Essays on Industrial Organization and Platform Economy

Qifan Huang

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Yuya Takahashi, Chair

Fahad Khalil

Alex Shcherbakov

Program Authorized to Offer Degree:
Economics

University of Washington

Abstract

Essays on Industrial Organization and Platform Economy

Qifan Huang

Chair of the Supervisory Committee:

Yuya Takahashi

Department of Economics

My dissertation consists of four chapters which address important questions in empirical industrial organization and platform economy.

Chapter 1 presents the joint work with Castiel Chen Zhuang [1] and Zhentong Lu [2]. This chapter studies the "Zero-Markup Drug Policy" (ZMDP) in China's pharmaceutical industry. The main motivation of the policy is to break the integration between drug prescription and dispensation so that the known agency problem between physicians and patients can be alleviated. This chapter estimates a structural model of China's prescription drug market and quantifies the general equilibrium effects of the ZMDP on drug prices, patient welfare, firm profitability and market structure. Our results suggest that: physicians' prescription decisions are more sensitive to patients' out-of-pocket costs than hospitals' drug markups, unless the coinsurance rate is above 35 percent; pricing is mostly dominated by provincial governments that are assumed to represent the joint welfare of patients and physicians; the ZMDP makes generic drugs relatively more favorable and thus more profitable; while total sales is negatively affected by the ZMDP, overall patient welfare improves by a sizable amount because of the lowered prices.

---

[1] Castiel Chen Zhuang is a Ph.D. student in the Department of Economics at the University of Washington. Email: cczhuang@uw.edu.

[2] Zhentong Lu is a senior economist at the Bank of Canada. Email: zlu@bankofcanada.ca.

Chapter 2 is a joint work with Castiel Chen Zhuang, which was originally published in *Economics of Transition and Institutional Change* [Huang and Zhuang, 2022]. This chapter tests the famous theory by Acemoglu and Pischke [1998]. We estimate a structural model where labor quality can be affected by firm's optimal training decision. We use the data of China's manufacturing enterprises in an era of privatization (2004-2007). Training increases both marginal labor product and wages, but the productivity premium is larger due to the labour market rigidity, which explains the voluntary provision of on-the-job training and supports the theory by Acemoglu and Pischke [1998]. Our results also indicate that: state-controlled enterprises' investment in training could be both privately and socially efficient; unions play a positive role in facilitating training and increasing workers' bargaining power; female workers and low-educated workers have higher training premium.

Chapter 3 is based on the joint work with Castiel Chen Zhuang and Saizi Xiao[3], which has been accepted by *Applied Economics* [Huang et al., 2022]. This chapter investigates how demand, pricing and income distribution in digital platforms respond to the two-sided Covid-19 shock. We focus on a live-streaming platform, where fans could send gifts to streamers. We resort to the generalized quantile regression in Powell [2019] to quantify the unconditional quantile treatment effects of Covid-19 on the virtual gifting. Our result suggests that the pandemic severity on the fans side instead of the anchors side increases virtual gifting. Based on this result, our theoretical model predicts that the platform would cross-subside the anchors. This prediction is consistent with the reality, where the platform spent 1 billion RMB in subsiding anchors. Our estimation results also suggest that Covid-19 leads to the de-polarization of the anchors' gift income distribution, while other researches find Covid-19 exacerbates income polarization. One possible explanation to this puzzle is that the digital platform serves as the sanctuary and provides flexible working opportunities for those relatively unlucky people.

Chapter 4 describes a joint work with Ying-Chin Chen [4] and Castiel Chen Zhuang . It provides

---

[3] Saizi Xiao is a research fellow at Malaysian Institute of Economic Research. Email: saizixiao@gmail.com.

[4] Ying-Chin Chen is a Ph.D. student in the Department of Economics at the University of Washington. Email: ycchen62@uw.edu.

new evidence to a classical economics question: how would market competition affect product variety? We propose a novel multi-task graph convolutional neural network approach to measure market competition intensity. To get a clean identification of the treatment effect, we leverage a natural experiment in a live-commerce platform: the account of a top live streamer with strong market power was suspended by accident. Our results indicate that market competition increases product variety, while retail prices are sticky and do not adjust to the exogenous competition shock.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to the chair of my supervisory committee, Professor Yuya Takahashi, for his unparalleled support and continued guidance. His extensive knowledge, deep insights and great patience has helped me shape and refine the dissertation. He is a great role model in my career.

I am extremely thankful to the members of my supervisory committee: Dr.Alex Shcherbakov, Professor Fahad Khalil and Professor Thomas Richardson. I benefited enormously from their unwavering support and invaluable advice. I am also extremely thankful to Professor Jing Tao for her utmost help at the first half of my PhD study, such as funding support and research guidance.

Additionally, I am truly thankful to my coauthors Castiel Chen Zhuang, Ying-chin Chen, Saizi Xiao and Zhentong Lu. Without them, this dissertation would not have been possible. I would like to extend my sincere gratitude to Professor Yanqin Fan and Professor Pat Bajari for their constructive comments and advice. I would also like to thank Professor Tim Althoff for awarding me teaching assistantship, which provides the crucial financial support to my study.

Finally and most importantly, I would like to thank my parents for being so supportive throughout my PhD study. Their unconditional support and love brings me the courage to chase my dream.

DEDICATION

To my dear parents.

Chapter 1

EXAMINING THE ZERO-MARKUP DRUG POLICY IN CHINA: A STRUCTURAL APPROACH

1.1 Introduction

Policymakers around the world have been regulating their pharmaceutical industries for decades to lower drug prices and make health care more affordable, and China is not an exception. Since 2010, there have been a series of important regulatory changes in China's pharmaceutical industry, which affect the pricing decisions of firms and the drug choices of hospitals, physicians, and patients. One influential reform is the "Zero-Markup Drug Policy" (ZMDP), which requires that hospitals cannot profit from dispensing drugs. Although previous work has documented some aggregate effects of the policy on equilibrium outcomes such as drug prices, little is known about the underlying mechanism of how it works, e.g., how it changes physician choices, drug prices, firm profitability, and consumer welfare. This paper tries to fill this void by first estimating a structural model of China's pharmaceutical industry and then quantifying the impacts on different parties in the market using counterfactual simulations.

One important feature of the demand for drugs in China is that there is an "expert-client" relationship such that a physician acts as a patient's agent. This relationship naturally generates agency problems because physicians concern both hospitals' profit from selling drugs and patients' welfare.[1] Since the 1950s, due to the lack of funding, the Chinese government explicitly allowed public hospitals to add a 15% markup to the wholesale prices of drugs when selling them to patients. Part of hospitals' profit became physicians' income. Consequently, Chinese physicians had been taking hospitals' drug markups into account when making drug choices for or with patients by

[1]The integration between drug prescription and dispensation has a long history in China, dating back to the Eastern Han Dynasty. Inspired by Zhang Zhongjing (A.D. 150–219), Chinese physicians started to "sit" in the pharmacies to provide services and name themselves after *zuotangyi* (on-site physicians). It cultivated the partnership of physicians and drug sellers. Sometimes, physicians may even open pharmacies themselves, known as *langzhong*. With the rapid transformation of the pharmaceutical and healthcare systems in Mao Era, on-site physicians flooded in public hospitals and became their employees.

deliberately prescribing more and expensive drugs. Although this agency problem was somewhat restricted by the 15% markup itself and a direct price upper limit regulation, these pricing constraints are not stringent enough to eliminate "distortions" in physicians' prescription decisions. To mitigate this incentive problem, in 2015, China started the ZMDP in prefecture-level public hospitals. This policy was then implemented nationwide in 2017. [2]

Eliminating hospitals' drug markups affects drug retail prices through three mechanisms. (1) *Direct effect:* if wholesale prices are fixed, it would directly lower retail prices. (2) *Dethronement effect:* it makes the expensive branded drugs less attractive than before (because of changes in physicians' incentive), which might decrease the market power of branded drugs and lower their prices (either wholesale or retail); on the contrary, the market power of generic drugs will likely increase, which leads to higher prices. (3) *Push-out effect:* removing markups also makes physicians less likely to prescribe in general (such as encouraging patients to go on a healthy diet instead), so the overall market power of prescription drugs decreases, which might lead to lower prices. Through these mechanisms, the ZMDP is very likely to reduce the retail prices of branded drugs. However, the overall impact on prices of generic drugs is ambiguous. The policy effects on retail prices translate into those on manufactures' profitability, patient welfare, etc., and quantifying these different aspects is the main goal of this paper.

One difficulty is how to single out the effect of ZMDP from those of other policy changes that happen around the same time and aim for the same objectives. To tackle this problem, we develop a structural model of demand and supply of China's pharmaceutical market that exploits some institutional features and help us tease out the impact of ZMDP. This structural model will be more sensitive to the ZMDP than to other policy changes. Using data on drug sales and observed binding constraints on the prices of lipid-lowering drugs between 2012 and 2018, we first estimate the structural parameters in the model and then simulate the counterfactual unregulated equilibrium outcome in the absence of ZMDP. The comparison between the actual and counterfactual market outcomes gives us a quantitative account of the impact of ZMDP.

Our first step is to estimate the demand for differentiated lipid-lowering drugs. We follow the standard approach in empirical IO [Berry, 1994, Berry et al., 1995, Iizuka, 2007, Berry and Jia,

[2]The pilot reform was launched earlier for county-level or township hospitals. But due to the lack of detailed data, in this research, we focus on prefecture-level public hospitals.

2010], and set up a two-type mixed nested logit model of the joint preference of a physician-patient pair based on observed drug characteristics, where the mixture captures the unobserved preference heterogeneity due to our partial observation on whether a hospital is subject to the ZMDP. We find that the estimated demand is strongly affected by hospital drug markups, implying that physicians do not fully represent the patients' interests. Also, physicians put more weight on patient welfare than hospitals' profits from drugs, as long as the coinsurance rate for drugs is low enough.

Once we have estimated demand and the implied substitution patterns, we explore the impact of the ZMDP on retail prices in a setting where competing drug manufacturers simultaneously negotiate with the provincial government about wholesale prices in a Nash bargaining game [Horn and Wolinsky, 1988, Crawford and Yurukoglu, 2012, Grennan, 2013, Gowrisankaran et al., 2015, Ho and Lee, 2017, Dubois et al., 2019a] given the observed constraints imposed by the regulator. The model allows us to separately identify costs and bargaining parameters, the latter of which captures how the provincial policymakers in China trade off between firm profits and patients' welfare.

Finally, given the estimated parameters of preference, production cost, and bargaining power, we can quantify how much the observed decline in China's prescription drug prices can be explained by ZMDP using counterfactual simulations. In particular, we calculate the new equilibrium prices in a hypothetical scenario in which ZMDP did not happen. Then, conditional on all other policies implemented in 2018, we can compare the counterfactual retail prices without the ZMDP and the observed actual retail prices in 2018. The market shares, revenues, profits and social welfare under our counterfactual scenario are also compared to the estimates under the actual situation.

Our results have a few implications. First, though the prescription choices of physicians can be influenced by drug markups, physicians are more sensitive to patient's medication costs than hospitals' profits from drug markups, as long as the coinsurance rate is lower than 35 percent. Second, if we assume that the provincial government represents the consumer welfare, and simplify the centralized drug procurement into a 1-to-1 bargaining process between each firm and the government, then the pricing equilibrium is mostly dominated by provincial governments due to their strong bargaining power. Third, branded drugs are more preferred than generic drugs in China, and the demand elasticity for generic drugs is about 23 percent more elastic than branded drugs on average, suggesting a higher market power of the latter. Fourth, ZMDP makes generic drugs relatively more favorable, and thus can increase their profitability. Lastly, overall drug demand are weakened by the ZMDP, but due to the reduced prices, overall patient welfare is improved by more than 12 percent.

Our work is related to several strands of literature. First, it builds upon the broad research on estimating demand for pharmaceuticals using various methods to estimate preferences for drugs and substitution patterns, from the log-log models [Berndt et al., 1995] to the discrete choice models such as logit [Berndt et al., 2003a], nested logit [Iizuka, 2007, Donohue and Berndt, 2013, Song et al., 2017], and random coefficient logit [Björnerstedt and Verboven, 2016, Dubois and Lasio, 2018, Dubois et al., 2019a]. Also, it relates to the research on physicians' agency problem. For example, Ho and Pakes [2014] investigate the agency problem in physicians' referral decisions, and Lu [2014] studies physicians' overprescription behaviours. Furthermore, the demand model in our paper is similar to Iizuka [2007], which shows Japanese physicians' prescription decisions respond to drug markups when diagnoses and drug sales are vertically integrated. Moreover, it relates to the empirical studies of double marginalization [Berto Villas-Boas, 2007, Bonnet and Dubois, 2010, Gayle, 2013].

Our paper belongs to the literature on the program evaluation of China's healthcare reforms such as the ZMDP [Zhou et al., 2015, Yi et al., 2015, Fu et al., 2018], Samming model and "two invoices" system [Meng et al., 2019], and Shenzhen's experiment with group purchasing organizations [Yang et al., 2020]. These existing studies are either case studies using data from only a sample city or are based on county-level hospitals. Case studies may fail to distinguish the effects of different components of a systemic reform, while studies that focus on county-level hospitals leave the effects in the cities unanswered. Our work fills the gap by evaluating the nationwide implementation of ZMDP among public hospitals in China using a nationally representative sample.

Finally, we contribute to the discussion on how a drug procurement system affects market outcomes. Baldi and Vannoni [2017] use the data on tender prices of selected drugs for hospital usage provided by 52 Italian local health service providers during 2009–2012 and find that centralized procurers pay lower prices than decentralized units, conditional on measures of institutional quality, corruption, and some other covariates. Duggan and Scott Morton [2006] study the effects of government drug procurement using the data on Medicaid prescription drug purchasing, and find that a set fraction of the average price paid by non-Medicaid consumers can not only increase the overall prices but also lead to the introduction of new products that are free of the Medicaid price regulation and thus more expensive. In this paper, we investigate the role that centralized procurement plays in lowering the prescription drug prices. Different from Duggan and Scott Morton [2006] , we focus on China's basic medical insurance that covers more than 95% of its population. Similar to Baldi and Vannoni [2017], we are also able to compare a centralized system with a decentralized system, through replacing the Nash bargaining between the government and firms by the Bertrand competition among firms.

The remainder of the paper is organized as follows. Section 1.2 describes the empirical setting, including the prescription drug market in China, the incentive problem, the regulatory efforts to solve the problem, the recent policy changes, the data used, and a reduced-form evidence of price drop. In Section 1.3, we present the structural model of the demand and supply for each market, as well as the identification and estimation strategy. Section 1.4 presents the estimation results of the structural model. In Section 1.5, we then provide the counterfactual price equilibrium and profitability calculations in the absence ZMDP in 2018, and then calculate the welfare change for patients. Finally, we conclude in Section 1.6.

1.2 Background and Data

1.2.1 Drug procurement reform in China

In 2009, the Chinese government formally initiated a nationwide centralized drug procurement (henceforth CDP) scheme after 9 years of development and experiment in 4 provinces since 2000. The scheme is outlined in two documents released in 2010, namely *Notice on the Issuance of the Centralized Drug Procurement in Health Facilities* [Ministry of Health, 2010] and *State Council Office's Notice on Establishing and Standardizing Essential Drug Procurement in Government-sponsored Primary Health Facilities* [State Council's General Office, 2010]. The new policy required that all public healthcare institutions could procure drugs only via their provincial governments' CDP platforms.

The procurement procedure can be described as follows. First, each hospital takes physicians' advice into account and submits a proposal of drug demand. Then, the provincial government evaluates those proposals and approves a list of drugs to participate in the procurement process. Finally, drug suppliers (e.g., manufacturers, domestic agencies of foreign pharmaceutical companies) compete on the drugs they would like to provide via a rather complicated bidding process.[3] The bidding process is not a standard scoring auction and the specific rules are different across provinces. Without detailed information, it's hard to exactly model this process. So in our empirical analysis, we proceed with a parsimonious model of bargaining between the governments and drug suppliers on drug prices a la Dubois et al. [2019a,b]. After December 2018, the procurement process is changed/enhanced,[4] and so the data in 2019 are only used to generate summary statistics but not for estimating the structural model.

The major players in the CDP scheme, who are the subjects of our research, are included in Figure 1.1. As mentioned in Ministry of Health [2010], the bargaining should be between pharmaceutical companies and provincial governments. Renegotiation between pharmaceutical companies and hospitals is prohibited.

1.2.2 The ZMDP and price regulations

We briefly summarize the major regulatory policy changes that may affect drug prices during 2014-2018 in Table 1.1. The key policy change during this period is the ZMDP, which was implemented only among prefecture-level public hospitals in May 2015 and later extended to all public hospitals nationwide in September 2017.

To explain the implications of the price regulations on the retail price of a drug, let us denote p^W as a wholesale price, which is the same for all hospitals in the same province and is decided by the bargain between the firm and the provincial government. Let p^R denote the retail price of the drug at the hospital. Before April 2014, the regulations require that:

$$\frac{p^R - p^W}{p^W} \leq 15\% \text{ and } \; p^R \leq p^{Highest}, \tag{1.1}$$

where $p^{Highest}$ is the price cap imposed by the provincial government (may be different across

[3] For example, one popular bidding framework is the so-called "two envelope" bidding, in which drug suppliers are required to submit prices in one envelope (termed a price envelope) and the information of their drugs (such as indications) and suppliers (such as reputation) in another envelope (termed a quality envelope). The government then groups suppliers according to their proposals. Next, for each group, the government does a quality screening and chooses qualified candidates based on the quality envelope. Within each group, if there are only a few candidates (e.g., two), the government would directly negotiate with them, otherwise the government may simply choose several low bids (not necessarily the lowest one) as the winning suppliers.

[4] Joint procurement was carried out by "4+7" large cities in December 2018 and then by 27 provinces in September 2019.

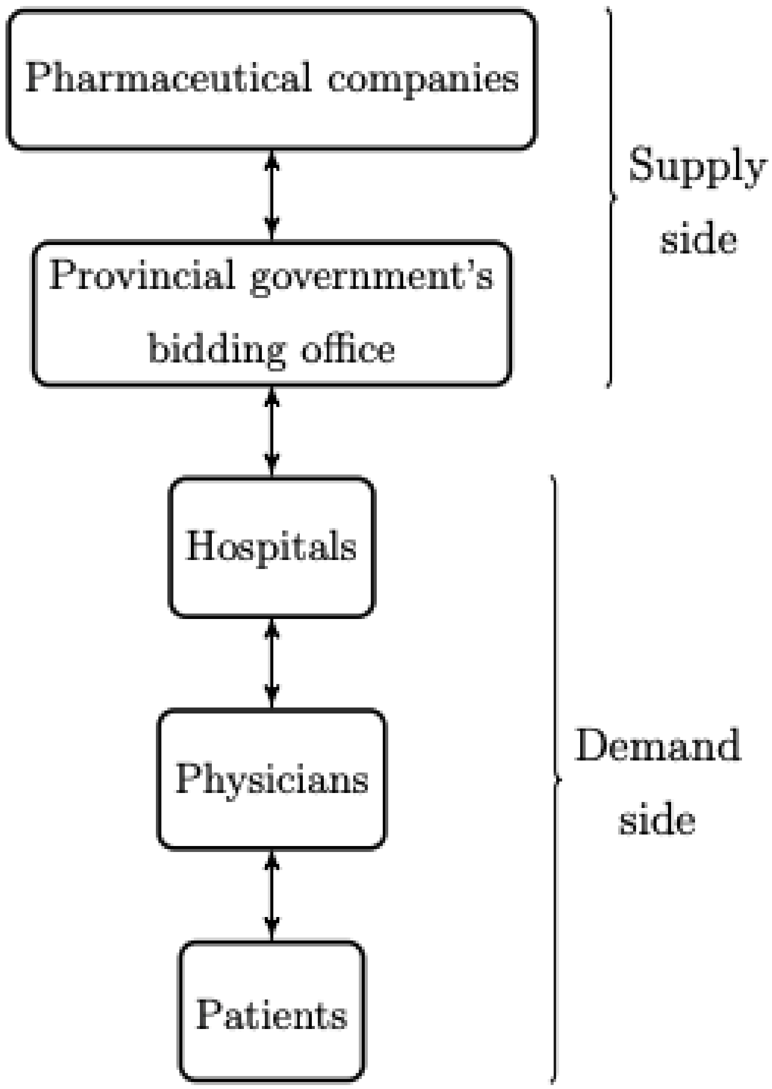

Figure 1.1: Basic structure of China's drug market (at provincial level)

Table 1.1: Major policy changes between 2014 and 2018

Time	Description
Apr 2014$^\diamond$	Remove the retail price caps for Lovastatin, Fenofibrate, Gemfibrozil, Xuezhiang, and Zhibituo.
May 2015*	Initiate the ZMDP among prefecture-level public hospitals (start of the phase-in period).
Jun 2015$^\diamond$	Remove retail price caps for all other lipid-lowering drugs.
2015–2017	Based on State Council's General Office [2015], the revenue from drugs should be no more than 30% of the total medical revenues in the urban public hospitals by 2017.
2016–2018	Encourage local governments to experiment with joint procurement. For example, Shanghai and Shenzhen experimented with some Group Purchasing Organizations (GPOs) in 2016; Beijing, Tianjin, and Hebei united in the procurement of medical supplies in 2017.
Mar 2016	Launch the Generic Consistency Evaluation (GCE) program to test the quality and efficacy of generic drugs. The deadline for chemical drugs that entered before October 2007 was set to December 2018 but then it was canceled/extended.
2017–2018	Based on State Council's Healthcare Reform Committee [2016], a "two invoices" system should be phased in among publicly owned medical institutions and implemented nationwide by 2018.
Sep 2017*	Zero hospital drug markup for all public hospitals.
Dec 2018	"4+7" large cities joint procurement of Atorvastatin and Rosuvastatin. Winners take 60%–70% public hospital market shares in those cities.

provinces). We can rewrite (1.1) as

$$p^R \leq \min\{p^{Highest}, 1.15p^W\}. \tag{1.2}$$

After June 2015, the price cap is removed, so we simply have $p^R \leq 1.15p^W$. Finally, it is replaced by $p^R = p^W$ since 2017Q4.

1.2.3 Data and descriptive statistics

We obtain quarterly data between 2012Q1 and 2019Q3 from the Pharmaceutical DataBase (PDB) on revenues and quantities of the prescription drugs in China's "national drug catalog"[5] treating hyperlipidemia in the sample hospitals. The sample covers around 700 hospitals in 24 provinces of China. Among these hospitals, about 79% are tertiary and about 20% are secondary.[6]

In the raw data, the same drug can come with different forms (e.g., tablets and capsules) and sizes (e.g., 5mg and 10mg). We aggregate drug products (defined by a molecule-firm pair) with the same name but with multiple forms and sizes by "standard unit", the recommended daily dose of a given molecule produced by a given firm.[7] We obtain aggregate sales of different drug products at the province-quarter (defined as a "market" later) level, and then compute quarterly wholesale prices as the ratio of total revenue to total quantity in standard units. Retail prices are not directly observed from the data. We calculate them by assuming that the price constraint (1.2) is binding.[8]

Drug characteristics (including standard units, indications, contraindications, and side effects) are manually collected from the package inserts provided by `yaozh.com` and various sources (most of which are publicly available). Information on price caps are from `yaozh.com` as well. Firm characteristics (such as the time a firm was first allowed to produce each drug in China, and the time each firm was certified by GSP for distribution) are obtained from `menet.com.cn`. We also manually collect the county-level minimum wages facing each manufacturer each quarter from the policy documents posted by local governments. Table 1.2 lists all of these variables.

[5] The "national drug catalog" is designed for the basic medical insurance, work-related injury insurance, and maternity insurance.

[6] Very few hospitals are either lower-level or not classed and thus are negligible. For more details, visit `http://pdb.pharmadl.com`.

[7] We treat firms that share the same parent company as one firm.

[8] That is, we assume that the hospitals set the highest possible retail prices, as hospitals typically add a 15% drug markup when they can. Anecdotal evidence suggests that this assumption almost certainly hold in reality.

Table 1.2: Definitions of main variables

Variable	Definition
Drug characteristics	
Dose	Amount (mg) of drug taken at one time
Frequency	How often each drug is taken every day
Standard unit	Daily dose = dose × frequency
# of indications	Number of indications
# of contraindications	Number of situations in which the drug should not be used with another drug (termed a drug contraindication) or by a patient (termed a patient contraindication)
Chinese	Dummy = 1 if the drug contains Chinese herbal medicine ingredients
Old Statins	Dummy = 1 for the first / second generation of Statins
New Statins	Dummy = 1 for the third generation of Statins
Fibrates	Dummy = 1 if the drug belongs to Fibrates
Niacin	Dummy = 1 if the drug belongs to Niacin
# of forms	Number of drug forms by each firm
# of sizes	Number of drug sizes by each firm
Firm characteristics	
First generic drug	Dummy = 1 if the drug is the first generic drug available in China
Branded	Dummy = 1 if the drug is branded
Time from entry	Number of quarters from entry in Chinese market
Foreign	Dummy = 1 if the firm is foreign-invested
Cost shifters	
Min wage	Minimum hourly wage of the county in which the manufacturer is located
Imported	Dummy = 1 if the drug is imported
GSP	Dummy = 1 if the firm has the GSP certification for distribution
Policy shocks	
Pilot rate	The proportion of cities in a province that pilot the systemic public hospital reform
Start GCE	Dummy = 1 if the firm has started the generic consistency evaluation
Market performance	
Retail price	Price charged by each hospital per standard unit
Wholesale price	Procurement price per standard unit
Hospital markup	Difference between retail price and wholesale price
Market share	The ratio of the sales volume of a firm/drug to the total market sales volume

We present the average wholesale prices in Figure 1.2 for some best-selling drugs, i.e., Atorvastatin, followed by Simvastatin, Rosuvastatin, and Fluvastatin since 2012. As shown, each of the best-selling prescription drugs treating hyperlipidemia experienced a modest decline in price before 2017Q3 and a drastic decline in price after the nationwide implementation of the ZMDP.

