. World Scientific, 2021.

[45] J. Chod. Inventory, risk shifting, and trade credit. Management Science, 63(10):3207–3225, 2016.

[46] T.I. Tunca and W. Zhu. Buyer intermediation in supplier finance. Management Science, 64(12):5631–5650, 2017.

[47] P. Kouvelis and W. Zhao. Financing the newsvendor: supplier vs. bank, and the structure of optimal trade credit contracts. Operations Research, 60(3):566–580, 2012.

[48] D.A. Iancu, N. Trichakis, and G. Tsoukalas. Is operating flexibility harmful under debt? Management Science, 63(6):1730–1761, 2016.

[49] P. Kouvelis and W. Zhao. Supply chain contract design under financial constraints and bankruptcy costs. Management Science, 62(8):2341–2357, 2016.

[50] S.A. Yang and J.R. Birge. Trade credit, risk sharing, and inventory financing portfolios. Management Science, 64(8):3667–3689, 2017.

[51] L. Dong, L. Ren, and D. Zhang. Financing small and medium-size enterprises via retail platforms. Working Paper, 2019.

[52] J. Chod, E. Lyandres, and S.A. Yang. Trade credit and supplier competition. Journal of Financial Economics, 131(2):484–505, 2019.

[53] J. Chod, N. Trichakis, and G. Tsoukalas. Supplier diversification under buyer risk. Management Science, 65(7):3150–3173, 2019.

[54] F. Xu, X. Guo, G. Xiao, and F. Zhang. Crowdfunding or bank financing: Effects of market uncertainty and word-of-mouth communication. Working Paper, 2019.

[55] S. Deng, C. Gu, G. Cai, and Y. Li. Financing multiple heterogeneous suppliers in assembly systems: Buyer finance vs. bank finance. Manufacturing & service operations management, 20(1):53–69, 2018.

[56] D. Gupta and Y. Chen. Retailer-direct financing contracts under consignment. Manufacturing & Service Operations Management, 22(3):528–544, 2020.

[57] B.J. Sopranzetti. The economics of factoring accounts receivable. Journal of Economics and Business, 50(4):339–359, 1998.

[58] L. Klapper. The role of factoring for financing small and medium enterprises. Journal of Banking & Finance, 30(11):3111–3130, 2006.

[59] S.K. Devalkar and H. Krishnan. The impact of working capital financing costs on the efficiency of trade credit. Production and Operations Management, 28(4):878–889, 2019.

[60] D.A. Wuttke, E.D. Rosenzweig, and H.S. Heese. An empirical analysis of supply chain finance adoption. Journal of Operations Management, 65(3):242–261, 2019.

[61] P. Kouvelis and F. Xu. A supply chain theory of factoring and reverse factoring. Management Science, 67(10):6071–6088, 2021.

[62] J. Li and W. Mann. Initial coin offering and platform building. Working Paper, 2018.

[63] L. W. Cong, Y. Li, and N. Wang. Tokenomics: Dynamic adoption and valuation. Review of Financial Studies (Forthcoming), 2020.

[64] L.W. Cong, Y. Li, and N. Wang. Token-based platform finance. Fisher College of Business Working Paper, page 028, 2020.

[65] R. Li and D. Shi. E-tailing with instant return credit. Working Paper, 2020.

[66] P. Kouvelis, L. Dong, and D. Turcic. Advances in supply chain finance and fintech innovations overview. Foundations and Trends® in Technology, Information and Operations Management, 14(1–2):1–4, 2020.

[67] I. Goldstein, W. Jiang, and G.A. Karolyi. To fintech and beyond. The Review of Financial Studies, 32(5):1647–1661, 2019.

[68] A.V. Thakor. Fintech and banking: What do we know? Journal of Financial Intermediation, 41:100833, 2020.

[69] G. Dorfleitner, J. Rad, and M. Weber. Pricing in the online invoice trading market: First empirical evidence. Economics Letters, 161:56–61, 2017.

[70] C.A. Parlour, U. Rajan, and H. Zhu. When fintech competes for payment flows. Available at SSRN, 2020.

[71] H. Lee, C. Tang, S.A. Yang, and Y. Zhang. Dynamic trade finance in the presence of information frictions and fintech. Manufacturing & Service Operations Management, 2022.

[72] G. Tsoukalas and B.H. Falk. Token-weighted crowdsourcing. Management Science, 66(9):3843–3859, 2020.

[73] V. Babich and G. Hilary. Om forum—distributed ledgers and operations: What operations management researchers should know about blockchain technology. Manufacturing & Service Operations Management, 22(2):223–240, 2020.

[74] J. Chod, N. Trichakis, G. Tsoukalas, Aspegren, and M. H., Weber. On the financing benefits of supply chain transparency and blockchain adoption. Management Science, 66(10):4378–4396, 2020.

[75] G.M. Hastig and M.S. Sodhi. Blockchain for supply chain traceability: Business requirements and critical success factors. Production and Operations Management, 29(4):935–954, 2020.

[76] L. Dong, Y. Qiu, and F. Xu. Blockchain-enabled deep-tier supply chain finance. *Manufacturing & Service Operations Management* (Forthcoming), 2022.

[77] Y. Cui, V. Gaur, and J. Liu. Supply chain transparency and blockchain design. Working Paper, 2022.

[78] J. Hou, B. Kazaz, and F. Xu. Accounts receivable tokenization in multitier supply networks. Working paper, 2022.

[79] Y. Cui, M. Hu, and J. Liu. Value and design of traceability-driven blockchains. Working Paper, 2021.

[80] L. Dong, P.P. Jiang, and F. Xu. Impact of traceability technology adoption in food supply chain networks. *Management Science* (Forthcoming), 2022.

[81] J. Chod, N. Trichakis, and S. A. Yang. Platform tokenization: Financing, governance, and moral hazard. Management Science forthcoming, 2021.

[82] J. Gan, G. Tsoukalas, and S. Netessine. Initial coin offerings, speculation, and asset tokenization. Management Science, 67(2):914–931, 2021.

[83] N. B. Keskin, C. Li, and J. S. J. Song. The blockchain newsvendor: Value of freshness transparency and smart contracts. Available at SSRN 3915358, 2021.

[84] Y. Bakos and H. Halaburda. Blockchains, smart contracts and connected sensors: Substitutes or complements? Available at SSRN, 2021.

[85] J.N. Barrot. Trade credit and industry dynamics: Evidence from trucking firms. The Journal of Finance, 71(5):1975–2016, 2016.

[86] B. Cant, A. Khadikar, A. Ruiter, J. B. Bronebakk, J. Coumaros, J. Buvat, and A. Gupta. Smart contracts in financial services: Getting from hype to reality. Capgemini Consulting, pages 1–24, 2016.

[87] O. Hart and J. Moore. Incomplete contracts and renegotiation. Econometrica: Journal of the Econometric Society, pages 755–785, 1988.

[88] J. Tirole. Incomplete contracts: Where do we stand? Econometrica, 67(4):741–781, 1999.

[89] P. Antras and C.F. Foley. Poultry in motion: a study of international trade finance practices. Journal of Political Economy, 123(4):853–901, 2015.

[90] S. Mateut. Reverse trade credit or default risk? explaining the use of prepayments by firms. Journal of Corporate Finance, 29:303–326, 2014.

[91] IMF. World economic outlook, october 2009, sustaining the recovery. International Monetary Fund, Washington, DC, 2009.

[92] G. F. Udell. What's in a relationship? the case of commercial lending. Business Horizons, 51(2):93–103, 2008.

[93] B.Z. Zhang, P. Baeck, T. Ziegler, J. Bone, and K. Garvey. Pushing boundaries: The 2015 uk alternative finance industry report. Available at SSRN 3621312, 2016.

[94] Austan Goolsbee and Chad Syverson. Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020. Journal of Public Economics, 193:104311, 2021.

[95] Nan Zhao, Raphael Thomadsen, Zack Wang, and Song Yao. Social-distancing and masking boost consumer spending instead of smothering it. Technical report, Working Paper, 2021.

[96] Haiqiang Chen, Wenlan Qian, and Qiang Wen. The Impact of the COVID-19 Pandemic on Consumption: Learning from High-Frequency Transaction Data. In AEA Papers and Proceedings, volume 111, pages 307–11, 2021.

[97] Brian Rongqing Han, Tianshu Sun, Leon Yang Chu, and Lixia Wu. COVID-19 and E-commerce Operations: Evidence From Alibaba. Manufacturing & Service Operations Management, 2022.

[98] Erik Brynjolfsson and Avinash Collis. How Should We Measure the Digital Economy. Harvard Business Review, 97(6):140–148, 2019.

[99] Paul Dolfen, Liran Einav, Peter J Klenow, Benjamin Klopack, Jonathan D Levin, Laurence Levin, and Wayne Best. Assessing the gains from e-commerce. Technical report, National Bureau of Economic Research, 2019.

[100] David Bounie, Youssouf Camara, and John W Galbraith. Consumers' Mobility, Expenditure and Online-Offline Substitution Response to COVID-19: Evidence from French Transaction Data. Available at SSRN 3588373, 2020.

[101] Feng Guo, Jingyi Wang, Fang Wang, Tao Kong, Xun Zhang, and Zhiyun Cheng. Measuring China's Digital Financial Inclusion: Index Compilation and Spatial Characteristics. China Economic Quarterly, 19(4):1401–1418, 2020.

[102] Els Gijsbrechts and Katrijn Gielens. "Logging Off": On Consumer Disadoption in Online Grocery Shopping. Technical report, Working Paper, 2021.

[103] Jaeung Sim, Daegon Cho, Youngdeok Hwang, and Rahul Telang. Frontiers: Virus Shook the Streaming Star: Estimating the COVID-19 Impact on Music Consumption. Marketing Science, 41(1):19–32, 2022.

[104] Raj Chetty, John N Friedman, Nathaniel Hendren, Michael Stepner, et al. The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data. Technical report, National Bureau of Economic Research, 2020.

[105] Guangying Chen, Seethu Seetharaman, and Luoyexin Shi. Shopping Behavior During the COVID-19 Pandemic: A Supermarket Perspective. Technical report, Working Paper, 2021.

[106] Yun Qiu, Xi Chen, and Wei Shi. Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China. Journal of Population Economics, 33(4):1127–1172, 2020.

[107] Hanming Fang, Long Wang, and Yang Yang. Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China. Journal of Public Economics, 191:104272, 2020.

[108] Solomon Hsiang, Daniel Allen, Sébastien Annan-Phan, Kendon Bell, Ian Bolliger, Trinetta Chong, Hannah Druckenmiller, Luna Yue Huang, Andrew Hultgren, Emma Krasovich, et al. The Effect of Large-Scale Anti-Contagion Policies on the COVID-19 Pandemic. Nature, 584(7820):262–267, 2020.

[109] Seth Flaxman, Swapnil Mishra, Axel Gandy, H Juliette T Unwin, Thomas A Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, et al. Estimating the Effects of Non-Pharmaceutical Interventions on COVID-19 in Europe. Nature, 584(7820):257–261, 2020.

[110] Meng Liu, Raphael Thomadsen, and Song Yao. Forecasting the Spread of COVID-19 under Different Reopening Strategies. Scientific reports, 10(1):1–8, 2020.

[111] John Jianqiu Bai, Shuili Du, Wang Jin, and Chi Wan. The Impact of Social Capital on Individual Responses to COVID-19 Pandemic: Evidence from Social Distancing. Working Paper, 2020.

[112] Ajay Shenoy, Bhavyaa Sharma, Guanghong Xu, Rolly Kapoor, Haedong Aiden Rho, and Kinpritma Sangha. God is in the Rain: The Impact of Rainfall-Induced Early Social Distancing on COVID-19 Outbreaks. Journal of Health Economics, 81:102575, 2022.

[113] David Holtz, Michael Zhao, Seth G Benzell, Cathy Y Cao, Mohammad Amin Rahimian, Jeremy Yang, Jennifer Allen, Avinash Collis, Alex Moehring, Tara Sowrirajan, et al. Interdependence and the Cost of Uncoordinated Responses to

COVID-19. Proceedings of the National Academy of Sciences, 117(33):19837–19843, 2020.

[114] Alexander W Bartik, Marianne Bertrand, Zoë B Cullen, Edward L Glaeser, Michael Luca, and Christopher T Stanton. How Are Small Businesses Adjusting to COVID-19? Early Evidence from a Survey. Technical report, National Bureau of Economic Research, 2020.

[115] Ruochen Dai, Dilip Mookherjee, Yingyue Quan, and Xiaobo Zhang. Industrial Clusters, Networks and Resilience to the Covid-19 Shock in China. Journal of Economic Behavior & Organization, 183:433–455, 2021.

[116] Xinyu Cao, Dennis Zhang, and Lei Huang. The Impact of COVID-19 Pandemic on Gig Economy Labor Supply. Working, 2020.

[117] Andrey Simonov, Szymon Sacher, Jean-Pierre Dubé, and Shirsho Biswas. Frontiers: The Persuasive Effect of Fox News: Noncompliance with Social Distancing During the COVID-19 Pandemic. Marketing Science, 41(2):230–242, 2022.

[118] Bradley Larsen, Marc J Hetherington, Steven H Greene, Timothy J Ryan, Rahsaan D Maxwell, and Steven Tadelis. Using Donald Trump's COVID-19 Vaccine Endorsement to Give Public Health a Shot in the Arm: A Large-Scale Ad Experiment. Technical report, National Bureau of Economic Research, 2022.

[119] Seth G Benzell, Avinash Collis, and Christos Nicolaides. Rationing Social Contact During the COVID-19 Pandemic: Transmission Risk and Social Benefits of US Locations. Proceedings of the National Academy of Sciences, 117(26):14642–14644, 2020.

[120] Shaun Larcom, Ferdinand Rauch, and Tim Willems. The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network. Quarterly Journal of Economics, 132(4):2019–2055, 2017.

[121] Jose Maria Barrero, Nicholas Bloom, and Steven J Davis. Why Working from Home Will Stick. Technical report, National Bureau of Economic Research, 2021.

[122] Seth Benzell, Avinash Collis, Christos Nicolaides, and Indranil Bardhan. Learning from covid-19: Hospital it, clinical trials and improvement in health outcomes. Clinical Trials and Improvement in Health Outcomes (December 1, 2020), 2020.

[123] John Jianqiu Bai, Erik Brynjolfsson, Wang Jin, Sebastian Steffen, and Chi Wan. Digital Resilience: How Work-from-home Feasibility Affects Firm Performance. Technical report, National Bureau of Economic Research, 2021.

[124] Lin William Cong, Xiaohang Yang, and Xiaobo Zhang. SMEs Amidst the Pandemic and Reopening: Digital Edge and Transformation. The SC Johnson College of Business Applied Economics and Policy Working Paper Series No, 2021.

[125] John B Burbidge, Lonnie Magee, and A Leslie Robb. Alternative transformations to handle extreme values of the dependent variable. Journal of the American Statistical Association, 83(401):123–127, 1988.

[126] Tomasz Piotr Wisniewski, Michal Polasik, Radoslaw Kotkowski, and Andrea Moro. Switching from cash to cashless payments during the covid-19 pandemic and beyond. Available at SSRN 3794790, 2021.

APPENDICES

Chapter A: Appendix for Chapter 1

Appendix A: Supplemental Materials

In Part A of the Online Appendix, we provide several supplemental materials to the main text.

A.1. Additional Results

This section provides additional results about the details of several propositions and the plots of relevant numerical examples that are omitted in the main text for concision.

A.1.1. Reselling Model ($\mathbb{C}$ mode)

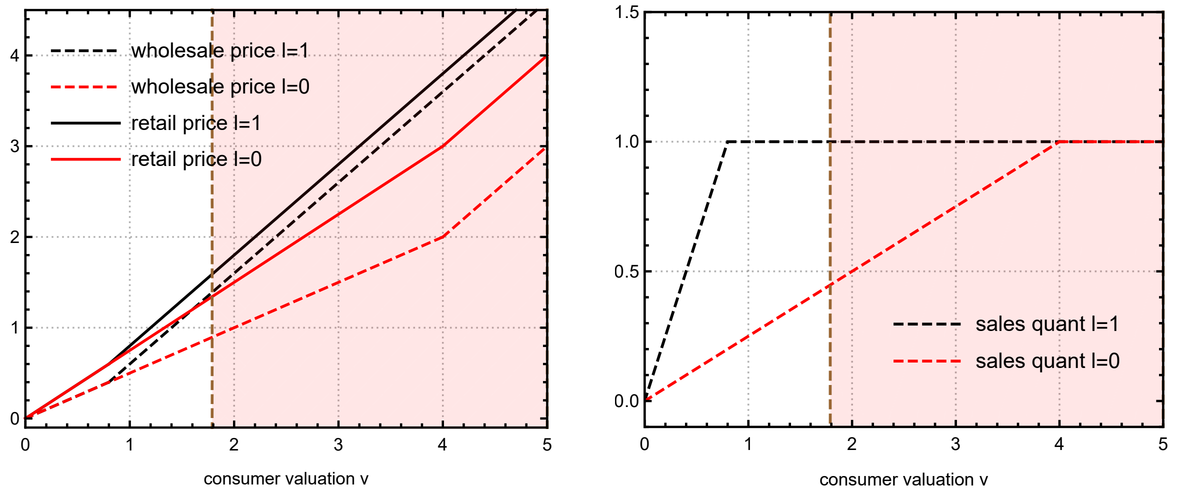


Figure A.1.: Wholesale price and sales quantity under the reselling model ($\mathbb{C}$ mode) w.r.t. v ($c = 0.2$, $\delta = 0.8$)

Figure A.1 depicts the wholesale price $\tilde{w}^*$, retail price $\tilde{p}^*$, and sales quantity $\mathcal{T}'(\tilde{p}^*)$ w.r.t. v under the GDPR policy, in which the shaded region represents the situation where the retailer does not collect consumer data. The consumers block the retailer

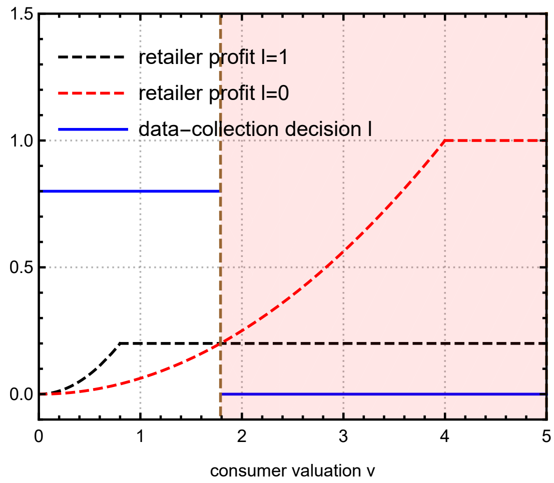


Figure A.2.: Retailer profit and data-disclosing decision under the reselling model ($\mathbb{C}$ mode) w.r.t. v ($c = 0.2$, $\delta = 0.8$)

from downside data exploitation, and the data monetization intensity is always 0. Not collecting consumer data increases the profit margin $\tilde{p}^* - \tilde{w}^*$.

Figure A.2 depicts the data collection decision $\tilde{\ell}^*$ and the retailer's profit under the GDPR policy. As consumer valuation v increases, switching to not collecting consumer data generates higher retailer profit, since it restricts the upstream supplier from charging a higher wholesale price.

A.1.2. Reselling Model: Welfare Comparison

In this part, we present the propositions and figures that are omitted in Section 1.5.

Proposition A.0.1 *If $v \leq 4\sqrt{1-\delta}$, the retailer collects consumer data regardless of the GDPR policy; if $4\sqrt{1-\delta} < v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the retailer only collects consumer data without the GDPR policy; if $v > \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the retailer never collects consumer data regardless of the GDPR policy.*

Proposition A.0.2 *If $v < \frac{4\sqrt{(1-\delta)\mathcal{M}}}{\mathcal{Z}}$, the GDPR policy increases the retailer profit; if $\frac{4\sqrt{(1-\delta)\mathcal{M}}}{\mathcal{Z}} < v < \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the GDPR policy reduces the retailer profit; otherwise, the GDPR policy does not affect the retailer profit.*

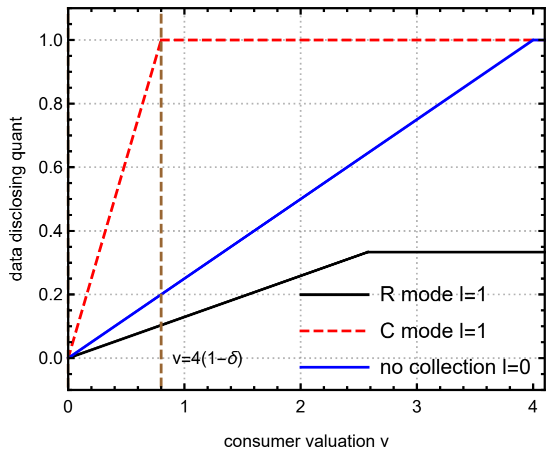
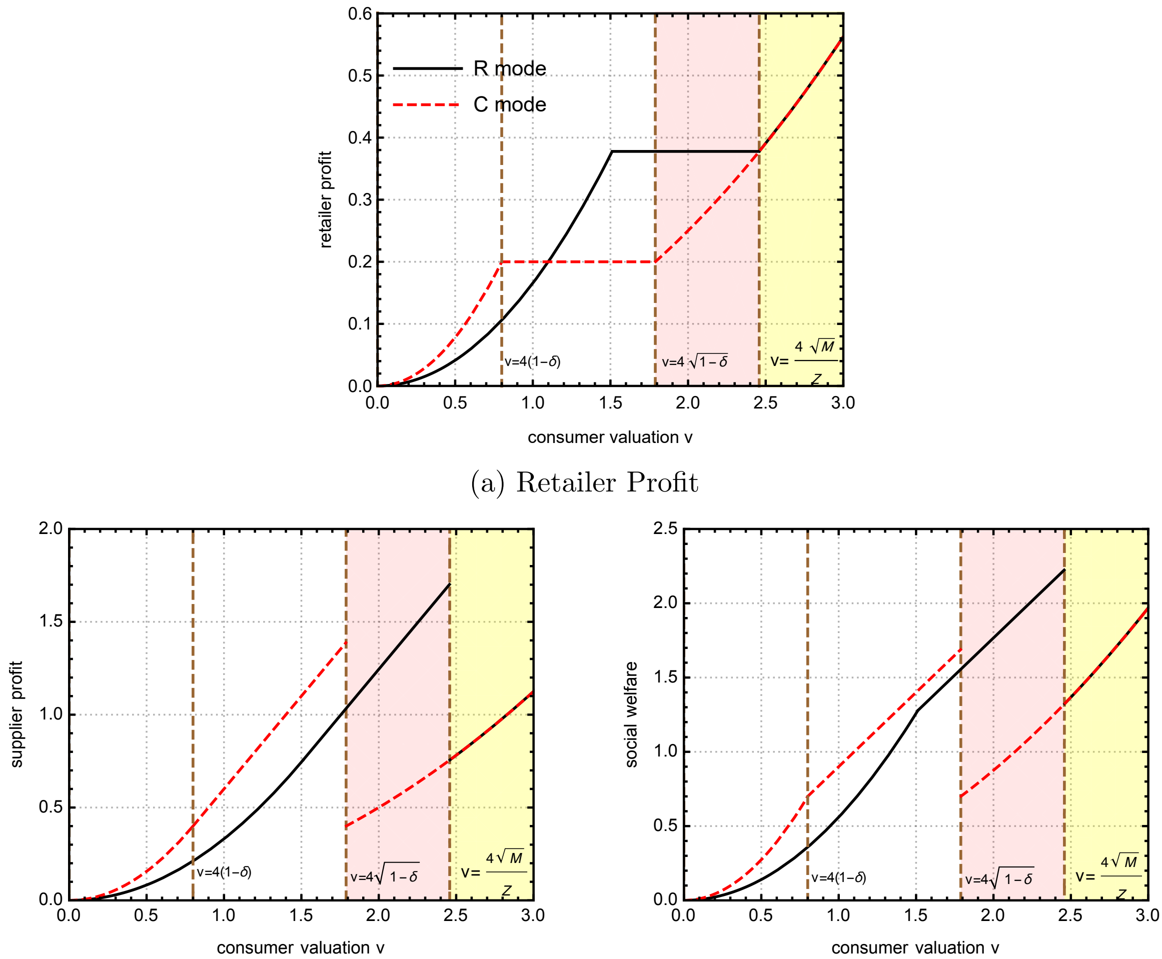