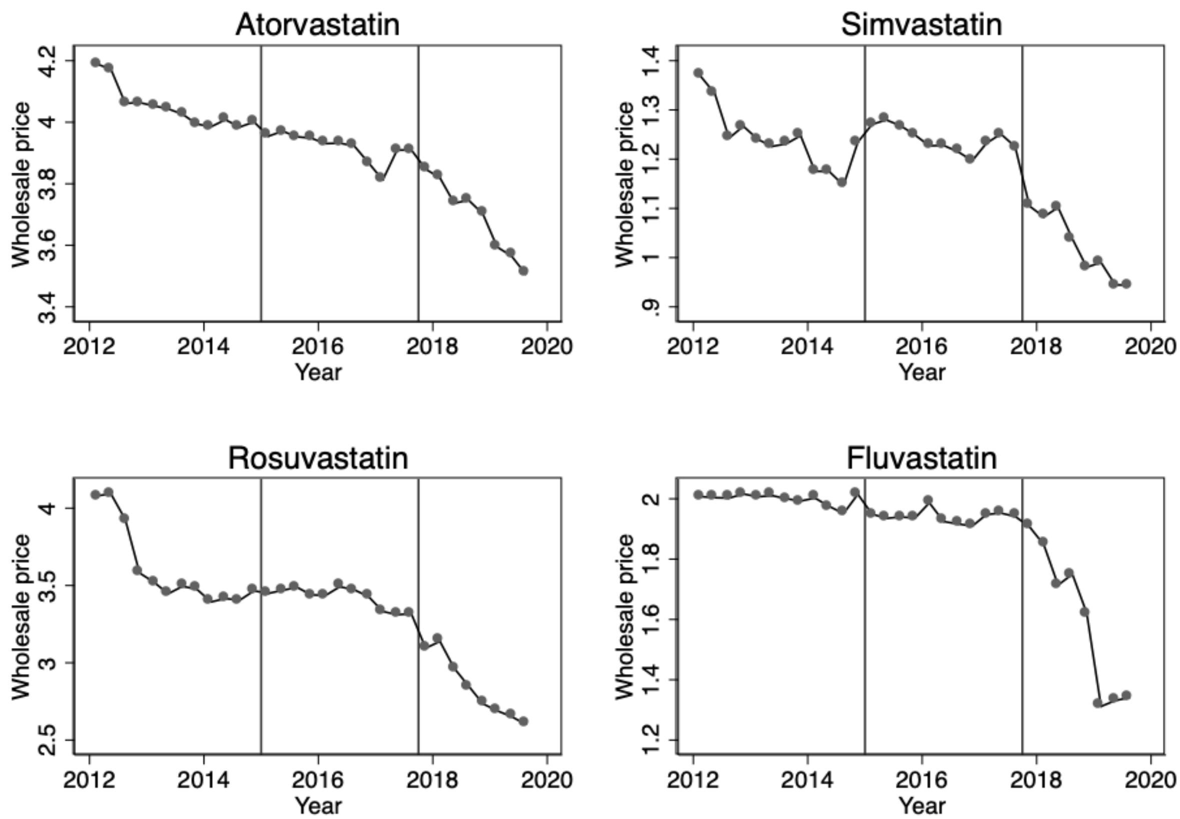

Figure 1.2: Average prices of top-selling lipid-lowering drugs in China

Finally, the summary statistics of drug characteristics and other variables used in this paper are shown in Table 1.3. Patient-day unit retail prices vary across molecules, but it is 3.67–3.70 CNY on average before ZMDP pilot kicked in (roughly speaking 1 CNY equaled $0.16 during 2012–2014). Physicians/hospitals in turn earned 0.43–0.46 CNY on average per patient-day (or 157–168 CNY per patient-year) by prescribing a lipid-lowering drug. Starting from 2017Q4, physicians could not

Table 1.3: Summary statistics

	Obs.	Mean	St. Dev.	Min	Max
Price and markup (CNY)					
(2012Q1–2014Q4)					
Retail price	10,178	3.70	2.74	0.06	14.94
Hospital markup	10,178	0.46	0.36	-5.07	1.95
(2015Q1–2017Q3)					
Non-pilot retail price	8,944	3.79	2.81	0.04	19.47
Non-pilot hospital markup	8,944	0.49	0.36	-0.19	2.54
Pilot retail price	8,944	3.30	2.45	0.03	16.93
Pilot hospital markup	8,944	0	0	0	0
(2017Q4–2018Q4)					
Retail price	4,025	3.28	2.49	0.12	14.4
Hospital markup	4,025	0	0	0	0
Product and firm features					
# of indications	23,147	3.05	0.97	1	4
# of contraindications	23,147	5.31	1.71	2	7
First generic drug	23,147	0.22	0.42	0	1
Branded	23,147	0.25	0.43	0	1
Time from entry	23,147	48.97	20.16	4	138
Foreign	23,147	0.28	0.45	0	1
Chinese	23,147	0.06	0.24	0	1
Old Statins	23,147	0.43	0.49	0	1
New Statins	23,147	0.28	0.45	0	1
Fibrates	23,147	0.19	0.39	0	1
Niacin	23,147	0.04	0.19	0	1
Cost shifters					
Min wage	23,147	18.71	13.61	6	80.39
Imported	23,147	0.22	0.42	0	1
GSP	23,147	0.72	0.45	0	1
Policy shocks					
Pilot rate	23,147	0.41	0.40	0	1
Start GCE	23,147	0.02	0.15	0	1

Note: Please refer to Table 1.2 for variable definitions.

earn such profits directly from dispensing drugs anymore. According to the number of quarters from first entry, we learn that some drugs are relatively new while some are quite old. And most of the drugs entered the Chinese market before our sample period.

1.2.4 Reduced-form analysis of price and quantity

As listed in Table 1.1, there are several regulatory changes that might affect demand or supply of lipid-lowering drugs in China during the period of study. Before showing how we structurally identify and estimate the impact of ZMDP on prices, demand, and profitability, we run reduced-form regressions of wholesale price and quantity over the three implementation phases of the ZMDP: pre-reform (2012–2014), partial reform (2015–2017Q3), and post-reform (2017Q4–2018) periods. Admittedly, we do not intend to explore any causal relationship with the reduced-form regressions as we may not completely rule out the effects of the "two invoices" system, GCE program, local joint procurement attempts, restricted revenue composition (e.g., revenues from drugs should account for less than 30 percent of total hospital revenues), and other regulations that were phased in during the same period of time. Nevertheless, the reduce-form regressions portray the trends of price and quantity over the three policy periods.

To make the price and quantity comparisons across different phases of the ZMDP meaningful, we control for the firm and drug fixed effects, and other characteristics of the drug products. We also try our best to control for measures of policy shocks that are uneven to different markets and firms in each quarter. Table 1.4 reports the results of the fixed-effect regression of the log wholesale price and quantity of drugs on the reform phase dummies and drug characteristics. When we control for the fixed effects, we see an evident price drop following the implementation of the ZMDP. In the first column, we can also see that branded drugs and those with more indications are more expensive; drugs with more contraindications are cheaper; older generic drug tends to have a lower price. Other policy shocks are also associated with lower wholesale prices. In the second column, we notice that the zero-markup drug policy was associated with lower quantity as well, seemingly indicating the "push-out" effect.[9]

This reduced-form evidence seems to confirm that vertical separation was associated with lower

[9] The results are similar when we jointly estimate both equations.

Table 1.4: Fixed-effect regressions of log wholesale price and quantity

	(1) log price	(2) log quantity
2012–2014 (pre-reform)	(reference group)	
2015–2017Q3 (partial reform)	-0.042***	-0.197***
	(0.005)	(0.027)
2017Q4–2018 (post-reform)	-0.074***	-0.246***
	(0.007)	(0.037)
# of indications	0.461***	1.848***
	(0.090)	(0.466)
# of patient contraindications	-0.239***	0.685***
	(0.035)	(0.178)
# of drug contraindications	-0.684***	-0.460*
	(0.054)	(0.278)
First generic drug	-0.083***	0.681***
	(0.020)	(0.103)
Branded	0.075**	3.209***
	(0.035)	(0.182)
Pilot rate	-0.033***	0.936***
	(0.006)	(0.033)
Start GCE	-0.220***	0.512***
	(0.016)	(0.083)
Firm fixed effect	Yes	Yes
Molecule fixed effect	Yes	Yes
Observations	23,147	23,147
R^2	0.551	0.394

Notes: (1) Standard errors in parentheses under each coefficient. (2) Dependent variables are the natural log of wholesale price in CNY, and the natural log of quantity in ”standard unit”. (3) Data for China in 2012–2018. (4) ***, **, * denote significance level at 1%, 5%, and 10%, respectively.

prices, conditional on other environmental changes being modest. Our structural estimation will allow us to interpret those results, test whether observed price changes are due to demand or supply conditions, and separate the effect of the ZMDP from other regulations and environmental changes. As a result, we can tell how much of the price drop can be explained by vertical separation. Moreover, our structural approach will allow us to perform simulations of counterfactual policies.

1.3 Model

In this section, we set up an empirical model of demand, supply and market equilibrium in the Chinese lipid-lowering drug market (focusing on hospital pharmacies).

1.3.1 Demand side

Given our data, we define a market as a province-quarter pair and label it by $t = 1, ..., T$. Each market t consists a set of competing products, labeled by $j = 1, ..., J_t$, which are defined as molecule-firm pairs.

A patient and her physician jointly decide on which drug product to buy and use. So we model the joint preference of a patient-physician pair, labeled by i, using a standard nested-logit random utility model, i.e., the utility that i obtains from choosing product j is

$$U_{ijt} = \underbrace{X_{jt}\theta_1 + \xi_{jt}}_{\equiv\delta_{jt}} - \alpha P_{ijt} + \gamma M_{ijt} + \zeta_{igt}(\lambda) + (1-\lambda)\varepsilon_{ijt}, \tag{1.3}$$

where

- δ_{jt} represents the mean utility, in which X_{jt} is a vector of observed product or market characteristics, including molecule dummies, the proportion of public hospital reform pilot cities in market t (to control for reform intensity), the indicator for the start of GCE process, and a constant term, and ξ_{jt} is an unobserved product-market level demand shock;

- P_{ijt} and M_{ijt} are the retail price and (hospital) markup that patient-physician i faces for drug product j in market t, and α is the dis-utility of price and γ measures the severity of the expert agency problem [Iizuka, 2007];

- Depending on whether the hospital associated with i is subject to the ZMDP, P_{ijt} and M_{ijt} differ across i's: $P_{ijt} = p_{jt}^{W}$ and $M_{ijt} = 0$ if i is subject to the ZMDP and $P_{ijt} = p_{jt}^{R}$ and $M_{ijt} = m_{jt}$ otherwise;[10] also, for the partial ZMDP periods, we do not observe whether i is subject to the ZMDP, so P_{ijt} and M_{ijt} become unobserved heterogeneity and we shall estimate the fraction of i's that are not subject to the ZMDP as a parameter $\phi \in [0, 1]$;

- ζ_{igt} is a random variable that is common to all products in nest g, whose distribution depends on λ. $\lambda \in [0, 1)$ is the "nesting parameter" capturing the within-group correlation between choices. Larger λ means nests matter more. ε_{ijt} is an i.i.d. idiosyncratic preference shock following the standard Type I extreme value distribution. In our empirical analysis, a group g is defined by a molecule and there are 17 of them (0 for outside goods, and 1–16 for the 16 molecules in Table A.1 except Jiaogulan).

Each decision maker i in market t maximizes her utility by choosing the best option in $\mathcal{J}_t$. Given nested-specification, the choice probability that i chooses j in t can be written as

$$\sigma_j\left(\delta_t, P_{it}, M_{it}\right) = \underbrace{\frac{\exp\left(\frac{\delta_{jt}-\alpha P_{ijt}+\gamma M_{ijt}}{1-\lambda}\right)}{\sum_{j\in g}\exp\left(\frac{\delta_{jt}-\alpha P_{ijt}+\gamma M_{ijt}}{1-\lambda}\right)}}_{\text{within-group share}} \underbrace{\frac{\left(\sum_{j\in g}\exp\left(\frac{\delta_{jt}-\alpha P_{ijt}+\gamma M_{ijt}}{1-\lambda}\right)\right)^{1-\lambda}}{\sum_{g\in\mathcal{G}_t}\left(\sum_{j\in g}\exp\left(\frac{\delta_{jt}-\alpha P_{ijt}+\gamma M_{ijt}}{1-\lambda}\right)\right)^{1-\lambda}}}_{\text{group share}}. \quad (1.4)$$

Thus we can obtain the aggregate market share $E\left[\sigma_j\left(\delta_t, P_{it}, M_{it}\right)\right]$ by integrating out the heterogeneous P_{it} and M_{it}. For the pre-2015 periods, all the i's are not subject to ZMDP and thus

$$E\left[\sigma_j\left(\delta_t, P_{it}, M_{it}\right)\right] = \sigma_j\left(\delta_t, p_t^R, M_{ijt}\right). \quad (1.5)$$

Also, between 2015Q1 and 2017Q3 (partial implementation of ZMDP), whether each i is subject to the ZMDP is an unobserved heterogeneity and thus

$$E\left[\sigma_j\left(\delta_t, P_{it}, M_{it}\right)\right] = \phi\sigma_j\left(\delta_t, p_t^R, M_{ijt}\right) + (1-\phi)\,\sigma_j\left(\delta_t, p_t^W, 0\right). \quad (1.6)$$

Finally, after 2017Q3 (full implementation of ZMDP), the market share equation is

$$E\left[\sigma_j\left(\delta_t, P_{it}, M_{it}\right)\right] = \sigma_j\left(\delta_t, p_t^W, 0\right). \quad (1.7)$$

[10] The variation in $p^{Highest}$ across markets will ensure non-colinearity between p^R and the markup m, which will help us identify how the retail price and hospital drug markup affect the utility of consumers separately.

With the above specified market share function, we can write the demand system as

$$s_{jt} = \bar{\sigma}_{jt}\left(\delta_t;\theta_2\right), \forall j,t \tag{1.8}$$

where s_{jt} is the observed market share of j in t, $\bar{\sigma}_{jt}\left(\delta_t;\theta_2\right) \equiv E\left[\sigma_j\left(\delta_t, P_{it}, M_{it}\right)\right]$, and $\theta_2 = (\theta_1, \alpha, \gamma, \lambda, \phi)$.

To estimate the model, we invert the demand systems[11], (1.5), (1.6) and (1.7), to obtain

$$X_{jt}\theta_1 + \xi_{jt} = \bar{\sigma}_{jt}^{-1}\left(s_t;\theta_2\right) \tag{1.9}$$

and assume the following identification condition

$$\mathbb{E}\left[Z_{jt}^d \xi_{jt}\right] = 0, \tag{1.10}$$

where Z_{jt}^d is a vector of exogenous variables, including arguably exogenous product characteristics, cost shifters ("Min wage" and "Imported" in Table 1.2) and BLP-type IVs: (1) the number of drugs and the sum of characteristics for other drugs sharing the same molecular class at market t (the crowdedness of the product space), and (2) the number of drugs and the sum of characteristics for other drugs sold by the same firm at market t (the ownership pattern).

Based on the moment condition (1.10), We estimate the demand model using GMM. Standard errors of the estimates are calculated according to the formulas provided in section A.1.

1.3.2 Supply side

As discussed earlier, the wholesale price of a drug is determined jointly by its pharmaceutical company and the local government, which typically have distinct objective functions. In particular, we assume that pharmaceutical firms try to maximize their profits while governments concern the welfare of patients and physicians, following the literrature convention [Crawford and Yurukoglu, 2012, Grennan, 2013, Gowrisankaran et al., 2015, Ho and Lee, 2017, Dubois et al., 2019a]. This is a parsimonious characterization of the trade-offs facing policymakers, who should balance producer profits against consumer welfare.

---

[11] We solve the following contraction mapping and obtain ξ_{jt}, whose validity has been proven by Iizuka [2007] and Berry and Jia [2010]:

$$\delta_{jt}^M = \delta_{jt}^{M-1} + (1-\lambda)\left\{\ln s_{jt} - \ln s_{jt}(\delta_{jt}^{M-1}, \theta_2)\right\}$$

where M is the iteration number.

To capture the clear conflict of interests between firms and governments, we model the determination of wholesale prices of drugs using a simultaneous "Nash-in-Nash" bargaining model [Dubois et al., 2019a], in which each drug's wholesale price is negotiated bilaterally between its firm and a local government given the equilibrium prices of other bargain pairs. Following Dubois et al. [2019a], we assume that bargaining takes place at product-by-product level.

In each market t, the profit function of a firm supplying a set of products $\mathcal{F}_t$ is

$$\Pi_{\mathcal{F}_t,t}\left(\mathbf{p}_t^W\right) = N_t \sum_{j\in\mathcal{F}_t} \left(p_{jt}^W - c_{jt}\right) \bar{\sigma}_{jt}\left(\delta_t;\theta_2\right) \tag{1.11}$$

where N_t is the market size of t. Note that we can write the profit function as a function of wholesale price only (given everything else) because the retail price is a fixed function of wholesale price (recall the discussion in Section 1.2.2).

For a given market t, the welfare is defined as the sum of the expected patient-physician joint utility produced by each drug available in market [Small and Rosen, 1981],

$$\Lambda_t\left(\mathbf{p}_t^W\right) = N_t E\left[\ln\left(\sum_{g\in\mathcal{G}_t}\left(\sum_{j\in g}\exp\left\{\frac{\delta_{jt} - \alpha P_{ijt} + \gamma M_{ijt}}{1-\lambda}\right\}\right)^{1-\lambda}\right)\right] \tag{1.12}$$

where the expectation is taken with respect to the heterogeneity in P_{it} and M_{it}.

In each market t, the equilibrium prices solve the Nash-in-Nash bargaining problem

$$\max_{p_{jt}^W}\left\{\left[\Pi_{\mathcal{F}_t,t}\left(\mathbf{p}_t^W\right) - \Pi_{\mathcal{F}_t\setminus\{j\},t}\left(\mathbf{p}_t^W\right)\right]^{\rho_j}\left[\Lambda_{\mathcal{J}_t,t}\left(\mathbf{p}_t^W\right) - \Lambda_{\mathcal{J}_t\setminus\{j\},t}\left(\mathbf{p}_t^W\right)\right]^{1-\rho_j}\right\},\ \forall j \tag{1.13}$$

where $\rho_j \in [0,1]$ represents the relative bargaining power of the firm in the bargaining of product j's price. The firm's objective is the change in profit generated by offering drug j in market t. The government's objective is the change in consumer welfare generated by the presence of drug j in market t. Note that we have assumed the bargaining power parameter of a product does not vary across markets. It's different from Dubois et al. [2019a], who studied seven countries while we only focus on China.

The first order condition of product j in market t is

$$c_{jt} = p_{jt}^W + \frac{1}{\underbrace{\frac{\partial \ln \bar{\sigma}_{jt}\left(\delta_t;\theta_2\right)}{\partial p_{jt}^W}}_{\text{Demand semi-elasticity}} + \frac{1-\rho_j}{\rho_j}\underbrace{\frac{\partial \ln \Lambda_{\mathcal{J},t}\left(\mathbf{p}_t^W\right)}{\partial p_{jt}^W}}_{\text{Welfare semi-elasticity}}}. \tag{1.14}$$

Note that (1.14) collapses to the first order condition of standard Bertrand-Nash equilibrium when ρ_j equals to 1, i.e., when government's preference is not taken into account.

Next, we parameterize the marginal cost as follows:

$$c_{jt} = \left(Z_{jt}^s\right)' \beta + \omega_{jt}, \tag{1.15}$$

where Z_{jt}^s includes a constant, the three cost shifters from Table 1.2, duration since entry, molecule and province-year dummies. Combining (1.14) and (1.15), we estimate the β and $(\rho_1, ..., \rho_J)$ based on the least square criteria, i.e.,

$$\min_{\beta \in R^{k_\beta}, (\rho_1,...,\rho_J) \in [0,1]^J} \sum_{j,t} \omega_{jt}^2. \tag{1.16}$$

Given that β enters the first order condition linearly, We simplify the optimization problem by concentrating out β in close-form

$$\tilde{\omega}_{jt}(\rho_j) = \left[1 - \left(Z_{jt}^s\right)' \left[Z_{jt}^s \left(Z_{jt}^s\right)'\right]^{-1} Z_{jt}^s\right] \tilde{c}_{jt}(\rho_j), \tag{1.17}$$

where

$$\tilde{c}_{jt}(\rho_j) \equiv p_{jt}^W + \frac{1}{\frac{\partial \ln \bar{\sigma}_{jt}(\delta_t;\theta_2)}{\partial p_{jt}^W} + \frac{1-\rho_j}{\rho_j} \frac{\partial \ln \Lambda_{\mathcal{J},t}\left(\mathbf{p}_t^W\right)}{\partial p_{jt}^W}}. \tag{1.18}$$

Then we solve the simplified optimization problem

$$\min_{(\rho_1,...,\rho_J) \in [0,1]^J} \sum_{j,t} \left[\tilde{\omega}_{jt}(\rho_j)\right]^2. \tag{1.19}$$

1.4 Estimation results

1.4.1 Demand estimation results

Demand estimation results are reported in Table 1.5. We can see that the physicians care both patients' and hospitals' interests, since the coefficients on retail price and hospital markup are significant. To make sense of the estimated coefficients, we illustrate how physicians trade off the markup and patients' out-of-pocket cost via a simple example. Suppose that patients on average pay 20% of the cost of medication. Since the coefficient of hospital drug markup is approximately 2.89 times of that (absolute value) of retail price, a patient-physician pair is willing to give up 1 dollar of markup for a reduction of drug price (to a patient) by 58 cents ($\approx 2.89 \times 0.2$). That is, a patient-physician puts a greater weight on patient welfare than hospital profit (derived from drug) unless the coinsurance rate is higher than 35%. This finding resembles Iizuka [2007]'s results on Japanese market, where Japanese physicians are willing to give up 1 dollar if patient's cost is reduced by 28 cents, suggesting that the agency problem of physicians in Japan might be lighter than China.

Other parameters in Table 1.5 also provide some interesting insights. For example, the number of indications significantly increases the demand, and branded drugs are also favored over generic ones. First mover advantage appears to exist in China's prescription drug market as the first generic drug marketed in China in its molecular class has a significantly higher demand. There is an upward trend in the demand for lipid-lowering drugs after entry, but the growth rate drops a little over time. Molecule dummies suggest that the demand for Statins is usually larger than drugs of other therapeutic class, except for Probucol. Public hospital reform seems to negatively impact the market share, while the generic consistency evaluation program may increase the market share (although not significant, probably due to a small sample issue because it's relatively new).

From the estimated demand model, we calculate the price elasticities and summarize them in Table 1.6 and Table 1.7. First, the mean own-price elasticity across products and markets in China in 2018 is -2.84 and ranges from -3.44 to -1.68 across markets. As expected, generics are more elastic than branded drugs (-2.98 versus -2.42), suggesting that even in 2018, after ZMDP is fully implemented, the branded drugs generally have higher market power in China. To see how the price elasticities change over time, Table 1.6 and Table 1.7 also report own- and cross-price elasticities for the main lipid-lowering drugs in China from 2012 to 2018. Own-price elasticity tends to be lower for both branded and generic drugs. Also, lipid-lowering drugs becomes less substitutable as indicated by lowering magnitudes of cross-price elasticities. These drugs are more substitutable within a molecular class (e.g., Atorvastatin produced by Pfizer versus Jialin, or Rosuvastatin produced by AstraZeneca versus Lunan) than between branded and generic groups.

The decreasing price sensitivity might seem a bit surprising given that retail prices are also decreasing, because standard oligopoly theory tells us that they should be inversely related. However, recall that the overall demand becomes much weaker (less prescriptions from physicians) after ZMDP so the market become more competitive, which explains the decreasing price.

Table 1.5: Demand estimation results

	Coef.	St. Err.
# of indications	5.085***	0.519
# of patient contraindications	-0.188*	0.102
# of drug contraindications	-2.266***	0.260
First generic drug	0.251***	0.057
Branded	1.148***	0.063
Time from entry	0.034***	0.007
(Time from entry)^2	-0.000***	0.000
Pilot rate	-0.042***	0.012
Start GCE	0.176	0.240
α	0.439***	0.080
γ	1.268***	0.264
λ	0.668***	0.005
ϕ	0.796*	0.440
Constant	-14.63***	1.100
Molecule dummies		
(Reference: Acipimox, Rosuvastatin, Simvastatin, Xuezhikang)		
Atorvastatin	0.146	0.151
Bezafibrate	-5.784***	0.347
Ezetimibe	-0.161	0.236
Fenofibrate	-7.855***	0.618
Fluvastatin	0.617***	0.210
Gemfibrozil	-11.991***	0.987
Inositol Nicotinate	-2.270	1.601
Lovastatin	-2.384***	0.321
Pitavastatin	1.904***	0.361
Pravastatin	2.082***	0.282
Probucol	6.436***	0.572
Zhibituo	-0.854***	0.258
Observations	23,147	
Objective function value	0.140	

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) See Table 1.2 for variable definitions. (3) α measures the disutility of price, γ measures physician's marginal utility from drug markup (or the severity of expert agency problem) if there is any, λ is the nesting parameter, and ϕ measures the average proportion of type 1 consumers between 2015Q1 and 2017Q3.

Table 1.6: Own-price elasticities for main lipid-lowering drugs, 2012–2018 (China)

	Branded							Generic		
Subclass	Fibrate	Statin	Statin	Statin	Statin	CAI	Statin	Statin	Niacin	Statin
Company	Fournier	Luye	Novartis	Pfizer	AstraZ	SGP	MSD	Jialin	Lunan	Lunan
Molecule	Feno.	Xuezhikang	Fluva.	Atorva.	Rosuva.	Ezetimibe	Simva.	Atorva.	Acipimox	Rosuva.
Drug name	Tcicor	Xuezhikang	Lescol	Lipitor	Crestor	Zetia	Zocor	–	–	–
Year	Estimate									
2012	-9.918	-3.206	-2.478	-7.414	-4.305	-7.473	-3.119	-10.713	-6.373	-9.149
2013	-9.772	-3.419	-2.396	-7.118	-6.506	-10.413	-3.040	-9.850	-4.956	-10.870
2014	-8.450	-3.119	-2.111	-7.114	-5.538	-9.878	-2.932	-9.100	-5.187	-9.038
2015	-6.442	-3.155	-1.790	-5.307	-5.310	-7.290	-2.261	-6.838	-4.598	-8.351
2016	-5.834	-2.855	-1.817	-4.690	-4.740	-6.319	-1.997	-5.315	-4.190	-7.485
2017	-4.612	-2.926	-1.517	-4.372	-4.536	-5.479	-1.638	-5.152	-3.735	-6.447
2018	-3.570	-2.020	-0.704	-2.818	-2.987	-3.267	-1.159	-3.261	-2.268	-3.708

Notes: (1) Each number is the estimated own-price elasticity of demand for the drug defined in the first few rows. (2) Company names: Luye stands for Luye Pharma Group, AstraZ is AstraZeneca, SGP is Schering-Plough, and MSD is Merck Sharp & Dohme. (3) Molecules: Feno. is Fenofibrate, Fluva. is Fluvastatin, Atorva. is Atorvastatin, Rosuva. is Rosuvastatin, and Simva. is Simvastatin. (4) Subclass: CAI stands for Cholesterol absorption inhibitors.

Table 1.7: Average cross-price elasticities among main lipid-lowering drugs, 2012–2018 (China)

	Branded							Generic		
Subclass	Fibrate	Statin	Statin	Statin	Statin	CAI	Statin	Statin	Niacin	Statin
Company	Fournier	Luye	Novartis	Pfizer	AstraZ	SGP	MSD	Jialin	Lunan	Lunan
Molecule	Feno.	Xuezhikang	Fluva.	Atorva.	Rosuva.	Ezetimibe	Simva.	Atorva.	Acipimox	Rosuva.
Drug name	Tcicor	Xuezhikang	Lescol	Lipitor	Crestor	Zetia	Zocor	–	–	–
Year	Estimate									
2012	0.027	0.021	0.048	1.363	0.718	0.009	0.080	0.342	0.013	0.080
2013	0.027	0.023	0.051	1.429	0.924	0.017	0.063	0.314	0.018	0.119
2014	0.022	0.018	0.042	1.300	0.823	0.011	0.051	0.316	0.010	0.112
2015	0.019	0.015	0.033	1.019	0.702	0.012	0.036	0.261	0.008	0.123
2016	0.015	0.014	0.022	0.836	0.597	0.015	0.026	0.244	0.007	0.136
2017	0.012	0.012	0.012	0.796	0.512	0.015	0.019	0.230	0.006	0.152
2018	0.007	0.007	0.007	0.494	0.279	0.011	0.009	0.154	0.004	0.113

Notes: (1) Each number is the average of the estimated cross-price elasticities of demand for the drug defined in the first few rows with respect to (the price changes) of the other drugs. (2) Company names: Luye stands for Luye Pharma Group, AstraZ is AstraZeneca, SGP is Schering-Plough, and MSD is Merck Sharp & Dohme. (3) Molecules: Feno. is Fenofibrate, Fluva. is Fluvastatin, Atorva. is Atorvastatin, Rosuva. is Rosuvastatin, and Simva. is Simvastatin. (4) Subclass: CAI stands for Cholesterol absorption inhibitors.

Table 1.8 shows that since 2015 the total revenue of all drugs keeps decreasing, which could be attributed to the push-out effect. Also, the total revenue of top 10% generic drugs is increasing in the meantime, which may be due to the dethronement effect (i.e., the market power of top generic drugs is increasing).

1.4.2 Supply side estimation

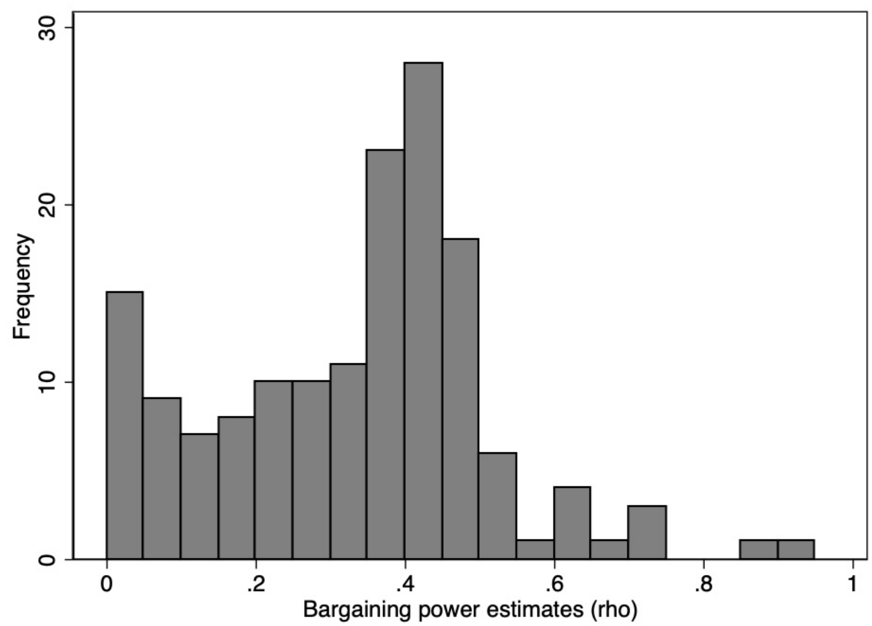

Figure 1.3: Distribution of bargaining parameters

We first present our estimates of bargaining power parameters ρ_j in Figure 1.3. It's not surprising to see that most firms/products have lower bargaining power than the provincial governments (indicated by $\rho_j < 0.5$), and only a small fraction of firms/products show higher bargaining power than the government.

To show the goodness of fit of the bargaining model, we predict the wholesale prices using our estimated marginal cost function $c_{jpt}(\rho_j)$, following Pakes [2017] and Wollmann [2018]. The predicted prices and actual prices are largely centering around a 45-degree line. The linear regression of actual prices on predicted prices without a constant gives a coefficient of 0.998, which is almost 1.

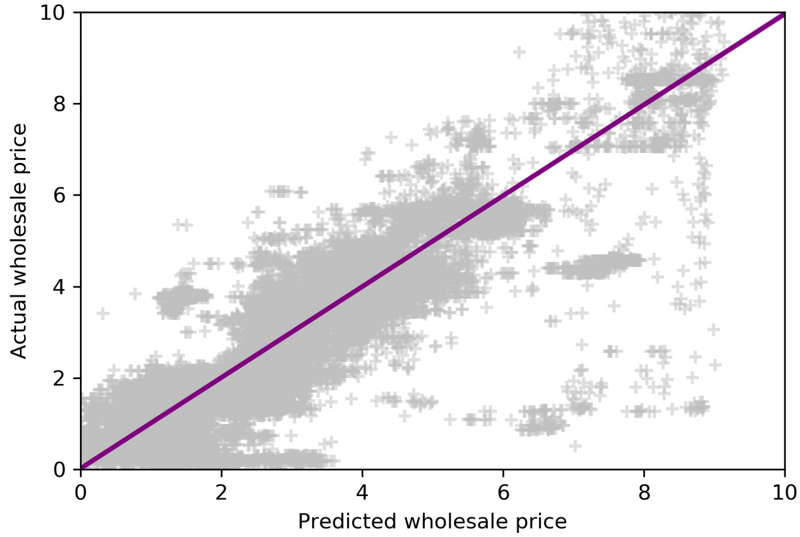

Figure 1.4: Predicted prices versus actual prices

We also look at the predicted price index and compare it with the actual one (like the one in Table A.2). As shown by Figure 1.5, predicted price index is rather close to (although slightly lower than) the actual one and captures the general declining trend over time.

Using our estimated demand parameters, bargaining power parameters, and pricing equilibrium, we can then estimate total revenue and profit of each market. Before showing the total revenues and total profits, we provide the distribution of estimated margins of each product in 2018 in Figure 1.6 and Figure 1.7. The average profit per each standard unit across products and markets (i.e., each observation is weighted by the corresponding amount of standard units sold) in 2018 is 0.63 CNY, ranging from nearly 0 to 2.15 CNY. Profit margin, or price-cost margin (also known as the Lerner index), is 0.27 on average, and most products exhibit a relatively low market power.

We noticed that branded drugs typically have a higher price-cost margin than generic drugs.[12]

[12]Price-cost margin is defined as the difference between wholesale price and marginal cost as a fraction of wholesale

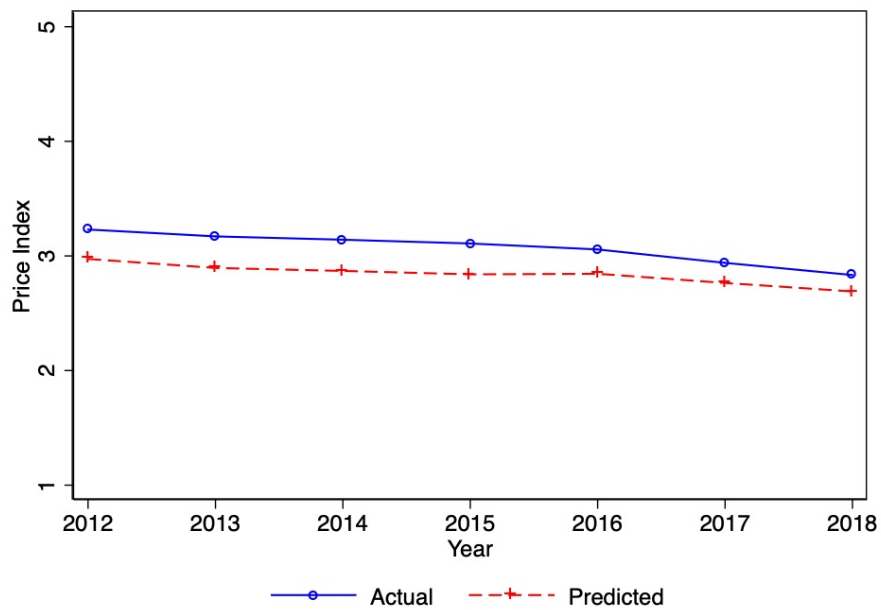

Figure 1.5: Predicted price index versus the actual one

As pointed out by Dubois and Lasio [2018], it is known in the industry that generic firms have lower marginal costs. Our calculations suggest that, in 2018, the (weighted) average cost of a standard unit of generic drugs is 2.48 CNY, compared to 2.91 CNY for branded drugs. The prices of generic drugs, however, are much lower than branded drugs (2.79 versus 3.69), suggesting lower margins of generic drugs.

We summarize the average revenue and profit per market in Table 1.9. In an average market, branded drugs take up the majority (61 percent) of the market share, and the top 10 percent best selling branded products account for 83 percent of the branded market share, indicating high market concentration. The total manufacture revenue of an average market (defined by a season-province pair) in 2018 is 2.46 billion CNY, and the total manufacture profit of a market is 0.45 billion CNY. Due to higher market shares and higher prices, branded drugs are more lucrative. The average revenue of all branded drugs is 110 percent higher than that of all generic drugs, while total profit is 321 percent higher.

price. Weighted averages are 0.28 versus 0.19.

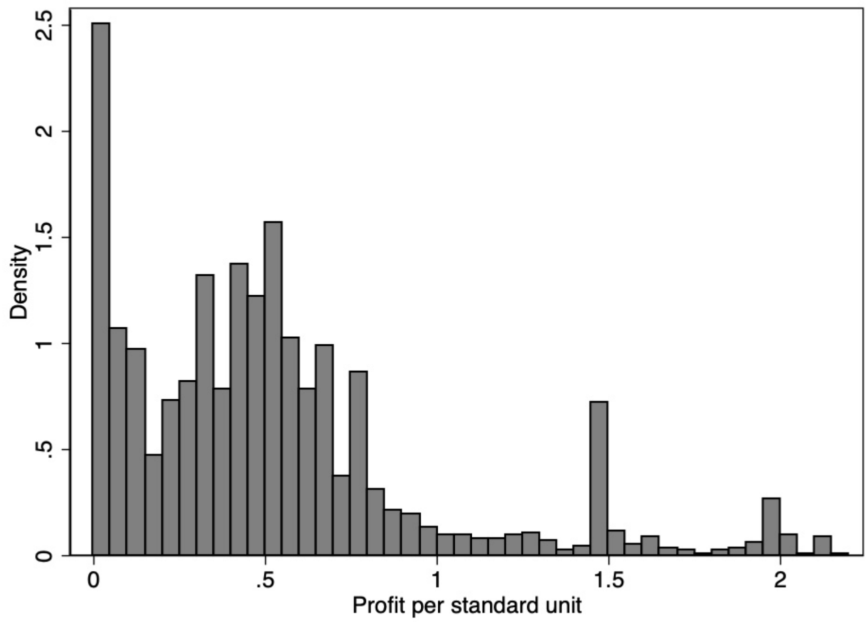

Figure 1.6: Estimated profit per standard unit in 2018

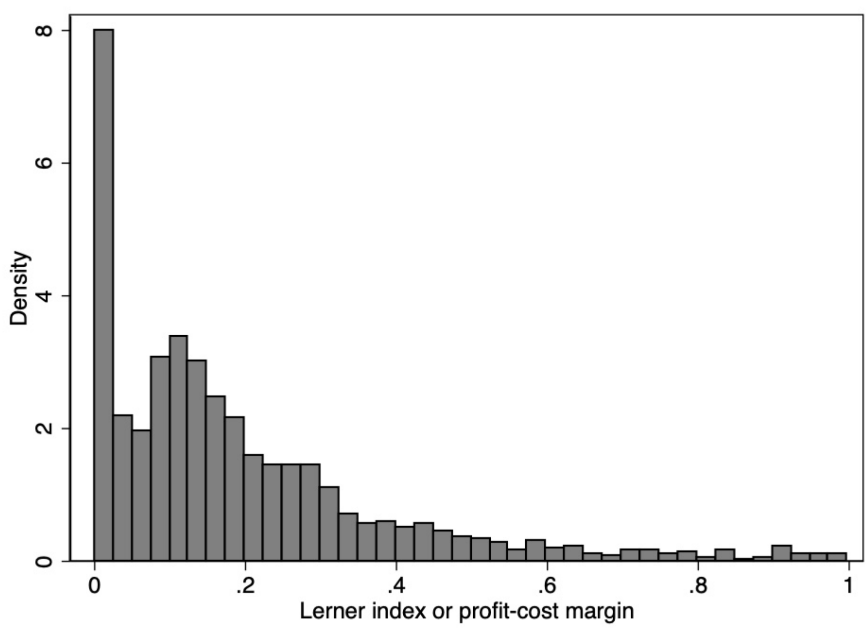

Figure 1.7: Lerner index in 2018

Table 1.8: Revenue per market in 2012-2018 (China)

Year		All firms	Bottom 90%	Top 10 %
	All drugs	24.62	6.17	18.45
2012	Branded	18.11	2.48	15.63
	Generic	6.51	3.69	2.82
	All drugs	25.20	6.02	19.18
2013	Branded	18.44	2.43	16.01
	Generic	6.76	3.59	3.17
	All drugs	25.42	5.72	19.70
2014	Branded	18.55	2.24	16.31
	Generic	6.87	3.48	3.39
	All drugs	25.31	5.89	19.42
2015	Branded	17.60	2.17	15.43
	Generic	8.05	4.06	3.99
	All drugs	25.33	6.18	19.15
2016	Branded	17.28	2.28	15.00
	Generic	8.05	3.9	4.15
	All drugs	25.03	5.87	19.16
2017	Branded	17.09	2.28	14.81
	Generic	7.94	3.59	4.35
	All drugs	24.59	5.70	18.89
2018	Branded	16.67	2.25	14.41
	Generic	7.92	3.44	4.48

Notes: (1) Market is defined by a specific quarter of a year in a province in China. (2) Revenue is sample estimation, which is just 20-30% of the real-world values. (3) Revenue is in 100 million CNY.

Table 1.9: Market Share, Revenue and Profit per market in 2018 (China)

		All firms	Bottom 90%	Top 10 %
	All drugs	38.99	11.97	27.02
Market share (%)	Branded	23.73	3.93	19.80
	Generic	15.26	8.04	7.22
	All drugs	24.59	5.70	18.89
Revenue	Branded	16.67	2.25	14.41
	Generic	7.92	3.44	4.48
	All drugs	4.47	0.87	3.60
Profit	Branded	3.61	0.27	3.34
	Generic	0.86	0.60	0.26

Notes: (1) Market is defined by a specific quarter of a year in a province in China. (2) Revenue and profit are in 100 million CNY.

1.5 Counterfactual: quantifying the effects of ZMDP

In this section, we examine how profit and consumer surplus were affected by ZMDP that breaks the integration between prescribing and dispensing drugs [Iizuka, 2007]. To avoid the complication of price caps that were in place during the transition periods, we conduct the counterfactual simulation based on on the data of the post-reform era, i.e., 2018. Specifically, we assume the absence of ZMDP such that pre-reform hospital markup, i.e., 15 percent of wholesale price, is restored. Then, we calculate counterfactual equilibrium prices, market shares, profits, etc., using the estimates and data of 2018 . [13]

Figure 1.8 compares the counterfactual retail prices to the actual prices, showing that the distribution of counterfactual prices shifts to the right, i.e., if the 15% hospital drug markup still existed, the average retail price would be higher. The profit from selling a standard unit of lipid lowering drug (defined by the difference between the retail price and the marginal cost) is shown

[13] We solve for new equilibrium prices using firms' first-order conditions. A fixed point algorithm was used to solve the system with a numerical tolerance level smaller than 10^{-6}.

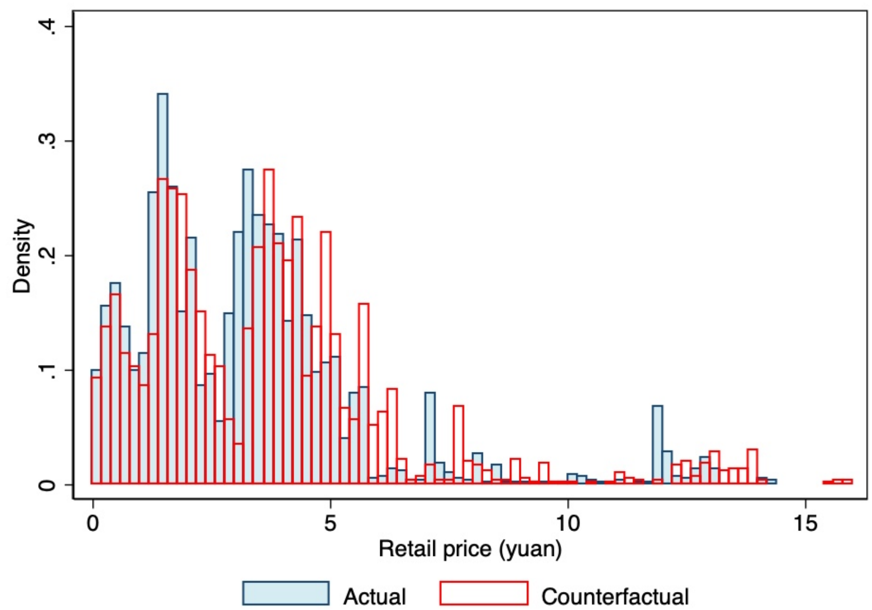

Figure 1.8: Counterfactual versus actual retail prices in 2018

in Figure 1.9. Again, it is clear that hospital drug markup would generate higher profits.

Based on the counterfactual market equilibrium, we calculate the implied market shares, revenues, and profits that are summarized in Table 1.10. The share of an average market could expand to nearly 51 percent of the hyperlipidemia population with the hypothetical 15 percent hospital drug markup in 2018. Compared to the estimates in Table 1.9, branded drugs could experience a much larger increase in market share, and become even more concentrated; this again suggests that, without ZMDP, top selling branded drugs are more preferred by physicians.

Comparing to actual data, the total revenue under the counterfactual scenario goes up by 37.5 percent, and the total profit increases by 8 percent. Generic drugs would lose some profit if ZMDP were removed in 2018. This is consistent with our conjecture – without the drug markup, the dethronement effect should lead to a higher relative market power of generic drugs compared to the branded drugs. Note that this result is also consistent with the estimated revenues in Table 1.8. Finally, we measure the changes in patients' welfare due to the counterfactual drug markup. This is done by assuming that the utility function fully represents patients' preference (so γ is fixed at

Table 1.10: Counterfactual share, profit, revenue, and surplus per market (2018)

		All firms	Bottom 90%	Top 10 %
Market share (%)	All drugs	50.58	13.52	37.06
	Branded	34.57	4.69	29.88
	Generic	16.01	8.83	7.18
Revenue	All drugs	33.81	7.53	26.28
	Branded	24.29	2.93	21.36
	Generic	9.52	4.61	4.92
Profit	All drugs	4.83	0.82	4.01
	Branded	4.09	0.26	3.83
	Generic	0.74	0.56	0.18
Patient surplus change			-12.24%	

Notes: (1) We use 2018 product attributes for all counterfactual exercises. The counterfactuals assume that there is a 15% drug markup just like 2012-2014. (2) Other cautions are in Table 1.9. (3) Revenue and profit are in 100 million CNY.

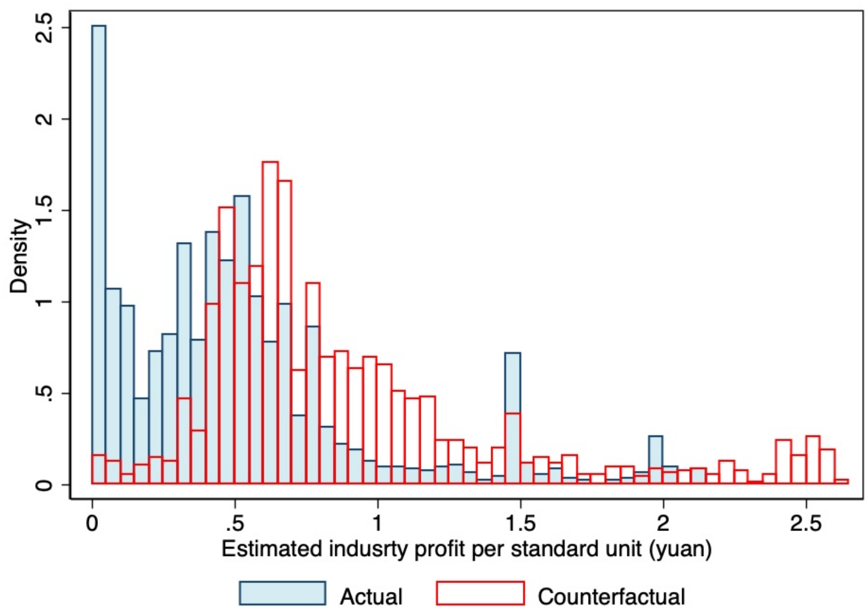

Figure 1.9: Counterfactual versus actual industry profit per unit in 2018

0). Our calculation suggests that, patient's welfare would drop by 12.24 percent if there were a 15 percent hospital drug markup. This results suggest that the ZMDP is overall beneficial to patients.

1.6 Concluding remarks

In this paper, we develop a structural model of China's prescription drug market to investigate the impact of the ZMDP. Using the data from PDB and various sources (e.g., *menet*, *yaozh*, package inserts, policy documents, etc.) on wholesale transactions of lipid-lowering drugs in a sample of hospitals during 2012–2018, we first estimate a mixed nested-logit demand model that accommodates the consumer heterogeneity due to zero-markup drug policy pilot programs. The demand estimation suggests that lipid-lowering drugs are highly differentiated. Brand-name drugs are preferred to their generic versions, which is in line with the literature. Moreover, physicians' prescription decisions are affected by the hospital markups, although they care more about patient welfare and choose drugs that have less out-of-pocket costs, unless coinsurance rate is high.

Under the assumption that prices are set according to Nash bargaining between each firm and the corresponding provincial government in China, we separately identify costs and bargaining parameters, the latter of which can be interpreted as the degree to which the government leaders choose to trade off between firm profits and immediate consumer welfare. Results suggest that most policymakers value consumer welfare more, and thus firms typically have a bargaining power parameter that is less than 0.5.