Figure A.3.: (Color Online) Comparison of the data-disclosing quantity under the reselling model w.r.t. v ($c = 0.8$, $\delta = 0.8$)



(a) Retailer Profit

(b) Supplier Profit

(c) Social Welfare

Figure A.4.: (Color Online) Comparison of the retailer's profit, the supplier's profit, and the social welfare under the reselling model w.r.t. v ($c = 0.8$, $\delta = 0.8$)

Proposition A.0.3 *If $v < 4\sqrt{1-\delta}$, the GDPR policy increases the supplier profit; if $4\sqrt{1-\delta} < v < \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the GDPR policy hurts the supplier profit; otherwise, the GDPR policy does not affect the supplier profit.*

Proposition A.0.4 *If $v < 4\sqrt{1-\delta}$, the GDPR policy increases the social welfare; if $4\sqrt{1-\delta} < v < \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, the GDPR policy hurts the social welfare; otherwise, the GDPR policy does not affect the social welfare.*

A.1.3. Agency Selling Model ($\mathbb{R}$ mode)

Proposition A.0.5 *Without the GDPR policy, the retailer always collects consumer data ($L^* = \delta$) and the optimal monetization intensity, retail price and fixed per-item fee are:*

$$(R^*, P^*, U^*) = \begin{cases} \left(\dfrac{v\lambda\delta\mathcal{Z}}{4\mathcal{M}'}, \dfrac{3v}{4} - \dfrac{vc\lambda^2\delta^2}{8\mathcal{M}'}, \dfrac{v}{2} - \dfrac{vc\lambda^2\delta^2}{4\mathcal{M}'}\right) & \text{if } v \leq \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}, \\ \left(\dfrac{\lambda\delta}{\mathcal{Z}}, v - 1 + \dfrac{2c\delta}{\mathcal{Z}}, v - 2 + \dfrac{4c\delta}{\mathcal{Z}}\right) & \text{if } v > \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}. \end{cases}$$

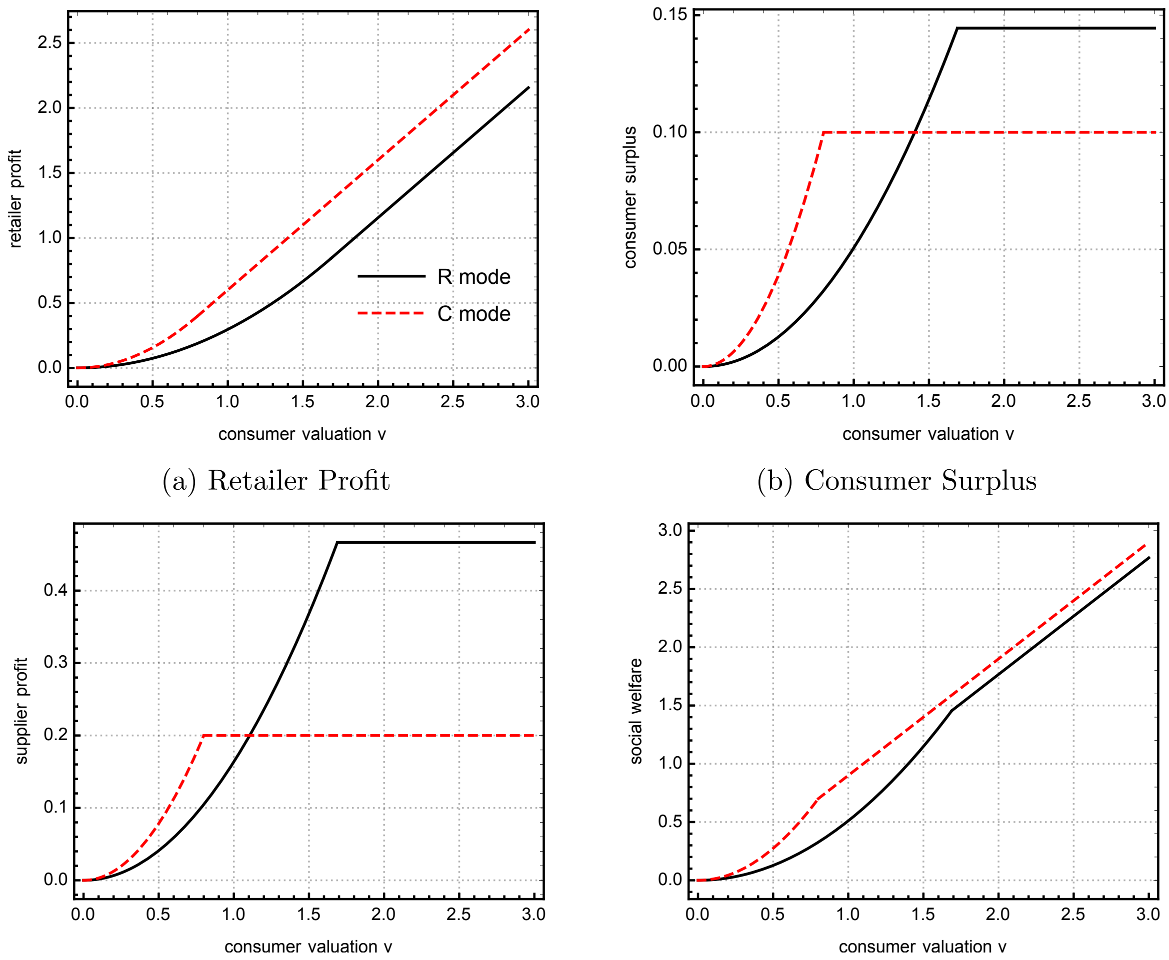
A.1.4. Agency Selling Model ($\mathbb{C}$ mode)

Proposition A.0.6 *With the GDPR policy, the retailer always collects data ($\tilde{L}^* = \delta$) and the optimal retail price and fixed per-item fee are:*

$$(\tilde{P}^*, \tilde{U}^*) = \begin{cases} (3v/4, v/2) & \text{if } v \leq 4(1-\delta), \\ (v - (1-\delta), v - 2(1-\delta)) & \text{if } v > 4(1-\delta). \end{cases}$$

A.1.5. Agency Selling Model: Welfare Comparison

Figure A.5 depicts the consumer surplus, retailer profit, supplier profit as well as social welfare with/without the GDPR policy. Figure A.6 shows the comparison of the consumer surplus, retailer profit as well as supplier profit between the GDPR policy and traditional data policy under the agency selling model. The shaded region represents where the GDPR policy generates higher surplus or profits.



(a) Retailer Profit

(b) Consumer Surplus

(c) Supplier Profit

(d) Social Welfare

Figure A.5.: (Color Online) Comparison of retailer profit, consumer surplus, supplier profit, and social welfare under the agency selling model w.r.t. v ($c = 0.8$, $\delta = 0.8$, $\lambda = 1$)

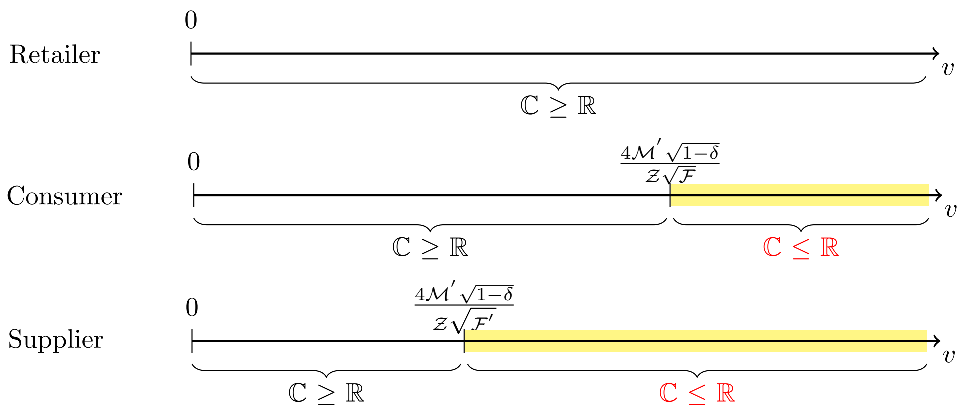


Figure A.6.: Comparison of retailer profit, consumer surplus, and supplier profit under two policies (agency selling model)

A.2. General Data Privacy Regulation (GDPR)

The General Data Protection Regulation (GDPR) is a regulation on data protection and privacy in the European Union. The GDPR is aimed primarily at giving control to individuals over their personal data and to simplify the regulatory environment for international business by unifying the regulation. The processing of individual data is based on consent. Consumer has eight rights: The right to information; The right of access; The right to rectification; The right to restrict processing; The right to erasure; The right to object; The right to an explanation; The right to data portability.

In this article, we mainly deal with the following two rights:

(1) **The right to information**. When a company, a government body, or an organisation collects and uses information, consumers have the right to get information about the reason for which the entity will use the data, the type of personal data the entity holds and whether data will be shared with third parties and who they are;

(2) **The right to object**. Consumers have the right to object to the collection, use, and storage of personal data by a company, government body, or organisation when the data is being used for direct marketing, automated decision making and scientific or historical research and statistics. Searching a website, individual consumer would receive a request of data collection. It clearly states for what purpose the consumer data is collected. Consumers could check each entry and decide the specific purpose to allow the website to collect data for.

A.3. Manage Your Cookie Settings: the LEGO Example

1. **Necessary Cookies (Always ON)**:
   Necessary cookies are required to enable technical site functionality and to provide the services explicitly requested by you. This includes as an example services such as your selected country and language, keeping you logged in, providing security and fraud prevention, having your digital shopping bag and wish list items stored while you browse, remembering volume settings, and you getting access to secure areas

of the website. This category of cookies cannot be disabled and does not require a consent (or in case of users outside of the EU/EEA – where appropriate - you have been deemed to give your consent by continuing to use this website and/or our services).

2. **Analytic Cookies (Upside Usage)**:

   These cookies are optional and collect information about how visitors use and experience our website in order to optimise design, operations, efficiency and to improve your user experience.

3. **LEGO Marketing Cookies (Downside Usage)**:

   We would also collect cookies to learn more about your interests, including which sites and ads you click on, which products and services you are interested in, or purchase, on this and other of our websites or apps. We use this data to show you more personal marketing and product recommendations on our websites or in our apps, in our membership and program offerings and to use the information about your interests and behavior on our website to make the content of any marketing messages we send the user more relevant based on your interests and site behavior.

4. **Third-Party Marketing Cookies (Downside Usage)**:

   We would like third parties to collect cookies allowing us to make targeted marketing/ads of our products and services on other websites, apps and on social media. If you allow this you will allow the listed third parties to set cookies tracking your interests and behavior including which products and services you are interested in, or purchase, on this and other websites, social media, apps and devices. Be aware, that these third parties are data controllers of the personal data tracked via the cookies and they will use the this data for their own purposes. These are the third parties we allow to place cookies on our websites and in our apps.

Online Appendix B: Proofs of Statements

We present all the proofs of the Lemmas and Propositions in Online Appendix B.

Proof of Lemma 1:

It can be verified that consumers with $\theta \leq v - p$ will choose to purchase but hide data. Consumers with $\theta \leq \mathcal{T}(s,p)$ will choose to purchase and disclose data at the same time. Consumers with $\theta \geq \lambda r(s)$ will find purchasing with data disclosure more profitable than purchasing with data hiding (the profit might be negative). Therefore, there exists two potential cases of consumer's purchasing and data-disclosing decisions, which are shown in Figure A.7.

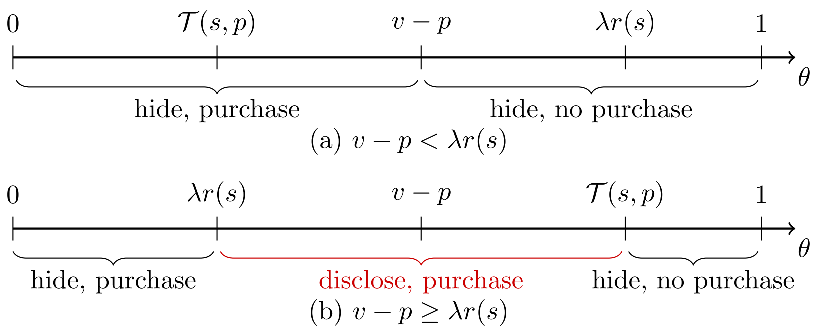


Figure A.7.: An illustrative example of consumer data disclosure and purchase decisions

Note that $\mathcal{T}(s,p) = \min\left\{\frac{v-p-\lambda r(s)\delta}{1-\delta}, 1\right\}$, from which we can show that the two cases are identical (i.e., $\mathcal{T}(s,p) = \lambda r(s) = v - p$) when $v - p = \lambda r(s)$. We have the first case if $v - p < \lambda r(s)$ and the second case if $v - p \geq \lambda r(s)$. The first case does not exist in equilibrium since nobody discloses data regardless of consumer's believed data-disclosing quantity s. As a result, the rational expectations equilibrium does not exist under the first case. Consumers automatically adjust their belief s and the first case would be transformed into the second case. Thus we only need to consider the second case, which proves the lemma.

Proof of Proposition 1.4.1:

We start with the discussion where the retailer collects data in the first stage. Recall from the analysis in the main text that $r(s) = \frac{\lambda\delta}{2c}s$. Adopting rational expectations equilibrium, we show that the number of consumers disclosing data is $s = \mathcal{T}(s,p) - \lambda r(s)$. From the above two equations, we can solve the data-disclosing quantity $s(p)$ and data monetization intensity $r(p)$:

$$\begin{cases} s(p) = \dfrac{2c(v-p)}{2c(1-\delta)+\lambda^2\delta}, \\ r(p) = \dfrac{\lambda\delta(v-p)}{2c(1-\delta)+\lambda^2\delta}. \end{cases}$$

We use backward induction to solve the retailer's pricing decision p. The retailer objective function is $\pi_r(w,p) = (p-w)\mathcal{T}(p)+\lambda r(p)\delta\left[\mathcal{T}(p) - \lambda r(p)\right] - cr(p)^2$. Inserting $r(p)$ into the objective function, it can be shown that the objective function is concave in p. Therefore, we derive the optimal retail price

$$p(w) = \frac{v+w}{2} - \frac{c\lambda^2\delta^2(v-w)}{2\mathcal{M}}. \tag{A.1}$$

Plugging the retailer's price $p(w)$ into the supplier's objective function

$$\pi_s(w) = w\mathcal{T}(p(w)) = w\min\left\{1, \frac{(v-w)\mathcal{Z}^2}{2\mathcal{M}}\right\},$$

we can derive the supplier's optimal wholesale price. Note that there exists two cases, depending on whether the sales quantity $\mathcal{T}(p(w))$ is binding at 1. We first analyze the no binding case. It can be readily shown that the optimal wholesale price $w^* = v/2$. We plug w^* into the no binding condition $\mathcal{T}(p(w)) < 1$, and we can show that the no binding equilibrium holds if $v < \frac{4\mathcal{M}}{\mathcal{Z}^2}$. Next consider the binding case where $\mathcal{T}(p(w)) = 1$. Under this condition, the supplier's profit becomes w, which is linearly increasing. Hence, the supplier would increase the wholesale price unless the binding condition is violated. In other words, the supplier charges the wholesale price until the market is just saturated at $\frac{(v-w)\mathcal{Z}^2}{2\mathcal{M}} = 1$. Therefore, we have the supplier's optimal wholesale price $w^* = v - \frac{2\mathcal{M}}{\mathcal{Z}^2}$. Next consider the no collection case (i.e., $\ell = 0$), under which the retailer sets the monetization intensity $r = 0$. Consumers with $\theta \leq v - p$ purchase the product and generate positive utility. With the market demand $v - p$ at price p, the retailer's objective function is $(p-w)(v-p)$. The optimal retail price $p(w) = \frac{v+w}{2}$. By plugging $p(w)$ into the supplier's objective function $w(v-p)$, we can solve $w^* = v/2$ and $p^* = 3v/4$. Similar to the data collection analysis, the above equilibrium corresponds to the non-binding situation, where the condition can be simplified as $v \leq 4$. If $v > 4$, we can derive the equilibrium with $w^* = v - 1$ and $p^* = v - 1$.

Finally, we analyze the first stage where the retailer decides whether or not to collect data. From the above analysis, if the retailer collects data, her profit can be written as:

$$\begin{cases} \dfrac{v^2 \mathcal{Z}}{16\mathcal{M}} & \text{if} \;\; v \leq \dfrac{4\mathcal{M}}{\mathcal{Z}^2}, \\ 1 + \dfrac{2c^2\delta}{\mathcal{Z}^2} - \dfrac{3c\delta}{\mathcal{Z}} & \text{if} \;\; v > \dfrac{4\mathcal{M}}{\mathcal{Z}^2}. \end{cases}$$

If the retailer does not collect data, her profit is:

$$\begin{cases} \dfrac{v^2}{16} & \text{if} \;\; v \leq 4, \\ 1 & \text{if} \;\; v > 4. \end{cases}$$

Observe that under the non-binding situation, collecting data yields a higher retailer profit: $\frac{v^2}{16} \leq \frac{v^2 \mathcal{Z}}{16\mathcal{M}}$; under the binding situation, not collecting data generates a higher retailer: profit $1 \geq 1 + \frac{2c^2\delta}{\mathcal{Z}^2} - \frac{3c\delta}{\mathcal{Z}}$; and as v increases, collecting data is "easier" to be binding: $4 \geq \frac{4\mathcal{M}}{\mathcal{Z}^2}$. Combining these results and comparing the retailer's profits (see point C' in Figure A.8b), we obtain the threshold $v = \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, at which the retailer is indifferent between collecting and not collecting data.

Proof of Proposition 1.4.2:

We first consider the sensitivity study of λ. Based on the discussion of Proposition 1.4.1, we divide the parameter region of v into four sub-regions: $v \leq \frac{4\mathcal{M}}{\mathcal{Z}^2}$, $\frac{4\mathcal{M}}{\mathcal{Z}^2} < v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, $\frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}} < v \leq 4$, and $v > 4$ as I, II, III, and IV, respectively. Note that for region III and region IV, the retailer does not collect data and thus $r^* = 0$, which leads to a trivial case that can be omitted. Recall that for I and II, the cutoff condition is $v = \frac{4\mathcal{M}}{\mathcal{Z}^2}$. We can verify that $\frac{4\mathcal{M}}{\mathcal{Z}^2}$ is increasing in λ. Hence, as λ increases, the equilibrium switches from II to I. In region II, the optimal monetization intensity $r^* = \frac{\lambda\delta}{\mathcal{Z}}$. For simplicity, we denote $\mathcal{H} = \sqrt{\frac{2c}{\delta}}$, and it can be verified that r^* is concave in λ and the first-order condition holds when $\lambda = \mathcal{H}$. As λ keeps increasing, the equilibrium is in region I, where $r^* = \frac{v\lambda\delta\mathcal{Z}}{4\mathcal{M}}$. We take the first derivative of r^* with respect to λ, and we have:

$$\frac{-v\delta[-8c^3(1-\delta) - 2c^2\lambda^2\delta(2-3\delta) + c\lambda^4\delta^2(2+3\delta) + \lambda^6\delta^3]}{4[4c^2(1-\delta) + c\lambda^2\delta(4-3\delta) + \lambda^4\delta^2]^2}.$$

It is clear that the denominator is always positive. We denote the second part of numerator $\mathcal{J} = -8c^3(1-\delta) - 2c^2\lambda^2\delta(2-3\delta) + c\lambda^4\delta^2(2+3\delta) + \lambda^6\delta^3$. Taking the derivative of $\mathcal{J}$ with respect to λ gives:

$$-4c^2\lambda\delta(2-3\delta) + 4c\lambda^3\delta^2(2+3\delta) + 6\lambda^5\delta^3. \tag{A.2}$$

We can verify that if $\mathcal{J}$ is positive, the derivative Equation A.2 is positive as well. In other words, once $r^* = \frac{v\lambda\delta\mathcal{Z}}{4\mathcal{M}}$ starts decreasing in λ (i.e., $\mathcal{J} > 0$), it never increases again. Thus we show $r^* = \frac{v\lambda\delta\mathcal{Z}}{4\mathcal{M}}$ is quasi-concave in λ. We can show that the global maximum is achieved at $\lambda = \mathcal{Y}$, where $\mathcal{Y} = \arg\max_\lambda \frac{v\lambda\delta\mathcal{Z}}{4\mathcal{M}}$. Note that $\mathcal{Y} \leq \mathcal{H}$. Combining the two regions, we can verify r^* is quasi concave in λ and derive the corresponding results.

We next turn to the sensitivity study of δ. It is straightforward to show r^* is increasing in δ for both region I and region II. This completes the proof.

Proof of Proposition 1.4.3:

Following the logic in the proof of Proposition 1.4.1, we analyze the model based on the retailer's data collection decision. First consider the case of collecting data. Note that $r^* = 0$ under the $\mathbb{C}$ mode. Consumers only allows upside exploitation and the corresponding utility of purchasing is $v - (1-\delta)\theta - p$. Hence, only consumers with $\theta \leq \frac{v-p}{1-\delta}$ purchase the product and obtain a positive utility. Given the wholesale price w and retail price p, the retailer's profit is $(p-w)\frac{v-p}{1-\delta}$, and we derive $p(w) = \frac{v+w}{2}$. We can rewrite the supplier's profit as $w\frac{v-p(w)}{1-\delta}$. Plug in $p(w)$ and we derive $w^* = v/2$. This represents the no binding situation. Using the same approach as in the discussion of $\mathbb{R}$ mode, the equilibrium holds if $\frac{v-p}{1-\delta} \leq 1$, where we can rewrite the no binding condition as $v \leq 4(1-\delta)$. If $v \geq 4(1-\delta)$, we can derive the equilibrium under the binding condition: $p^* = v - (1-\delta)$ and $w^* = v - 2(1-\delta)$.

Next consider the case of no data collection under the $\mathbb{C}$ mode, which is similar to the analysis under the $\mathbb{R}$ mode. Following the same procedure, we find that under the non-binding situation, collecting data yields a higher retailer profit $\frac{v^2\mathcal{Z}}{16\mathcal{M}} \leq \frac{v^2}{16(1-\delta)}$; under the binding situation, not collecting data generates a higher retailer profit: $1+\frac{2c^2\delta}{\mathcal{Z}^2}-\frac{3c\delta}{\mathcal{Z}} \geq 1-\delta$; and as v increases, collecting data is "easier" to be binding: $4 \geq 4(1-\delta)$. Thus we can derive the threshold for data collection: $v = 4\sqrt{1-\delta}$, where the retailer is indifferent between collecting and not collecting data (see point B' in Figure A.8b).

Proof of Proposition 1.5.1:

To analyze the consumer surplus under the $\mathbb{R}$ mode and $\mathbb{C}$ mode, we have three scenarios to consider: (a) collecting data under the $\mathbb{R}$ mode, (b) collecting data under the $\mathbb{C}$ mode, and (c) no data collection, which is the same under both modes. Based on the analysis of Proposition 1.4.1 and Proposition 1.4.3, we can derive the consumer surplus under scenario (a) as follows:

$$\begin{cases} \dfrac{v^2\mathcal{Z}^2\mathcal{F}}{32\mathcal{M}^2} & \text{if } \ v \leq \dfrac{4\mathcal{M}}{\mathcal{Z}^2}, \\ \dfrac{\mathcal{F}}{2\mathcal{Z}^2} & \text{if } \ v > \dfrac{4\mathcal{M}}{\mathcal{Z}^2}. \end{cases}$$

The consumer surplus under scenario (b) is given by:

$$\begin{cases} \dfrac{v^2}{32(1-\delta)} & \text{if } \ v \leq 4(1-\delta), \\ \dfrac{1-\delta}{2} & \text{if } \ v > 4(1-\delta). \end{cases}$$

And the consumer surplus under scenario (c) is given by:

$$\begin{cases} \dfrac{v^2}{32} & \text{if } \ v \leq 4, \\ \dfrac{1}{2} & \text{if } \ v > 4. \end{cases}$$

We can verify that as v increases from 0, scenario (b) is the first to become binding, followed by scenario (a) and scenario (c), i.e., $4 \geq \frac{4\mathcal{M}}{\mathcal{Z}^2} \geq 4(1-\delta)$. Under the non-binding situation, scenario (b) yields the highest consumer surplus, followed by scenario (a) and scenario (c), i.e., $\frac{v^2}{32(1-\delta)} \geq \frac{v^2\mathcal{Z}^2\mathcal{F}}{32\mathcal{M}^2} \geq \frac{v^2}{32}$. Under the binding situation, scenario (c) yields the highest consumer surplus, followed by scenario (a) and scenario (b), i.e.,

$\frac{1-\delta}{2} \leq \frac{\mathcal{F}}{2\mathcal{Z}^2} \leq \frac{1}{2}$. This analysis can be depicted in Figure A.8a. Point A represents that, if the retailer collects data under both modes, the consumer surplus is the same between the $\mathbb{C}$ mode and $\mathbb{R}$ mode. Point B represents the threshold where the consumer surplus is indifferent between collecting and not collecting data under the $\mathbb{C}$ mode. Point C represents the threshold where the consumer surplus is indifferent between collecting and not collecting data under the $\mathbb{R}$ mode. We can obtain the corresponding horizontal axis coordinates of points A, B and C as follows: $v = \frac{4\mathcal{M}\sqrt{1-\delta}}{\mathcal{Z}\sqrt{\mathcal{F}}}$, $v = 4\sqrt{1-\delta}$ and $v = \frac{4\sqrt{\mathcal{F}}}{\mathcal{Z}}$. From Propositions 1.4.1 and 1.4.3, we know that under the $\mathbb{C}$ mode, the threshold where the retailer is indifferent between collecting and not collecting data is $v = 4\sqrt{1-\delta}$ (see point B' in Figure A.8b), which is the same as point B. Note that before reaching point B, the retailer collects data under both modes, and the $\mathbb{C}$ mode yields a higher consumer surplus if v is between 0 and point A. Under the $\mathbb{R}$ mode, the threshold where the retailer is indifferent between collecting and not collecting data is $v = \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$ (see point C' in Figure A.8b), which is to the right of point C. In other words, under the $\mathbb{R}$ mode, even if the consumer surplus becomes worse-off with data collection (i.e., when v is between point C and point C'), the retailer still collects data. Hence, the $\mathbb{R}$ mode reduces the consumer surplus compared to the $\mathbb{C}$ mode, since the retailer already switches to not collecting data under the $\mathbb{C}$ mode (point B). Combining all the above results proves the proposition.

Proof of Proposition 1.5.2:

The proof is straightforward by combining the proofs of Proposition 1.5.1 (consumer surplus), Proposition A.0.2 (retailer profit), and Proposition A.0.3 (supplier profit).

Proof of Proposition 1.5.3:

The case of no data collection has been discussed in Proposition 1.4.1 and Proposition 1.4.3. The proof is similar and therefore omitted.

Proof of Proposition 1.6.1:

We know from Proposition A.0.5 and Proposition A.0.6 that the retailer always collects consumer data under the agency selling model. We first analyze the consumer surplus. Given the optimal decisions in Propositions A.0.5 and A.0.6, we can derive the consumer surplus. Under the $\mathbb{R}$ mode, we have the consumer surplus under the non-binding and binding situations:

$$\begin{cases} \dfrac{v^2\mathcal{Z}^2\mathcal{F}}{32\mathcal{M}'^2} & \text{if } \ v \le \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}, \\ \dfrac{\mathcal{F}}{2\mathcal{Z}^2} & \text{if } \ v > \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}. \end{cases}$$

Under the $\mathbb{C}$ mode, we have the following consumer surplus:

$$\begin{cases} \dfrac{v^2}{32(1-\delta)} & \text{if } \ v \le 4(1-\delta), \\ \dfrac{1-\delta}{2} & \text{if } \ v > 4(1-\delta). \end{cases}$$

We can verify that the following conditions hold: (1) $\frac{4\mathcal{M}'}{\mathcal{Z}^2} \ge 4(1-\delta)$, which means that as v increases, the $\mathbb{C}$ mode is easier to be binding; (2) $\frac{v^2}{32(1-\delta)} \ge \frac{v^2\mathcal{Z}^2\mathcal{F}}{32\mathcal{M}^2}$, meaning that under the non-binding situation, the $\mathbb{C}$ mode generates a higher consumer surplus than the $\mathbb{R}$ mode; and (3) $\frac{1-\delta}{2} \le \frac{\mathcal{F}}{2\mathcal{Z}^2}$, which means that under the binding situation, the $\mathbb{R}$ mode generates a higher consumer surplus than the $\mathbb{C}$ mode. The above analysis is illustrated in Figure A.9a. Point A represents the threshold where the consumer surplus is indifferent between the $\mathbb{C}$ mode and the $\mathbb{R}$ mode. We derive the horizontal axis coordinate of A as $v = \frac{4\sqrt{1-\delta}\mathcal{M}'}{\mathcal{Z}\sqrt{\mathcal{F}}}$.

We next turn to the supplier profit. Following the similar discussion as in the analysis of consumer surplus, we have the supplier profit under the $\mathbb{R}$ mode as follows:

$$\begin{cases} \dfrac{v^2\mathcal{Z}^2\mathcal{F}'}{16\mathcal{M}'^2} & \text{if } \ v \le \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}, \\ 1 - \dfrac{2c\delta}{\mathcal{Z}} & \text{if } \ v > \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}. \end{cases}$$

The supplier profit under the $\mathbb{C}$ mode is:

$$\begin{cases} \dfrac{v^2}{16(1-\delta)} & \text{if } \ v \le 4(1-\delta), \\ 1-\delta & \text{if } \ v > 4(1-\delta). \end{cases}$$

Again, we can verify that as v increases, the $\mathbb{C}$ mode is easier to be binding, i.e., $\frac{4\mathcal{M}'}{\mathcal{Z}^2} \ge 4(1-\delta)$. Under the non-binding situation, the $\mathbb{C}$ mode generates a higher consumer surplus than the $\mathbb{R}$ mode, i.e., $\frac{v^2}{16(1-\delta)} \geq \frac{v^2 \mathcal{Z}^2 \mathcal{F}'}{16\mathcal{M}'^2}$. Under the binding situation, the $\mathbb{R}$ mode generates a higher consumer surplus than the $\mathbb{C}$ mode, i.e., $1 - \delta \leq 1 - \frac{2c\delta}{\mathcal{Z}}$. The above analysis is depicted in Figure A.9c. Point A' represents the threshold where the supplier's profit is indifferent between the $\mathbb{C}$ mode and the $\mathbb{R}$ mode. We can derive the horizontal axis coordinate of A' as $v = \frac{4\sqrt{1-\delta}\mathcal{M}'}{\mathcal{Z}\sqrt{\mathcal{F}'}}$.

Finally, we turn to the retailer's profit. As before, the retailer always collects data under the two data policies and we have the retailer profit under the $\mathbb{R}$ mode as follows:

$$\begin{cases} \dfrac{v^2 \mathcal{Z}^2}{8\mathcal{M}'} & \text{if} \ \ v \leq \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}, \\ v - 2 + \dfrac{c\delta(8c + 5\delta\lambda^2)}{\mathcal{Z}^2} & \text{if} \ \ v > \dfrac{4\mathcal{M}'}{\mathcal{Z}^2}. \end{cases}$$

The retailer profit under the $\mathbb{C}$ mode is:

$$\begin{cases} \dfrac{v^2}{8(1-\delta)} & \text{if} \ \ v \leq 4(1-\delta), \\ v - 2 + 2\delta & \text{if} \ \ v > 4(1-\delta). \end{cases}$$

We can verify that as v increases, the $\mathbb{C}$ mode is easier to be binding, i.e., $\frac{4\mathcal{M}'}{\mathcal{Z}^2} \geq 4(1-\delta)$. Under the non-binding situation, the $\mathbb{C}$ mode generates a higher consumer surplus than the $\mathbb{R}$ mode, i.e., $\frac{v^2}{8(1-\delta)} \geq \frac{v^2 \mathcal{Z}^2}{8\mathcal{M}'}$. Under the non-binding situation, the $\mathbb{C}$ mode also generates a higher consumer surplus than the $\mathbb{R}$ mode, i.e., $v - 2 + 2\delta \geq v - 2 + \frac{c\delta(8c+5\delta\lambda^2)}{\mathcal{Z}^2}$. All these results are summarized in Figure A.9b. We can see that the $\mathbb{C}$ mode always generates a higher retailer profit than the $\mathbb{R}$ mode. This completes the proof.

Proof of Proposition 1.6.2:

The proof is similar to that of Proposition 1.6.1 and thus omitted.

Proof of Proposition A.0.1:

Based on Proposition 1.4.1 and Proposition 1.4.3, we know that the retailer collects data if $v \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$ under the $\mathbb{R}$ mode; the retailer collects data if $v \leq 4\sqrt{1-\delta}$ under the $\mathbb{C}$ mode. It can be verified that $4\sqrt{1-\delta} \leq \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$, which proves the proposition.

Proof of Proposition A.0.2:

Following the proof of Proposition 1.5.1, we analyze the following three scenarios for the retailer profit: (a) collecting data under the $\mathbb{R}$ mode, (b) collecting data under the $\mathbb{C}$ mode, and (c) no data collection. Plug in the optimal decisions derived from Propositions 1.4.1 and 1.4.3, we can rewrite the retailer's profit under scenario (a) as:

$$\begin{cases} \dfrac{v^2\mathcal{Z}^2}{16\mathcal{M}} & \text{if } \ v \leq \dfrac{4\mathcal{M}}{\mathcal{Z}^2}, \\ 1 + \dfrac{2c^2\delta}{\mathcal{Z}^2} - \dfrac{3c\delta}{\mathcal{Z}} & \text{if } \ v > \dfrac{4\mathcal{M}}{\mathcal{Z}^2}. \end{cases}$$

The retailer profit under scenario (b) is:

$$\begin{cases} \dfrac{v^2}{16(1-\delta)} & \text{if } \ v \leq 4(1-\delta), \\ (1-\delta) & \text{if } \ v > 4(1-\delta). \end{cases}$$

The retailer profit under scenario (c) is given by:

$$\begin{cases} \dfrac{v^2}{16} & \text{if } \ v \leq 4, \\ 1 & \text{if } \ v > 4. \end{cases}$$

We can verify that as v increases from 0, scenario (b) is the first to be binding, followed by scenario (a) and scenario (c), i.e., $4 \geq \frac{4\mathcal{M}}{\mathcal{Z}^2} \geq 4(1-\delta)$. Under the non-binding situation, scenario (b) yields the highest retailer profit, followed by scenario (a) and scenario (c), i.e., $\frac{v^2}{16(1-\delta)} \geq \frac{v^2\mathcal{Z}^2}{16\mathcal{M}} \geq \frac{v^2}{16}$. Under the binding situation, scenario (c) generates the highest retailer profit, followed by scenario (a) and scenario (b), i.e., $\frac{1-\delta}{2} \leq 1 + \frac{2c^2\delta}{\mathcal{Z}^2} - \frac{3c\delta}{\mathcal{Z}} \leq \frac{1}{2}$. All the above analysis is depicted in Figure A.8b. Point A' represents that, if the retailer collects data under both modes, her profit is the same between the $\mathbb{C}$ mode and the $\mathbb{R}$ mode. Point B' represents the threshold where the retailer's profit is indifferent between collecting and not collecting data under the $\mathbb{C}$ mode. Point C' represents the threshold where the retailer's profit is indifferent between collecting and not collecting data under the $\mathbb{R}$ mode. We can derive that the corresponding horizontal axis coordinates of point A', B' and C' are: $v = \frac{4\sqrt{(1-\delta)\mathcal{M}}}{\mathcal{Z}}$, $v = 4\sqrt{1-\delta}$, and $v = \frac{4\sqrt{\mathcal{M}}}{\mathcal{Z}}$. Note that point B' is the threshold where the retailer is indifferent between collecting and not collecting data under the $\mathbb{C}$ mode, and point C' is the threshold where the retailer is indifferent between collecting and not collecting data under the $\mathbb{R}$ mode. The above results together prove the proposition.

Proof of Proposition A.0.3:

Following the proof of Proposition 1.5.1, we analyze the following three scenarios for the supplier profit: (a) collecting data under the $\mathbb{R}$ mode, (b) collecting data under the $\mathbb{C}$ mode, and (c) no data collection. Plug in the optimal decisions derived from Propositions 1.4.1 and 1.4.3, we can rewrite the supplier profit under scenario (a) as follows:

$$\begin{cases} \dfrac{v^2\mathcal{Z}^2}{8\mathcal{M}} & \text{if } \ v \leq \dfrac{4\mathcal{M}}{\mathcal{Z}^2}, \\ v - 2 + \dfrac{2c\delta(4c+3\delta\lambda^2)}{\mathcal{Z}^2} & \text{if } \ v > \dfrac{4\mathcal{M}}{\mathcal{Z}^2}. \end{cases}$$

The supplier profit under scenario (b) is given by:

$$\begin{cases} \dfrac{v^2}{8(1-\delta)} & \text{if } \ v \leq 4(1-\delta), \\ v - 2(1-\delta) & \text{if } \ v > 4(1-\delta). \end{cases}$$

The supplier profit under scenario (c) is:

$$\begin{cases} \dfrac{v^2}{8} & \text{if } \ v \leq 4, \\ v - 2 & \text{if } \ v > 4. \end{cases}$$

We can verify that as v increases from 0, scenario (b) is the first to be binding, followed by scenario (a) and scenario (c), i.e., $4 \geq 4\frac{\mathcal{M}}{\mathcal{Z}^2} \geq 4(1-\delta)$. Under the nonbinding situation, scenario (b) yields the highest supplier profit, followed by scenario (a) and scenario (c), i.e., $\frac{v^2}{8(1-\delta)} \geq \frac{v^2\mathcal{Z}^2}{8\mathcal{M}} \geq \frac{v^2}{8}$. Under the binding situation, scenario (b) generates the highest supplier profit, followed by scenario (a) and scenario (c), i.e., $v - 2(1-\delta) \geq v - 2 + \frac{2c\delta(4c+3\delta\lambda^2)}{\mathcal{Z}^2} \geq v - 2$. The above analysis is depicted in Figure A.8c. Note that not collecting data always generates the lowest supplier profit under both modes. Hence the supplier always prefers data collection. When data is collected, the $\mathbb{C}$ mode always generates the highest supplier profit. Recall that under the $\mathbb{C}$, the retailer is more likely to switch to not collecting data, which hurts the supplier's profit. The proof is complete by combining the above results.

Proof of Proposition A.0.4:

We can prove this result by combining the analyses of the consumer surplus, retailer profit and supplier profit in Propositions 1.5.1, Proposition A.0.2, and A.0.3.

Proof of Proposition A.0.5:

Similar to the reselling model, we first consider the case where the retailer collects data. Recall that, given the fixed fee u and the supplier's direct sales price p, the last two stages (i.e., consumer purchasing decision and retailer's data monetization decision) are the same as that in the baseline reselling model. Hence we have:

$$r(p) = \frac{\lambda\delta(v-p)}{2c(1-\delta)+\lambda^2\delta}, \qquad s(p) = \frac{2c(v-p)}{2c(1-\delta)+\lambda^2\delta}.$$

Next we investigate the third stage, in which the supplier's objective function is: $(p - u)\frac{v-p-\lambda r(p)\delta}{1-\delta}$. Plug $r(p)$ into the supplier's profit, and we can verify that the objective function is concave in p. Thus the optimal sales price is $p(u) = \frac{v+u}{2}$. Then we go back to the second stage in which the retailer decides the fixed per-transaction fee u to maximize the objective function: $u\frac{v-p-\lambda r(p)\delta}{1-\delta} + \lambda r(p)\delta s(p) - cr^2(p)$. Plugging $s(p)$ and $r(p)$ into the retailer's objective function, we can solve the optimal per-item fee $U^* = \frac{v}{2} - \frac{vc\lambda^2\delta^2}{4\mathcal{M}'}$. Note that similar to the baseline reselling model, the above equilibrium applies to the non-binding situation, where the sales quantity $\frac{v-p-\lambda r(p)\delta}{1-\delta} < 1$. We plug the optimal decisions into the threshold and rewrite it as: $v \leq \frac{4\mathcal{M}'}{\mathcal{Z}^2}$. Next we consider the binding situation, where the upstream retailer would charge the fixed fee u until the market is just saturated. Following the analysis in the reselling model, we derive the equilibrium fixed fee and retail price under the binding situation as follows:

$$\begin{cases} U^* = v - 2 + \dfrac{4c\delta}{\mathcal{Z}}, \\ P^* = v - 1 + \dfrac{2c\delta}{\mathcal{Z}}. \end{cases}$$

We then analyze the no data collection case. Under this condition, we have $r^* = 0$. Only consumers with $\theta \leq v - p$ will purchase the product. Thus the supplier's objective function becomes $(p - u)(v - p)$, from which we obtain $p(u) = \frac{v+u}{2}$. We work backward to the retailer's objective function $u(v - p)$ and derive $U^* = v/2$ and $P^* = 3v/4$. Again this equilibrium holds under non-binding situation, where $(v - p) < 1$. The threshold can be simplified as $v < 4$. If $v \geq 4$, we have $U^* = v - 2$ and $P^* = v - 1$. It can be readily verified that the retailer's profit under the no collection case is strictly dominated by the data collection case. Thus the retailer always collects consumer data under the $\mathbb{R}$ mode.

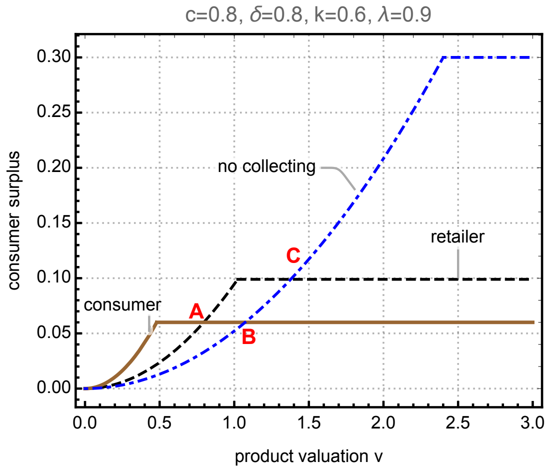
Proof of Proposition A.0.6:

We start with the retailer's data collection decision. Under the $\mathbb{C}$ mode, we have $r = 0$ since consumers only allows upside exploitation, and thus data monetization yields no profit for the retailer. The consumers' utility is $v - (1-\delta)\theta - p$ and only those with $\theta \leq \frac{v-p}{1-\delta}$ disclose data and purchase the product. The supplier's objective function is $(p - u)\frac{v-p}{1-\delta}$, from which we derive the optimal sales price $p(u) = \frac{v+u}{2}$. The retailer's objective function is $u\frac{v-p(u)}{1-\delta}$ and we can solve $\tilde{U}^* = v/2$. Again, this non-binding equilibrium holds when $\frac{v-p(u)}{1-\delta} \leq 1$, which we can simplify to $v \leq 4(1-\delta)$. When $v > 4(1-\delta)$, following the same logic in the proof of proposition A.0.5, we have

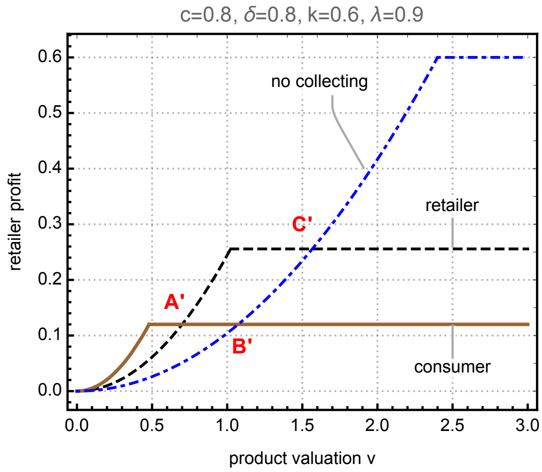
$$\tilde{U}^* = v - 2(1-\delta), \qquad \tilde{P}^* = v - (1-\delta).$$

The no data collection case is similar to the proof of Proposition A.0.5. It is straightforward to check that the retailer's profit with no data collection is strictly dominated by collecting data. Hence under the $\mathbb{C}$ mode, the retailer always collects consumer data.