We then perform a counterfactual analysis by removing the ZMDP in 2018 and then quantify its impact on firms' profitability and patients' welfare. Our calculations indicate that, ZMDP leads to an increase in patients' welfare by about 12 percent. Moreover, ZMDP benefits generic drugs more than branded ones and make the branded drugs less concentrated. Overall, this counterfactual exercise confirms that ZMDP largely achieves its policy goal of reducing drug prices and increasing patients' welfare.

Finally, we close our paper by mentioning some caveats and limitations of our work left for future research. First, our results are based on a static model and does not include dynamic considerations, such as investment, entry and exit, etc., so they only reflect a short-term evaluation of the policy effects. Also, we only focus on lipid-lowering drugs in a selected sample of hospitals in this paper, a more comprehensive investigation that covers more hospitals and types of drugs would be helpful to understand ZMDP to a greater extent.

Chapter 2

TRAINING, PRODUCTIVITY, AND WAGES: AN INVESTIGATION OFCHINAS MANUFACTURING ENTERPRISES IN A PRIVATIZATION ERA

2.1 Introduction

The accumulation of human capital is generally believed to be the driver of economic growth, especially in developing and transition economies. As a result, economists have studied human capital investment intensively. However, most scholars focus on the school system and emphasize the importance of education attainment, but work-related training provided by firms is less explored in the literature due to at least two reasons: (i) the limited availability of detailed firm-level data, and (ii) the potential bias in econometric estimation due to the unobserved firm heterogeneity and the endogeneity problems [Konings and Vanormelingen, 2015, Liu and Lu, 2016]. Three important questions waiting for more discussions in the context of economic transition are: (i) is on-the-job training important in increasing productivity, (ii) would training increase labor costs, and (iii) why would firms provide training to their workers voluntarily? Our paper tries to address all three of them amid a massive privatization of labor market.

We begin with jointly investigating the first two questions. According to Becker [1964]'s human capital theory, wages can be a proxy for productivity, and education or training can increase both.[1] However, according to the theory of "compensating wage differentials" conceived by Adam Smith, training can be a "pleasant" characteristic of a job and compensate for lower wages.[2] As a result, it's theoretically unclear whether training can increase wages or not. Furthermore, training may be a signaling tool (of firms to attract workers) in the labor market and is not necessarily productive. Thus, whether training can increase productivity is an empirical question. Empirical

[1] In a perfect and complete labor market where laborers flow freely, if education or training can increase labor productivity and if worker contribution is only rewarded by wages, firms have to raise wages to attract/keep better educated or trained employees.

[2] Both theory and empirical studies suggest that an increase in minimum wages can lead to lower on-the-job training [Hashimoto, 1982, Ho and Lee, 2017]. That is, training and wages can be complementary.

studies typically are limited to using industry-level data [Conti, 2005, Dearden et al., 2006] and mostly are done in developed countries [Konings and Vanormelingen, 2015]. In developing countries, there are quite few empirical studies of the two questions using firm-level data. To the best of our knowledge, this is among the first few attempts to use a large firm-level data of a transition economy to estimate the productivity and wage effects of training.

Then, our answer to the third question is related to the first two. Acemoglu and Pischke [1998] suggest that it is exactly the positive gap between the increasing rate of productivity and that of wages (costs) brought by employee training (i.e., a positive training premium) that gives firms a motivation to offer training. In most countries, employee training is not mandated by the government and it's up to the firms to decide whether and how to invest in training. Therefore, it's important to verify and understand this incentive.

This paper contributes to the literature in a few ways. First, we provide a new perspective of economic transition—the transition of training decisions of firms. We explain why privatization can raise productivity, as suggested by Brown et al. [2006] and Estrin et al. [2009], by investigating the relationship between productivity, wages and training amid a massive privatization of labor market. Our results suggest that, the gap between productivity and wage increases could be widen by training amid privatization, and can incentivize firms to further invest in the human capital, which in turn can raise productivity more. The combination of the research of training and the transition context can generate new insights and policy implications.

Second, China is an interesting case for the research of training and economic transition. China, as one of the largest and fastest-growing developing countries, has been experiencing massive and rapid changes since its market reforms. According to the World Bank, China's average annual GDP growth rate from 1978 to 2018 is 9.5%. In the early 1990s, the Chinese central government started to introduce far-ranging reforms into the state-owned sector, transforming the stated-owned enterprises into corporations. In 1998, the government started to "grasp the large and let go the small", and according to our data (see Table B1), it is proven successful. The number of state-controlled manufacturing enterprises fell by over a half within just 5 years after its initiation. In the 2003 Third Plenum, the Chinese government started to call for "putting people first" and "establishing a human capabilities-based perspective on growth". In the eleventh "5-year plan", the Chinese government again emphasized that "accelerating the development of education was the basic path to convert the enormous pressure of population in China into the comparative advantage of abundant human resources". As suggested by Fleisher et al. [2005], China experienced a slow increase in returns to schooling during transition, compared to Central and Eastern European countries and Russia.[3] Thus, it would be interesting to know how training contributed to productivity and if there was any return to training amid China's transition. As argued by Liu and Lu [2016], since the Chinese education (especially professional education) system was not that effective and qualified skilled workers were unabundant, training was expected to play critical roles in the process of China's structural change. However, it is surprising that there is little research on its effects in China, and we will complement the literature by doing so. The lesson from China can be applicable to other transition economies and developing countries.

Third, we modify the most recent production function estimation framework by Ackerberg et al. [2015] to estimate the effect of training and total factor productivity (TFP) jointly in a single structural equation, which is a novel and efficient way of studying the effect of training. The only significant work similar to ours is Konings and Vanormelingen [2015]. In the previous literature, most studies adopt a two-step procedure by estimating TFP first and then study the effect of training on TFP such as Liu and Lu [2016]. In our estimation, we deal with the endogeneity of training by using intermediate goods (raw materials) as a proxy for time-varying firm-specific productivity shocks, and we introduce heterogeneity into the structural model, which can help us further understand how the relationship varies with context and population of interest.

Last but not least, we contribute to the literature by using a large firm-level data set with nearly a million observations (compared to most existing studies that used industry-level data with only a thousand observations). The scalability of firm-level data enables us to get more accurate estimations.

The remaining of the paper is structured as follows. Section 2.2 briefly introduces the empirical framework of the paper. Section 2.3 discusses the identification strategy and estimation procedures. Sections 2.4 and 2.5 describe data and discuss results, respectively. Section 2.6 concludes the paper with policy implications.

---

[3] For example, Münich et al. [2005] found drastic increases in returns to education in the Czech Republic.

2.2 Empirical Framework

To estimate the impacts of training on productivity and wages simultaneously, we adopt Hellerstein et al. [1999]'s framework that has been widely applied to compare returns to characteristics of workers such as gender and race on both productivity and wages. We extend this framework a la Konings and Vanormelingen [2015] by allowing for continuous labor characteristics. The following subsections briefly outline the empirical model, as well as the endogeneity and heterogeneity issues.

2.2.1 Training Affects Marginal Product of Labor

Following Konings and Vanormelingen [2015], we assume a Cobb-Douglas production function for firm i in period t:

$$Y_{it} = A_{it}\hat{L}_{it}^{\beta_l} K_{it}^{\beta_k} e^{\mu_{it}} \tag{2.1}$$

where Y_{it} is the total output of firm i in period t, $\hat{L}_{it}$ represents the effective labor endowment, K_{it} is capital endowment, A_{it} is the total factor productivity (TFP) that represents technical efficiency and assumed to be predictable (or even observed) by firms when they make input decisions, μ_{it} represents shocks to productivity or production that are not predictable by firms. Note that, both A_{it} and μ_{it} are not directly observed by econometricians.

Workers are assumed to be distinguished by their skill level, and it enters the effective labor input as in a Mincer [1974] wage equation, then as proven by Konings and Vanormelingen [2015], $\hat{L}_{it}$ satisfies:

$$\ln(\hat{L}_{it}) = \ln(L_{it}) + \phi_T C_{it} + u_{it} \tag{2.2}$$

where L_{it} is the number of employees, C_{it} is the average training cost (represents intensity), ϕ_T is the percentage increase in marginal product of labor brought by training, and u_{it} represents the "unobserved" labor quality (e.g., health conditions, inter-personal skills, etc)—this is assumed to be partially predictable (or observed) by firms (e.g., during a job interview the firm can observe a candidate's inter-personal skills), but not observed by econometricians. Thus, the term "unobserved" is from the perspective of a researcher instead of a firm. To proceed with estimation, we take the natural logs of (1) (and use lowercases to denote logs), plugging in equation (2) to get the following estimation equation:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_l \phi_T C_{it} + \beta_k k_{it} + \rho_{it} + \mu_{it} \tag{2.3}$$

where ρ_{it} includes both the technological progress term A_{it} and the unobserved labor quality term u_{it}:

$$\rho_{it} = log(A_{it}) + \beta_l u_{it} := f_t^{-1}(m_{it}, l_{it}, c_{it}, k_{it}) \tag{2.4}$$

Following Konings and Vanormelingen [2015], we assume ρ_{it} is a monotonically increasing inverse material demand function. We assume labor input, capital and training to be set before the choice of material input. The main parameter of interest in (3) is ϕ_T, which reflects the effect of training on labor productivity (termed "training productivity effect").

2.2.2 Training Affects Average Wages

As proven by Konings and Vanormelingen [2015], under this framework, the average wage paid by firm i in period t has the following form:

$$\ln(W_{it}) = \ln(W_0) + \lambda_T C_{it} + \alpha_\rho \rho_{it} \tag{2.5}$$

where W_0 refers to the average "base" wage level. We are interested in the parameter λ_T, which measures how wages change in response to training. Note that, wages are affected by unobserved labor quality, and α_ρ represents the return to the labor quality. We can rewrite (5) as:

$$w_{it} = w_0 + \lambda_T C_{it} + \alpha_\rho \rho_{it} \tag{2.6}$$

Similar to Hellerstein et al. [1999], we add industry (indexed by j) and year specific effects to the estimation equation as well as observed firm characteristics X_{it} such as the capital-labor ratio and an additive i.i.d. error term ϵ_{it}:

$$w_{it} = w_0 + \lambda_T C_{it} + X_{it}\gamma + \alpha_\rho \rho_{it} + \chi_j + \chi_t + \epsilon_{it} \tag{2.7}$$

2.2.3 Endogeneity and Heterogeneity

Due to the fact that the components ρ (technological progress and labor quality) in (3) and (6) are predictable by firms but not observed by econometricians, and the fact that the firm's profit-maximizing choice variables (e.g., training) can correlate with these components, we cannot apply OLS directly to obtain consistent estimates of the parameters. If we apply OLS by ignoring those unobserved components, they will enter the error term and render the estimates of productivity and wage effects inconsistent. Such endogeneity issue is dicussed as early as Marschak and Andrews Jr. [1944] in the literature. We will show how we can obtain consistent estimates for the parameters using a control function approach in Section 2.3.

To address heterogeneity across firms (for example, to see how the effects vary by firm ownership, union representation, etc.), we estimate (3) and (6) simultaneously in subgroups of firms separately. To exploit the heterogeneous effects within a firm (termed "worker heterogeneity"), we decompose the effects based on the proportions of different types of workers. For example, training can affect male and female workers differently. If we assume that training resources are distributed evenly across different genders within a firm,[4] the effects can be decomposed using the following formulae:

$$\phi_T = \phi_{T,\text{ type 1}} \times \text{ratio of type 1} + \phi_{T,\text{ type 2}} \times \text{ratio of type 2}$$

$$\lambda_T = \lambda_{T,\text{ type 1}} \times \text{ratio of type 1} + \lambda_{T,\text{ type 2}} \times \text{ ratio of type 2}$$

Similarly, we can decompose the effects based on education and professional levels of workers within a firm. Potential bias will be discussed in the result section.

2.3 Estimation

Identifying the differential effect of training on both productivity and wages requires consistent estimation of $\phi_T = \frac{\beta_l \phi_T}{\beta_l}$ in (3) and λ_T in (6). In this section, we describe how we are able to do so by estimating a production function using the procedure of Ackerberg et al. [2015] (ACF in short).

The key identification challenge is ρ_{it}, which represents a firm's TFP and labor quality. To control for this, we rely on a few assumptions. First, we assume that the capital stock satisfies a

[4]If more training resources are allocated to male workers, then the effects on female workers will be underestimated. We will take into account this potential bias when discussing our results.

dynamic investment process. That is, investment in period $t-1$ becomes capital stock in period t. Then, we assume ρ_{it} to follow a first order Markov process, which is standard in the literature:

$$\rho_{it} = E[\rho_{it}|I_{i,t-1}] + \xi_{it} = E[\rho_{it}|\rho_{i,t-1}] + \xi_{it} = g(\rho_{i,t-1}) + \xi_{it} \tag{2.8}$$

where $I_{i,t-1}$ refers to all the information a firm can have access to in period $t-1$, and ξ_{it} reflects the changes of technical progress and labor quality that will affect the firm's choices in period t. According to the Markov process, a firm's expected ρ_{it} in period t only depends on information in period $t-1$ and before. As k_{it} is pre-determined in period $t-1$, it can be supposed that $E[\xi_{it}|k_{it}] = 0$. Third, similar to Konings and Vanormelingen [2015], we further assume $E[\xi_{it}|l_{it}] = 0$ and $E[\xi_{it}|c_{it}] = 0$. That is, recruitment and training plans are decided when firms made financial budget in period $t-1$. If there is any labor/training adjustment within period t, we assume that it happens before intermediate goods are purchased. Our estimation of the production function can then be divided into the following two stages.

In the first stage, we rely on the firm's demand for raw materials (intermediate goods) in period t, which according to the assumptions above can be expressed as a function of the firm's TFP and unobserved labor quality, capital stock, labor input, and training cost in period t. Variables in previous periods (e.g., $t-1$) have no influence on the demand for raw materials. That is to say:

$$m_{it} = f_t(\rho_{it}, l_{it}, c_{it}, k_{it}) \tag{2.9}$$

If m_{it} is a monotonic and increasing function of ρ_{it}, conditional on capital, labor and training, we can then invert the demand function to express ρ_{it} as:

$$\rho_{it} = f_t^{-1}(m_{it}, l_{it}, c_{it}, k_{it}) \tag{2.10}$$

Combining Equations (9) and (3), we can get:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_l \phi_T c_{it} + \beta_k k_{it} + f_t^{-1}(m_{it}, l_{it}, c_{it}, k_{it}) + \mu_{it} \tag{2.11}$$

In this stage we cannot identify all coefficients, since l_{it}, c_{it} and k_{it} are included in the inverse function $f^{-1}(\cdot)$ as well. Nevertheless, the level of the whole Equation (10) can be estimated non-parametrically. We will use cubic polynomials of m_{it}, l_{it}, c_{it} and k_{it}. By the end of the first stage, we obtain an estimate of:

$$\Theta_{it} = y_{it} - \mu_{it} = \beta_l l_{it} + \beta_l \phi_T c_{it} + \beta_k k_{it} + \rho_{it} \tag{2.12}$$

Notice that, if we have an estimate of Θ_{it}, then ρ_{it} can be estimated given the estimates of coefficients in Equation (11), i.e.,

$$\rho_{it}(\beta_l, \beta_K, \phi_T) = \Theta_{it} - (\beta_l l_{it} + \beta_l \phi_T c_{it} + \beta_k k_{it}) \tag{2.13}$$

Therefore, in the second stage, we choose a set of candidate coefficient vectors ($\tilde{\beta}_l$, $\tilde{\beta}_K$, $\tilde{\phi}_T$) and get $\tilde{\rho}_{it}$ using Equation (12). According to Equation (7), we can regress $\tilde{\rho}_{it}$ on $\tilde{\rho}_{i,t-1}$ to obtain estimates of $\tilde{\xi}_{it}$. At the same time, we evaluate the sample analogue of the moment conditions following Equation (7):

$$\frac{1}{T}\frac{1}{N}\Sigma_t \Sigma_i \tilde{\xi}_{it} \begin{pmatrix} k_{it} \\ l_{it} \\ c_{it} \end{pmatrix} = 0 \tag{2.14}$$

After non-linear search using a computer, we can obtain the GMM estimators ($\hat{\beta}_l$,$\hat{\beta}_K$,$\hat{\phi}_T$), and hence $\hat{\rho}_{it}$. This completes the estimation of the production function.

After estimating the production function based on the above ACF procedure, We can then use the estimates in Equation (3) to help with the consistent estimation of Equation (6), in which we're interested in the effect of training on wages, λ_T.

$$w_{it} = w_0 + \lambda_T c_{it} + X_{it}\gamma + \gamma_\rho \hat{\rho}_{it} + \chi_j + \chi_t + \epsilon_{it} \tag{2.15}$$

where $\hat{\rho}_{it}$ is estimated from production Equation (3). If $\hat{\rho}_{it}$ includes factors other than labor quality, then the estimated λ_T can be upward biased. We will discuss the potential impact on interpreting the results and thus our conclusions.

To confirm whether $\hat{\phi}_T$ is significantly larger than $\hat{\lambda}_T$ in terms of their value, the standard error of each coefficient is calculated by applying block bootstrap. We also do a Chow test to investigate if there is heterogeneity in $\hat{\phi}_T - \hat{\lambda}_T$.

2.4 Data Source and Descriptive Statistics

To answer our research question, we resort to the Chinese Industrial Enterprises Database (CIED). Although the database spans from 1998 to 2013 (see Table B1), information about training is only available in 2004–2007, which happens to be the privatization period we are interested in.

This database has been exploited by well-known scholars such as Hsieh and Klenow [2009], Song et al. [2011], and Brandt et al. [2017]. The CIED is constructed by China's National Bureau of Statistics and the data sources are mainly the annual or quarterly reports submitted to local bureau of statistics. The database contains all industrial enterprises that are "large-scale"[5] and non-state-owned or state-owned. In the database, about 90% of the enterprises are manufacturing firms, which are what we're going to focus on in this research.

Table 2.1: Descriptive Statistics of Main Variables

	Total	2004	2005	2006	2007
Gross output (thousand yuan)	126,635	–	103,428	124,682	151,942
Wages or salary (thousand yuan)	15.80	12.82	14.45	16.44	19.49
Training cost (thousand yuan)	0.135	0.114	0.126	0.144	0.155
Training rate (%)	43.60	44.87	42.40	43.53	43.64
Male employees (%)	63.13	72.61	58.93	60.00	61.07
Well-educated employees in 2004 (%)	11.53	11.53	11.56	11.53	11.51
Highly-skilled employees in 2004 (%)	30.10	30.10	30.10	30.09	30.10
Union rate in 2004 (%)	46.94	47.02	46.89	46.91	46.94
SOEs rate (%)	3.40	3.68	3.49	3.25	3.20
Number of employees	272	253	272	279	283
Capital-labor ratio	90.37	86.07	79.95	93.51	101.98
Observations	674,106	167,642	169,456	168,594	168,414

Notes: (1) SOEs = state-owned enterprises. (2) Well-educated employees are those with at least 15 years of education (e.g., college, graduate); highly-skilled employees are those with at least intermediate professional titles. (3) Total output is not available in 2004, while well-educated employees, highly-skilled employees, and union establishment status are not updated after 2004.

Between 2004 and 2007, around 168 thousand firms met our inclusion criteria (manufacturing firms established before 2004 that did not go out of business before 2007 and had reported information we need) every year, which gives us a total of 674 thousand observations to construct an unbalanced panel. The ACF procedure makes use of the 2005-2007 panel, which includes about

[5] That is, the main business income of an enterprise was larger than 5 million RMB, and this standard was revised to 20 million RMB in 2011.

506 observations. The descriptive statistics of the main variables of interest are shown in Table 2.1. The gross output of manufacturing firms grew at an annual rate of around 21% between 2005 and 2007, which was faster than other industries on average at the time in China. The average size of a manufacturing firm grew modestly, while capital grew faster, showing the transition to a more capital-intensive economy. Wages increased at an increasing rate from about 13% to about 19% during 2004-2007, but it was slower than the increasing rate of gross output. Training costs were about 0.8% of employees' annual income, and less than half of the firms actually provided training.[6] There were more male workers in these manufacturing firms, among which only around 3.4% were state-owned and 47% had a union representation in 2004. In 2004, only 11.5% of the workers in the manufacturing industry had a college degree or above, and 30.1% of the workers had an intermediate or advanced professional titles. These facts show that training could potentially play an important role in a firm's production decision amid privatization.

To get a better idea of the trends in data before doing any estimation, we plot the average log deflated gross output (in million RMB yuan) and average log wage per worker-year (in thousand RMB yuan) against different levels of training cost per worker-year (in RMB yuan) in Figure 2.1. Note that, this is an illustrative graph and is not trying to replicate any regression or estimation hereinafter. From this descriptive figure, we can see that both output levels and wages increase with training spending in general (at an average rate of 12% and 5% per 1 yuan spent on each worker each year, respectively). In the next section, we will resort to the empirical strategy discussed above to investigate the effects of training on output, labor productivity, and wages. Specifically, we will explore if there is a gap between the effects on labor productivity and wages.

2.5 Results

2.5.1 Overall Effects

We begin by taking a look at the overall effects of training on productivity and wages. Table 2.2 shows the results of estimating Equations (3) and (6) for all active firms. The first column reports the results by applying ordinary least squares (OLS), and the second column presents the coefficient

[6] Due to the limit of our data, we are not able to distinguish general training from firm-specific training and thus we treat it as general training.

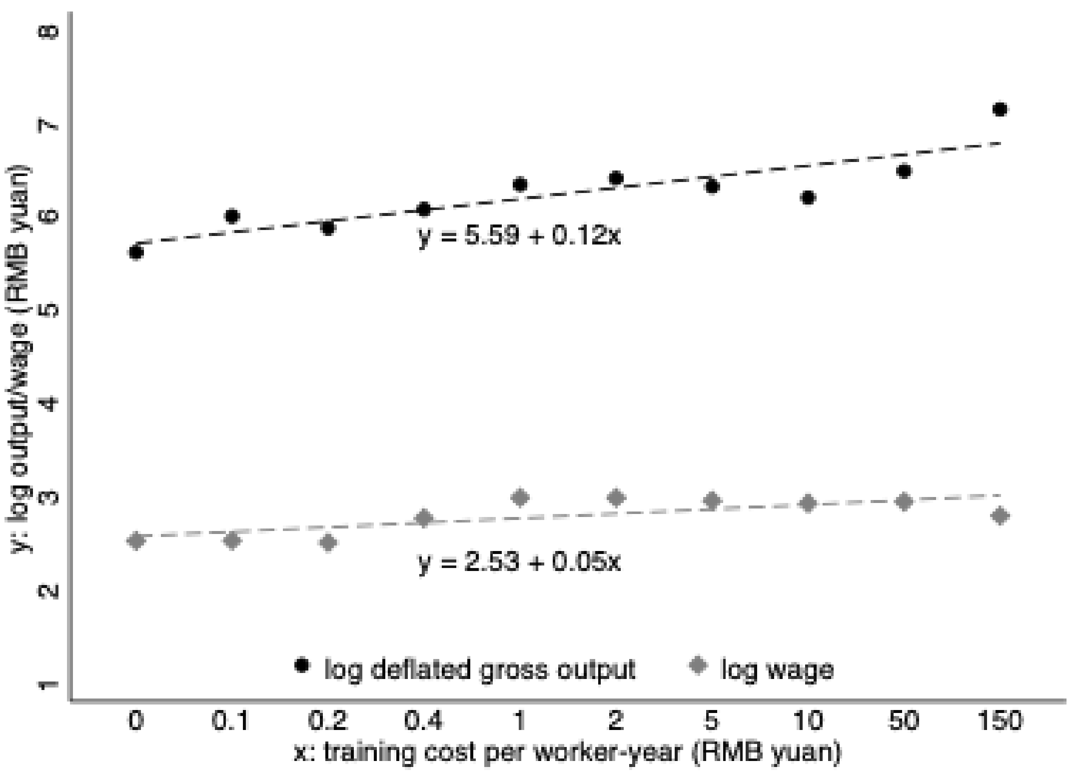

Figure 2.1: Gross Output, Wage, and Their Relations with Training

Notes: (1) Each scatter represents the average log output/wage level between two adjacent training levels (not equally spaced as the distribution is highly right-skewed). (2) Each dashed line is a linear fit of the corresponding group of scatters.

estimates obtained by following Section 2.3, where we control for the endogeneity of input and training choices of firms by a control function approach using information on raw materials.

For the production function, the OLS results suggest that training has a positive and statistically significant effect on productivity. Specifically, a 1,000 RMB (about 121 US$ in 2004) increase in the average training cost per worker-year is associated with a 8.9% rise in gross output. The ACF approach that controls for the endogeneity of inputs and training lowers the coefficient estimates significantly (indicated by the Chow test), and implies that a 1,000 RMB increase in the average training cost per worker-year can lead to a 4.2% rise in gross output. This is still a pretty decent effect. Note that, the marginal product of labor increases by 9.4% (ϕ_T =0.042/0.447) in response to a 1,000 RMB increase in the average training cost per worker-year.[7]

For the wage equation, the natural log of average wage is regressed on the average training cost together with year and industry dummies. The coefficient on training cost drops from 0.067 to 0.043 once we control for endogeneity by including the estimated ρ (technological progress and labor quality) in the ACF model, suggesting that the average wage can increase by 4.3% in response to a 1,000 RMB increase in the firm's training investment per worker-year. It's interesting that the "compensating wage differentials" theory (predicting a negative effect) is outweighed by the human capital theory (predicting a positive effect) in overall during China's transition to a more privatized market economy. This result compensates Münich et al. [2005] and Fleisher et al. [2005]'s findings by suggesting that returns to on-the-job training could outweigh its compensating effect during transition in China and thus could be pretty significant.