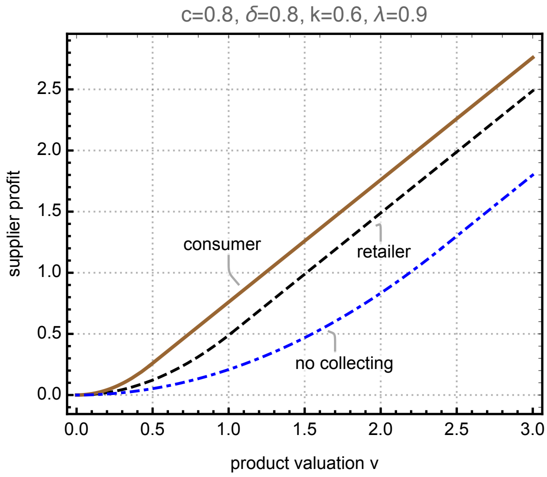
Figure Reference in Appendix B



(a) Consumer Surplus

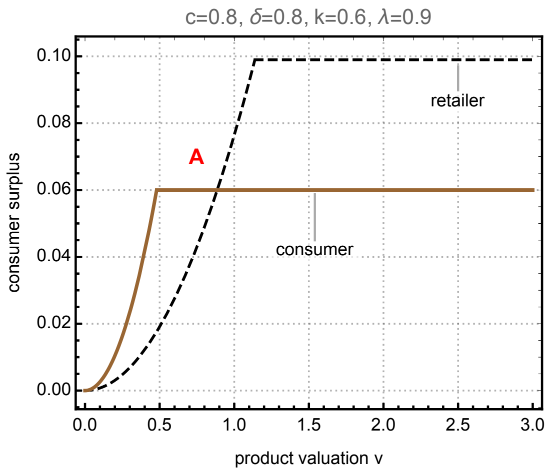


(b) Retailer Profit

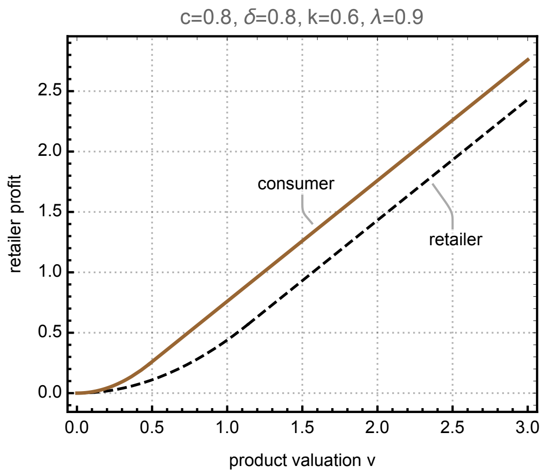


(c) Supplier Profit

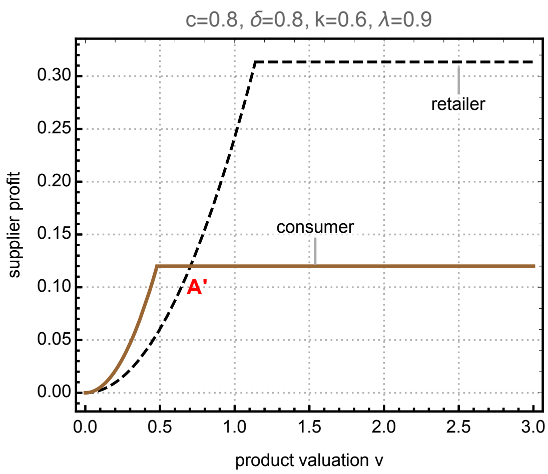
Figure A.8.: (Color Online) The relationship of consumer surplus, retailer profit, and supplier profit among three scenarios under the reselling model w.r.t. v ($c = 0.8$, $\delta = 0.8$, $\lambda = 0.9$)



(a) Consumer Surplus



(b) Retailer Profit



(c) Supplier Profit

Figure A.9.: (Color Online) The relationship of consumer surplus, retailer profit, and supplier profit among three scenarios under the agency selling model w.r.t. v ($c = 0.8$, $\delta = 0.8$, $\lambda = 0.9$)

Chapter B: Appendix for Chapter 2

Appendix A: Supplemental Materials

In Appendix A, we provide several supplemental materials to the main text.

A.1. Model Extension: Fire Sale of Accounts Receivable

In this section, we present mathematical details of the fire sale model. We denote the fire sale premium as $\eta_{\mathcal{S}}$. On the one hand, as an emergency tool, the fire sale premium cannot be too low. So we assume $\eta_{\mathcal{S}} \geq \underline{\eta}_{\mathcal{S}}$, where $\underline{\eta}_{\mathcal{S}} = \min\left\{\ln\left[\frac{\rho_s e^{\eta_{\mathcal{F}} t_2}}{1-(1-\rho_s)e^{\eta_{\mathcal{F}} t_2}}\right]/t_2', \eta_{\mathcal{I}}\right\}$. This assumption avoids the discussion of unrealistic situation where the supplier strategically uses fire sale as a regular financing method over factoring and invoice trading. The detailed discussion of $\underline{\eta}_{\mathcal{S}}$ is provided in the proof of Proposition B.0.1 and Proposition B.0.2. On the other hand, the fire sale premium cannot be too high. Otherwise, fire sale is too expensive, leading to the failure of the bank loan repayment due to insufficient funds raised by fire sale. Under this situation, the fire sale model reduces to our main model where the supplier directly goes bankrupt. We assume $\eta_{\mathcal{S}} \leq \bar{\eta}_{\mathcal{S}}$, where $\bar{\eta}_{\mathcal{S}} = \ln\left[\frac{1-\rho_r}{1-q_{\mathcal{BS}}^{t*} z(q_{\mathcal{BS}}^{t*})}\right]/t_2'$. Note that $q_{\mathcal{BS}}^{t*}$ is the optimal order quantity given the supplier adopts factoring under the baseline trade finance model. The detailed discussion of $\bar{\eta}_{\mathcal{S}}$ is provided in the proof of Proposition B.0.1. We investigate the baseline model, the BDF (as an alternative pre-shipment financing) model, and the invoice trading (as an alternative post-shipment financing) model in order.

A.1.1. The Baseline Trade Finance Model

We first explore the no smart contract case. Applying backward induction approach, we start with the supplier's factoring decision given (w, q, r). If factoring is adopted, the supplier's profit is still $\pi^t_{\mathcal{BS}}(w,q,r) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2}wq - cqe^{rt_c}$, which is irrelevant to fire sale. If factoring is not adopted, instead of going bankrupt, the supplier is able to use fire sale against liquidity shock. The supplier's profit with fire sale is $(1-\rho_r)e^{-\eta_{\mathcal{S}} t'_2}wq - cqe^{rt_c}$. If liquidity shock occurs, we show $(1-\rho_r)e^{-\eta_{\mathcal{S}} t'_2}wq - cqe^{rt_c} \geq 0$ in the proof of Proposition B.0.1, which means using fire sale generates positive profit, and therefore is better than going bankrupt upon liquidity shock. If liquidity shock does not occur, holding accounts receivable on hand generates the supplier profit $(1-\rho_r)(wq - cqe^{rt_c})$, which is higher than the supplier's profit of fire sale $(1-\rho_r)e^{-\eta_{\mathcal{S}} t'_2}wq - cqe^{rt_c}$. As a quick summary, when factoring is not adopted, the supplier uses fire sale if and only if liquidity shock occurs. Therefore, the supplier's profit of not adopting factoring is $\pi^n_{\mathcal{BS}}(w,q,r) = \rho_s[(1-\rho_r)e^{-\eta_{\mathcal{S}} t'_2}wq - cqe^{rt_c}] + (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$. The supplier adopts factoring if and only if $\pi^t_{\mathcal{BS}}(w,q,r) \geq \pi^n_{\mathcal{BS}}(w,q,r)$. Then we study the bank loan rate decision. Similar as the discussion of the main model, we examine two scenarios: (i) If adopting factoring is exogenously given, the bank loan is secured with the corresponding loan rate $r'_t = \eta_{\mathcal{B}}$; (ii) If not adopting factoring is exogenously given, the bank loan rate r'_n is determined by $e^{\eta_{\mathcal{B}} t_c} = [\rho_s + (1-\rho_s)(1-\rho_r)]e^{r'_n t_c}$. We denote:

$$\bar{w}' = \frac{(1-\rho_s)\rho_r c e^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}}{[1-(1-\rho_s+\rho_s e^{-\eta_{\mathcal{S}} t'_2})e^{\eta_{\mathcal{F}} t_2}](1-\rho_r)}.$$

Given (w,q), the bank loan rate and the supplier's factoring decision are summarized as follows.

Lemma 3 *If $w \leq \bar{w}'$, the bank offers the loan rate r'_n and the supplier does not adopt factoring; otherwise, the bank offers the loan rate r'_t and the supplier adopts factoring.*

Next consider the retailer's order quantity decision. The retailer's profit $\Pi_{\mathcal{BS}}(w,q) = (1-\rho_r)[S(q) - qw]$, from which we derive $q^*(w) = \bar{F}^{-1}(w)$. We denote the supplier's effective unit production cost of adopting factoring as $c^t_{\mathcal{BS}} = \frac{ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}}{1-\rho_r}$, and the supplier's effective unit production cost of not adopting factoring as $c^n_{\mathcal{BS}} = \frac{ce^{\eta_{\mathcal{B}} t_c}}{(1-\rho_s+\rho_s e^{-\eta_{\mathcal{S}} t'_2})(1-\rho_r)}$. Plugging $w = \bar{F}(q)$ and r'_t into the supplier's objective function of adopting factoring, we have $\pi^t_{\mathcal{BS}}(q) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2}\bar{F}(q)q - cqe^{\eta_{\mathcal{B}} t_c}$. We derive scenario (i) optimal decisions from $\bar{F}(q^{t*}_{\mathcal{BS}})[1-$

$q^{t*}_{\mathcal{BS}} z(q^{t*}_{\mathcal{BS}})] = c^t_{\mathcal{BS}}$ and $w^{t*}_{\mathcal{BS}} = \bar{F}(q^{t*}_{\mathcal{BS}})$. Plugging $w = \bar{F}(q)$ and r'_n into the supplier's objective function of not adopting factoring, we have $\pi^n_{\mathcal{BS}}(q) = (1 - \rho_s + \rho_s e^{-\eta_{\mathcal{S}} t'_2})(1 - \rho_r)\bar{F}(q)q - cqe^{\eta_{\mathcal{B}} t_c}$. We derive scenario (ii) optimal decisions from $\bar{F}(q^{n*}_{\mathcal{BS}})[1 - q^{n*}_{\mathcal{BS}} z(q^{n*}_{\mathcal{BS}})] = c^n_{\mathcal{BS}}$ and $w^{n*}_{\mathcal{BS}} = \bar{F}(q^{n*}_{\mathcal{BS}})$. Then we solve the supplier's optimal wholesale price. Similar as the main model, we define β'_1 as the solution to $\pi^n_{\mathcal{BS}}(w^{n*}_{\mathcal{BS}}, q^{n*}_{\mathcal{BS}}) = \pi^t_{\mathcal{BS}}(\bar{w}', q^*(\bar{w}'))$ and β'_2 as the solution to $\bar{w}' = w^{t*}_{\mathcal{BS}}$. We summarize the equilibrium in the next proposition.

Proposition B.0.1 *(i) The equilibrium bank loan rate $r^{o*}_{\mathcal{BS}}$, wholesale price $w^{o*}_{\mathcal{BS}}$ and order quantity $q^{o*}_{\mathcal{BS}}$ are summarized as follows:*

$$(r^{o*}_{\mathcal{BS}}, w^{o*}_{\mathcal{BS}}, q^{o*}_{\mathcal{BS}}) = \left\{ \begin{array}{ll} (r'_n, w^{n*}_{\mathcal{BS}}, q^{n*}_{\mathcal{BS}}) & \text{if } \rho_s < \beta'_1 \\ \left(r'_t, \bar{w}', \bar{F}^{-1}(\bar{w}')\right) & \text{if } \beta'_1 \le \rho_s < \beta'_2 \\ \left(r'_t, w^{t*}_{\mathcal{BS}}, q^{t*}_{\mathcal{BS}}\right) & \text{if } \rho_s \ge \beta'_2 \end{array} \right\}.$$

(ii) The supplier adopts factoring if $\rho_s \ge \beta'_1$. The commitment issue exists when $\rho_s < \beta'_2$.

As previously mentioned, fire sale increases the supplier's profit of not adopting factoring from $\pi^n_{\mathcal{B}}(w^{n*}_{\mathcal{B}}, q^{n*}_{\mathcal{B}}, r_n)$ to $\pi^n_{\mathcal{BS}}(w^{n*}_{\mathcal{BS}}, q^{n*}_{\mathcal{BS}}, r'_n)$, while the supplier's profit of adopting factoring is not affected. As a result, $\bar{w}'$ increases (commitment issue is more severe), and $\pi^t_{\mathcal{BS}}(\bar{w}', q^*(\bar{w}'))$ decreases. So $\beta'_1 \ge \beta_1$ and $\beta'_2 \ge \beta_2$, which means that with fire sale, the supplier is less likely to adopt factoring and the commitment issue is more likely to arise. Now we consider the value of smart contract. If adopting factoring is committed, loan rate r'_t is charged and the supplier's profit $\pi^t_{\mathcal{BS}}(w, q) = (1 - \rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{\eta_{\mathcal{B}} t_c}$. If not adopting factoring is committed, loan rate r'_n is charged and the supplier's profit $\pi^n_{\mathcal{BS}}(w, q) = (1 - \rho_s + \rho_s e^{-\eta_{\mathcal{S}} t'_2})(1 - \rho_r)wq - cqe^{\eta_{\mathcal{B}} t_c}$. The supplier always adopts factoring by $\eta_{\mathcal{S}} \ge \underline{\eta}_{\mathcal{S}}$. We have the equilibrium $(r'_t, w^{t*}_{\mathcal{BS}}, q^{t*}_{\mathcal{BS}})$. Smart contract resolves the commitment frictions and strictly increases the supplier and the retailer's profits when $\rho_s < \beta'_2$.

A.1.2. Alternative Pre-Shipment Financing: Buyer Direct Financing

We first analyze the case that BDF is exogenously given. We start with the supplier's factoring decision. At time t_1, the supplier receives accounts receivable with face value $wq-cq$. Note that the supplier has no bank loan since the production cost is covered by the retailer BDF. If factoring is adopted, the supplier's profit $\pi_{\mathcal{ES}}(w,q) = (1-\rho_r)e^{-\eta_{\mathcal{E}} t_2}(wq-cq)$. If factoring is not adopted, the supplier can use fire sale to offset the liquidity shock. Fire sale generates the discounted value $(1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'}(wq-cq)$. It can be verified that the supplier uses fire sale if and only if liquidity shock occurs. Hence, the supplier's profit of not adopting factoring is $(1-\rho_s+\rho_s e^{-\eta_{\mathcal{S}} t_2'})(1-\rho_r)(wq-cq)$. By $\eta_{\mathcal{S}} \geq \underline{\eta}_{\mathcal{S}}$, we can show that the supplier always adopts factoring and the corresponding profit $\pi_{\mathcal{ES}}(w,q) = (1-\rho_r)e^{-\eta_{\mathcal{E}} t_2}(wq-cq)$, which is independent of fire sale. Same as our main model, the retailer's profit $\Pi_{\mathcal{ES}}(w,q) = (1-\rho_r)[S(q)-(wq-cq)-cqe^{r_v t_c}]$, where we derive $q^*(w) = \bar{F}(1-w-c(e^{r_v t_c}-1))$. Plugging $w = 1-\bar{F}(q)-c(e^{r_v t_c}-1)$ into the supplier's profit $\pi_{\mathcal{ES}}(w,q)$, the equilibrium $(w^*_{\mathcal{ES}}, q^*_{\mathcal{ES}})$ can be derived from $\bar{F}(q^*_{\mathcal{ES}})[1-q^*_{\mathcal{ES}} z(q^*_{\mathcal{ES}})] = ce^{r_v t_c}$ and $w^*_{\mathcal{ES}} = \bar{F}(q^*_{\mathcal{ES}})-c(e^{r_v t_c}-1)$. Next, we endogenize the retailer's BDF offering decision. The retailer offers BDF if and only if $\Pi_{\mathcal{ES}}(w^*_{\mathcal{ES}}, q^*_{\mathcal{ES}}) \geq \Pi_{\mathcal{BS}}(w^{o*}_{\mathcal{BS}}, q^{o*}_{\mathcal{BS}})$. Therefore, there exists a threshold $\tilde{r}^o_v$ such that the retailer offers BDF only when $r_v \leq \tilde{r}^o_v$.

Now we turn to the discussion of smart contract. First we characterize the BDF offering decision with smart contract. If BDF is offered, the retailer's profit is $\Pi_{\mathcal{ES}}(w^*_{\mathcal{ES}}, q^*_{\mathcal{ES}})$. If BDF is not offered, the retailer's profit is $\Pi_{\mathcal{BS}}(w^{t*}_{\mathcal{BS}}, q^{t*}_{\mathcal{BS}})$. We can see that when smart contract is available, fire sale is never used, and thus $\Pi_{\mathcal{ES}}(w^*_{\mathcal{ES}}, q^*_{\mathcal{ES}})$ and $\Pi_{\mathcal{BS}}(w^{t*}_{\mathcal{BS}}, q^{t*}_{\mathcal{BS}})$ are unaffected. So we have the same threshold $\tilde{r}^s_v = \bar{r}^s_v$ such that the retailer offers BDF only when $r_v \leq \tilde{r}^s_v$. Smart contract increases the supply chain output with and without fire sale, which means that not offering BDF becomes more profitable for the retailer. To offer the BDF, smaller cost of capital is required when smart contract is available. Our results continue to hold that smart contract switches the retailer to not offering BDF when $\tilde{r}^s_v \leq r_v \leq \tilde{r}^o_v$, and the supplier's profit is hurt.

A.1.3. Alternative Post-Shipment Financing: On-Demand Invoice Trading

Similar as the main model, we focus on $\eta_{\mathcal{I}} \leq \bar{\eta}_{\mathcal{I}}$ to be consistent with the reality that companies prefer invoice trading to factoring. Given (w, q, r), we start with the supplier's post-shipment financing decision at time t'_1. Note that the supplier has three options: selling accounts receivable via invoice trading, selling via fire sale, and holding accounts receivable on hand. If liquidity shock occurs, the supplier's profit with invoice trading is $(1-\rho_r)e^{-\eta_{\mathcal{I}} t'_2} wq - cqe^{rt_c}$; the supplier's profit with fire sale is $(1-\rho_r)e^{-\eta_{\mathcal{S}} t'_2} wq - cqe^{rt_c}$; the supplier's profit with holding the accounts receivable on hand is 0. We can see that invoice trading leads to the highest profit. If liquidity shock does not occur, the supplier's profit with invoice trading is still $(1-\rho_r)e^{-\eta_{\mathcal{I}} t'_2} wq - cqe^{rt_c}$; the supplier's profit with fire sale is $(1-\rho_r)e^{-\eta_{\mathcal{S}} t'_2} wq - cqe^{rt_c}$; the supplier's profit with holding accounts receivable on hand is $(1-\rho_r)(wq - cqe^{rt_c})$. Clearly, holding accounts receivable on hand generates the highest profit. Hence, the supplier sells accounts receivable via invoice trading only when liquidity shock occurs. Fire sale does not alter the game structure. We denote the loan rate as r'_i, the optimal wholesale price as $w^*_{\mathcal{IS}}$, and the optimal order quantity as $q^*_{\mathcal{IS}}$. The results are summarized in the next proposition.

Proposition B.0.2 *Fire sale does not change the equilibrium. We have* $(w^*_{\mathcal{IS}}, q^*_{\mathcal{IS}}, r'_i) = (w^*_{\mathcal{I}}, q^*_{\mathcal{I}}, r_i)$. *The supplier liquidates via invoice trading if and only if liquidity shock occurs. There exists a threshold* $\tilde{\eta}_{\mathcal{IS}}$. *If* $\eta_{\mathcal{I}} \leq \tilde{\eta}_{\mathcal{IS}}$, *invoice trading benefits the supplier. If* $\eta_{\mathcal{I}} > \tilde{\eta}_{\mathcal{IS}}$, *invoice trading can reduce the supplier profit, where smart contract can be used to increase the supply chain value.*

A.2. The Value of Digitalization

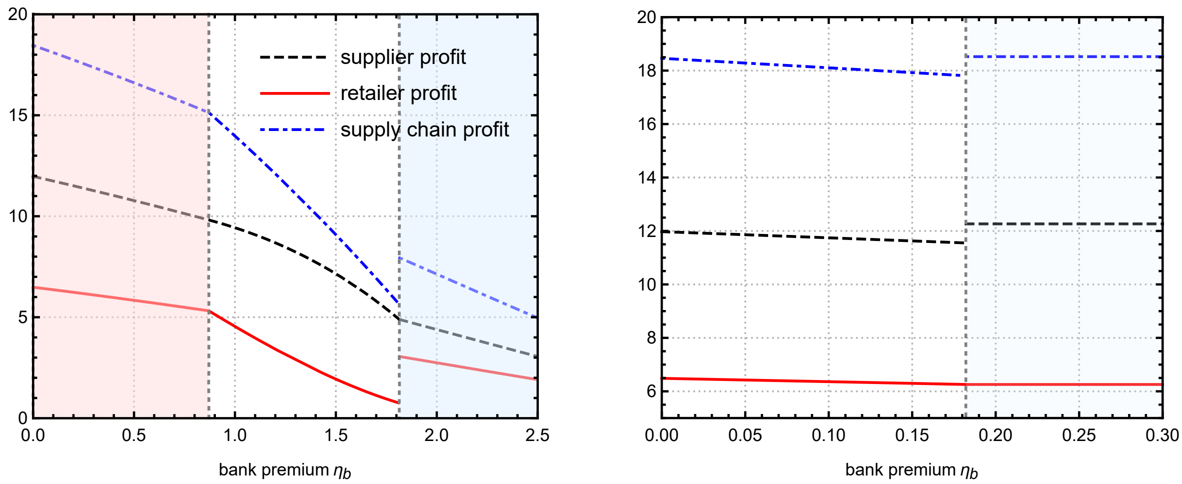
We find digitalization does not necessarily increase the supply chain value. The results are given in the following proposition.

Proposition B.0.3 (i) *Under the baseline trade finance model, digitalization can lead to price distortion and hurt the supply chain.*

(ii) Under BDF, digitalization might discourage the retailer from offering BDF and hurt the supply chain.

(iii) Under invoice trading, digitalization always benefits the supply chain.

Under the baseline trade finance model, the wholesale revenue becomes more weighted than bank loan cost as the bank premium $\eta_{\mathcal{B}}$ decreases. If $\eta_{\mathcal{B}}$ is sufficiently large, wholesale revenue is insignificant compared to the bank loan cost. To reduce the expected loan cost, the supplier has a strong incentive not to adopt factoring, leading to an extremely high safe loan threshold $\bar{w}$. As a result, the supplier gives up factoring adoption. As $\eta_{\mathcal{B}}$ decreases from the sufficiently high level, the wholesale revenue becomes more weighted. To secure the wholesale revenue, the supplier is encouraged to adopt factoring. However, to convince the bank of factoring adoption, the supplier has to severely distort the wholesale price. When digitalization switches the supplier to factoring adoption, the overpricing behavior tremendously reduces the output quantity and hurts the supply chain profit (see Figure B.1).



(a) Baseline model (b) BDF available

Figure B.1.: Profits under baseline trade finance model and BDF available as pre-shipment financing model
($\rho_s = 0.2$, $\rho_r = 0.1$, $c = 0.2$, $r_v = 0.8$, $\eta_{\mathcal{F}} = 0.4$, $\eta_{\mathcal{E}} = 0.1$, $t_1 = 0.1$, $t_2 = 0.2$ and $t_2' = 0.15$)

Under the BDF (as an alternative pre-shipment financing) scheme, the bank loan cost $cqe^{\eta_{\mathcal{B}} t_c}$ decreases as $\eta_{\mathcal{B}}$ decreases, and thus POF becomes more appealing to the retailer due to the reduction of the supply chain financing cost. As a result, digitalization discourages the retailer from offering BDF to the supplier, which reallocates the financing cost to the supplier. As discussed in Section 2.5, reallocating financing cost to the cost-sensitive supplier enormously increases the wholesale price, which in turn reduces the output quantity and hurts the supply chain profit (see Figure B.1). Lastly, under the invoice trading (as an alternative post-shipment financing) scheme, digitalization cannot resolve the commitment trap. Nevertheless, the decreased premium $\eta_{\mathcal{B}}$ indeed alleviates the bank financing burden for the supplier, and thus benefits the supplier as well as the whole supply chain.

A.3. Additional Results

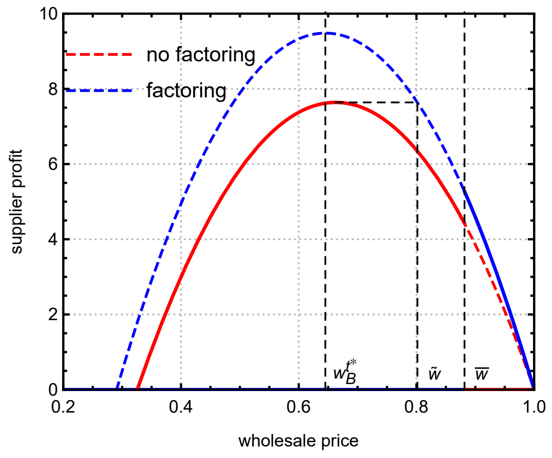
This section provides additional results about the details of several propositions and the plots of relevant numerical examples that are omitted in the main text for expositional clarity.

A.3.1. Discussions of Baseline Trade Finance Model

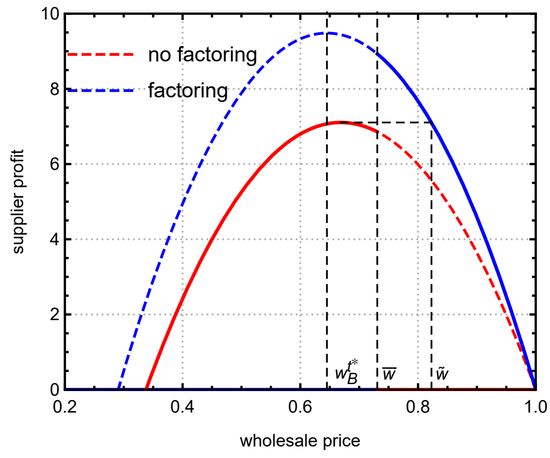
In Figure B.2, we plot the supplier's profits and factoring decisions under the three cases in Proposition 2.4.1. Case 1 represents the low liquidity risk situation $\rho_s < \beta_1$, where the supplier chooses not to adopt factoring in response to an extremely high safe loan threshold (i.e., $\bar{w} > \tilde{w}$). Case 2 represents the medium liquidity risk situation $\beta_1 \leq \rho_s < \beta_2$, where the safe loan threshold is not too high (i.e., $\bar{w} \leq \tilde{w}$) and factoring is preferable to the supplier. However, to convince the bank of factoring adoption, overpricing is needed since $\bar{w} > w_{\mathcal{B}}^{t*}$. Case 3 represents the high liquidity risk situation $\rho_s \geq \beta_2$, where the bank requires a low safe loan threshold due to the high default risk of not adopting factoring. The supplier charges the optimal price $w_{\mathcal{B}}^{t*}$ without the overpricing concerns (i.e., $\bar{w} \leq w_{\mathcal{B}}^{t*}$).

A.3.2. The Value of Factoring under Baseline Trade Finance Model

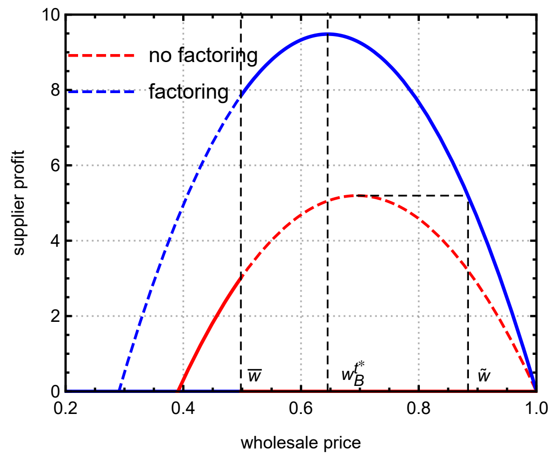
We define $\bar{q} = \bar{F}^{-1}(\bar{w})$ and $\bar{\beta}$ as the unique solution to $\Pi_{\mathcal{B}}(\bar{w}, \bar{q}) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$.



(a) Low liquidity risk ($\rho_s = 0.16$)



(b) Medium liquidity risk ($\rho_s = 0.19$)



(c) High liquidity risk ($\rho_s = 0.3$)

Figure B.2.: The supplier's profits of adopting and not adopting factoring with overpricing issue
($c = 0.2$, $\rho_r = 0.2$, $\eta_{\mathcal{F}} = 0.3$, $\eta_{\mathcal{B}} = 0.3$, $t_1 = 0.1$, and $t_2 = 0.2$)

Proposition B.0.4 *Factoring always benefits the supplier, but reduces the retailer's profit when $\beta_1 < \rho_s < \bar{\beta}$.*

It is clear that factoring gives the supplier another option to handle the liquidity shock. Hence, the existence of factoring always benefits the supplier. However, the retailer might get hurt (see Figure B.3). Recall that when $\beta_1 \leq \rho_s < \beta_2$, the supplier adopts factoring and overprices the wholesale price. When $w_{\mathcal{B}}^{n*} < \bar{w} \leq \tilde{w}$, the existence of factoring increases the wholesale price from $w_{\mathcal{B}}^{n*}$ to $\bar{w}$. The order quantity is reduced and the retailer becomes worse-off.

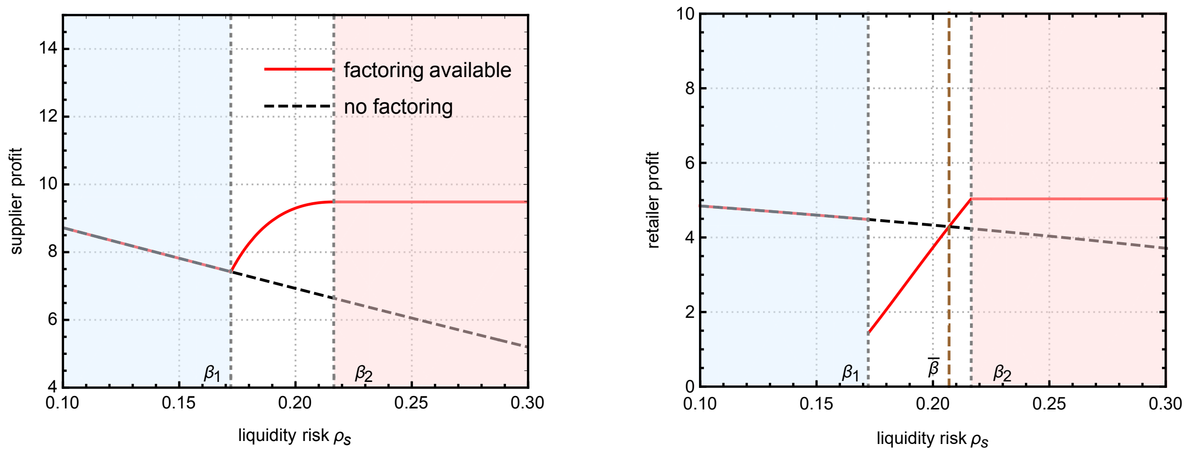


Figure B.3.: The value of factoring for the supplier and the retailer
($c = 0.2$, $\rho_r = 0.2$, $\eta_{\mathcal{F}} = 0.3$, $\eta_{\mathcal{B}} = 0.3$, $t_1 = 0.1$, and $t_2 = 0.2$)

A.3.3. Discussions of BDF

Lemma 4 *Given BDF, the supplier always adopts factoring, and the unique equilibrium* $(q^*_{\mathcal{E}}, w^*_{\mathcal{E}})$ *is derived from:* $\bar{F}(q^*_{\mathcal{E}})[1 - q^*_{\mathcal{E}} z(q^*_{\mathcal{E}})] = ce^{r_v t_c}$, $w^*_{\mathcal{E}} = \bar{F}(q^*_{\mathcal{E}}) - c(e^{r_v t_c} - 1)$.

Proposition B.0.5 *When* $r_v \leq \bar{r}^o_v$, *BDF is adopted, and it benefits both the supplier and the retailer.*

A.3.4. Discussions of Invoice Trading

We can denote $c_{\mathcal{I}} = ce^{\eta_{\mathcal{B}} t_c} / \left[1 - \rho_s + \rho_s e^{-\eta_{\mathcal{I}} t'_2}\right]$ as the supplier's effective unit production cost under invoice trading.

Proposition B.0.6 *With invoice trading, the supplier sells accounts receivable if and only if liquidity shock occurs. The equilibrium order quantity and wholesale price* $(q^*_{\mathcal{I}}, w^*_{\mathcal{I}})$ *are derived from:* $\bar{F}(q^*_{\mathcal{I}})[1 - q^*_{\mathcal{I}} z(q^*_{\mathcal{I}})] = c_{\mathcal{I}}$ *and* $w^*_{\mathcal{I}} = \bar{F}(q^*_{\mathcal{I}})$.

Proposition B.0.7 *Under the invoice trading scheme, (i) if* $\eta_{\mathcal{I}} \leq \tilde{\eta}_{\mathcal{I}}$, *smart contract has no value on the supply chain; (ii) if* $\eta_{\mathcal{I}} > \tilde{\eta}_{\mathcal{I}}$, *smart contract benefits the supply chain. The supplier does not use invoice trading but adopts factoring instead.*

Appendix B: Proofs of Statements

We present the proofs of the Lemmas and Propositions in Appendix B.

Proof of Lemma 2:

Given the wholesale price w, the order quantity q and the bank loan rate r, the supplier makes the factoring decision at time t_1. The supplier's profit of adopting factoring is $\pi_{\mathcal{B}}^t(w,q,r) = (1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq - cqe^{rt_c}$ and the profit of not adopting factoring is $\pi_{\mathcal{B}}^n(w,q,r) = (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$. The supplier adopts factoring if and only if $\pi_{\mathcal{B}}^t(w,q,r) \geq \pi_{\mathcal{B}}^n(w,q,r)$, which can be written as: $(1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq - cqe^{rt_c} \geq (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$. We define

$$\bar{w}(r) = \frac{[1-(1-\rho_s)(1-\rho_r)]ce^{rt_c+\eta_{\mathcal{F}}t_2}}{[1-e^{\eta_{\mathcal{F}}t_2}(1-\rho_s)](1-\rho_r)},$$

in which $\bar{w}(r)$ is a function of the bank loan rate r. The condition of the supplier's factoring decision can be simplified as $w \geq \bar{w}(r)$. It is clear that $\bar{w}(r)$ is monotone increasing in r. Note that we have shown in the main text that $r_t = \eta_{\mathcal{B}}$ and r_n can be derived from $e^{\eta_{\mathcal{B}}t_c} = (1-\rho_s)(1-\rho_r)e^{r_nt_c}$. So we have $r_n = \eta_{\mathcal{B}} - \frac{\ln[(1-\rho_s)(1-\rho_r)]}{t_c}$. We plug the loan rates r_t and r_n in the threshold $\bar{w}(r)$. For simplicity, we define:

$$\begin{cases} \bar{w}_t = \bar{w}(r_t) = \dfrac{[1-(1-\rho_s)(1-\rho_r)]ce^{\eta_{\mathcal{B}}t_c+\eta_{\mathcal{F}}t_2}}{[1-e^{\eta_{\mathcal{F}}t_2}(1-\rho_s)](1-\rho_r)}, \\ \bar{w}_n = \bar{w}(r_n) = \dfrac{[1-(1-\rho_s)(1-\rho_r)]ce^{\eta_{\mathcal{B}}t_c+\eta_{\mathcal{F}}t_2}}{[1-e^{\eta_{\mathcal{F}}t_2}(1-\rho_s)](1-\rho_r)^2(1-\rho_s)}. \end{cases}$$

It is easy to check that $\bar{w}_n \geq \bar{w}_t$ since $r_n \geq r_t$. If the wholesale price satisfies $w \geq \bar{w}_n \geq \bar{w}_t$, the supplier adopts factoring and the bank anticipates the factoring decision and offers the loan rate r_t; if the wholesale price satisfies $w < \bar{w}_t \leq \bar{w}_n$, the supplier does not adopt factoring and the bank anticipates the factoring decision and offers the loan rate r_n, which also forms a equilibrium; if $\bar{w}_t \leq w < \bar{w}_n$, the supplier's factoring decision depends on the received loan rate. Specifically, if the bank offers the loan rate r_t, the supplier adopts factoring since $\bar{w}_t \leq w$. If the bank offers the loan rate r_n, the supplier does not adopt factoring since $w < \bar{w}_n$. Note that the bank is indifferent between offering loan rate r_t or

r_n, since both rates are competitively priced and generate the same profits for the bank. The fully competitive banking system would offer the lowest loan rate that is achievable. When $\bar{w}_t \leq w < \bar{w}_n$, the bank offers r_t and the supplier adopts factoring since $\bar{w}_t \leq w$. Combing the above analysis, we can show that the supplier adopts factoring if and only if $w \geq \bar{w}_t$, where $\bar{w}_t$ equals $\bar{w}$ that is specified in the Lemma.

Proof of Proposition 2.4.1:

We solve the equilibrium by backward induction. (i) Given the wholesale contract (w, q), the bank's loan rate decision at time 0 and the supplier's factoring decision at time t_1 have been characterized in Lemma 2. (ii) Given the wholesale price w, the retailer's profit function is $\Pi_{\mathcal{B}}(w, q) = S(q) - wq$. Note that the retailer profit is only affected by the supplier's wholesale price but not the bank loan rate. We take the derivative with respect to q. Then,

$$\frac{\partial \Pi_{\mathcal{B}}(w, q)}{\partial q} = 1 - F(q) - w,$$

which is strictly decreasing in q. The retailer's optimal order quantity $q^* = \bar{F}^{-1}(w)$ is derived from the first order condition $1 - F(q) - w = 0$. (iii) Lastly, we derive the supplier's wholesale price decision. Note that the supplier determines the wholesale price not only considering the sales revenue, but the financing cost. We have discussed two scenarios in the main text: scenario (i) where the supplier adopts factoring; scenario (ii) where the supplier does not adopt factoring. By equation 2.4, the supplier's profit under scenario (i) is $\pi^t_{\mathcal{B}}(w, q) = (1 - \rho_r)e^{-\eta_{\mathcal{F}} t_2} qw - cqe^{\eta_{\mathcal{B}} t_c}$. Note that the retailer's optimal order decision $q^*(w)$ and the wholesale price w is a bijection. Therefore, the supplier's optimal wholesale price decision is transformed into the retailer's optimal order quantity decision. In other words, the supplier determines the optimal wholesale price by choosing the corresponding order quantity that the retailer would decide with the upfront wholesale price. Plugging the retailer's response function $q^*(w) = \bar{F}^{-1}(w)$ into the supplier's objective function under scenario (i), the supplier's profit function can be written as $\pi^t_{\mathcal{B}}(q) = (1 - \rho_r)e^{-\eta_{\mathcal{F}} t_2} q\bar{F}(q) - cqe^{\eta_{\mathcal{B}} t_c}$. The optimal quantity $q^{t*}_{\mathcal{B}}$ is derived from the first order condition $[1 - F(q^{t*}_{\mathcal{B}})]\,[1 - q^{t*}_{\mathcal{B}} z(q^{t*}_{\mathcal{B}})] = c^t_{\mathcal{B}}$. The cumulative demand function $F(\xi)$ and the generalized failure rate $\xi z(\xi)$ are monotone increasing in ξ by IGFR assumption. So $[1-F(q_{\mathcal{B}}^{t*})][1-q_{\mathcal{B}}^{t*}z(q_{\mathcal{B}}^{t*})]$ is monotone decreasing in $q_{\mathcal{B}}^{t*}$. We have the unique equilibrium $(q_{\mathcal{B}}^{t*}, w_{\mathcal{B}}^{t*})$.

Next we derive the equilibrium of scenario (ii). Equation 2.5 gives the supplier's profit of not adopting factoring $\pi_{\mathcal{B}}^{n}(w,q) = (1-\rho_s)(1-\rho_r)qw - cqe^{\eta_{\mathcal{B}}t_c}$. By similar arguments, we can show the unique equilibrium $(q_{\mathcal{B}}^{n*}, w_{\mathcal{B}}^{n*})$ exists and can be derived from: $[1-F(q_{\mathcal{B}}^{n*})]\,[1-q_{\mathcal{B}}^{n*}z(q_{\mathcal{B}}^{n*})] = c_{\mathcal{B}}^{n}$ and $\bar{F}(q_{\mathcal{B}}^{n*}) = w_{\mathcal{B}}^{n*}$. Since $c_{\mathcal{B}}^{t} \leq c_{\mathcal{B}}^{n}$, it is clear that $q_{\mathcal{B}}^{t*} \geq q_{\mathcal{B}}^{n*}$ and $w_{\mathcal{B}}^{t*} \leq w_{\mathcal{B}}^{n*}$. The supplier's profit $\pi_{\mathcal{B}}^{t}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$ of scenario (i) is higher than $\pi_{\mathcal{B}}^{n}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$ of scenario (ii).

The strategic supplier anticipates that the bank sets the loan rate contingent on the wholesale price w. When $\bar{w} \leq w_{\mathcal{B}}^{t*} \leq \tilde{w}$, the supplier chooses the optimal wholesale price $w_{\mathcal{B}}^{t*}$, gets the loan rate r_t, and adopts factoring. When $w_{\mathcal{B}}^{t*} < \bar{w} \leq \tilde{w}$, by Lemma 2, the supplier has to overprice the wholesale price in order to convince the bank of factoring adoption (i.e., $w \geq \bar{w}$). Note the supplier's profit of adopting factoring $\pi_{\mathcal{B}}^{t}(w,q)$ is quasi-concave in q. Hence, the profit of adopting factoring $\pi_{\mathcal{B}}^{t}(w,q)$ is decreasing as w deviates more to the right of $w_{\mathcal{B}}^{t*}$ ($w > w_{\mathcal{B}}^{t*}$). To convince the bank of factoring adoption, the supplier must overprice the wholesale price to satisfy the condition of Lemma 2. It is clear that the optimal wholesale price is $\bar{w}$, with the least overpricing. With $\bar{w}$, the bank offers loan rate r_t and the supplier adopts factoring. When $w_{\mathcal{B}}^{t*} \leq \tilde{w} < \bar{w}$, by previous discussion, the supplier has to overprice the wholesale price to $\bar{w}$ in order to convince the bank of factoring adoption (scenario (i) case). Alternatively, the supplier can give up convincing the bank and simply charges the optimal wholesale price $w_{\mathcal{B}}^{n*}$ of not adopting factoring (scenario (ii) case). By definition, we know that $\pi_{\mathcal{B}}^{n}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*}) = \pi_{\mathcal{B}}^{t}(\tilde{w}, q^{*}(\tilde{w}))$. So we have $\pi_{\mathcal{B}}^{t}(\bar{w}, q^{*}(\bar{w})) < \pi_{\mathcal{B}}^{t}(\tilde{w}, q^{*}(\tilde{w})) = \pi_{\mathcal{B}}^{n}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$, which means giving up convincing the bank generates higher profit than overpricing with wholesale price $\bar{w}$. The supplier stops overpricing. The bank offers the loan rate r_n and the supplier does not adopt factoring afterwards.

Now we investigate how $\bar{w}$, $\tilde{w}$ and $w_{\mathcal{B}}^{t*}$ change in ρ_s. First, $w_{\mathcal{B}}^{t*}$ is unchanged in ρ_s. The equilibrium $(q_{\mathcal{B}}^{t*}, w_{\mathcal{B}}^{t*})$ is determined by $[1-F(q_{\mathcal{B}}^{t*})]\,[1-q_{\mathcal{B}}^{t*}z(q_{\mathcal{B}}^{t*})] = c_{\mathcal{B}}^{t}$. The effective unit production cost of factoring $c_{\mathcal{B}}^{t} = \frac{ce^{\eta_{\mathcal{B}}t_c + \eta_{\mathcal{F}}t_2}}{1-\rho_r}$ is not affected by ρ_s, and thus the optimal wholesale price $w_{\mathcal{B}}^{t*}$ is unaffected by ρ_s either. Second, $\tilde{w}$ is increasing in ρ_s. Following the similar discussion, $\pi_{\mathcal{B}}^n(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$ is decreasing in ρ_s because the unit production cost $c_{\mathcal{B}}^n = \frac{ce^{\eta_{\mathcal{B}} t_c}}{(1-\rho_s)(1-\rho_r)}$ is increasing in ρ_s. $\pi_{\mathcal{B}}^t(w, q_{\mathcal{B}}^{t*}(w))$ is unaffected. Thus as ρ_s increases, to maintain the equality $\pi_{\mathcal{B}}^n(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*}) = \pi_{\mathcal{B}}^t(\tilde{w}, q^*(\tilde{w}))$, $\tilde{w}$ increases and deviates more from the optimal solution so as to reduce $\pi_{\mathcal{B}}^t(\tilde{w}, q^*(\tilde{w}))$. Third, $\bar{w}$ is decreasing in ρ_s. We take the derivative of $\bar{w}$ in ρ_s and we have:

$$\frac{\partial \bar{w}}{\partial \rho_s} = \frac{(1-\rho_r)ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}[1-e^{\eta_{\mathcal{F}} t_2}(1-\rho_s)](1-\rho_r) - [1-(1-\rho_s)(1-\rho_r)]ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}e^{\eta_{\mathcal{F}} t_2}(1-\rho_r)}{[1-e^{\eta_{\mathcal{F}} t_2}(1-\rho_s)]^2(1-\rho_r)^2},$$

which is simplified as:

$$\frac{\partial \bar{w}}{\partial \rho_s} = \frac{-(1-\rho_r)\rho_r ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}}{[1-e^{\eta_{\mathcal{F}} t_2}(1-\rho_s)]^2(1-\rho_r)^2} < 0.$$

Hence, $\bar{w}$ is decreasing in ρ_s. By the definition that β_1 (resp., β_2) is the unique value of ρ_s such that $\bar{w} = \tilde{w}$ (resp., $\bar{w} = w_{\mathcal{B}}^{t*}$), we derive that $\rho_s < \beta_1$ is equivalent to $\bar{w} > \tilde{w}$; $\beta_1 \leq \rho_s < \beta_2$ is equivalent to $w_{\mathcal{B}}^{t*} < \bar{w} < \tilde{w}$; $\rho_s \geq \beta_2$ is equivalent to $w_{\mathcal{B}}^{t*} \geq \bar{w}$.