From Table 2.2, we can clearly see that, in general, while the labor productivity effect and the wage effect of training are both positive, the former is significantly larger than the latter, and the difference is statistically significant (indicated by the Wald test of the equality of ϕ_T and λ_T). The labor productivity effect is twice as high as the wage effect for the same increase in a firm's training spending, which supports Acemoglu and Pischke [1998]'s hypothesis that explains why firms would provide general training to employees willingly.

[7] We do not observe which workers are trained within a firm, so this effect includes both the direct effect on trained workers and indirect (spillover) effect on untrained workers within the same firm. There might also be spillover effects between firms and our measure includes those—nevertheless, they are not expected to be the major part of our effect.

Table 2.2: Overall Effects of Training on Gross Output and Wages

	(1) OLS	(2) ACF	(3) Chow Test: (1) = (2)
Production Function			
Training ($\beta_l \phi_T$)	0.089***	0.042***	–
	(0.006)	(0.005)	
$ln(L)$ (β_l)	0.501***	0.447***	–
	(0.002)	(0.010)	
$ln(K)(\beta_k)$	0.290***	0.261***	–
	(0.001)	(0.005)	
Wage Equation			
Training (λ_T)	0.067***	0.043***	158.51***
	(0.005)	(0.004)	[0.000]
$ln(K/L)$ (γ)	–	0.080***	–
		(0.001)	
TFP (γ_ρ)	–	0.167***	–
		(0.001)	
ϕ_T	0.179***	0.094***	41.44***
	(0.011)	(0.011)	[0.000]
$\phi_T - \lambda_T$	0.111***	0.050***	28.71***
	(0.007)	(0.011)	[0.000]
Wald Test: $\phi_T = \lambda_T$			
χ^2	283.00***	19.26***	–
p-value	0.000	0.000	
Observations	506,464	506,464	506,464
Firms	171,717	171,717	171,717

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) In ACF, we control for the endogeneity of input and training choices. (3) Standard errors in parentheses are computed using a block bootstrap procedure and are robust against heteroskedasticity and intragroup correlation, and p-values are in square brackets. (4) ϕ_T is computed as the ratio of the training coefficient over the labor coefficient. (5) Chow tests are achieved by jointly estimating the OLS and ACF models in a single procedure.

2.5.2 *Firm Heterogeneity*

In this subsection, we explore the potential heterogeneous effects of training across different firms. One important source of heterogeneity can be the ownership. One concern is that different firms may have different production technology and wage structure, and thus training can affect productivity and wages differently. In China, state-owned enterprises (SOEs) are systematically different from other types of firms, and so the differential effects of training could just reflect the productivity-wage gaps of these two different ownership structures. To address this heterogeneity, we now assume different production technologies, wage grids, as well as training effects across different types of firms.

Table 2.3 reports results for SOEs and non-SOEs separately. As for the production technology, the non-SOEs show more decreasing returns to scale as the average sum of exponents of the inputs in the Cobb-Douglas function is smaller for non-SOEs than SOEs. As we can see, the labor productivity effect of training is much larger for SOEs (significant at 5%), and the increase in wages due to training (higher skills or better performances of the workers) is also slightly higher. As a result, the training premium (the gap between productivity effect and wage effect) is larger for SOEs. Both SOEs and non-SOEs have significant training effects as well as the training premiums, which explain why firms could provide training willingly regardless of their ownership. Due to also the fact that SOEs account for only a small portion of our sample, our main result does not seem to be driven by ownership.

According to the hypothesis of Acemoglu and Pischke [1998], SOEs should have slightly more private incentives to provide training than non-SOEs, as the training premium is higher for SOEs. Interestingly, we do observe a larger proportion of firms among the SOEs that provide training than non-SOEs (67% vs 43%) as well as a larger average training spending per worker-year (192 vs 132 RMB yuan) in our data. This suggests that during China's massive privatization process the motivation for the remaining SOEs to provide training could also be profit maximization. In fact, it is also considered socially efficient for SOEs to provide more training because their training productivity effect, ϕ_T, is higher. In China, SOEs often follow the government to undertake some political/national tasks (for example, the "scientific developmentalist viewpoint"). Those political actions during transition in China are observed to coincide with private incentives according to our

Table 2.3: ACF Estimations by Registered Type

	(1) SOEs	(2) Non-SOEs	(3) Chow Test: (1)=(2)
Production Function			
Training ($\beta_l \phi_T$)	0.091**	0.041***	–
	(0.029)	(0.008)	
$ln(L)$ (β_l)	0.588***	0.441***	–
	(0.015)	(0.004)	
$ln(K)(\beta_k)$	0.255***	0.263***	–
	(0.044)	(0.003)	
Wage Equation			
Training (λ_T)	0.051***	0.043***	5.85**
	(0.022)	(0.001)	[0.016]
$ln(K/L)$ (γ)	0.079***	0.080***	–
	(0.002)	(0.001)	
TFP (γ_ρ)	0.164***	0.167***	–
	(0.008)	(0.001)	
ϕ_T	0.154***	0.094***	5.27**
	(0.045)	(0.019)	[0.022]
$\phi_T - \lambda_T$	0.103**	0.050**	5.19**
	(0.043)	(0.020)	[0.023]
Wald Test: $\phi_T = \lambda_T$			
χ^2	5.64**	6.17**	–
p-value	[0.018]	[0.013]	
Observations	16,785	489,679	506,464
Firms	6,589	165,128	171,717

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) Standard errors in parentheses are computed as in Table 2.2, and p-values are in square brackets. (3) ϕ_T is computed as the ratio of the training coefficient over the labor coefficient. (4) To carry out the Chow test, we estimate the parameters for SOEs and those for non-SOEs in a joint procedure.

results.

Another dimension of firm heterogeneity can come from union representation. As suggested by Yao and Zhong [2013], the unions' role in providing training services to employees are well documented in the literature, and they verified that in Chinese firms. However, does this fact coincide with private incentives? Table 2.4 shows the results for firms that established a union in 2004 and firms that did not. It turns out that firms that established a union (earlier) experienced both a larger productivity effect and a larger wage effect of training. The training premium is also larger for firms with earlier union establishment, while both of them are positive and statistically significant. Indeed, the private incentives of firms would favor the role of union in promoting training during privatization. Admittedly, this could also suggest that our main result partially reflects the role of union. Nevertheless, the firms without (early) union establishment also show sufficient private incentives for providing training, and thus at least union is not the only driver of our main result.

Finally, we investigate sector heterogeneity by reporting results for firms in each two-digit sector code separately. Due to the space limit, we only report the effects of training on marginal product of labor and wages in Table B2. As we can see, for most of the sectors, both the productivity effect and wage effect go down when controlling for the possible endogeneity of labor and training choices. For 28 out of 29 sectors, the productivity effect is larger than the wage effect with the ACF procedure, suggesting a positive training premium.[8] The largest productivity gains from training are seen in the beverage products, furniture, and artware sectors.

2.5.3 Worker Heterogeneity

Similarly, we address concerns that the worker heterogeneity is not fully controlled by our method. In this section, we decompose the effects by gender, education, and professional levels based on Section 2.2.3.

Table 2.5 shows that, if we assume that training resources are allocated evenly between men

[8]Similar to Konings and Vanormelingen [2015], due to the relatively low number of observations and the nonlinear search over the parameters, the difference is also not significant for quite a few sectors with the ACF procedure—only 11 out of 29 sectors have a statistically significant (and positive) difference at the 10% level; nevertheless, when the wage effect is larger than the productivity effect (only one case), the difference is not significant. For the OLS results, all 29 sectors have positive training premiums and 26 of them are statistically significant at the 10% level.

Table 2.4: ACF Estimations by Union Representation

	(1) With Union	(2) Without Union	(3) Chow Test: (1)=(2)
Production Function			
Training ($\beta_l \phi_T$)	0.083***	0.036***	–
	(0.003)	(0.007)	
$ln(L)$ (β_l)	0.524***	0.432***	–
	(0.005)	(0.001)	
$ln(K)(\beta_k)$	0.293***	0.225***	–
	(0.002)	(0.001)	
Wage Equation			
Training (λ_T)	0.050***	0.043***	93.03***
	(0.002)	(0.001)	[0.000]
$ln(K/L)$ (γ)	0.085***	0.074***	–
	(0.001)	(0.001)	
TFP (γ_ρ)	0.165***	0.166***	–
	(0.001)	(0.001)	
ϕ_T	0.159***	0.083***	31.68***
	(0.003)	(0.017)	[0.000]
$\phi_T - \lambda_T$	0.109***	0.040**	28.98***
	(0.005)	(0.018)	[0.000]
Wald Test: $\phi_T = \lambda_T$			
χ^2	394.46***	4.97**	–
p-value	0.000	0.026	
Observations	237,601	268,863	506,484
Firms	80,481	91,236	171,717

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) Standard errors in parentheses are computed as in Table 2.2, and p-values are in square brackets. (3) ϕ_T is computed as the ratio of the training coefficient over the labor coefficient. (4) To carry out the Chow test, we estimate the parameters for firms with union establishment and those without union establishment in a joint procedure.

and women within a firm, then the productivity effect of training is larger for women than for men (although not significantly different), and so does the wage effect of training (significant). As a result, the training premium for female workers are not so different from that for male workers—according to Acemoglu and Pischke [1998], firms should not have less incentives to provide training when they have more female workers. However, in our data, we observe that the provision of training is slightly higher among firms with higher proportion of male workers, unconditional on other characteristics of firms. For male-dominated firms (male ratio > 50%), the per capita training cost is 145 RMB yuan, while the number is only 102 RMB yuan for female-dominated firms. To reconcile the observed facts and profit-oriented incentives of firms, we might say that we overestimated the effects of training on female workers, which could be due to the fact that training resources might in fact lean more toward female workers within a firm, while in our model we assume that female workers use the same training resources as male workers to generate the (insignificantly) larger effects.

Of course, the reason for the observed association between more male workers and more training spending could go beyond simply the profit-maximizing incentive. It's observed that gender discrimination (mostly "male preference") existed in the labor market in the context of training support [Halldén, 2015]. That is, it's likely that a firm would pay more to train their male workers or would give more training chances to workers simply due to the perception about gender. If this is the case, then the effects on female workers could be underestimated, and female workers could have an even larger productivity effect of training than estimated. As a result, the training of female workers is likely to be under-invested from both the private and social points of view.

We then explore the dimension of education and professional levels in Columns (2) and (3) of Table 2.5. Similar trends are observed. On the one hand, a significantly larger training premium is observed among the lower-educated, which indicates higher private incentives to train lower-educated workers; on the other hand, in our data, we observe that firms with more than 10% of workers who received higher education (15 years or above) spent more on training (on average 205 RMB yuan), and so did firms with more than 30% of senior/advanced workers (on average 172 RMB yuan).[9] To reconcile private incentives with these observations, training resources could actually

[9]These trends are robust even when we control for other characteristics of the firms such as ownership.

Table 2.5: Worker Type Decompositions

	(1) Gender (Type 1: Female)	(2) Education (Type 1: ¡15 yrs)	(3) Professional Level (Type 1: Low)
Production Function			
$ln(L)$ (β_l)	0.452*** (0.011)	0.447*** (0.009)	0.451*** (0.005)
$ln(K)(\beta_k)$	0.259*** (0.004)	0.263*** (0.005)	0.260*** (0.006)
Training Type 1 ($\beta_l \phi_{T,\text{ type 1}}$)	0.053*** (0.011)	0.043*** (0.007)	0.047*** (0.010)
Training Type 2 ($\beta_l \phi_{T,\text{ type 2}}$)	0.038*** (0.012)	0.037 (0.028)	0.037*** (0.013)
Wage Equation			
$ln(K/L)$ (γ)	0.080*** (0.001)	0.080*** (0.001)	0.080*** (0.001)
TFP (γ_ρ)	0.167*** (0.002)	0.167*** (0.001)	0.167*** (0.001)
Training Type 1 ($\lambda_{T,\text{ type 1}}$)	0.059*** (0.005)	0.005 (0.008)	0.034*** (0.004)
Training Type 2 ($\lambda_{T,\text{ type 2}}$)	0.044*** (0.007)	0.150*** (0.043)	0.064*** (0.013)
$\phi_{T,\text{ type 1}}$	0.116*** (0.025)	0.097*** (0.014)	0.104*** (0.021)
$\phi_{T,\text{ type 2}}$	0.084*** (0.028)	0.082 (0.063)	0.081*** (0.029)
$\phi_{T,\text{ type 1}} - \lambda_{T,\text{ type 1}}$ (A)	0.057** (0.029)	0.092*** (0.010)	0.070*** (0.017)
$\phi_{T,\text{ type 2}} - \lambda_{T,\text{ type 2}}$ (B)	0.040* (0.021)	-0.068 (0.052)	0.017 (0.031)
Wald Test: χ^2 [p-value]			
$\phi_{T,\text{ type 1}} = \lambda_{T,\text{ type 1}}$	4.00** [0.046]	87.79*** [0.000]	16.92*** [0.000]
$\phi_{T,\text{ type 2}} = \lambda_{T,\text{ type 2}}$	3.63* [0.057]	1.71 [0.191]	0.29 [0.588]
$\phi_{T,\text{ type 1}} = \phi_{T,\text{ type 2}}$	0.58 [0.445]	0.05 [0.829]	0.28 [0.595]
$\lambda_{T,\text{ type 1}} = \lambda_{T,\text{ type 2}}$	8.18*** [0.004]	8.44*** [0.004]	3.91** [0.048]
A = B	0.19 [0.665]	9.36*** [0.002]	1.76 [0.185]
Observations	506,484	506,484	506,484
Firms	171,717	171,717	171,717

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (3) Standard errors in parentheses are computed as in Table 2.2, and p-values are in square brackets. (4) ϕ_T is computed as the ratio of the training coefficient over the labor coefficient. (5) The compositions of workers by education and professional levels are based on information in 2004.

lean toward these workers within a firm so that the effects for them are overestimated. If we continue to assume that training resources are allocated evenly within a firm, then for some reason those firms might under-invest in these workers from the private (profit-maximizing) perspective. From the social perspective, it seems that investing more in training these workers (instead of the higher-educated and skilled) can yield a larger output, and so the social efficiency might be improved. Nevertheless, in either case, firms have invested or should invest more on training lower-educated and perhaps also junior workers to obtain better effects. Admittedly, due to the limitation of our data (lacking employee information), the interpretations of the results in this section depend on our assumptions. We hope that future research can extend this section by merging employee-level data with firm-level data.

2.6 Conclusion

This paper is among the first to use a large firm-level penal data set to estimate the effects of on-the-job training on productivity and wages in a transition economy. Based on a control function approach, we jointly estimate the production function and wage equation for Chinese manufacturing enterprises amid a privatization era (2004-2007), which happens to be the first few years after the proposal of the "scientific developmentalist viewpoint", taking the endogeneity in input and training choices of firms into account. Our results support the theory of Acemoglu and Pischke [1998] in general. That is, there existed a positive gap between training productivity effect and wage effect, which could be part of the reason why firms were willing to provide training. There could still be under-investment in training (especially for specific types of workers) from a social point of view. We can clearly see that better trained workers have a higher marginal product of labor, which can benefit not only their current employers but also future employers, if any. It's common that firms only consider the cost of training but not its externalities, and so they typically would not have incentives to provide enough training.

We also find several interesting context-specific results: the provision of training by state-owned manufacturing enterprises might be both privately and socially efficient in China; Chinese unions played a role in promoting the training benefits of workers; during privatization, it could be more privately and socially efficient for manufacturing firms to prioritize training resources to lower-educated workers (and perhaps female and junior workers), if they had not done so.

This opens a window for government intervention such as training subsidies. In fact, training (or personnel education) has always been one of the tax deductible items in China since the 1990s, which is meant to encourage the provision of training. However, it remained a very small proportion of a firm's tax deductible expenses until recently. In 2018, the Chinese government raised the standard of tax deduction regarding training from 2.5% of wages to 8% of wages, which broke the decades long standard. It would be interesting to see how this policy change can affect the provision of training and thus productivity and wages—which is beyond the scope of this research and requires future studies using more recent (and preferably employer-employee pairwise) data. Another limitation of this research is the inability to distinguish general training from firm-specific training, and we hope that it can be addressed in future research. Moreover, it would be interesting to see research done in another context of transition economy.

Chapter 3

UNCONDITIONAL QUANTILE TREATMENT EFFECT OF COVID-19 IN A TWO-SIDED MARKET: EVIDENCE FROM A LIVE-STREAMING PLATFORM

3.1 Introduction

Our economy is made of myriad markets, many of which are "two-sided" nowadays. That is, an intermediate firm enters a market and provides services to two distinct groups of participants of the market to encourage their interaction. Then, the involvement of one group can affect that of the other group, which is known as the cross-group externality [Rochet and Tirole, 2003]. Two-sided markets are now increasingly based on digital platforms, such as live-streaming platforms. On such a platform, the two distinct groups of end users are usually known as "fans" and "anchors". On the one hand, fans watch live shows produced by anchors, send virtual gifts to reward the anchors they admire, and buy products from their stores; anchors, on the other hand, use the platform to broadcast their shows, receive virtual gifts with monetary value, and do live commerce to sell products and gain profits. A firm that provides different functions of a platform to two groups of end users typically believe that having more users on one end increases the participation of users on the other end. This paper investigates how the digital platform and the emerging group of flexible workers (e.g., live streamers) can be affected by an economy-wide shock which could affect the both sides of the platform. Particularly, we study how virtual gifting, the major income source of the online anchor and the platform, can be affected by an exogenous and positive demand shock Covid-19. We are also interested in its implication for the polarization of labor market, which has recently become a common issue in the United States, Germany, and other countries [Acemoglu and Autor, 2011, Bachmann et al., 2021, David and Dorn, 2013, Vom Lehn, 2020]. The recent study [Cortes and Forsythe, 2020] shows Covid-19 leads to the increase in job polarization and job loss among disadvantaged groups. Since early 2020, the world has been gripped by the COVID-19 virus. China was the first country to report the outbreak and then control it at the national level. This provides us with a unique chance to study the impacts of an unexpected, universal, and potentially positive shock (to both groups of end users) in a two-sided market.

We first present the two-sided nature of a digital platform using an illustrative theoretical model by extending Rochet and Tirole [2003]. In this model, we model the COVID-19 pandemic as a "two-sided" shock (affects both online anchors and fans) to the digital platform. The key insight from our theoretical analysis is that when the shock to fans is positive and large enough, the platform tends to cross-subside anchors. This theory prediction is consistent with the empirical fact in reality. We estimate a reduced-form model in a quasi-experimental setting. We first obtain a unique data set from one of the largest live-streaming platforms via web spiders. At the same time, we keep track of the daily number of COVID-19 cases in each city and that of the related local policies. Then, we control for macroeconomic conditions of each city. Under conditional exogeneity, we run an ordinary least square (OLS) regression to estimate an average treatment effect (ATE). In our data, we observe a large variation in the value of virtual gifts per show, indicating a wide range of popularity among live streamers. To further inspect and incorporate the heterogeneity into our model, we allow the effects of COVID-19 to vary by side as well as along the distribution of total gift value per show by adopting the general quantile regression (GQR) framework [Powell, 2019]. Our empirical results confirm a positive ATE of COVID-19 on virtual gifting. The COVID-19 severity in fans' cities increases virtual gifting from fans to anchors, while the COVID-19 severity in anchors' cities does not affect virtual gifting on average. In the GQR estimation, we further find that this positive effect is mostly driven by the bottom of the per-show gifting volume distribution, suggesting that anchors who receive fewer gifts can benefit more from a positive demand shock in their fans' cities. In robustness checks, we estimate a model of single binary treatment instead of a model of multiple continuous treatments. This thus allows us to apply the semiparametric estimator developed by Firpo [2007]. Our alternative settings all confirm that the impact of COVID-19 on the virtual gifting behaviors on the platform is heterogeneous and "pro-small". Understanding heterogeneous effects of an economy-wide shock in a two-sided market can help us design policies that aim to mitigate the polarization in a labor market and help with the growth of smaller producers that rely on the other side of the market. We document how the least popular anchors as well as the most popular ones converge to the "middle class" when a positive shock to the platform demand is introduced. In addition to the total demand, we also explore how COVID-19 affects the quasi-demand of anchors and fans. We find the results are consistent with the main conclusion.

This research complements the literature in a few ways. First, it provides new insights and empirical evidence to the literature on two-sided markets, featured by the work of Armstrong [2006], Rochet and Tirole [2003, 2006a], and Rysman [2009]. In the context of a two-sided market, very few empirical studies have discussed external shocks that can affect both sides simultaneously. For example, Baker et al. [2020] used the data from an online savings platform called Saver Life but did not distinguish the COVID-19 shocks from two sides. Chang and Meyerhoefer [2021], Glaeser et al. [2020] and Raj et al. [2020] also studied COVID-19 impacts using data from two-sided markets (online food shopping services, Yelp, and Uber Eats) but they also did not model shocks as two-sided shocks. We show that, by distinguishing the shocks from two sides, we can obtain new insights about the effects of shocks in a two-sided market. Second, this paper is among one of the very few studies that investigate a positive effect of COVID-19. The current COVID-19 literature is mostly focusing on the negative impacts of this pandemic, while soft-pedalling the fact that the COVID-19 pandemic also provides opportunities for new economies and businesses. Some studies explore the macroeconomic impacts of COVID-19 and their plausible mechanisms[Baker et al., 2020, Binder, 2020, Chetty et al., 2020, Coibion et al., 2020a, Jordà et al., 2020]. They typically find that the economic impacts of COVID-19 are much worse in comparison to the previous pandemics such as the 1918 Spanish Influenza [Keogh-Brown et al., 2010, Barro et al., 2020]. Some document the negative impacts of the pandemic on physical health, mental health, and/or well-being [Adams-Prassl et al., 2020, Brodeur et al., 2020, de Pedraza et al., 2020]. Additionally, researchers also study its negative impacts on various (mostly one-sided) markets. For example, Baker et al. [2020] and Kirk and Rifkin [2020a] examine the impacts of COVID-19 on the overall goods market based on the theory of demand shocks [Lorenzoni, 2009]; other researchers suggest the negative impacts of COVID-19 on labor markets, e.g., reduced working hours, increased unemployment, income losses, job offerings [Aum et al., 2020, Bui et al., 2020, Coibion et al., 2020b, Forsythe et al., 2020]. Although this latest pandemic has brought a series of negative shocks to the macroeconomic environments, health, and various market outcomes, there are some positive influences that are less studied and documented. This paper complements a large—and rising—group of studies studying the (mostly negative) impacts of COVID-19. Third, our work is related to the literature on estimating quantile treatment effects. Similar to Autor and Kerr [2017], we are among the first to investigate the effects of multiple treatments on a type of earnings at different quantiles. The major difference between of our papers is that we are estimating the effects at unconditional quantiles, instead of conditional quantiles, which can make the interpretation of the results less straightforward sometimes. This is one of the first applications of Powell [2019]'s GQR framework, other than Powell and Goldman [2021]. Last, this work discusses polarization in a special labor market where laborers are live streamers and content creators, and their efforts are rewarded by virtual gifting from fans and revenue from sales. If we can regard less popular anchors as low-skill workers, and the most popular anchors as high-skill workers, then we find that they tend to converge to the middle class during COVID-19. Our explanation is that the digital platform serves as the sanctuary for them.

The remainder of the paper is organized as follows. In Section 3.2 , we introduce an illustrative model to analyze the effects of two-sided COVID-19 shocks theoretically and concluded that while there may be a clear price prediction, the predicted change in transaction volume is ambiguous—this motivates our empirical investigation in Section 3.3 in which we present the statistical model and estimation framework. In Section 3.4, we illustrate the data sources and basic characteristics. In Section 5, we provide estimation results and discussions, along with robustness checks. Finally, we note our concluding remarks in Section 6.

3.2 An Illustrative Theoretical Model

In this section, we use a simple model to show that COVID-19 shocks (either positive or negative shocks) to the demand for a two-sided platform have ambiguous effects on transaction volume. It is thus important to investigate the effects empirically. We do not intend to structurally recover the parameters and/or functions introduced below, and our empirical investigation will focus on figuring out the relationship between COVID-19 cases and the equilibrium transaction volume.

Our theoretical model modifies the benchmark model in Rochet and Tirole [2003]. In a live-streaming platform like Kuaishou, fans could give virtual gifts to online anchors. Let's call one dollar of such gift as one transaction. For each transaction, the platform charges the anchor a fee of p^A and a fan a fee of p^F, which could be explained as commission fee or subsidy (negative price). Benefits b^A and b^F correspond to the utility of anchors and fans gained from the gifting behavior. They are two random variables. We assume the utility functions of anchors and fans from a transaction are:

$$U^A = \left(b^A + \theta^A - p^A\right) N^F$$

$$U^F = \left(b^F + \theta^F - p^F\right) N^A$$

Where θ^A and θ^F are Covid-19 shocks to anchors and fans.