Proof of Proposition B.0.4:

Factoring allows the supplier to choose between adopting and not adopting factoring, whereas the supplier has no such option without the existence of factoring. Hence, it is clear that factoring benefits the supplier due to the increased choice set.

We have shown that the retailer's response function $q^*(w) = \bar{F}^{-1}(w)$, which is a bijection between w and q. Note that $q^*(w)$ is decreasing in w. We plug the response function in the retailer objective function. The retailer's profit function can be written as $\Pi_{\mathcal{B}}(q) = S(q) - q\bar{F}(q)$. Then, we take the derivative of the retailer's profit $\Pi_{\mathcal{B}}(q)$ with regard to q and we have:

$$\frac{\partial \Pi_{\mathcal{B}}(q)}{\partial q} = 1 - F(q) - \bar{F}(q) + qf(q) = qf(q) \geq 0.$$

We can see the retailer's profit is increasing in q. Therefore, factoring benefits the retailer if and only if the equilibrium order quantity (resp., wholesale price) is larger (resp., smaller) compared with no factoring case. Hence, if $\bar{w} \leq w_{\mathcal{B}}^{n*}$ (the same as $\bar{q} \geq q_{\mathcal{B}}^{n*}$), factoring benefits the retailer, since the supplier chooses the wholesale price $\max[\bar{w}, w_{\mathcal{B}}^{t*}]$, which is no larger than $w_{\mathcal{B}}^{n*}$. If $\bar{w} > w_{\mathcal{B}}^{n*}$ (the same as $\bar{q} < q_{\mathcal{B}}^{n*}$), factoring hurts the retailer's profit, since the supplier chooses the wholesale price $\bar{w}$ or $w_{\mathcal{B}}^{n*}$, which is no less than or equal to $w_{\mathcal{B}}^{n*}$. $\bar{w} = w_{\mathcal{B}}^{n*}$ is the threshold where retailer is indifferent between supplier adopting factoring or not. By the proof of Proposition 2.4.1, there exists a unique $\rho_s = \bar{\beta}$, which is the solution to $\Pi_{\mathcal{B}}(\bar{w}, \bar{q}) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{n*}, q_{\mathcal{B}}^{n*})$ and it is straightforward to show that $\beta_1 < \bar{\beta} < \beta_2$.

Proof of Proposition 2.4.2:

First we derive the equilibrium with smart contract. Adopting the backward induction method, we start with the supplier's factoring decision and the bank's loan rate decision. Different from the no smart contract case where the bank's loan rate decision is made before the supplier's factoring decision, two decisions are equivalent with smart contract. Specifically, knowing the factoring decision equals knowing the bank loan rate. We have shown in the main text that with such equivalence, $\pi_{\mathcal{B}}^n(w,q) \leq \pi_{\mathcal{B}}^t(w,q)$ for $\forall(w,q)$. So the bank offers rate r_t and the supplier always adopts factoring. Next we derive the retailer's best response function to the supplier's wholesale price with smart contract. Similar as the proof of Proposition 2.4.1, for a given wholesale price w, we have shown the retailer's profit is $\Pi_{\mathcal{B}}(w,q) = S(q) - qw$. Taking the derivative with respect to q and we have:

$$\frac{\partial \Pi_{\mathcal{B}}(w,q)}{\partial q} = 1 - F(q) - w,$$

which is strictly decreasing in q. Hence, given the wholesale price w, we derive the retailer's optimal order quantity $q^*(w) = \bar{F}^{-1}(q)$ from the first order condition. Next, we derive the supplier's optimal wholesale price with smart contract. Note that the retailer's optimal order decision $q^*(w)$ and the wholesale price w is a bijection and by IGFR, the supplier's optimal wholesale price is equivalent to choosing the corresponding order quantity as shown in the proof of Proposition 2.4.1. Plugging the retailer's best response $\bar{F}(q) = w$ in the supplier profit, we rewrite the supplier's profit function as

$\pi^t_{\mathcal{B}}(q) = (1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}q\bar{F}(q) - cqe^{\eta_{\mathcal{B}}t_c}$. The optimal quantity $q^{t*}_{\mathcal{B}}$ is derived from the first order condition:

$$\left[1 - F(q^{t*}_{\mathcal{B}})\right]\left[1 - q^{t*}_{\mathcal{B}}z(q^{t*}_{\mathcal{B}})\right] = c^t_{\mathcal{B}}.$$

By definition of IGFR, the cumulative demand function $F(\xi)$ and the generalized failure rate $\xi z(\xi)$ are monotone increasing in ξ. We have the unique equilibrium $(q^{t*}_{\mathcal{B}}, w^{t*}_{\mathcal{B}})$. As the unit production cost c, the retailer's credit risk ρ_r, or the supplier's liquidity risk ρ_s increases, the effective production cost $c^t_{\mathcal{B}} = \frac{ce^{\eta_{\mathcal{B}}t_c+\eta_{\mathcal{F}}t_2}}{1-\rho_r}$ increases. Hence, the optimal order quantity $q^{t*}_{\mathcal{B}}$ decreases, the optimal wholesale price $w^{t*}_{\mathcal{B}}$ increases and both supplier's profit and retailer's profit decrease.

Next we characterize the value of smart contract. Note that $c^t_{\mathcal{B}} \le c^n_{\mathcal{B}}$, which means smart contract reduces the supply chain cost, leading to higher profits for the supplier as well as the retailer. It is straightforward that smart contract generates strictly higher profits for the supplier and the retailer when $\rho_s < \beta_2$ (equivalent to $\bar{w} > w^{t*}_{\mathcal{B}}$), since the supplier either overprices $\bar{w} > w^{t*}_{\mathcal{B}}$ or switches to no factoring $w^{n*}_{\mathcal{B}} > w^{t*}_{\mathcal{B}}$ without smart contract. Both cases lead to reduced profits for the supplier and the retailer due to commitment frictions.

Proof of Lemma 4:

When BDF is offered, the retailer's profit is $\Pi_{\mathcal{E}}(w,q) = S(q) - (qw - cq) - cqe^{r_v t_c}$. Taking derivative with respect to q, we have:

$$\frac{\partial \Pi_{\mathcal{E}}(w,q)}{\partial q} = 1 - F(q) - w - c(e^{r_v t_c} - 1),$$

which is strictly decreasing in q. Thus, under the BDF scheme, the retailer's response function is $q^*(w) = \bar{F}^{-1}[w + c(e^{r_v t_c} - 1)]$. With BDF, the supplier always adopts factoring. The supplier profit is $\pi_{\mathcal{E}}(w,q) = e^{-\eta_{\mathcal{E}}t_2}(1-\rho_r)(qw - cq)$. Plugging the retailer's response function $q^*(w)$ in the supplier's objective function, we have $\pi_{\mathcal{E}}(q) = e^{-\eta_{\mathcal{E}}t_2}(1-\rho_r)q[1 - F(q)] - e^{-\eta_{\mathcal{E}}t_2}(1-\rho_r)cqe^{r_v t_c}$. Taking the derivative in q and we get:

$$\frac{\partial \pi_{\mathcal{E}}(q)}{\partial q} = e^{-\eta_{\mathcal{E}}t_2}(1-\rho_r)[1 - F(q)][1 - qz(q)] - e^{-\eta_{\mathcal{E}}t_2}(1-\rho_r)ce^{r_v t_c},$$

which is strictly decreasing in q following the similar logic as previously discussed. The optimal quantity $q_{\mathcal{E}}^*$ is derived from the first order condition $[1 - F(q_{\mathcal{E}}^*)][1 - q_{\mathcal{E}}^* z(q_{\mathcal{E}}^*)] = ce^{r_v t_c}$. Hence, we solve the unique equilibrium $(q_{\mathcal{E}}^*, w_{\mathcal{E}}^*)$. By our assumption $\min[(1 - \rho_s)(1 - \rho_r)e^{-\eta_{\mathcal{B}} t_c}, e^{-r_v t_c}] \geq c$, we have $q_{\mathcal{E}}^* \geq q_{\mathcal{B}}^{t*} \geq q_{\mathcal{B}}^{o*}$, and the wholesale price $w_{\mathcal{E}}^* = \bar{F}(q_{\mathcal{E}}^*) - c[e^{r_v t_c} - 1] \leq \bar{F}(q_{\mathcal{E}}^*) \leq \bar{F}(q_{\mathcal{B}}^{o*}) = w_{\mathcal{B}}^{o*}$. Thus, when BDF is adopted, the supplier charges a lower wholesale price and the retailer orders more products compared with the POF scheme.

Proof of Proposition 2.5.1:

Recall that we have $\Pi_{\mathcal{B}}(w, q) = S(q) - qw$ and $\Pi_{\mathcal{E}}(w, q) = S(q) - (qw - cq) - cqe^{r_v t_c}$. (i) When BDF is adopted, from retailer's objective function $\Pi_{\mathcal{E}}(w, q)$, we derive retailer's response curve $\bar{F}(q) = w + c(e^{r_v t_c} - 1)$. Plugging the response function in the retailer's objective function, we rewrite it as $\Pi_{\mathcal{E}}(q) = S(q) - q[\bar{F}(q) - c(e^{r_v t_c} - 1)] + cq - cqe^{r_v t_c} = S(q) - q\bar{F}(q)$. (ii) When BDF is not adopted, from retailer's objective function $\Pi_{\mathcal{B}}(q)$, we derive retailer's response curve $\bar{F}(q) = w$. We plug the response function into the retailer's objective function and simplify the objective function $\Pi_{\mathcal{B}}(q) = S(q) - q\bar{F}(q)$. It is clear that given the order quantity q, the retailer's profits are the same with/without BDF. So retailer's BDF offering decision only depends on the equilibrium output quantity. Proposition 2.4.1 indicates that $(q_{\mathcal{B}}^{o*}, w_{\mathcal{B}}^{o*})$ is not affected by r_v, and therefore, $\Pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$ does not change with r_v. Lemma 4 shows that $q_{\mathcal{E}}^*$ is decreasing in r_v. Thus $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ is decreasing in r_v as well. As r_v increases from 0 to $+\infty$, the retailer's profit under BDF $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ decreases from higher than that under POF $\Pi_{\mathcal{B}}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$ to lower than it (drop down to $-\infty$). The retailer chooses whether to offer BDF or not by comparing which one generates higher profit (more order quantity). Therefore, there exists a threshold $\bar{r}_v^o$ such that the retailer offers BDF if and only if $r_v \leq \bar{r}_v^o$ under the situation where the smart contract is adopted.

$\bar{r}_v^o$ is the solution to $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$. $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ is unaffected by ρ_s. The sensitivity analysis of $\Pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$ is studied in Proposition 2.4.1. So the sensitivity of $\bar{r}_v^o$ in ρ_s can be checked. Similar analysis applies to ρ_r and we skip that part.

Proof of Proposition B.0.5:

From the proof of Proposition 2.5.1, we have characterized retailer's BDF offering decisions. It is clear that when BDF is offered, we have the equilibrium results $(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ and the corresponding profits. When BDF is not offered, we have the equilibrium results $(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$, and the corresponding profits. When $r_v \leq \bar{r}_v^o$, retailer offers BDF, which gives her a higher profit. So BDF increases the retailer profit by $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) - \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$. As for supplier's profit, if BDF is offered, the supplier always accepts it and his profit is enhanced by $\pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) - \pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*})$.

Proof of Proposition 2.5.2:

Similar as the proof of Proposition 2.5.1, we analyze the situation where the smart contract is adopted. By the similar arguments, $\Pi_{\mathcal{B}}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$ does not change with r_v, and $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ is decreasing in r_v. Hence, there exists a threshold $\bar{r}_v^s$ such that the retailer offers BDF if and only if $r_v \leq \bar{r}_v^s$ under the situation where the smart contract is adopted. Note that $\Pi_{\mathcal{B}}(w_{\mathcal{B}}^{o*}, q_{\mathcal{B}}^{o*}) \leq \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$ and $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ is decreasing in r_v, thus we have $\bar{r}_v^s \leq \bar{r}_v^o$. If $r_v < \bar{r}_v^s$, the retailer always offers BDF with/without the smart contract. Hence, both the supplier profit and retailer profit are not affected by smart contract. If $r_v \geq \bar{r}_v^o$, the retailer never offers BDF with/without the smart contract. The analysis goes back to the baseline trade fiance model. Therefore, smart contract benefits both the supplier and retailer. If $\bar{r}_v^s \leq r_v < \bar{r}_v^o$, the retailer only offers BDF if the smart contract is not adopted. When BDF is not offered, the equilibrium output quantity is higher, which benefits the retailer. The financing cost is transferred to the supplier, which hurts the supplier. Hence, smart contract increases the retailer profit, but reduces the supplier profit.

Based on the proof of Proposition 2.5.1, we have shown that under the BDF scheme, the supplier's profit $\pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ and the retailer's profit $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*)$ are decreasing in r_v, whereas under the POF scheme, the supplier's profit and retailer's profit are not affected by r_v. When $r_v = 0$, BDF generates higher supply chain profit compared with POF, whereas when $r_v = -\infty$, BDF generates negative supply chain profit. Hence, there exists a threshold $\bar{r}_v^k$, where $\Pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) + \pi_{\mathcal{E}}(w_{\mathcal{E}}^*, q_{\mathcal{E}}^*) = \Pi_{\mathcal{B}}(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*}) + \pi_{\mathcal{B}}^t(w_{\mathcal{B}}^{t*}, q_{\mathcal{B}}^{t*})$ if $r_v = \bar{v}_v^k$. If $r < \bar{r}_v^k$, the BDF increases the supply chain profit, but retailer does not offer it if the smart contract is adopted. Therefore, smart contract reduces the supply chain profit.

Proof of Proposition B.0.6:

Let us start with the discussion of the assumption $\eta_{\mathcal{I}} \leq \bar{\eta}_{\mathcal{I}}$, which guarantees: (i) Given $\forall(w,q,r)$, the supplier adopts invoice trading rather than factoring; (ii) Invoice trading is sufficient to repay the bank loan. First, we characterize part (i). Given $\forall(w,q,r)$, the supplier's profit of factoring is $\pi^t_{\mathcal{B}}(w,q,r) = (1-\rho_r)e^{-\eta_{\mathcal{F}}t_2}wq - cqe^{\eta_{\mathcal{B}}t_c}$, and the supplier's profit of invoice trading is $\pi_{\mathcal{I}}(w,q,r) = [1-\rho_s+\rho_s e^{-\eta_{\mathcal{I}}t'_2}](1-\rho_r)wq - [(1-\rho_s)(1-\rho_r)+\rho_s]cqe^{rt_c}$. To guarantee the selection of invoice trading, the following condition must be satisfied $\pi_{\mathcal{I}}(w,q,r) \geq \pi^n_{\mathcal{B}}(w,q,r)$, for $\forall(w,q,r)$, which can be simplified and written as $(1-\rho_s)\rho_r ce^{\eta_{\mathcal{B}}t_c} \geq [e^{-\eta_{\mathcal{F}}t_2} - (1-\rho_s) - \rho_s e^{-\eta_{\mathcal{I}}t'_2}](1-\rho_r)$. Hence we derive the part (i) condition:

$$e^{\eta_{\mathcal{I}}t'_2} \leq \frac{\rho_s(1-\rho_r)}{(1-\rho_r)e^{-\eta_{\mathcal{F}}t_2} - (1-\rho_s)(1-\rho_r) - (1-\rho_s)\rho_r ce^{\eta_{\mathcal{B}}t_c}}.$$

Next, we characterize part (ii). To guarantee the repayment of the bank loan after invoice trading, the following condition must be satisfied $(1-\rho_r)e^{-\eta_{\mathcal{I}}t'_2}w^*_{\mathcal{I}}q^*_{\mathcal{I}} - cq^*_{\mathcal{I}}e^{rt_c} \geq 0$. In other words, selling accounts receivable via invoice trading collects sufficient cash to pay off the bank loan at the equilibrium $(w^*_{\mathcal{I}}, q^*_{\mathcal{I}})$. To simplify the exposition, we define q^*_1, q^*_2, function H, and function K as follows. Note that the retailer's response function $\bar{F}(q) = w$ is the same throughout the following proof, and thus we do not mention it anymore. Given $\forall\ r$ that is fixed, the supplier's profit of invoice trading is $\pi_{\mathcal{I}}(w,q) = [1-\rho_s+\rho_s e^{-\eta_{\mathcal{I}}t'_2}](1-\rho_r)wq - [(1-\rho_s)(1-\rho_r)+\rho_s]cqe^{rt_c}$, in which the optimal order quantity q^*_1 is derived. The supplier's profit consists of two cases. That is, holding accounts receivable on hand when liquidity shock does not occur and sells accounts receivable when liquidity shock occurs. Given r, the supplier's profit of holding accounts receivable on hand is $H = (1-\rho_r)(wq - cqe^{rt_c})$, in which the optimal order quantity q^*_2 is derived. We define $K = (1-\rho_r)e^{-\eta_{\mathcal{I}}t'_2}wq - cqe^{rt_c}$, which is the net revenue of selling accounts receivable via invoice trading. Note that $q^*_2 \geq q^*_1$, since the effective unit production cost of holding accounts receivable on hand ce^{rt_c} is less than that under $\pi_{\mathcal{I}}(w,q)$, which is

$$\frac{[(1-\rho_s)(1-\rho_r)+\rho_s]ce^{rt_c}}{(1-\rho_r)(1-\rho_s+\rho_s e^{-\eta_{\mathcal{I}}t'_2})}.$$

If $K \geq 0$ at q_2^*, we can guarantee that selling accounts receivable via invoice trading is sufficient to repay the bank loan at the equilibrium. As q increases from 0 to q_1^*, both H and K are positive. As q increases to q_2^*, H and K are still positive. As q increases across q_2^* to the quantity where $K = 0$, invoice trading becomes useless and the supplier's profit becomes $(1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$. Note that the optimal quantity q_2^* is already passed, and thus the profit is decreasing in q. So q_1^* is the optimal order quantity under invoice trading, where $K \geq 0$ is also satisfied. q_2^* is derived from the first order condition $1-F(q_2^*)-q_2^* f(q_2^*) = ce^{rt_c}$. $K \geq 0$ at q_2^* can be written as $(1-\rho_r)e^{-\eta_{\mathcal{I}} t_2'}[1-F(q_2^*)]-ce^{rt_c} \geq 0$. Plug $1 - F(q_2^*) - q_2^* f(q_2^*) = ce^{rt_c}$ into $(1 - \rho_r)e^{-\eta_{\mathcal{I}} t_2'}[1 - F(q_2^*)] - ce^{rt_c} \geq 0$, we have $q_2^* f(q_2^*) > [1-F(q_2^*)][1-(1-\rho_r)e^{-\eta_{\mathcal{I}} t_2'}]$. So we can derive the part (ii) of the assumption. Now we derive the equilibrium. Adopting backward induction, we start with supplier's post-shipment financing decision. The supplier adopts invoice trading, but the trading decision can be made at two time points: (a) making trading decision at the beginning of the payment period t_1; (b) making trading decision at time t_1'. It is clear that scenario (a) is dominated by scenario (b), because the delay of trading decision does not affect the supplier's hedging against liquidity shock, while saves the premium discounting cost from time t_1 to time t_1'. In the main text we have also shown that the supplier sells accounts receivable if and only if liquidity shock occurs. Given the wholesale price w, order quantity q and the loan rate r, the supplier profit is $\pi_{\mathcal{I}}(w, q, r) = [1 - \rho_s + \rho_s e^{-\eta_{\mathcal{I}} t_2'}](1 - \rho_r)wq - [(1 - \rho_s)(1 - \rho_r) + \rho_s]cqe^{rt_c}$. By previous discussion, factoring is never adopted. Next consider the bank loan rate. With the supplier's post-shipment financing decision, the bank loan is secured if liquidity shock occurs, since selling accounts receivable guarantees the bank loan repayment. If liquidity shock does not occur, the bank loan suffers from the retailer default risk. Hence, the bank loan rate is decided by the equation: $e^{\eta_{\mathcal{B}} t_c} = [\rho_s + (1-\rho_s)(1-\rho_r)]e^{r_i t_c}$, in which $\rho_s + (1-\rho_s)(1-\rho_r)$ can be decomposed as two cases. With probability ρ_s, liquidity shock occurs. The supplier sells accounts receivable and the bank loan can be paid back. With probability $1 - \rho_s$, liquidity shock does not occur. The supplier does not sell accounts receivable and the bank loan can only be repaid if there's no credit default $1 - \rho_r$ at the end of the payment period. The retailer's response function is still $\bar{F}(q) = w$ as discussed. Hence, plugging the bank rate $r = r_i$ and the retailer's response function $\bar{F}(q) = w$ in the supplier's profit $\pi_{\mathcal{I}}(w, q, r)$, the supplier's profit can be written as $\pi_{\mathcal{I}}(q) = \left[1 - \rho_s + \rho_s e^{-\eta_{\mathcal{I}} t_2'}\right](1 - \rho_r)\bar{F}(q)q - cqe^{\eta_{\mathcal{B}} t_c}$. Taking the derivative of the supplier profit $\pi_{\mathcal{I}}(q)$ in q, we have:

$$\frac{\partial \pi_{\mathcal{I}}(q)}{\partial q} = \left[1 - \rho_s + \rho_s e^{-\eta_{\mathcal{I}} t_2'}\right](1 - \rho_r)[1 - F(q)][1 - qz(q)] - ce^{\eta_{\mathcal{B}} t_c},$$

which is strictly decreasing in q following the IGFR. Note that we define

$$c_{\mathcal{I}} = ce^{\eta_{\mathcal{B}} t_c} / \left[1 - \rho_s + \rho_s e^{-\eta_{\mathcal{I}} t_2'}\right]$$

. Hence there's a unique optimal order quantity $q_{\mathcal{I}}^*$ from the equation $\bar{F}(q_{\mathcal{I}}^*)[1 - q_{\mathcal{I}}^* z(q_{\mathcal{I}}^*)] = c_{\mathcal{I}}$, and a unique optimal wholesale price $w_{\mathcal{I}}^*$ from the equation $\bar{F}(q_{\mathcal{I}}^*) = w_{\mathcal{I}}^*$.

Proof of Proposition 2.6.1:

The supplier's profit of factoring $\pi_{\mathcal{B}}^t(q) = (1 - \rho_r)e^{-\eta_{\mathcal{F}} t_2}\bar{F}(q)q - cqe^{\eta_{\mathcal{B}} t_c}$. If $\eta_{\mathcal{I}} \to +\infty$, clearly $\pi_{\mathcal{I}}(q) < \pi_{\mathcal{B}}^t(q)$ for $\forall q$. Invoice trading always hurts the supplier. If $\eta_{\mathcal{I}} \to 0$, $\pi_{\mathcal{I}}(q) \to (1 - \rho_r)\bar{F}(q)q - cqe^{\eta_{\mathcal{B}} t_c}$. Invoice trading has the same bank loan cost $cqe^{\eta_{\mathcal{B}} t_c}$ as factoring, but generates higher wholesale revenue $(1 - \rho_r)\bar{F}(q)q$ than factoring $(1 - \rho_r)e^{-\eta_{\mathcal{F}} t_2}\bar{F}(q)q$ for $\forall q$. So $\pi_{\mathcal{I}}(q) > \pi_{\mathcal{B}}^t(q)$, under which invoice trading always benefits the supplier. So there exists a threshold $\tilde{\eta}_{\mathcal{I}}$ such that when $\eta_{\mathcal{I}} \leq \tilde{\eta}_{\mathcal{I}}$, the invoice trading always benefits the supplier. When $\eta_{\mathcal{I}} > \tilde{\eta}_{\mathcal{I}}$, the bank loan costs $cqe^{\eta_{\mathcal{B}} t_c}$ are the same for invoice trading and factoring, but the wholesale revenue of invoice trading $\left[1 - \rho_s + \rho_s e^{-\eta_{\mathcal{I}} t_2'}\right](1 - \rho_r)\bar{F}(q)q$ is also smaller than that of factoring $(1 - \rho_r)e^{-\eta_{\mathcal{F}} t_2}\bar{F}(q)q$. If c is larger, the bank loan cost effect dominates the wholesale revenue, so invoice trading benefits the supplier. If c is smaller, the wholesale revenue dominates the bank loan cost, so invoice trading hurts the supplier. So there exists a threshold $\tilde{c}$, such that invoice trading hurts the supplier profit if and only if $c < \tilde{c}$.

Proof of Proposition B.0.7:

As we have discussed in the proof of Proposition 2.6.1, invoice trading benefits the supply chain if and only if $\eta_{\mathcal{I}} \leq \tilde{\eta}_{\mathcal{I}}$. When $\eta_{\mathcal{I}} \leq \tilde{\eta}_{\mathcal{I}}$, the supplier always chooses invoice trading and there exists no commitment frictions. So smart contract adds no extra value to the supply chain. When $\eta_{\mathcal{I}} > \tilde{\eta}_{\mathcal{I}}$, the supplier adopts invoice trading due to the commitment trap, while factoring could generate higher supplier profit. So smart contract resolves the commitment issue and adds value to the supply chain.

Proof of Lemma 3:

Given the wholesale price w, the order quantity q and the bank loan rate r, the supplier makes the factoring decision at time t_1. The supplier's profit of adopting factoring is $\pi^t_{\mathcal{BS}}(w,q,r) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{rt_c}$ and the profit of not adopting factoring is $\pi^n_{\mathcal{BS}}(w,q,r) = \rho_s[(1-\rho_r)e^{-\eta_S t_2'} wq - cqe^{rt_c}] + (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$. The supplier adopts factoring if and only if $\pi^t_{\mathcal{BS}}(w,q,r) \geq \pi^n_{\mathcal{BS}}(w,q,r)$, which can be written as: $(1-\rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{rt_c} \geq [1-\rho_s+\rho_s e^{-\eta_{\mathcal{I}} t_2'}](1-\rho_r)wq - [(1-\rho_s)(1-\rho_r)+\rho_s]cqe^{rt_c}$. We define

$$\bar{w}(r) = \frac{(1-\rho_s)\rho_r c e^{rt_c + \eta_{\mathcal{F}} t_2}}{[1-(1-\rho_s+\rho_s e^{-\eta_S t_2'})e^{\eta_{\mathcal{F}} t_2}](1-\rho_r)},$$

in which $\bar{w}(r)$ is a function of the bank loan rate r. The condition of the supplier's factoring decision can be simplified as $w \geq \bar{w}(r)$. It is clear that $\bar{w}(r)$ is monotone increasing in r. Note that we have shown that $r_t' = \eta_{\mathcal{B}}$ and r_n' can be derived from $e^{\eta_{\mathcal{B}} t_c} = [(1-\rho_s)(1-\rho_r)+\rho_s]e^{r_n' t_c}$. So we have $r_n' = \eta_{\mathcal{B}} - \frac{\ln[(1-\rho_s)(1-\rho_r)+\rho_s]}{t_c}$. We plug the loan rates r_t' and r_n' in the threshold $\bar{w}(r)$. For simplicity, we define:

$$\begin{cases} \bar{w}_t = \bar{w}(r_t') = \dfrac{[1-(1-\rho_s)(1-\rho_r)]ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}}{[1-e^{\eta_{\mathcal{F}} t_2}(1-\rho_s)](1-\rho_r)}, \\ \bar{w}_n = \bar{w}(r_n') = \dfrac{[1-(1-\rho_s)(1-\rho_r)]ce^{\eta_{\mathcal{B}} t_c + \eta_{\mathcal{F}} t_2}}{[1-e^{\eta_{\mathcal{F}} t_2}(1-\rho_s)](1-\rho_r)[(1-\rho_s)(1-\rho_r)+\rho_s]}. \end{cases}$$

It is easy to check that $\bar{w}_n \geq \bar{w}_t$ since $r_n' \geq r_t'$. If the wholesale price satisfies $w \geq \bar{w}_n \geq \bar{w}_t$, the supplier adopts factoring and the bank anticipates the factoring decision and offers the loan rate r_t'; if the wholesale price satisfies $w < \bar{w}_t \leq \bar{w}_n$, the supplier does not adopt factoring and the bank anticipates the factoring decision and offers the loan rate r_n', which also forms a equilibrium; if $\bar{w}_t \leq w < \bar{w}_n$, the supplier's factoring decision depends on the received loan rate. Specifically, if the bank offers the loan rate r_t', the supplier adopts factoring since $\bar{w}_t \leq w$. If the bank offers the loan rate r_n', the supplier does not adopt factoring since $w < \bar{w}_n$. Note that the bank is indifferent between offering loan rate r_t' or r_n', since both rates are competitively priced and generate the same profits for the bank. The fully competitive banking system would offer the lowest loan rate that is achievable. When $\bar{w}_t \leq w < \bar{w}_n$, the bank offers r_t' and the supplier adopts factoring since $\bar{w}_t \leq w$. Combing the above analysis, we can show that the supplier adopts factoring if and only if $w \geq \bar{w}_t$, where $\bar{w}_t$ equals $\bar{w}'$ that is specified in the Lemma.

Proof of Proposition B.0.1:

Let us start with the discussion of the assumption

$$\eta_{\mathcal{S}} \geq \underline{\eta}_{\mathcal{S}}$$

, where $\underline{\eta}_{\mathcal{S}} = \min\left\{\ln\left[\frac{\rho_s e^{\eta_{\mathcal{F}} t_2}}{1-(1-\rho_s)e^{\eta_{\mathcal{F}} t_2}}\right]/t_2', \eta_{\mathcal{I}}\right\}$. We only discuss the first part $\eta_{\mathcal{S}} \geq \ln\left[\frac{\rho_s e^{\eta_{\mathcal{F}} t_2}}{1-(1-\rho_s)e^{\eta_{\mathcal{F}} t_2}}\right]/t_2'$, which is relevant under the baseline trade finance model. From subsection A.1.1, we have the supplier's profit of factoring $\pi_{\mathcal{BS}}^t(q) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{\eta_{\mathcal{B}} t_c}$ and the supplier's profit of not adopting factoring $\pi_{\mathcal{BS}}^n(q) = (1-\rho_s+\rho_s e^{-\eta_{\mathcal{S}} t_2'})(1-\rho_r)wq - cqe^{\eta_{\mathcal{B}} t_c}$. Similar as the discussion of the main model, we assume $e^{-\eta_{\mathcal{F}} t_2} \geq (1-\rho_s+\rho_s e^{-\eta_{\mathcal{S}} t_2'})$ to avoid the trivial case where factoring is never used. From the above inequality, we derive $\eta_{\mathcal{S}} \geq \ln\left[\frac{\rho_s e^{\eta_{\mathcal{F}} t_2}}{1-(1-\rho_s)e^{\eta_{\mathcal{F}} t_2}}\right]/t_2'$. Next consider the assumption $\eta_{\mathcal{S}} \leq \ln\left[\frac{1-\rho_r}{1-q_{\mathcal{S}}^{t*} z(q_{\mathcal{S}}^{t*})}\right]/t_2'$, which guarantees that fire sale raises sufficient fund to repay the bank loan. To simplify the exposition, we define q_a^*, q_b^*, function H', and function K' as follows. Note that the retailer's response function $\bar{F}(q) = w$ is the same throughout the following proof, and thus we do not mention it anymore. Given $\forall\ r$ that is fixed, the supplier's profit of not adopting factoring $\pi_{\mathcal{BS}}^n(w,q,r) = \rho_s[(1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'} wq - cqe^{rt_c}] + (1-\rho_s)(1-\rho_r)(wq - cqe^{rt_c})$, in which the optimal order quantity q_a^* is derived. The supplier's profit consists of two cases. That is, holding accounts receivable on hand when liquidity shock does not occur and fire sale when liquidity shock occurs. Given r, the supplier's profit of holding accounts receivable on hand is $H' = (1-\rho_r)(wq - cqe^{rt_c})$, in which the optimal order quantity q_b^* is derived. We define $K' = (1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'} wq - cqe^{rt_c}$, which is the net revenue of fire sale. Note that $q_b^* \geq q_a^*$, since the effective unit production cost of holding accounts receivable on hand ce^{rt_c} is less than that under $\pi_{\mathcal{BS}}^n(w,q)$, which is

$$\frac{[(1-\rho_s)(1-\rho_r)+\rho_s]ce^{rt_c}}{(1-\rho_r)(1-\rho_s+\rho_s e^{-\eta_{\mathcal{S}} t_2'})}.$$

If $K' \geq 0$ at q_b^*, we can guarantee that fire sale is sufficient to repay the bank loan at the equilibrium. As q increases from 0 to q_a^*, both H' and K' are positive. As q increases to q_b^*, H' and K' are still positive. As q increases across q_b^* to the quantity where $K'=0$, fire sale becomes useless and the supplier's profit becomes $(1-\rho_s)(1-\rho_r)(wq-cqe^{rt_c})$. Note that the optimal quantity q_b^* is already passed, and thus the profit is decreasing in q. So q_a^* is the optimal order quantity, where $K' \geq 0$ is also satisfied. q_b^* is derived from the first order condition $1-F(q_b^*)-q_b^* f(q_b^*) = ce^{rt_c}$. $K' \geq 0$ at q_b^* can be written as $(1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'}[1-F(q_b^*)] - ce^{rt_c} \geq 0$. Plug $1-F(q_b^*)-q_b^* f(q_b^*) = ce^{rt_c}$ into $(1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'}[1-F(q_b^*)] - ce^{rt_c} \geq 0$, we have $q_b^* f(q_b^*) > [1-F(q_b^*)][1-(1-\rho_r)e^{-\eta_{\mathcal{S}} t_2'}]$. So we have $\eta_{\mathcal{S}} \leq \ln\left[\frac{1-\rho_r}{1-q_{\mathcal{S}}^{t*} z(q_{\mathcal{S}}^{t*})}\right]/t_2'$.

Now we derive the optimal decisions with fire sale option, similar as the proof of Proposition 2.4.1. We start with supplier's factoring decision. If the supplier adopts factoring, the supplier profit is still $\pi_{\mathcal{BS}}^t(w,q,r) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{rt_c}$. If the supplier does not adopt factoring, the profit becomes

$$\pi_{\mathcal{BS}}^n(w,q,r) = (1-\rho_s)(1-\rho_r)(wq-cqe^{rt_c}) + \rho_s\left(\frac{(1-\rho_r)wq}{e^{\eta_{\mathcal{S}} t_2'}} - cqe^{rt_c}\right).$$

It differs from the Proposition 2.4.1 in that the supplier is able to adopt fire sale of accounts receivable at a premium $\eta_{\mathcal{S}}$ when liquidity shock occurs. Similar as the two scenarios we discussed in Proposition 2.4.1, if the supplier commits to factoring, the bank loan is secured and the loan rate r_t' is decided by $cqe^{\eta_{\mathcal{B}} t_c} = cqe^{rt_c}$. If the supplier commits to not adopting factoring, the bank loan rate r_n' is decided by $cqe^{\eta_{\mathcal{B}} t_c} = (1-\rho_s)(1-\rho_r)cqe^{r_n' t_c} + \rho_s cqe^{r_n' t_c}$. Notice that when liquidity shock occurs with probability ρ_s, fire sale guarantees the bank loan repayment. Plugging r_t' in the supplier profit of adopting factoring, $\pi_{\mathcal{BS}}^t(w,q) = (1-\rho_r)e^{-\eta_{\mathcal{F}} t_2} wq - cqe^{\eta_{\mathcal{B}} t_c}$. Note that the retailer's response function is still $\bar{F}(q) = w$. We derive the corresponding optimal decisions $(w^{t*}_{\mathcal{BS}}, q^{t*}_{\mathcal{BS}})$. Plugging r'_n and $\bar{F}(q) = w$ in the supplier profit of not adopting factoring, $\pi^n_{\mathcal{BS}}(w,q) = \left[1-\rho_s+\frac{\rho_s}{e^{\eta_{\mathcal{S}}t'_2}}\right](1-\rho_r)wq - cqe^{\eta_{\mathcal{B}}t_c}$. We derive the optimal decisions $(w^{n*}_{\mathcal{BS}}, q^{n*}_{\mathcal{BS}})$.

Given the wholesale contract (w,q) and the loan rate r, the supplier adopts factoring if and only if $\pi^t_{\mathcal{BS}}(w,q,r) \geq \pi^n_{\mathcal{BS}}(w,q,r)$. Following the same logic as Proposition 2.4.1, we rewrite it as $w \geq \bar{w}'$, where

$$\bar{w}' = \frac{(1-\rho_s)\rho_r c e^{\eta_{\mathcal{B}}t_c+\eta_{\mathcal{F}}t_2}}{\left[1-\left((1-\rho_s)+\frac{\rho_s}{e^{\eta_{\mathcal{S}}t'_2}}\right)e^{\eta_{\mathcal{F}}t_2}\right](1-\rho_r)}.$$

If $w \geq \bar{w}'$, the supplier commits to adopting factoring and the bank offers loan rate r'_t. If $w < \bar{w}'$, the supplier commits to not adopting factoring and the bank offers loan rate r'_n. The supplier anticipates the bank loan decision and makes the wholesale price decision. We skip this part since the proof is the same as Proposition 2.4.1.

Proof of Proposition B.0.2:

Let us start with the discussion of the assumption $\eta_{\mathcal{S}} \geq \underline{\eta}_{\mathcal{S}}$, where

$$\underline{\eta}_{\mathcal{S}} = \min\left\{\ln\left[\frac{\rho_s e^{\eta_{\mathcal{F}}t_2}}{1-(1-\rho_s)e^{\eta_{\mathcal{F}}t_2}}\right]/t'_2, \eta_{\mathcal{I}}\right\}$$

. We discuss the second part $\eta_{\mathcal{S}} \geq \eta_{\mathcal{I}}$, which is relevant under the invoice trading scheme. This condition guarantees that the supplier sells accounts receivable via invoice trading instead of fire sale.

Fire sale does not affect the supplier's post-shipment financing decision, and thus the equilibrium is the same as Proposition B.0.6. However, $\tilde{\eta}_{\mathcal{IS}} \leq \tilde{\eta}_{\mathcal{I}}$, since fire sale saves the supplier's cost of bankruptcy, leading to the increase of the supplier's profit. As a result, the invoice trading premium needs to be lower so as generate the same profit under invoice trading.

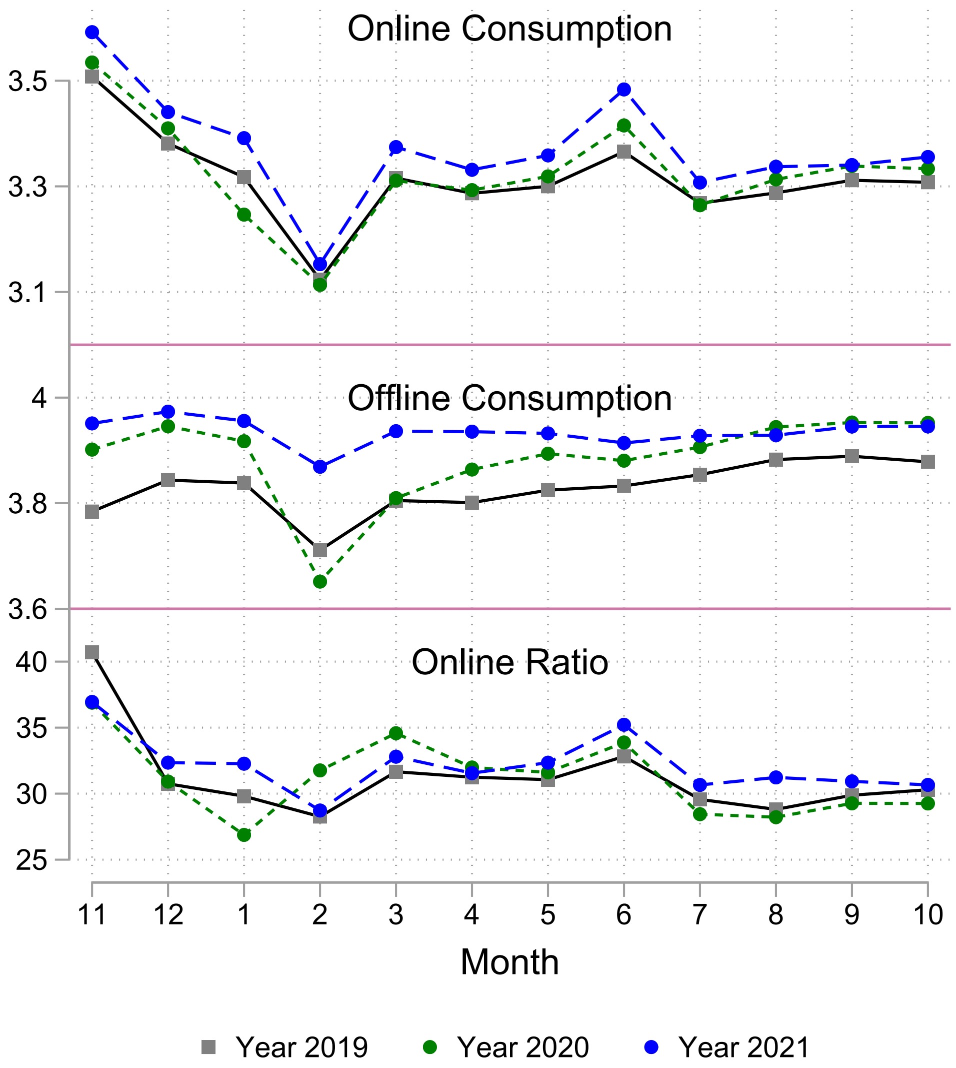
Proof of Proposition B.0.3:

The comparative analysis has been covered in previous proofs, so we skip it.

Chapter C: Appendix for Chapter 3

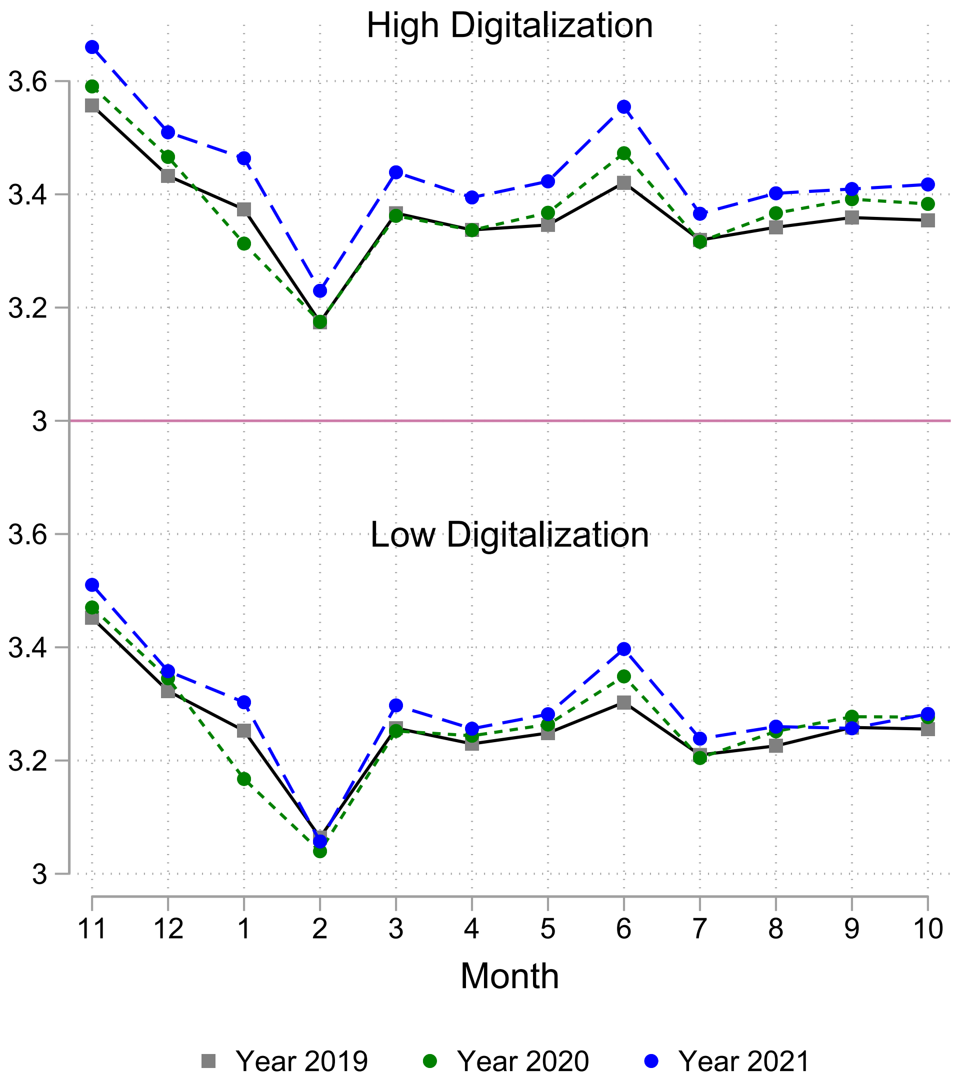
Appendix A: Figures

Figure C.1.: Online Consumption, Offline Consumption, and Online Ratio (2020 vs. 2019)



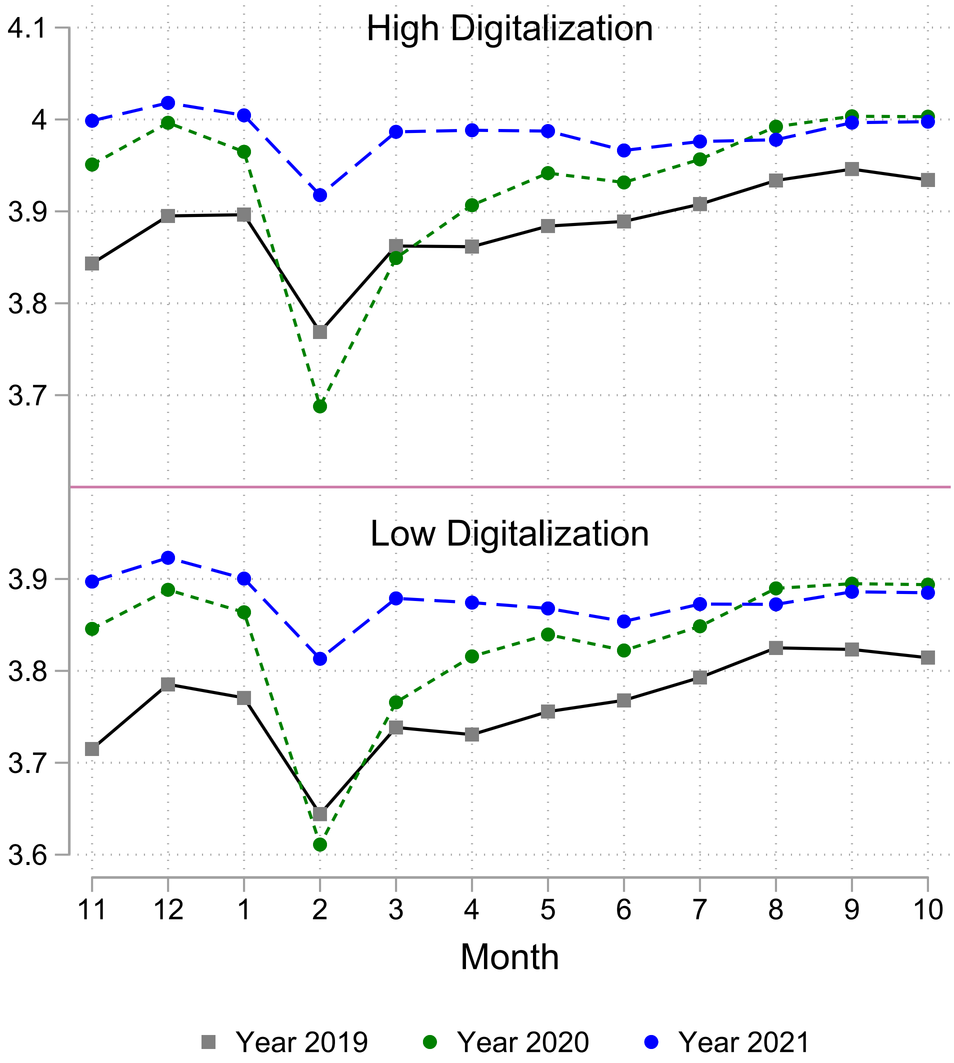
Notes: Online and offline consumption are in CNY and are with inverse hyperbolic sine transformation. Online ratio is in percentage points.