We normalize the utility of outside options are 0, so the "quasi-demand" functions of anchors and fans are as follows:

$$D^A\left(p^A, \theta^A\right) = \Pr\left(\left(b^A + \theta^A - p^A\right) N^F \geq 0\right) = \Pr\left(b^A + \theta^A - p^A \geq 0\right) = N^A$$

$$D^F\left(p^F, \theta^F\right) = \Pr\left(\left(b^F + \theta^F - p^F\right) N^A \geq 0\right) = \Pr\left(b^F + \theta^F - p^F \geq 0\right) = N^F$$

We assume they are both log-concave. Here another key assumption is there are network externalities in the utility/surplus functions but not in the "quasi-demand" functions. Following Rochet and Tirole [2003], the actual platform demand is the product of $D^A\left(p^A, \theta^A\right)$ and $D^F\left(p^F, \theta^F\right)$:

$$Q(p^A, p^F, \theta^A, \theta^F) = D^A(p^A, \theta^A) D^F(p^F, \theta^F)$$

For simplicity we assume the platform is a private monopoly, which chooses prices to maximize total profit:

$$\pi = \left(p^A + p^F - c\right) D^A(p^A, \theta^A) D^F(p^F, \theta^F)$$

The equilibrium is specified as:

$$\varphi(p, \theta) = \begin{cases} \lambda^A\left(p^A, \theta^A\right) - \frac{1}{p^A + p^F - c} = 0 \\ \lambda^F\left(p^F, \theta^F\right) - \frac{1}{p^A + p^F - c} = 0 \end{cases}$$

where $\lambda^A\left(p^A, \theta^A\right) = \frac{-\left(D^A\left(p^A, \theta^A\right)\right)'}{\left(D^A\left(p^A, \theta^A\right)\right)}$, $\lambda^F\left(p^F, \theta^F\right) = \frac{-\left(D^F\left(p^F, \theta^F\right)\right)'}{\left(D^F\left(p^F, \theta^F\right)\right)}$, $p = (p^A, p^F)$ and $\theta = (\theta^A, \theta^F)$.

We have the following propositions:

Proposition 1 *If we assume the quasi-demand functions of anchors and fans are both log-concave, we have:*

(1) If more COVID-19 cases implies a higher θ^F (a positive demand shock to fans), then, ceteris paribus, p^A decreases and p^F increases. Vice versa.

(2) If more COVID-19 cases implies a higher θ^A (a positive demand shock to anchors), then, ceteris paribus, p^A increases and p^F decreases. Vice versa.

Proposition 2 *Assume b^A and b^F both follow uniform distribution over $[b_1, b_2]$, which means their quasi-demand functions are linear. If more COVID-19 cases implies a higher θ^A or a higher θ^F , then, ceteris paribus, the welfare of both anchors and fans increases.*

The proofs are in section C.1 and section C.2. Proposition 1 says facing a strong positive demand shock to fans, the platform would cross subside anchors and vice versa. Proposition 2 says due to the network externality, the positive demand shock to fans would also increase the welfare of anchors and vice versa. The illustrative model could help us to rationalize the empirical facts. But due to the data restriction, We do not intend to structurally recover the demand functions.

3.3 Statistical Model and Estimation Method

In our main specification, we resort to the generalized quantile regression (GQR) in Powell [2013, 2019] to estimate unconditional quantile treatment effects of multiple continuous treatments. Let's denote $\mathbf{d}_i$, $\mathbf{x}_i$, $y_{i,\mathbf{d}_i}$ as the treatment vector, covariate vector and potential earning outcome of live video i. We impose the following linear structure on potential outcome $y_{i,\mathbf{d}_i}$:

$$y_{i,\mathbf{d}_i} = \mathbf{d}_i' \beta \left(\mathbf{u}_i^* \right),$$

where $\mathbf{u}_i^* \sim U(0,1)$ is a rank variable determining placement in the unconditional distribution of $y_{i,\mathbf{d}_i}$ and β refers to the marginal quantile treatment effect(s). In addition, we denote

$$\tau_{\mathbf{x}_i} \equiv \Pr \left(y_{i,\mathbf{d}_i} \leq \mathbf{d}_i' \beta(\tau) \mid \mathbf{x}_i \right)$$

and its estimate $\hat{\tau}_{\mathbf{x}_i}$. We assume with probability one:

(1) **Monotonicity of Potential Outcome:** $\mathbf{d}_i' \beta \left(\mathbf{u}_i^* \right)$ is increasing in $\mathbf{u}_i^*$

(2) **Conditional Independence:**
(a) $\Pr \left(\mathbf{u}_i^* \leq \tau \mid \mathbf{d}_i, \mathbf{x}_i \right) = \Pr \left(\mathbf{u}_i^* \leq \tau \mid \mathbf{x}_i \right)$
(b) $\mathbb{E} \left[\mathbf{d}_i (\hat{\tau}_{\mathbf{x}_i} - \tau_{\mathbf{x}_i}) \right] = 0$

(3) **Full Rank:** If $\Pr \left(\mathbf{d}' c = 0 \mid \mathbf{x}_i \right) = 1$ for all $\mathbf{x}_i$, then $c = 0$

(4) **Continuity:** $y_{i,\mathbf{d}_i}$ is continuously distributed conditional on $\mathbf{d}_i$ and $\mathbf{x}_i$

The above assumption implies the following moment conditions and the proof is shown in Powell [2013, 2019]:

$$\mathbb{E}\left\{\mathbf{d}_i\left[\mathbf{1}\left(y_{i,\mathbf{d}_i} \leq \mathbf{d}_i'\beta(\tau)\right) - \hat{\tau}_{\mathbf{x}_i}\right]\right\} = 0$$
$$\mathbb{E}\left[\mathbf{1}\left(y_{i,\mathbf{d}_i} \leq \mathbf{d}_i'\beta(\tau)\right) - \tau\right] = 0$$

Let's define $\mathbf{d}_i = (1, \tilde{\mathbf{d}}_i)$. Let $\gamma(\tau, \tilde{\beta})$ be the τ^{th} quantile of the distribution of $y_{i,\mathbf{d}_i} - \tilde{\mathbf{d}}_i'\tilde{\beta}$. Following Powell [2013, 2019], We rewrite $\tau_{\mathbf{x}_i}$ as a more parametric form $F(\mathbf{x}_i'\delta)$. The form of $F(\cdot)$ is flexible, for example we could use logit or probit specification here. The estimation framework is Generalized Method of Moments (GMM). The estimation procedure is as follows:

For each $\tilde{b}$ (candidate values of $\tilde{\beta}$, which are treatment coefficients except for the constant γ and we denote $\beta = (\tilde{\beta}, \gamma)$) in the grid:

(1) solve $\hat{\gamma}(\tau, \tilde{b})$ from sample moment $\frac{1}{N}\sum_i^N \mathbf{1}\left(y_i - \tilde{\mathbf{d}}_i'\tilde{b} \leq \hat{\gamma}(\tau, \tilde{b})\right) = \tau$ where $\hat{\gamma}(\tau, \tilde{b})$ is the estimated constant

(2) solve $\hat{\delta}(b, \tau)$ by likelihood maximization:

$$\underset{\delta(b,\tau)}{\operatorname{argmax}} \sum_{i=1}^{N} \mathbf{1}\left(y_{i,\mathbf{d}_i} \leq \mathbf{d}_i'b\right) \ln F\left(\mathbf{x}_i'\delta(b,\tau)\right) + \mathbf{1}\left(y_{i,\mathbf{d}_i} > \mathbf{d}_i'b\right) \ln\left(1 - F\left(\mathbf{x}_i'\delta(b,\tau)\right)\right).$$

Then calculate $\hat{\tau}_{\mathbf{x}_i} = \hat{F}(\mathbf{x}_i'\hat{\delta}(b, \tau))$

(3) calculate the objective function $\hat{g}(b)'\hat{A}\hat{g}(b)$, where $\hat{A}$ is a weight matrix and

$$g_i(b) = \mathbf{d}_i\left[\mathbf{1}\left(y_{i,\mathbf{d}_i} \leq \hat{\gamma}(\tau, \tilde{b}) + \tilde{\mathbf{d}}_i'\tilde{b}\right) - \hat{\tau}_{\mathbf{x}_i}(b)\right]$$
$$\hat{g}(b) = \tfrac{1}{N}\textstyle\sum_{i=1}^{N} g_i(b)$$

(4) keep step 1–3 until b minimizes the objective function

For more details of GQR, please refer to Powell [2013, 2019].

In our robustness checks, we also resort to the localized debiased machine learning (LDML) developed by Kallus et al. [2020] and a semiparametric estimator developed by Firpo [2007]. Details are not presented here.

3.4 Data, Variables, and Summary Statistics

For our empirical investigation, we obtained a unique live-streaming data set via web crawler. The data set contains snapshots of around 5% of all the live videos on the Kuaishou platform (with more than 300 million daily active users) from October 2019 until July 2020. It covers information on anchors, their fans, listings (of goods and gifts) and transactions between them.

Figure 3.1 features a typical user interface of the Kuaishou platform. In Figure 3.1a, "Anchor 1" is trying to persuade his fans to purchase a bottle of olive oil by showing how to cook shrimps with it; in the comment section, "Fan 1" is asking how big is the stone pot; fans can click "Suscribe" to follow the activities of this anchor, "Watch more" to switch to other live videos, "Purchase" to view the anchor's online store (Figure 3.1b), and/or "Gift" to send a virtual gift (Figure 3.1c) to the anchor that can be later converted to real money. The upper right corner of the screen shows who are currently watching this live video, from which we obtain the information on average attendance.

To compensate our Kuaishou data, we gathered daily data of COVID-19 cases from thousands of authoritative news pieces, local policy documents from websites of local governments, and macroeconomic indicators (such as GDP and population) from national and local bureaus of statistics.

In Figure 3.2, we show how our sample points are distributed across China. The blue circles are proportional to the total numbers of live-streaming videos hosted in each province during our study period. As we can see, most live-streaming videos are hosted in coastal areas. The green-scale shades illustrate how many confirmed COVID-19 cases are accumulated during the study period, and there seems to be a positive correlation between the number of live videos and the number of confirmed cases, except for Hubei in which Wuhan is located.

To further illustrate our data, we present summary statistics of our key variables by the Wuhan Lockdown period in Table 3.1. First, let's talk about the anchors: while there are more male than female anchors, we observe a slight decrease in the proportion of female anchors (from 34.8% before the lockdown to 31.9% after the lockdown) in our study period; on the other hand, the average number of fans per anchor increases during and after the pandemic. As for fans, we also observe slight decreases in the proportion of female fans. In either the cities of anchors or the cities of fans, we observe increases in the average number of policies per month; of course, the average number

(a) Main Interface (b) Click "Purchase" (c) Click "Gift"

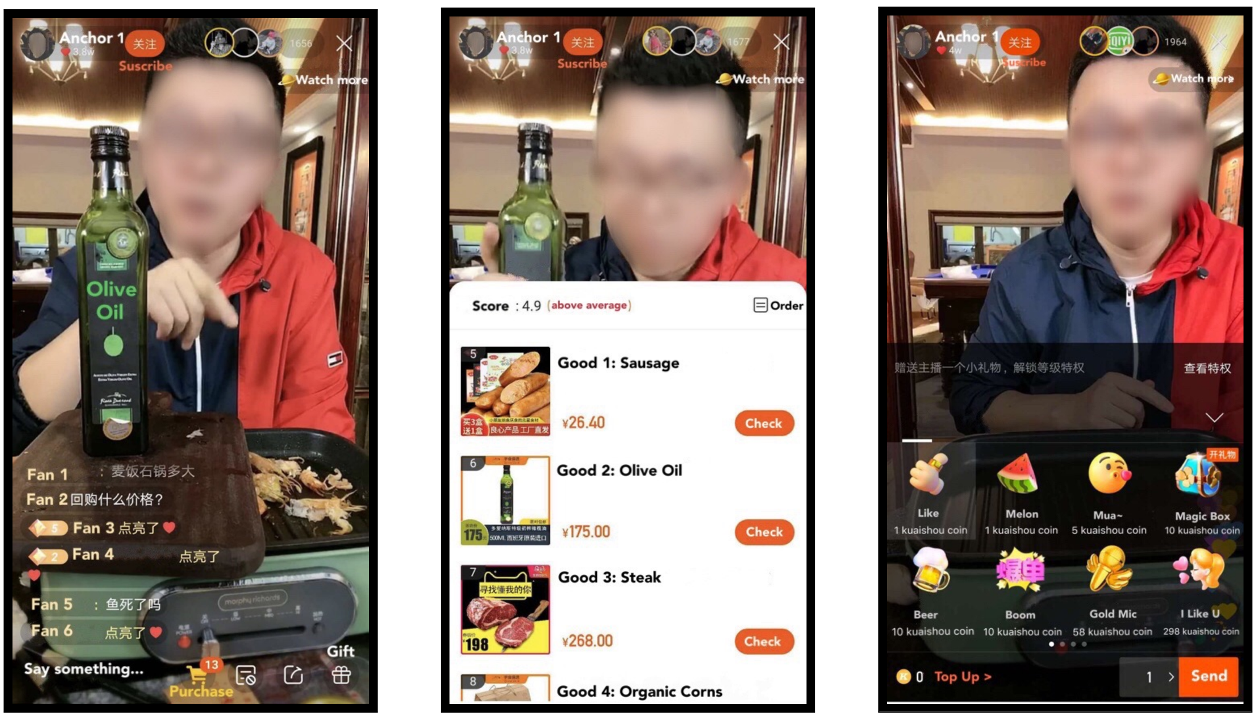

Figure 3.1: User Interface of Kuaishou

Notes: (1) Faces of the example anchor have been blurred. (2) Screenshots have been edited to replace some Chinese words by their English translations. Additional English words are added to provide information about the functions of some buttons. (3) 1 kuaishou coin = 0.1 RMB yuan.

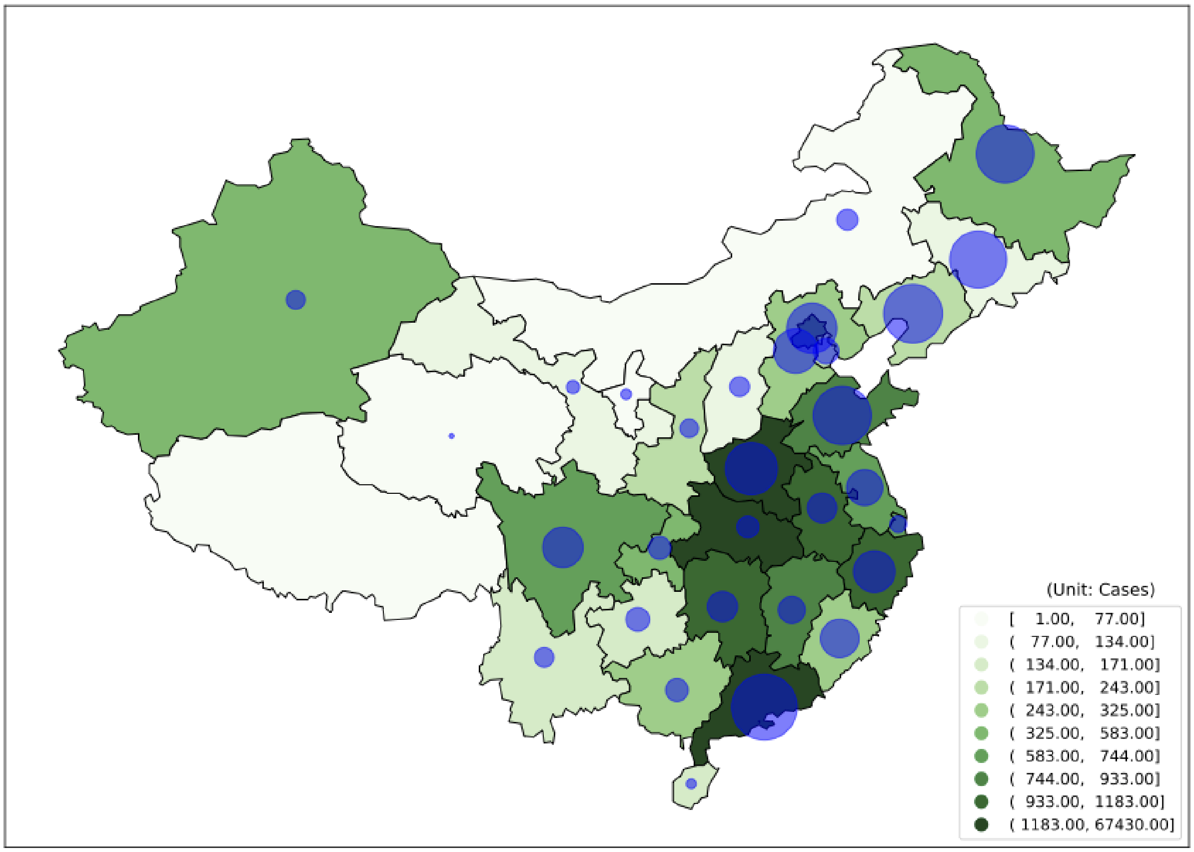

Figure 3.2: COVID-19 cases (green) and live videos (blue), Oct 2019–Jul 2020

Notes: (1) The size of the blue circle represents the total number of live-streaming videos in the province; the larger it is, the more live videos; for example, there are in total 12,523 live videos in Beijing, but only 1,765 in Xinjiang. (2) The darkness of the shade in each area represents the COVID-19 intensity measured by total confirmed cases; the darker it is, the more cases.

Table 3.1: Summary Statistics (Oct 2019–Jul 2020) by Lockdown Period

	(1) Before		(2) Wuhan Lockdown		(3) After	
	Mean	SD	Mean	SD	Mean	SD
Anchors						
Female	0.348	0.476	0.337	0.473	0.319	0.466
Total fans (1M)	5.423	4.624	6.103	5.281	6.866	5.782
City characteristics:						
—*COVID-19 cases (14 days)*	0.087	4.230	105.996	1394.173	5.244	26.407
—*Lockdown policies (30 days)**	0.643	1.997	22.539	6.947	32.097	15.588
—*Reopening policies (30 days)**	0.000	0.000	10.313	9.557	16.478	7.408
—*Population (100K)*	84.831	58.499	87.783	61.224	88.555	62.163
—*Area (10K* km^2*)**	1.839	2.375	1.823	2.584	1.751	2.020
—*Quarterly GDP (1T yuan)**	0.239	0.309	0.199	0.269	0.236	0.257
Fans						
Female ratio	0.466	0.252	0.454	0.242	0.436	0.221
City characteristics:						
—*COVID-19 cases (14 days)*	0.039	0.466	59.210	133.770	3.407	5.920
—*Lockdown policies (30 days)**	0.645	1.950	22.155	5.346	31.055	12.854
—*Reopening policies (30 days)**	0.000	0.000	10.331	9.455	16.112	5.688
—*Population (100K)*	74.166	10.281	74.213	9.995	78.229	14.749
—*Area (10K* km^2*)**	2.509	1.483	2.510	1.350	3.049	11.093
—*Quarterly GDP (1T yuan)**	0.198	0.065	0.147	0.054	0.245	0.192
Live Streaming						
Length (hours)	2.106	1.642	2.221	1.642	1.969	1.474
Average attendance (1K)	4.624	11.360	7.090	14.914	7.045	15.209
Log of gift value+1 (yuan)	9.555	2.860	10.015	2.662	9.329	3.135
Observations	71,565		47,205		75,059	

Notes: (1) Before the Wuhan lockdown, there were already suspicious COVID-19 related cases, and the government already started some quarantine measures to prevent spreading of these cases. (2) Wuhan lockdown is defined as the period between January 23, 2020 and April 8, 2020. (3) * denotes the proneness variables. (4) The monthly summary statistics is provided in Table C1.

of newly confirmed cases per fortnight reached the peak during the lockdown period; nevertheless, the quarterly GDP was in a slump during the lockdown period. Finally, we characterize the live streaming: during the lockdown, the videos were on average longer, and there were more viewers on average, and the value of the gifts from fans to anchors was on average higher; after the lockdown, the time length and gift value both dropped below the pre-lockdown levels, while average attendance remained relatively stablely high.

To get an idea of the (unconditional) distribution of the gift earnings of these anchors, Figure 3.3 provide histograms of the log values of gifts per live video before, during, and after Wuhan Lockdown. We notice that the distribution became more concentrated (the kurtosis is higher) and left-skewed (the skewness is more negative). Histograms by month (see Figure C1) show similar trends.

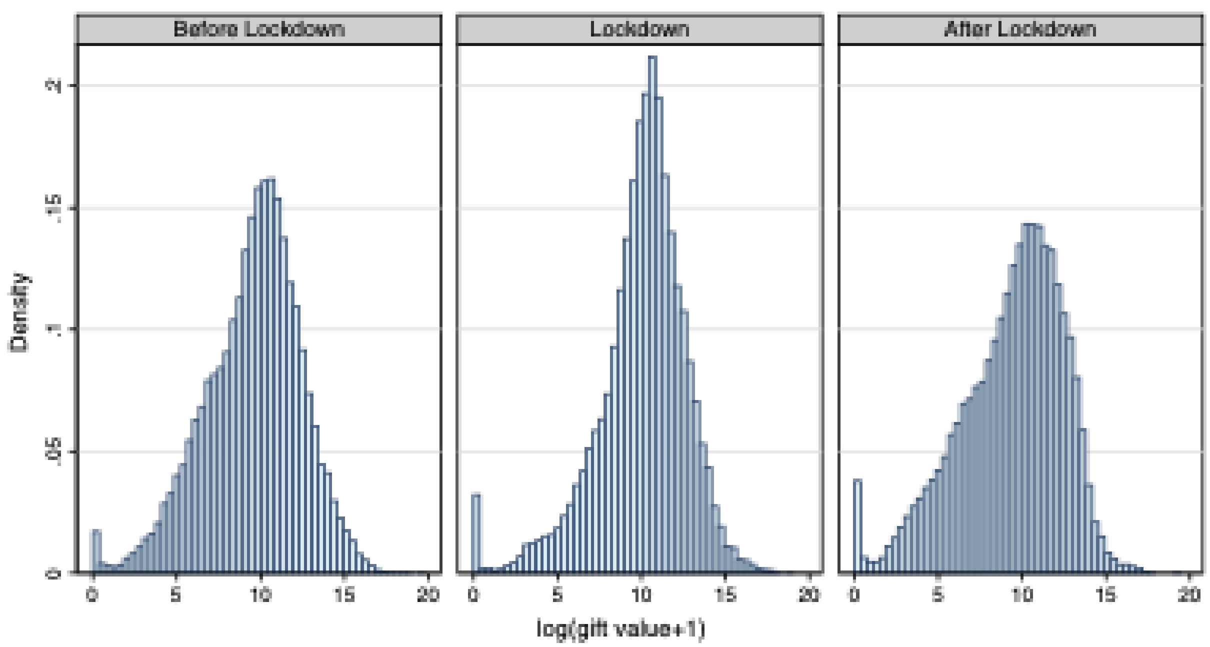

Figure 3.3: Histograms of gift values per live streaming video around Wuhan lockdown

3.5 Results and Discussions

3.5.1 Main Specification

The first row of Table 3.6 reports the average treatment effects (ATEs) of the COVID-19 intensity (measured by the number of newly confirmed cases in the past 14 days per 10,000 people) in an anchor's city and its average in their fans' cities. Interestingly, on average, the severity of the COVID-19 pandemic in an anchor's city does not seem to affect the equilibrium of virtual gifting, while that from the fans side increases virtual gifting. Particularly, an additional case of COVID-19 in the past 2 weeks per 10,000 people in the fans' cities increases the total value of virtual gifts received by the anchor per show by 0.12 to 0.15 log points, based on columns (1) to (3) of Table 3.6.

Based on the summary statistics in Section 3.4, there are very likely heterogeneous COVID-19 impacts on gift income along its distribution. Understanding this heterogeneity is meaningful since we do observe wide-ranging earnings among anchors, and they are representing a special and growing workforce. Therefore Table 3.6 further shows the estimated unconditional quantile treatment effects (QTEs) of COVID-19 from both the anchors and fans sides. The estimates suggest that the effects of COVID-19 from the anchors side are insignificant and likely homogeneous. On the contrary, we find significantly positive effects of COVID-19 shocks from the fans side on virtual gifting, especially at the lower quantiles of the earning distribution. Based on columns (1) to (3), an additional case of COVID-19 in the past 2 weeks per 10,000 people in the fans' cities increases the total value of virtual gifts received by the anchor per show by 0.31 to 0.35 log points at the 20th percentile of the gifting distribution, and by around 0.18 log points at the 40th percentile. Both QTEs are higher than the ATE (and the QTE at the median).

Our main result suggests that anchors who earn less from virtual gifts benefit more under the COVID-19 shocks (considered as positive demand shocks) from the cities of their fans. This tends to be a de-polarization process of this special labor market—the "high-skill" anchors (who manage to obtain a lot of virtual gifts from fans) and the "low-skill" anchors (who are only able to receive a few virtual gifts per show) are now converging to the middle-skill class. This result also corresponds to the middle panel of Figure 3.3 where the gifting distribution becomes more concentrated around the center.