Figure C.2.: Dynamics of Online Consumption by Digitalization (2020 vs. 2019)



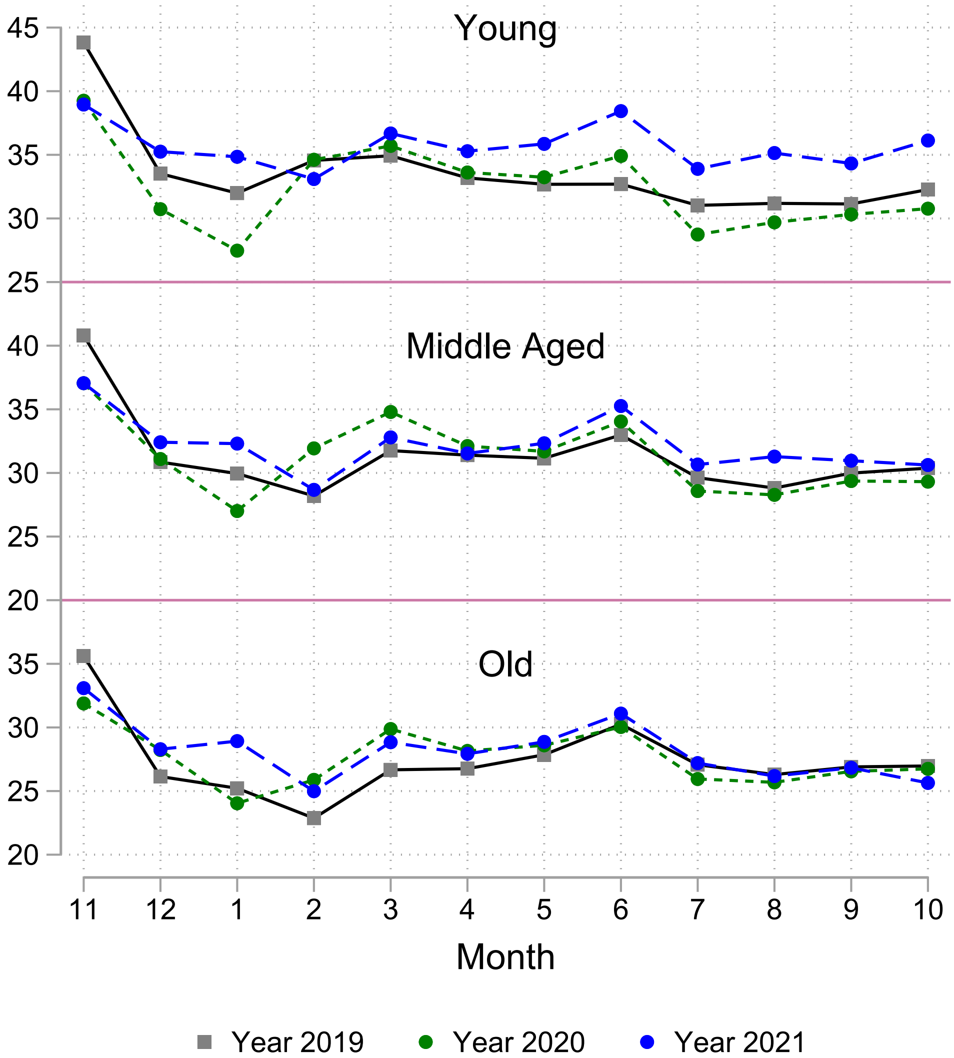
Notes: Digitization segments are based on a median split of all cities according to the PKU-DFIIC index. Online consumption is in CNY and are with inverse hyperbolic sine transformation.

Figure C.3.: Dynamics of Offline Consumption by Digitalization (2020 vs. 2019)



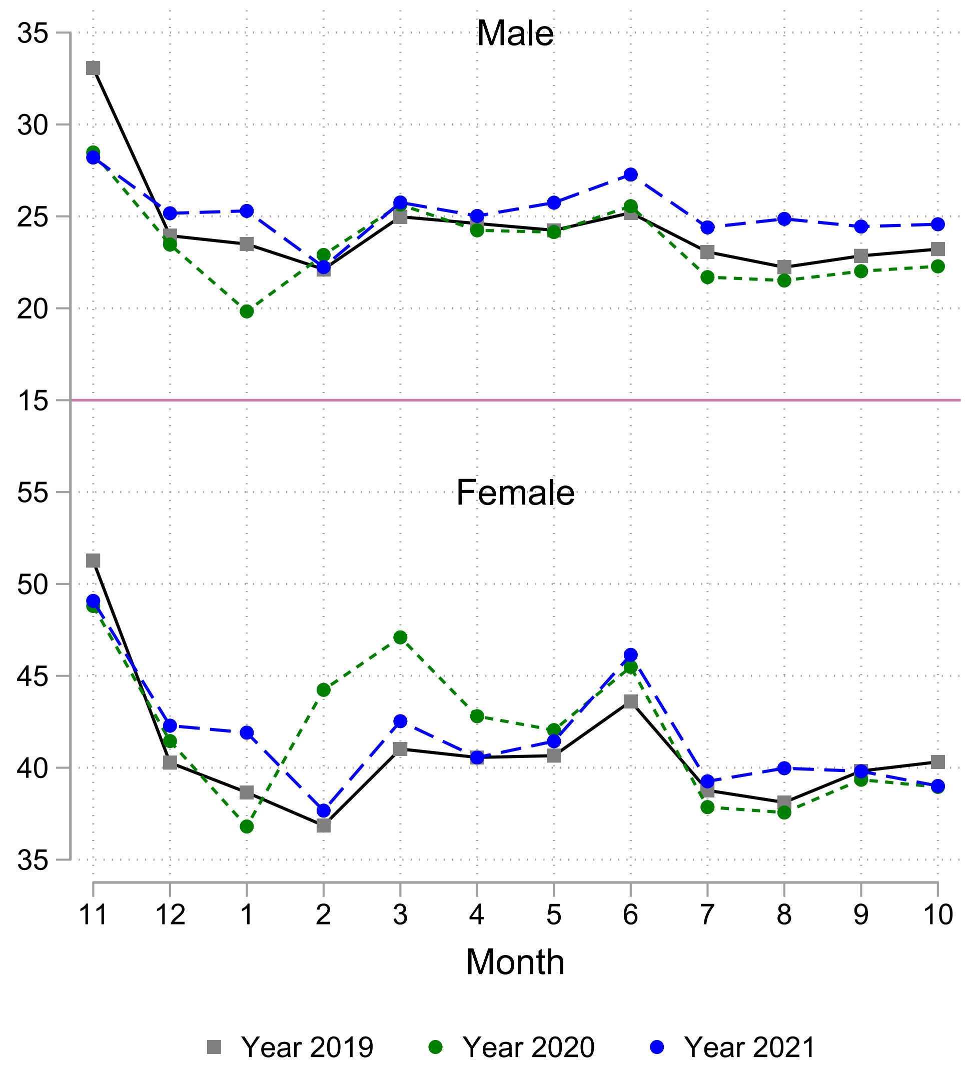
Notes: Digitization segments are based on a median split of all cities according to the PKU-DFIIC index. Offline consumption is in CNY and are with inverse hyperbolic sine transformation.

Figure C.4.: Dynamics of Online Consumption Ratio by Age (2020 vs. 2019)



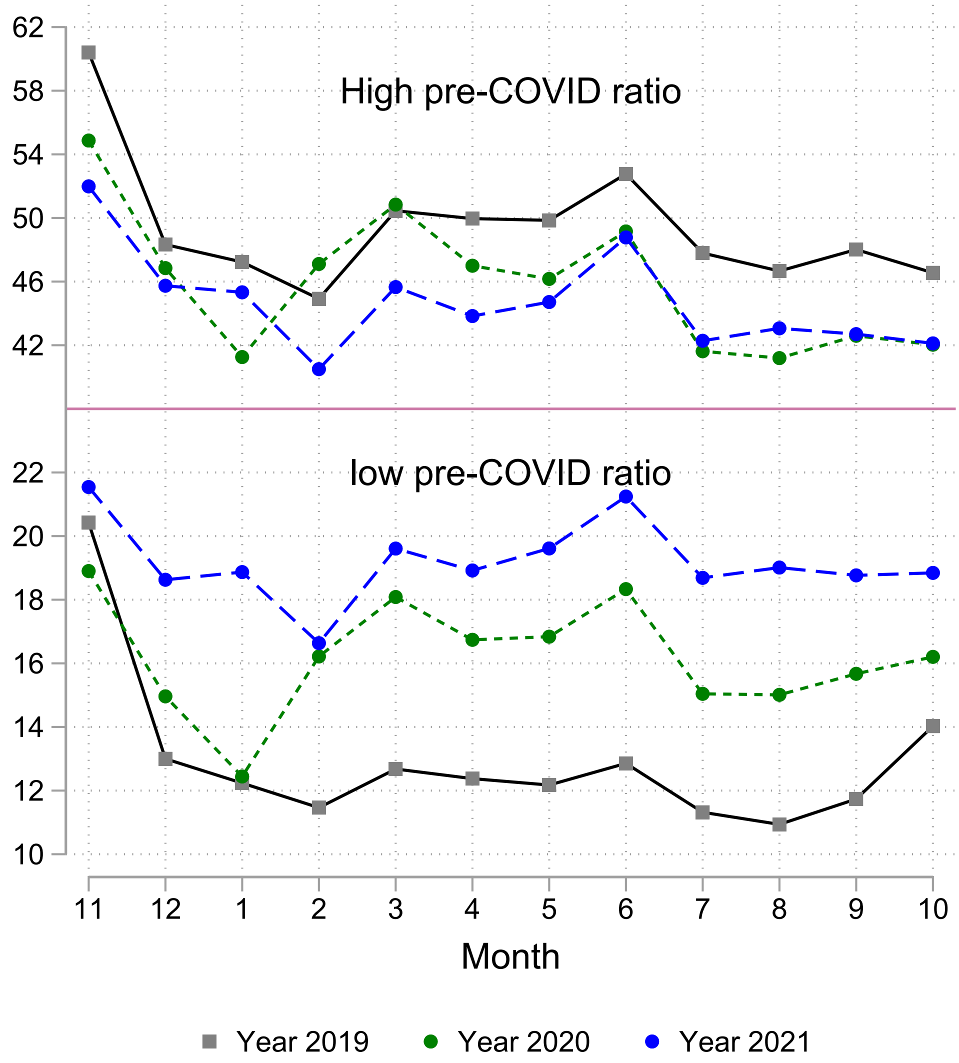
Notes: Young: below 23 years old. Med: 23-50 years old. Old: at least 50 years old. Online ratio is in percentage points.

Figure C.5.: Dynamics of Online Consumption Ratio by Gender (2020 vs. 2019)



Notes: Online ratio is in percentage points.

Figure C.6.: Dynamics of Online Consumption Ratio by Pre-COVID Consumption (2020 vs. 2019)



Notes: The two segments are based on a median split of consumers using their transaction data in year 2019. Online ratio is in percentage points.

Appendix B: Tables

Table C.1: Summary Statistics

	Mean	SD	Observations
Individual Outcomes			
Online Monthly Consumpion (yuan)	1065.42	3455.39	4485436
Offline Monthly Consumption (yuan)	3627.22	10048.87	4485436
Online Ratio (%)	31.19	33.83	4485436
Individual Charateristics			
Female	0.42	0.49	190330
Age	33.28	8.86	190330
City Covariates			
GDP per capita 2018 (yuan)	64884.28	36860.16	225
Digitalization Index (standardized)	0	1	225
Digitalization = High	0.5	0.5	225

Table C.2: Impacts of Covid on Consumption Behaviors: Cross-year Diff-in-diff

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
Year20 $\times$ *During*	-0.2086 (0.0351) [0.0000]	-0.5557 (0.0209) [0.0000]	4.3619 (0.1687) [0.0000]
Year20 $\times$ *Post1*	0.0668 (0.0140) [0.0000]	-0.1579 (0.0134) [0.0000]	2.0057 (0.1266) [0.0000]
Year20 $\times$ *Post2*	-0.0138 (0.0124) [0.0000]	-0.1514 (0.0114) [0.0000]	1.0401 (0.1115) [0.0000]
Base-year Mean (yuan or %)	1048.39	3377.40	31.22
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Individual Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Month Fixed Effects	Y	Y	Y

Note: Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption.

Table C.3: Impacts of Covid on Consumption Behaviors: Cross-region Diff-in-diff

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
$HighCovid \times During$	-0.4004	-0.2965	0.8233
	(0.1080) [0.0002]	(0.0485) [0.0000]	(0.3346) [0.0139]
$HighCovid \times Post1$	0.0244	-0.1303	1.3485
	(0.0198) [0.2166]	(0.0316) [0.0000]	(0.3244) [0.0000]
$HighCovid \times Post2$	0.0431	-0.0790	0.7724
	(0.0235) [0.0661]	(0.0269) [0.0033]	(0.2464) [0.0017]
$MedCovid \times During$	-0.0813	-0.1260	0.7017
	(0.0352) [0.0209]	(0.0297) [0.0000]	(0.2304) [0.0023]
$MedCovid \times Post1$	-0.0143	-0.0544	0.4136
	(0.0202) [0.4793]	(0.0231) [0.0187]	(02426) [0.0882]
$MedCovid \times Post2$	-0.0026	-0.0364	0.3306
	(0.0213) [0.9028]	(0.0218) [0.0952]	(0.2090) [0.1137]
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Log GDP per capita	Y	Y	Y
Individual Fixed Effects	Y	Y	Y
Date Fixed Effects	Y	Y	Y

Note: Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption. We control for interactions between log GDP per capita in 2019 and date fixed effects.

Table C.4: Digitalization and Consumption Dynamics: Baseline Results

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
$Hdigital \times Year20 \times During$	0.2157	0.0893	0.4936
	(0.1032) [0.0366]	(0.0508) [0.0787]	(0.5290) [0.3508]
$Hdigital \times Year20 \times Post1$	-0.0258	-0.0256	-0.0315
	(0.0508) [0.6115]	(0.0297) [0.3876]	(0.2945) [0.9149]
$Hdigital \times Year20 \times Post2$	0.0124	-0.0726	0.5840
	(0.0354) [0.7265]	(0.0287) [0.0114]	(0.2778) [0.0355]
Base-year Mean (yuan or %)	1048.39	3377.40	31.22
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Individual Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Month Fixed Effects	Y	Y	Y

Note: The sample is based on cities in the bottom four quintiles of GDP per capita. The full regression specification is shown in equation 3.3. Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption. We control for interactions between log GDP per capita in 2019 and date fixed effects.

Table C.5: Impacts of Covid on Consumption Behaviors: Heterogeneity by Age Groups

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
Young $\times$ *Year20* $\times$ *During*	-1.0564	-0.5334	1.2061
	(0.0661) [0.0000]	(0.0465) [0.0000]	(0.4732) [0.0108]
Young $\times$ *Year20* $\times$ *Post1*	-0.2349	-0.4257	3.6268
	(0.0710) [0.0009]	(0.0431) [0.0000]	(0.5521) [0.0000]
Young $\times$ *Year20* $\times$ *Post2*	-0.3796	-0.3647	2.3261
	(0.0801) [0.0000]	(0.0474) [0.0000]	(0.6265) [0.0002]
Med $\times$ *Year20* $\times$ *During*	-0.3959	-0.1550	1.3661
	(0.0458) [0.0000]	(0.0302) [0.0000]	(0.3175) [0.0000]
Med $\times$ *Year20* $\times$ *Post1*	-0.1085	-0.1492	1.6617
	(0.0488) [0.0264]	(0.0372) [0.0001]	(0.3924) [0.0000]
Med $\times$ *Year20* $\times$ *Post2*	-0.1821	-0.0602	0.8908
	(0.0572) [0.0014]	(0.0380) [0.1128]	(0.4492) [0.0473]
Base-year Mean (yuan or %)	1048.39	3377.40	31.22
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Individual Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Month Fixed Effects	Y	Y	Y

Note: Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption. Young: age$\leq$ 22; Middle: 22<age$\leq$50; Old: age>50.

Table C.6: Impacts of Covid on Consumption Behaviors: Heterogeneity by Gender

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
Female × *Year20* × *During*	0.0653	-0.1591	2.9713
	(0.0218) [0.0028]	(0.0166) [0.0000]	(0.1850) [0.0000]
Female × *Year20* × *Post1*	-0.0862	0.2550	-0.6681
	(0.0217) [0.0001]	(0.0192) [0.0000]	(0.2069) [0.0012]
Female × *Year20* × *Post2*	-0.0712	0.3449	-1.7937
	(0.0251) [0.0046]	(0.0217) [0.0000]	(0.2214) [0.0000]
Year20 × *During*	-0.2358	-0.4897	3.1228
	(0.0371) [0.0000]	(0.0183) [0.0000]	(0.1440) [0.0000]
Year20 × *Post1*	0.1028	-0.2643	2.2831
	(0.0197) [0.0000]	(0.0115) [0.0000]	(0.1163) [0.0000]
Year20 × *Post2*	0.0160	-0.2774	1.7890
	(0.0186) [0.3910]	(0.0118) [0.0000]	(0.1239) [0.0000]
Base-year Mean (yuan or %)	1048.39	3377.40	31.22
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Individual Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Month Fixed Effects	Y	Y	Y

Note: Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption.

Table C.7: Impacts of Covid on Consumption Behaviors: Heterogeneity by Pre-Covid Online Consumption Ratio

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
High × *Year20* × *During*	-0.2132	0.1500	-0.3533
	(0.0253) [0.0000]	(0.0174) [0.0000]	(0.2263) [0.1184]
High × *Year20* × *Post1*	-0.4106	0.5465	-4.4355
	(0.0259) [0.0000]	(0.0244) [0.0000]	(0.2238) [0.0000]
High × *Year20* × *Post2*	-0.3109	0.5445	-3.8318
	(0.0256) [0.0000]	(0.0252) [0.0000]	(0.2350) [0.0000]
Year20 × *During*	-0.1013	-0.6314	4.5397
	(0.0393) [0.0100]	(0.0178) [0.0000]	(0.1203) [0.0000]
Year20 × *Post1*	0.2734	-0.4328	4.2326
	(0.0223) [0.0000]	(0.0126) [0.0000]	(0.1205) [0.0000]
Year20 × *Post2*	0.1439	-0.4080	2.9795
	(0.0198) [0.000]	(0.0143) [0.0000]	(0.1245) [0.0000]
Base-year Mean (yuan or %)	1048.39	3377.40	31.22
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Individual Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Month Fixed Effects	Y	Y	Y

Note: Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption. *High* is a dummy for pre-Covid online consumption ratio being above median.

Table C.8: Digitalization and Consumption Dynamics: Baseline Results (All Cities)

Dependent Variable	(1) Online Consumption	(2) Offline Consumption	(3) Online Ratio (%)
Hdigit × *Year20* × *During*	0.1770	0.0720	0.5154
	(0.0908) [0.0512]	(0.0496) [0.1469]	(0.4712) [0.2740]
Hdigit × *Year20* × *Post1*	-0.0193	-0.0113	-0.1230
	(0.0492) [0.6941]	(0.0279) [0.6858]	(0.2685) [0.6469]
Hdigit × *Year20* × *Post2*	0.0221	-0.0658	0.4983
	(0.0352) [0.5300]	(0.0284) [0.0205]	(0.2725) [0.0675]
Base-year Mean (yuan or %)	1048.39	3377.40	31.22
Observations	4485436	4485436	4485436
No. of Clusters (Cities)	225	225	225
Individual Fixed Effects	Y	Y	Y
Year Fixed Effects	Y	Y	Y
Month Fixed Effects	Y	Y	Y

Note: The sample is based on all cities.Standard errors clustered at city level in parentheses; p-values in brackets. Dependent variables are (1) The inverse hyperbolic sine transformation of online consumption; (2) the inverse hyperbolic sine transformation of offline consumption; (3) individuals' ratio of online consumption to total consumption. We control for interactions between log GDP per capita in 2019 and date fixed effects.