Table 3.2: Treatment Effects of Two-Sided COVID-19 on Gifting Behavior

	(1)	(2)	(3)
ATE			
Anchor side	0.000	0.003	0.001
	(0.003)	(0.004)	(0.003)
Fans side	0.133**	0.121**	0.148**
	(0.004)	(0.057)	(0.052)
QTE at Percentile			
20			
Anchor side	-0.002	0.001	0.001
	(0.001)	(0.002)	(0.002)
Fans side	0.350***	0.312***	0.311***
	(0.011)	(0.041)	(0.043)
40			
Anchor side	-0.000	0.000	-0.001
	(0.001)	(0.002)	(0.003)
Fans side	0.180***	0.180***	0.175***
	(0.009)	(0.044)	(0.060)
60			
Anchor side	0.001*	0.001	-0.000
	(0.000)	(0.001)	(0.001)
Fans side	0.038***	0.049	0.060
	(0.007)	(0.030)	(0.038)
80			
Anchor side	0.001	-0.000	-0.001
	(0.001)	(0.004)	(0.005)
Fans side	-0.084***	-0.042	-0.033
	(0.008)	(0.087)	(0.094)
Additional Proneness variables			
#Lockdown policies		✓	✓
#Reopening policies		✓	✓
Quarterly GDP		✓	✓
City areas/sizes		✓	✓
Province dummies			✓
Time dummies			✓
Observations	194505	194505	194505

Notes: (1) The outcome is log(gift value + 1), while the exposures are COVID-19 intensities (cases/population) in the anchor's city and fans' cities. (2) ATE = average treatment effect; QTE = quantile treatment effect. (3) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (4) The ATE is obtained by linear regression adjustment. (5) We use the "linear" model to predict the probability that the outcome is below the quantile function conditional on proneness variables. (6) Numerical optimization proceeds via a Nelder-Mead algorithm. (7) Conventional standard errors are in parentheses for QTEs, while robust standard errors clustered on the anchor are in the square brackets for ATEs.

3.5.2 Robustness Checks

To check the robustness of our results, we attempted several similar settings. We first convert our multiple continuous treatments into a single binary treatment. By doing so, we will be able to resort to a recent method called localized debiased machine learning (LDML) developed by Kallus et al. [2020] as well as a more traditional and well-established method developed by Firpo [2007] in estimating the QTE. The single binary treatment is defined by comparing the COVID-19 intensity of the two sides—if the COVID-19 intensity is higher on the fans side, the treatment variable will be 1; otherwise, it will be 0. According to the results in Table 3.6, the treatment effect should be positive and higher at lower quantiles. Indeed, Table 3.3 confirm this prediction.

Table 3.3: Effects of Single Binary Treatment: Fans Side > Anchor Side

ATE	QTE at Percentile			
	20	40	60	80
A. Unadjusted	C. Firpo [2007]			
0.405***	1.606***	1.297***	0.435***	-0.022
[0.106]	(0.052)	(0.047)	(0.037)	(0.035)
B. Adjusted	D. LDML			
0.660***	0.571***	0.596***	0.263***	0.077
[0.190]	(0.046)	(0.035)	(0.051)	(0.096)
	E. GQR			
	1.415	0.776***	0.354	0.156
	(4.043)	(0.219)	(0.317)	(0.690)

Notes: (1) N=194,505. (2) ATE = average treatment effect; QTE = quantile treatment effect; LDML = localized debiased machine learning; GQR = generalized quantile regression. (3) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (4) The adjusted ATE is obtained by linear regression adjustment. (5) The LDML part is based on Kallus et al. [2020] using random forests to fit the conditional expectation functions and spliting the data into 5 folds. (6) The GQR part is based on Powell [2019] using the "linear" model to predict the probability the outcome is below the quantile function conditional on the proneness variables in Table 3.1, and the numerical optimization proceeds via a Nelder-Mead algorithm. (7) Conventional standard errors are in parentheses for QTEs, while robust standard errors clustered on the anchor are in the square brackets for ATEs. (8) For QTEs at other percentiles, refer to Table C3; for QTEs by LDML using other methods such as neural networks, refer to Table C4.

We implemented other robustness checks in section C.4. By changing the linear probability model to a logit/probit model (Table C5 and Table C6), the numerical optimization method from a Nelder-Mead algorithm to an adaptive MCMC procedure (Table C7) and a grid search procedure (Table C8), and by adding province fixed effects (Table C9), main results remained unchanged.

3.5.3 Platform Pricing

From Table 3.6 we could conclude that θ^A is around zero and θ^F is positive. So our prediction from the theory model is that the platform should cross-subside anchors. In reality, p^A includes commission fee, tax and platform subsidy&benefits (negative p^A). According to a research report from Essence Securities, commission fee and tax remains fixed (about 0.6 dollars in total for every dollar worth of gift transaction) throughout our study period, however, we do observe a drastic increase in platform subsidy. For example, between Feb 7th 2020 and March 31st 2020, the platform launched a special program, which spent 1 billion RMB in subsiding streamers. In addition, the platform also launched a training program to help streamers. The prices for fans p^F are more implicit and subtle, including various forms of gift discounts, advertisement density per live video etc. Due to the data limitation, it's hard to track the change of p^F. But as far as we know, there is no significant fans subsidy policy throughout our study period. In a word, the platform pricing in reality is mostly consistent with our theory prediction and the estimation results.

3.5.4 Informal Investigation of Quasi-demand

To better understand how COVID-19 shocks affect the equilibrium volume of gifts, we proceed with an informal investigation of the quasi-demand in equilibrium. We cannot directly observe the quasi-demand (or engagement) and need to approximate them. Here we use the live video length as the proxy for anchors' engagement and the average number of audiences [1] as the proxy for fans' engagement. The results are shown in Table 3.4 and Table 3.5. We could see the basic conclusions still hold. The Covid-19 shock increases the quasi-demand of both anchors and fans and the main driver is the shock to the fans side.

Previously Table 3.6 has shown the depolarization of anchors' gift income distribution, however, other researches [Cortes and Forsythe, 2020] find Covid-19 exacerbates income polarization,

---

[1] It's defined as $\frac{\text{total number of audiences}}{\text{live video length}}$

where low-skill workers are more likely to lose their jobs under the Covid-19 shock. To better understand such contradiction, we further examine the heterogeneity of streamers' quasi-demand or engagement. We find a similar conclusion here, where the labor participation of bottom streamers is more responsive to the Covid-19 shock. One possible story to resolve this puzzle is that digital platforms are sanctuaries for those relatively low-skilled or unlucky people. When their outside options are deteriorated, the bottom streamers would rationally spend more time on the live-streaming platform.

Table 3.4: The Impact of COVID-19 on Anchors' Quasi-demand in Equilibrium

	(1)	(2)	(3)
COVID-19 intensity (Anchor)	-0.000**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
COVID-19 intensity (Fans)	0.481***	0.493***	0.497***
	(0.060)	(0.060)	(0.060)
Female		0.103***	0.102***
		(0.004)	(0.004)
#Lockdown policies (Anchor)			-0.002***
			(0.000)
#Lockdown policies (Fans)			0.002***
			(0.000)
Constant	8.611***	8.577***	8.587***
	(0.002)	(0.003)	(0.004)
Time dummies			✓
Observations	194,505		

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) The dependant variable is the logarithm of the live video length.

3.6 Conclusion

In this paper, we studied the heterogeneous (quantile) treatment effects of two-sided COVID-19 shocks on earnings. We first modified Rochet and Tirole [2003]'s model to theoretically analyze the effects. The total platform demand depends on the quasi-demand of anchors and fans. One

Table 3.5: The Impact of COVID-19 on Fans' Quasi-demand in Equilibrium

	(1)	(2)	(3)
COVID-19 intensity (Anchor)	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
COVID-19 intensity (Fans)	2.175***	2.187***	2.031***
	(0.107)	(0.107)	(0.107)
Female		0.097***	0.102***
		(0.008)	(0.004)
#Lockdown policies (Anchor)			0.018***
			(0.001)
#Lockdown policies (Fans)			-0.009***
			(0.001)
Constant	7.404***	7.371***	7.205***
	(0.004)	(0.005)	(0.006)
Time dummies			✓
Observations	194,505		

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) The dependant variable is the logarithm of the average number of audiences.

Table 3.6: The Heterogeneous Impact of COVID-19 on Anchors' Quasi-demand in Equilibrium

	(1)	(2)	(3)
Percentile			
20			
Anchor side	-0.000***	-0.000***	0.001
	(0.000)	(0.000)	(0.001)
Fans side	0.513***	0.496***	0.151***
	(0.097)	(0.098)	(0.055)
40			
Anchor side	-0.000***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)
Fans side	0.477***	0.469***	0.183***
	(0.064)	(0.063)	(0.036)
60			
Anchor side	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Fans side	0.424***	0.434***	0.178***
	(0.050)	(0.047)	(0.027)
80			
Anchor side	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Fans side	0.268***	0.346***	0.145***
	(0.045)	(0.047)	(0.027)
Additional Control variables			
Female		✓	✓
#Lockdown policies			✓
#Reopening policies			✓
Time dummies			✓
Observations	194505	194505	194505

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) The dependant variable is the logarithm of the live video length. (3) Here we just use the usual quantile regression instead of GQR, which means we don't explore the treatment effect.

prediction is that under the fans side shock, the platform tends to cross-subside the anchors. With the uniform distribution assumption, the welfare of both anchors and fans would increase. In reality, we do observe that the live-streaming platform subsided anchors, which is consistent with our theory model. To empirically estimate how the transaction volume (in our setting, gifting behaviors or gift earnings) in a two-sided market can be affected by COVID-19 shocks, we obtained a unique live-streaming dataset from Kuaishou using web spiders and investigated the effects empirically. To accommodate multiple continuous treatments at different quantiles, we resort to Powell [2013, 2019]'s GQR framework. Results suggest that, COVID-19 shocks from the fans side encouraged gifting behaviors, while COVID-19 shocks from the anchor side did not eventually increase gift earnings. We also find COVID-19 shocks lead to the de-polarization of the anchor gift income distribution. An informal quasi-demand analysis revealed that COVID-19 shocks increased the labor participation of the bottom streamers.

Admittedly, there are a few directions that need to be considered in future research. Firstly, we do not directly test the mediating effect of platform prices, and we hope that future studies should try to obtain information on the transaction fees or develop a way to estimate these fees and see if they are consistent with the theoretical model. Secondly, we mainly focus on the heterogeneous treatment effects on transactions along the distribution of transactions, while there might be sources of heterogeneity from other dimensions such as race, education, and/or age. However, due to the limitation of our data, we have yet to further explore these other dimensions. Thirdly, our approximation of the quasi-demand might be problematic. An important drawback is we cannot measure some unobserved engagement (e.g. anchors' efforts). More advanced data crawling&mining techniques can be adopted to acquire and extract more information on live-streaming videos.

With anchors on live-streaming platforms becoming a trending type of labor force and platform-based markets becoming increasingly mainstream, we believe that there shall be more theoretical and empirical studies using information on both sides of a platform and incorporate the two-sided market structure in the emerging industries, when studying the effects of shocks or policies. The trend of polarization can also be different in a two-sided labor market, which can be a future direction of research.

Chapter 4

MARKET COMPETITION AND PRODUCT ASSORTMENT IN A LIVE E-COMMERCE PLATFORM: A DEEP LEARNING APPROACH

4.1 Introduction

How does the market structure affect product variety? This is a classic question in economic theory. However, different theoretical models have very different predictions [Stole, 2007]. This paper aims to provide new empirical evidence to this topic.

In the past few years, a new retailing format has become popular in China: live streamers started to sell all kinds of products during their live broadcasting. The categories of the products they sell are widely ranged, from groceries to appliances. Thus, we could consider such live-streaming platform as a retail market. This paper studies the relationship between market competition and product variety in Kuaishou, which is the second largest live-streaming platform in China. We collect the video and personal information of the top streamers[1] between February and June in 2020.

The main challenges are (1) how to measure market competition and (2) how to address the endogeneity of such measurement and get a clean identification of the treatment effect. We use a novel multi-task graph convolutional network (GCN) approach [Kipf and Welling, 2016a] to get the embedding (latent) vector of each streamer and use embedding similarity to proxy the intensity of market competition. The market could be represented as a streamer competition graph and GCN is the state-of-the art deep learning approach to extract useful graph information. We also resort to multi-task learning to incorporate widely accepted economic theories into the model. To address endogeneity, we leverage a natural experiment, where a top streamer (Xinba) with strong market power quits the market by accident. This event brings an exogenous shock to competition among remaining streamers. The intensity of the shock has both longitudinal and cross-sectional variations, which depend on the product portfolio of each streamer.

[1] Around 1,000 most popular streamers based on the number of fans.

Our work is related to several strands of literature. First, this paper contributes to the empirical research about product variety or assortment. Most studies [McManus, 2007, Draganska et al., 2009, Sweeting, 2010, 2013, Fan, 2013, Illanes and Moshary, 2020] find that a decrease in market competition (e.g., mergers and acquisitions) leads to lower product variety. As pointed out by Berry and Waldfogel [2001], the impact of market competition on product variety is theoretically ambiguous, but they find that lower market competition increases product variety empirically. Similarly, Wollmann [2018] finds after acquisitions, firms would increase product variety as long as the sunk cost of introducing products is low enough. Our paper leverages a natural experiment to explore how prices and product variety would respond to an exogenous market competition shock. We discover that when facing lower market competition, retailers on a social e-commerce platform on average would reduce product variety but would not change prices.

Second, this paper contributes to the recent literature on uniform pricing. Recent studies conclude that uniform pricing widely exists among the retail chains in the United States [Adams and Williams, 2019, DellaVigna and Gentzkow, 2019, Hitsch et al., 2019, Illanes and Moshary, 2020]. We find a similar phenomenon exists in the digital retailing platform of an emerging country, where prices are sticky and do not adjust to exogenous shocks.

Last, we contribute to a growing marketing literature on using representation learning to study competition. For example, Ruiz et al. [2020] proposes a hierarchical latent variable model of ordered shopping baskets, which is called as SHOPPER. Chen et al. [2020] introduces Product2Vec to transform products into low-dimensional embedding vectors and adds it to the traditional discrete choice model. Gabel et al. [2019] proposes an exploratory approach by leveraging a neural network language model to derive product embedding and model market competition. Our contribution is to add multi-task learning and graph neural network to the literature.

The remainder of the paper is organized as follows: section 4.2 introduces the data sources, section 4.3 describes our graph convolutional network, section 4.4 shows the evaluation of our neural network results, section 4.5 presents our regression design and the estimation results, and section 4.6 concludes the paper.

4.2 Data

Our data is from Kuaishou, the second largest short video and live streaming platform in China. Nowadays, many streamers start to recommend and sell products during their live broadcasting. Live commerce is becoming an important format in China's online retail market. A recent research done by iRsearch estimates that the market size of China's live commerce in 2021 is about 320 billion dollars.

We collect the information of the top 1,000 subscription streamers in terms of the number of fans and their live videos, which includes (1) the general information of each streamer, e.g. gender, location, number of fans and self-introduction, (2) the general information of each live video, e.g., the number of watchers and comments, and (3) the sales information in each live video, including product names, categories, prices and quantities. Some of the popular streamers on Kuaishou has their own apprentices and form their own "families". In the families, streamers sometimes show up in videos of each others and what they sell can be decided by the "parent" of the family. Thus, we manually collect this information and include the family dummy variables to control the family fixed effects.

We focus on the cosmetic products that are sold by the top 1,000 streamers in our study period. 317 out of the 1,000 streamers have sold at least one product during this time span, and thus we consider only the 317 streamers as our retailers. Note that, if a streamer ever sells any product, the streamer usually sells more than just one products in each video. A streamer can also have multiple live streaming videos in a week selling overlapping products. In this paper, we are interested in three dependant variables. Their definitions are as follows: (1) product variety, which is the number of distinct cosmetic products that have been put on the shelf (may or may not have a deal) within a week by a streamer, (2) average cosmetic price, which is the simple average of the prices of the distinct cosmetic products deployed in each week, and (3)market share, which is the ratio of the weekly sales (price multiply by quantity) of cosmetic products of a streamer to that of all the 317 streamers. See Table 4.1 for summary statistics.

Distinct products are determined by the names provided in the live-streaming videos. As the exact string of the same product sold by different streamers can be slightly different, we use the Python library "FuzzyWuzzy" to calculate string similarity and group similar strings together.

Then, we manually check whether those strings are really mapped to the same product.

Our cosmetic data include 16 subcategories such as lip products and eye liners. We use a vector to record the weekly sales information of each streamer. For example, if streamer i sold lipstick A twice (i.e., in two videos), lipstick B once and eye liner A once within a week, the vector would be $h_i = (3, 1, 0, \ldots, 0)$. Each element of the feature indicates the frequency of the sales of a specific cosmetic subcategory. See Table 4.2 for the sales information of all subcategories.

We use the account suspension of Xinba, the most popular streamer on Kuaishou, as an exogenous shock to identify the effect of competition on market structure. Figure 4.1 shows that Xinba accounts for the majority of the large weekly market shares. The account suspension is due to a random war of words between Xinba and another popular streamer Sanda. The war of words between these two streamers were live streamed and involved "inappropriate" contents. As a result, their accounts were both suspended during April and June in 2020 by the platform manager.

Table 4.1: Summary Statistics of Some Key Variables

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Number of fans (mil)	7	5.7	8.5	3.7	8.4	61
Average attendance (k)	7.9	16.1	0.0	0.8	7.6	188.6
Product Variety	4.704	4.611	0	1	7	54
Average Cosmetic Price (CNY)	40.381	42.465	0	16.3	51.2	688
Market share (of sales)	0.010	0.041	0.000	0.0001	0.005	0.696

Notes: (1) N = 2,233; (2) CNY = Chinese Yuan; (3) Pctl = Percentile.

4.3 Multi-Task Graph Convolutional Network

We construct an undirected weighted graph, which summarizes the sales information of the live streamers. The nodes represent the live streamers $i = 1 \ldots N$. The edge between any two nodes measures their similarity, which is a proxy of the competition relationship. To obtain edge weights, at first we create a 16-dimensional vector h_{it} for each streamer-week pair, counting the number of live commerce events for 16 different cosmetic subgroups, as we described in section 4.2. Then the

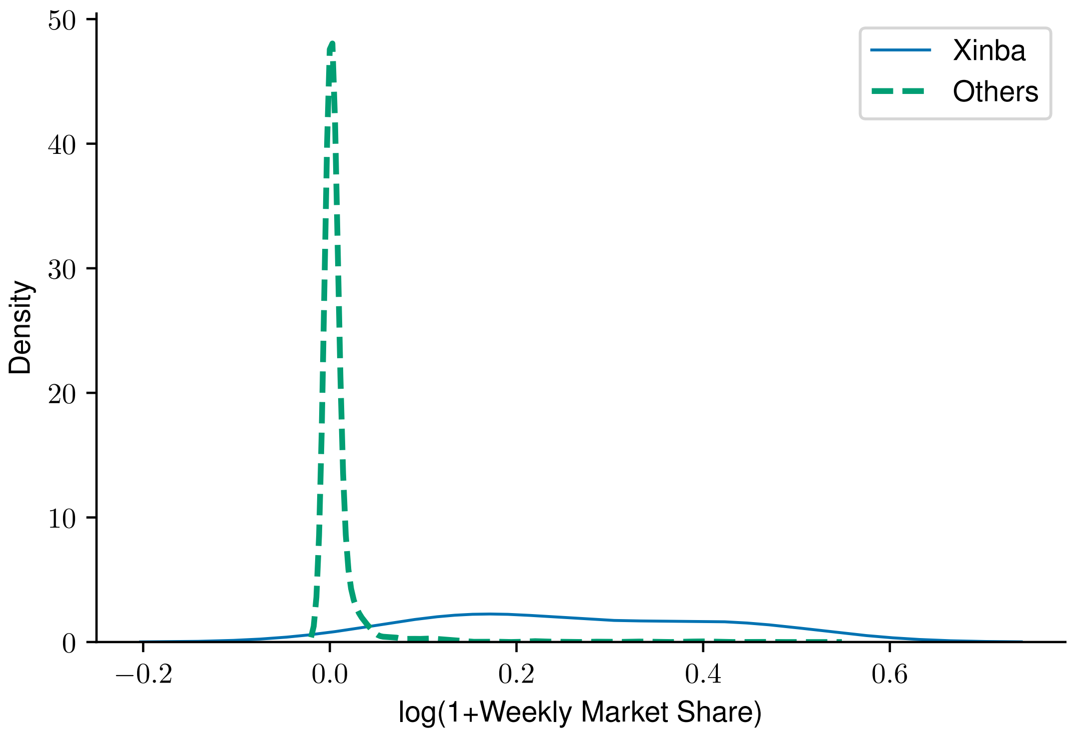

Figure 4.1: Market Shares between Xinba versus All Other Streamers

Table 4.2: Sales Information of All Subcategories

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Lip	2.548	5.299	0	0	3	62
Eye liner	0.932	2.657	0	0	1	30
Eye shadow	0.923	2.319	0	0	1	30
Cheek	0.116	0.517	0	0	0	8
Detailing	0.876	3.236	0	0	0	60
Foundation	2.929	5.046	0	0	3	52
Setup	0.012	0.155	0	0	0	5
Cleaning	0.578	1.653	0	0	0	28
Eye lash	0.566	1.474	0	0	0	16
Eye brow	0.523	1.538	0	0	0	25
Tool	0.390	1.678	0	0	0	21
Pre	0.669	2.071	0	0	1	26
Hyalironic	0.056	0.545	0	0	0	10
Double eye lid	0.035	0.452	0	0	0	9
Nail	0.011	0.152	0	0	0	5
Combo	0.031	0.296	0	0	0	7

Notes: (1) N = 2,233; (2) Pctl = Percentile.

edge weight between node i and j at week t is their cosine similarity: $e_{ijt} = \frac{h_{it} \cdot h_{jt}}{\|h_{it}\| \|h_{jt}\|}$.

There are several reasons why we do not directly use this naive specification as our competition index: (1) this index may fail to capture the implicit competition relationship when the similarity score is 0, (2) the naive pairwise comparison does not take the complicated network structure into account, and (3) this index does not make use of some widely accepted economic theories—e.g., the degree of market competition is negatively correlated with market share.

To address these problems, we use the so-called multi-task graph convolutional network approach. Graph convolutional network (GCN) is a state-of-the-art deep learning approach to make use of our complex graph data. It has been widely used for graph representation learning [Berg et al., 2017, Kipf and Welling, 2016a,b, Schlichtkrull et al., 2018]. GCN could help to address problems (1) and (2). Our goal is to use GCN to learn a set of weekly $N \times K$ streamer feature matrices H_t (N is the number of nodes, while K is the number of input features) and use it to construct a weekly market competition index. To address problem (3), we resort to multi-task learning, which adds an additional economic theory-guided loss to the original GCN loss function. In our case, it is the distance between the actual market share and the market share prediction based on the degree of market competition.

4.3.1 Network Architecture

Following the propagation rule introduced in Kipf and Welling [2016a], the ℓ-th convolutional layer is defined as:

$$H_t^{(\ell+1)} = \sigma\left(\tilde{D}_t^{-\frac{1}{2}}(A_t + I)\tilde{D}_t^{-\frac{1}{2}} H_t^{(\ell)} W^{(\ell)}\right)$$

where A_t is the adjacency matrix, I is the identity matrix, $\tilde{D}_t$ is the diagonal node degree matrix of $A_t + I$, $W^{(\ell)}$ is the weight matrix for the ℓ-th convolutional layer (no subscript t), $\sigma(\cdot)$ is the ReLu activtion function. We have two convolutional layers in total. We randomly initialize $H_t^{(0)}$ for $t = 0, 1, ..., 21$.

In the final stage we have two tasks:

- *Link Prediction*

  We reconstruct the adjacency matrix $\hat{A}_t = \text{CosineSimilarity}\left(H_t^{(L)}, H_t^{(L)}\right)$ which is our prediction of edge weights.

- *Market Share Prediction*

  For simplicity, our prediction of the market share vector is:

$$\hat{s}_t = \text{Softmin}\left(H_t^{(L)}\left(H_t^{(L)}\right)^T \mathbf{1} - \text{diag}\left(H_t^{(L)}\left(H_t^{(L)}\right)^T\right)\right)$$

For simplicity, we do not introduce any new parameter in the final stage. But potentially we could also add several fully connected layers here.

4.3.2 Loss Function

Our loss function consists of two parts: link prediction loss $\mathcal{L}_l$ and market share prediction loss $\mathcal{L}_m$. We define

$$\mathcal{L}_l = \sum_{t=0}^{21}\sum_{i,j=1}^{N}\left(A_{i,j,t} - \hat{A}_{i,j,t}\right)^2$$

and

$$\mathcal{L}_s = \sum_{t=0}^{21}\sum_{i=1}^{N}\left(s_{i,t} - \hat{s}_{i,t}\right)^2.$$

For multitask learning, it is important to specify the weights of different task losses. Here we use the task-dependant uncertainty weighting approach in Kendall et al. [2018]: σ_1 and σ_2 are uncertainty weighting parameters that are automatically learned from the network training:

$$\mathcal{L} = \frac{1}{2\sigma_1^2}\mathcal{L}_l + \frac{1}{2\sigma_2^2}\mathcal{L}_s + \log(\sigma_1\sigma_2).$$

4.4 Model Evaluation

In addition to the GCN loss function, we use a economics-related metric to evaluate our model. According to a well-established economic theory, the degree of market competition that a streamer faces is negatively correlated with its market share. The market competition index of streamer i in market t is defined as follows:

$$Comp_{i,t} = \frac{e^{\sum_j h_{i,t}\cdot h_{j,t}}}{\sum_i e^{\sum_j h_{i,t}\cdot h_{j,t}}}$$

We run the following regression to test the correlation between competition and market share:

$$y_{i,t} = \beta_0 + \beta_1 Comp_{i,t} + \mathbf{X}'_{i,t}\beta + \epsilon_{i,t}. \tag{4.1}$$

The result is shown in Table 4.3. The results show that the impact of competition on market share is negative at the 5% significance level.

Table 4.3: The Correlation between Competition and Market Share

	Dependent variable: market share			
	(1)	(2)	(3)	(4)
Comp	-0.2258**	-0.2076**	-0.2078**	-0.2151**
	(0.096)	(0.096)	(0.095)	(0.095)
Female		0.0049***	0.0057***	0.0056***
		(0.002)	(0.002)	(0.002)
Number of Fans			0.0000***	0.0000***
			(0.000)	(0.000)
Number of Videos				-0.0000**
				(0.000)
Constant	0.0076***	0.0041***	-0.0034*	-0.0021
	(0.001)	(0.001)	(0.002)	(0.002)
Family FE	Yes	Yes	Yes	Yes
Observations	2,233	2,233	2,233	2,233

Notes: (1) *p<0.1, **p<0.05, and ***p<0.01; (2) higher Comp means more competition.

4.5 DiD Design and Results

To better leverage the exogenous event in which Xinba's account is suspended by accident, we define our treatment variable as the average competition level between a streamer and Xinba before Xinba's account suspension, i.e.,

$$\text{Comp\_Xinba}_i = \frac{1}{13}\sum_{t=0}^{12} \frac{h_{i,t} \cdot h_{\text{Xinba},t}}{\|h_{i,t}\|\|h_{\text{Xinba},t}\|}.$$

4.5.1 Main Specification

Table 4.4: The Treatment Effect of Market Competition Shock on Product Variety

	Dependent variable: product variety				
	(1)	(2)	(3)	(4)	(5)
Comp_Xinba $\times$ I_{post}	2.5594***	2.6049***	2.7669***	2.7150***	2.7015***
	(0.818)	(0.820)	(0.805)	(0.805)	(0.805)
Comp_Xinba	0.2150	0.1802	-0.3170	-0.2706	-0.2658
	(0.679)	(0.684)	(0.673)	(0.673)	(0.673)
I_{post}	-1.5469**	-1.5914***	-1.6090***	-1.6051***	-1.5970***
	(0.610)	(0.612)	(0.600)	(0.600)	(0.600)
Female			2.0098***	2.1237***	2.0365***
			(0.214)	(0.214)	(0.214)
Number of Fans				0.000**	0.000**
				(0.000)	(0.000)
Average Number of Watchers per Video					-0.000
					(0.000)
Constant	4.4954***	4.5226***	3.3961***	3.1197***	3.1082***
	(0.513)	(0.517)	(0.521)	(0.531)	(0.539)
Family FE	No	Yes	Yes	Yes	Yes
Observations	2,233	2,233	2,233	2,233	2,233

Notes: (1) *p<0.1, **p<0.05, and ***p<0.01;

We adopt a difference-in-differences (DiD) strategy similar to **?**. We compare the relative change in product variety and price in the post-suspension period relative to the pre-suspension period between streamers who were similar to Xinba (and thus competed with Xinba) and those who were not. Our measure of the treatment intensity is continuous and thus can capture more variation in the data. We assume that a streamer's product variety and average price are proportional to our measure of competition between the streamer and Xinba in the main estimating equation. Thus, we have

$$y_{i,t} = \alpha + \beta \,\text{Comp\_Xinba}_i \times \text{I}_{post,t} + \gamma \,\text{Comp\_Xinba}_i + \rho \,\text{I}_{post,t} + \mathbf{X}'_{i,t}\Phi + \epsilon_{i,t} \tag{4.2}$$

where i indexes streamers and t represents each week (i.e., market), which are from 0 to 21, and $\text{I}_{post,t}$ is an indicator variable that equals one for the periods after week 13 (when Xinba's account was suspended). The outcome of interest, denoted by $y_{i,t}$, is either the weekly market share or the average cosmetic price. We also include family fixed effects and time-varying streamer-specific and video-specific characteristics in $\mathbf{X}_{i,t}$.

We are interested in the coefficient β in Equation (4.2). As shown in Table 4.4 and Table 4.5, there is strong evidence of a positive effect of market competition on product assortment, but no clear evidence of an effect on price change.

4.5.2 Flexible Estimation (Pre-Trend)

In Equation (4.2), we estimate the average effect of the competition between Xinba and other streamers on product variety and average price. Our main specification requires that during our study period the major event that affects competition is Xinba's exit of the market. However, as we all know, the digital world is ever-changing, and various events can happen (especially when our study period coincides with the COVID-19 pandemic). Thus, we adopt a more flexible estimating equation that takes the form below:

$$y_{i,t} = \alpha + \sum_{j=0}^{21} \beta_j \,\text{Comp\_Xinba}_i \times \text{I}_{j,t} + \sum_{j=0}^{21} \rho_j \,\text{I}_{j,t} + \mathbf{X}'_{i,t}\Phi + \epsilon_{i,t} \tag{4.3}$$

The difference between Equation (4.3) and Equation (4.2) is that we interact our competition measure Comp_Xinba_i with each of the week/market fixed effects. As a result, β_js reveal the correlation between competition and the outcomes of our interest in each week/market. We are interested in the pattern of the estimated β_js over the weeks. Particularly, we would like to see if

Table 4.5: The Treatment Effect of Market Competition Shock on Prices

	Dependent variable: prices				
	(1)	(2)	(3)	(4)	(5)
Comp_Xinba $\times$ I_{post}	-3.6128	-3.6986	-3.0458	-3.8756	-2.4645
	(7.588)	(7.507)	(7.483)	(7.468)	(7.431)
Comp_Xinba	-4.9487	-3.6881	-5.6912	-4.9485	-5.4489
	(6.294)	(6.263)	(6.261)	(6.249)	(6.214)
I_{post}	0.0303	0.4242	0.3533	0.4158	-0.4324
	(5.660)	(5.599)	(5.579)	(5.566)	(5.537)
Female			8.0965***	8.4780***	8.1787***
			(1.986)	(1.984)	(1.974)
Number of Fans				0.000***	0.000
				(0.000)	(0.000)
Average Number of Watchers per Video					0.0004***
					(0.000)
Constant	44.9075***	42.6169***	38.0784***	33.6561***	34.8559***
	(4.755)	(4.731)	(4.845)	(4.998)	(4.976)
Family FE	No	Yes	Yes	Yes	Yes
Observations	2,233	2,233	2,233	2,233	2,233

Notes: (1) *p<0.1, **p<0.05, and ***p<0.01;

there is any discontinuity in the pattern around the week when Xinba's account was suspended for a period of time.

One thing that we would like to emphasize is that, although we choose week 0 as the benchmark market and set its coefficient to be 0, we are not interested in the individual magnitudes of the point estimates themselves. To better illustrate the pattern, we draw Figure 4.2 based on Table 4.6. As shown in the figure, there are no clear trends of the estimated interaction effects before Xinba's account suspension—they go up and down drastically centering around 0. After Xinba's account suspension, the effects start to center around a much larger value, 3. In week 15, which is two weeks after Xinba's account suspension, we observe a larger effect.

We can clearly see that, the relationship between direct competition with Xinba and product

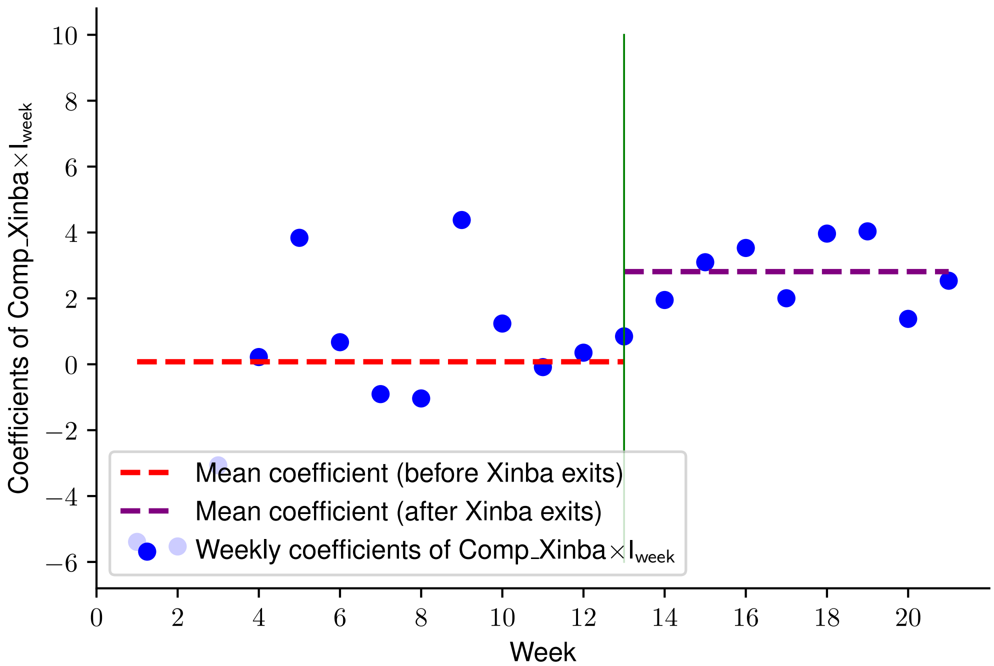

Figure 4.2: Flexible Estimates: The Impact of Competition Shock on Product Variety

variaty is not significant over time during week 0 to week 12 and then almost steadily increases in magnitude from week 13 to near the end of our observation window. After Xinba's account suspension, product assortment of streamers who directly competed with Xinba starts to diminish relative to streamers who did not compete with Xinba. The effect on product variety appears to lag behind Xinba's exit by approximately one to two weeks.

Table 4.6: Flexible Estimates: The Impact of Competition Shock on Product Variety

	Dependent variable: product variety	
	(1)	(2)
Comp_Xinba $\times$ I_1	-5.5324	-5.3918
	(3.973)	(3.978)

Continued on next page

Table 4.6 – continued from previous page

	(1)	(2)
Comp_Xinba × I_2	-5.6723	-5.5299
	(3.973)	(3.978)
Comp_Xinba × I_3	-3.2367	-3.0689
	(3.269)	(3.278)
Comp_Xinba × I_4	0.1897	0.2192
	(2.867)	(2.876)
Comp_Xinba × I_5	3.7863	3.8277
	(2.399)	(2.402)
Comp_Xinba × I_6	0.6980	0.6701
	(2.604)	(2.606)
Comp_Xinba × I_7	-0.9079	-0.9057
	(2.146)	(2.150)
Comp_Xinba × I_8	-1.0357	-1.0371
	(2.542)	(2.543)
Comp_Xinba × I_9	4.3163	4.3790*
	(2.633)	(2.635)
Comp_Xinba × I_{10}	1.4239	1.2360
	(1.709)	(1.725)
Comp_Xinba × I_{11}	0.0421	-0.0863
	(1.606)	(1.628)
Comp_Xinba × I_{12}	0.3383	0.3513
	(2.372)	(2.373)
Comp_Xinba × I_{13}	0.9464	0.8460
	(1.956)	(1.960)
Comp_Xinba × I_{14}	1.9632	1.9510
	(1.604)	(1.605)
Comp_Xinba × I_{15}	3.0938**	3.0975**
	(1.407)	(1.417)
Comp_Xinba × I_{16}	3.5010***	3.5302***
	(1.288)	(1.291)
Comp_Xinba × I_{17}	2.0168	2.0040

Continued on next page

Table 4.6 – continued from previous page

	(1)	(2)
	(1.335)	(1.336)
Comp_Xinba × I_{18}	3.9668***	3.9660***
	(1.272)	(1.274)
Comp_Xinba × I_{19}	3.9904***	4.0344***
	(1.168)	(1.174)
Comp_Xinba × I_{20}	1.2790	1.3780
	(1.521)	(1.525)
Comp_Xinba × I_{21}	2.5381*	2.5358*
	(1.297)	(1.298)
I_1	4.4220	4.3319
	(3.233)	(3.236)
I_2	3.6709	3.5665
	(3.236)	(3.239)
I_3	2.4392	2.3429
	(2.781)	(2.785)
I_4	-0.1966	-0.2069
	(2.403)	(2.407)
I_5	-2.0340	-2.0761
	(2.001)	(2.003)
I_6	-0.5737	-0.5505
	(2.271)	(2.273)
I_7	0.8196	0.8066
	(1.975)	(1.977)
I_8	1.0953	1.0927
	(2.211)	(2.212)
I_9	-3.0447	-3.1083
	(2.298)	(2.300)
I_{10}	-1.1625	-0.9930
	(1.684)	(1.695)
I_{11}	0.1304	0.2184
	(1.631)	(1.647)

Continued on next page

Table 4.6 – continued from previous page

	(1)	(2)
I_{12}	-0.2179	-0.2534
	(2.152)	(2.153)
I_{13}	-0.7121	-0.6499
	(1.851)	(1.854)
I_{14}	-0.9269	-0.9226
	(1.644)	(1.644)
I_{15}	-1.6777	-1.7002
	(1.519)	(1.526)
I_{16}	-1.8824	-1.9323
	(1.442)	(1.444)
I_{17}	-1.8403	-1.8392
	(1.467)	(1.468)
I_{18}	-2.3644*	-2.3647*
	(1.428)	(1.430)
I_{19}	-2.0225	-2.0562
	(1.384)	(1.388)
I_{20}	-0.9584	-1.0360
	(1.583)	(1.586)
I_{21}	-1.1707	-1.1659
	(1.441)	(1.442)
Constant	4.5882***	4.5882***
	(1.113)	(1.114)
Family FE	No	Yes
Observations	2,233	2,233

4.6 Conclusion

This paper proposes a novel approach to measure market competition, leveraging the recent methodological advances in machine learning. Through a natural experiment, we identify a treatment effect of market competition on product variety. Our results show that an online streamer would typically rely on product assortment rather than price adjustment when facing exogenous market competition shocks. More studies with the new tool in other contexts should be done to investigate the relationship between market competition and product assortment.

BIBLIOGRAPHY

Qifan Huang and Castiel Chen Zhuang. Training, productivity and wages: An investigation of china's manufacturing enterprises in a privatization era. *Economics of Transition and Institutional Change*, 30(2):269–288, 2022.

Daron Acemoglu and Jörn-Steffen Pischke. Why do firms train? Theory and evidence. *The Quarterly Journal of Economics*, 113(1):79–119, 1998.

Qifan Huang, Saizi Xiao, and Castiel Chen Zhuang. The heterogeneous impact of covid-19 in a two-sided market: Evidence from a live-streaming platform. *Applied Economics*, Forthcoming, 2022.

David Powell. Quantile treatment effects in the presence of covariates. *Review of Economics and Statistics*, (Just Accepted):1–39, 2019.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262, 1994.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, pages 841–890, 1995.

Toshiaki Iizuka. Experts' agency problems: evidence from the prescription drug market in Japan. *The RAND Journal of Economics*, 38(3):844–862, 2007.

Steven T Berry and Panle Jia. Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43, 2010.

Henrick Horn and Asher Wolinsky. Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, 19(3):408–419, 1988.

Gregory S. Crawford and Ali Yurukoglu. The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643–685, 2012.

Matthew Grennan. Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–177, 2013.

Gautam Gowrisankaran, Aviv Nevo, and Robert Town. Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105(1):172–203, 2015.

Kate Ho and Robin S. Lee. Insurer competition in health care markets. *Econometrica*, 85(2): 379–417, 2017.

Pierre Dubois, Ashvin Gandhi, and Shoshana Vasserman. Bargaining and international reference pricing in the pharmaceutical industry. Working paper, Havard University, 2019a.

Ernst R. Berndt, Linda Bui, David R. Reiley, and Glen L. Urban. Information, marketing, and pricing in the US antiulcer drug market. *American Economic Review*, 85(2):100–105, 1995.

Ernst R. Berndt, Robert S. Pindyck, and Pierre Azoulay. Consumption externalities and diffusion in pharmaceutical markets: Antiulcer drugs. *Journal of Industrial Economics*, 51(2):243–270, 2003a.

Julie M. Donohue and Ernst R. Berndt. Information content of advertising: Empirical evidence from the OTC analgesic industry. *International Journal of Industrial Organization*, 31(5):355–367, 2013.

Minjae Song, Sean Nicholson, and Claudio Lucarelli. Mergers with interfirm bundling: a case of pharmaceutical cocktails. *The RAND Journal of Economics*, 48(3):810–834, 2017.

Jonas Björnerstedt and Frank Verboven. Does merger simulation work? evidence from the Swedish analgesics market. *American Economic Journal: Applied Economics*, 8(3):125–164, 2016.

Pierre Dubois and Laura Lasio. Identifying industry margins with price constraints: Structural estimation on pharmaceuticals. *American Economic Review*, 108(12):3685–3724, 2018.

Kate Ho and Ariel Pakes. Hospital choices, hospital prices, and financial incentives to physicians. *American Economic Review*, 104(12):3841–84, 2014.

Fangwen Lu. Insurance coverage and agency problems in doctor prescriptions: evidence from a field experiment in China. *Journal of Development Economics*, 106:156–167, 2014.

Sofia Berto Villas-Boas. Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, 74(2):625–652, 2007.

Céline Bonnet and Pierre Dubois. Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance. *The RAND Journal of Economics*, 41(1):139–164, 2010.

Philip G Gayle. On the efficiency of codeshare contracts between airlines: is double marginalization eliminated? *American Economic Journal: Microeconomics*, 5(4):244–73, 2013.

Zhongliang Zhou, Yanfang Su, Benjamin Campbell, Zhiying Zhou, Jianmin Gao, Qiang Yu, Jiuhao Chen, and Yishan Pan. The financial impact of the 'Zero-Markup Policy for Essential Drugs' on patients in county hospitals in western rural China. *PLoS ONE*, 10(3):e0121630, 2015.

Hongmei Yi, Grant Miller, Linxiu Zhang, Shaoping Li, and Scott Rozelle. Intended and unintended consequences of China's Zero Markup Drug Policy. *Health Affairs*, 34(8):1391–1398, 2015.

Hongqiao Fu, Ling Li, and Winnie Yip. Intended and unintended impacts of price changes for drugs and medical services: Evidence from China. *Social Science & Medicine*, 211:114–122, 2018.

Zhaolin Meng, Min Zhu, Yuanyi Cai, Xiaohong Cao, and Huazhang Wu. Effect of a typical systemic hospital reform on inpatient expenditure for rural population: the Sanming model in China. *BMC Health Services Research*, 19(1):231, 2019.

Yan Yang, Jia hui Zhou, Xuan Zou, Xin yu Liu, Min xing Chen, Jiang jiang He, Li xuan Gong, and Chun lin Jin. Study on the effectiveness of GPO in reducing drug costs in Shenzhen [In Chinese]. *Chinese Journal of Health Policy*, 13(1):57–61, 2020.

Simona Baldi and Davide Vannoni. The impact of centralization on pharmaceutical procurement prices: the role of institutional quality and corruption. *Regional Studies*, 51(3):426–438, 2017.

Mark Duggan and Fiona M. Scott Morton. The distortionary effects of government procurement: Evidence from Medicaid prescription drug purchasing. *Quarterly Journal of Economics*, 121(1): 1–30, 2006.

Ministry of Health. Notice on the issuance of the centralized drug procurement in health facilities [In Chinese]. http://www.nhc.gov.cn/yaozs/s3573/201007/ea413230b3714b45b5b724f7bae84884.shtml, July 2010.

State Council's General Office. State council office's notice on establishing and standardizing essential drug procurement in government-sponsored primary health facilities [In Chinese]. http://www.gov.cn/xxgk/pub/govpublic/mrlm/201012/t20101208_63095.html, July 2010.

Pierre Dubois, Yassine Lefouili, and Stéphane Straub. Pooled procurement of drugs in low and middle income countries. Working Paper 508, Center for Global Development, 2019b.

State Council's General Office. Guiding opinions of the general office of the state council on urban public hospital comprehensive reform pilot [In Chinese]. http://www.gov.cn/zhengce/content/2015-05/17/content_9776.htm, May 2015.

State Council's Healthcare Reform Committee. Opinion on the implementation of the 'two invoices' system in the procurement of pharmaceutical products by public medical institutions (trial) [In Chinese]. http://www.nmpa.gov.cn/WS04/CL2196/324173.html, December 2016.

Kenneth A Small and Harvey S Rosen. Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, pages 105–130, 1981.

Ariel Pakes. Empirical tools and competition analysis: Past progress and current problems. *International Journal of Industrial Organization*, 53:241–266, 2017.

Thomas G Wollmann. Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406, 2018.

Jozef Konings and Stijn Vanormelingen. The impact of training on productivity and wages: firm-level evidence. *Review of Economics and Statistics*, 97(2):485–497, 2015.

Qing Liu and Ruosi Lu. On-the-job training and productivity: Firm-level evidence from a large developing country. *China Economic Review*, 40:254–264, 2016.

Gary S. Becker. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (1st Edition)*. NBER, 1964.

Masanori Hashimoto. Minimum wage effects on training on the job. *American Economic Review*, 72(5):1070–1087, 1982.

Gabriella Conti. Training, productivity and wages in Italy. *Labour Economics*, 12(4):557–576, 2005.

Lorraine Dearden, Howard Reed, and John Van Reenen. The impact of training on productivity and wages: Evidence from British panel data. *Oxford Bulletin of Economics and Statistics*, 68 (4):397–421, 2006.

J. David Brown, John S. Earle, and Telegdym Almos. The productivity effects of privatization: longitudinal estimates from Hungary, Romania, Russia and Ukraine. *Journal of Political Economy*, 114:61–99, 2006.

Saul Estrin, Jan Hanousek, Evzen Kocenda, and Jan Svejnar. The effects of privatization and ownership in transition economies. *Journal of Economic Literature*, 47(3):699–728, 2009.

Belton M. Fleisher, Klara Sabirianova, and Xiaojun Wang. Returns to skills and the speed of reforms: Evidence from central and eastern europe, china, and russia. *Journal of Comparative Economics*, 33(2):351–370, 2005.

Daniel Münich, Jan Svejnar, and Katherine Terrell. Returns to human capital under the communist wage grid and during the transition to a market economy. *Review of Economics and Statistics*, 87(1):100–123, 2005.

Daniel A. Ackerberg, Kevin Caves, and Garth Frazer. Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451, 2015.

Judith K. Hellerstein, David Neumark, and Kenneth R. Troske. Wages, productivity, and worker characteristics: Evidence from plant-level production functions and wage equations. *Journal of Labor Economics*, 17(3):409–446, 1999.

Jacob A. Mincer. *Schooling, Experience, and Earnings.* NBER, 1974.

Jacob Marschak and William H. Andrews Jr. Random simultaneous equations and the theory of production. *Econometrica*, 12(3/4):143–205, 1944.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

Zheng Song, Kjetil Storesletten, and Fabrizio Zilibotti. Growing like China. *American Economic Review*, 101(1):196–233, 2011.

Loren Brandt, Johannes van Biesebroeck, Luhang Wang, and Yifan Zhang. WTO accession and performance of Chinese manufacturing firms. *American Economic Review*, 107(9):2784–2820, 2017.

Yang Yao and Ninghua Zhong. Unions and workers' welfare in Chinese firms. *Journal of Labor Economics*, 31(3):633–667, 2013.

Karin Halldén. Taking training to task: Sex of the immediate supervisor and men's and women's time in initial on-the-job training. *Work and Occupations*, 42(1):73–102, 2015.

Jean-Charles Rochet and Jean Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.

Daron Acemoglu and David Autor. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4, pages 1043–1171. Elsevier, 2011.

Ronald Bachmann, Gökay Demir, and Hanna Frings. Labor market polarization, job tasks and monopsony power. *Journal of Human Resources*, pages 0219–10011R1, 2021.

H David and David Dorn. The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, 103(5):1553–97, 2013.

Christian Vom Lehn. Labor market polarization, the decline of routine work, and technological change: A quantitative analysis. *Journal of Monetary Economics*, 110:62–80, 2020.

Guido Matias Cortes and Eliza Forsythe. The heterogeneous labor market impacts of the covid-19 pandemic. *Available at SSRN 3634715*, 2020.

Sergio Firpo. Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, 75 (1):259–276, 2007.

Mark Armstrong. Competition in two-sided markets. *RAND Journal of Economics*, 37(3):668–691, 2006.

Jean-Charles Rochet and Jean Tirole. Two-sided markets: a progress report. *The RAND journal of economics*, 37(3):645–667, 2006a.

Marc Rysman. The economics of two-sided markets. *Journal of Economic Perspectives*, 23(3): 125–143, 2009.

Scott R. Baker, Robert A. Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis. How does household spending respond to an epidemic? consumption during the 2020 COVID-19 pandemic. *Review of Asset Pricing Studies*, raaa009, 2020.

Hung-Hao Chang and Chad D Meyerhoefer. Covid-19 and the demand for online food shopping services: Empirical evidence from taiwan. *American Journal of Agricultural Economics*, 103(2): 448–465, 2021.

Edward L. Glaeser, Ginger Zhe Jin, Benjamin T. Leyden, and Michael Luca. Learning from deregulation: The asymmetric impact of lockdown and reopening on risky behavior during COVID-19. NBER Working Paper No. 27650, 2020.

Manav Raj, Arun Sundararajan, and Calum You. Covid-19 and digital resilience: Evidence from uber eats, 2020.

Carola Binder. Coronavirus fears and macroeconomic expectations. *Review of Economics and Statistics*, (Early Access):1–10, 2020.

Raj Chetty, John N. Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. How did COVID-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data. NBER Working Paper No. 27431, 2020.

Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. The cost of the COVID-19 crisis: Lockdowns, macroeconomic expectations, and consumer spending. NBER Working Paper No. 27141, 2020a.

Òscar Jordà, Sanjay R. Singh, and Alan M. Taylor. Longer-run economic consequences of pandemic. NBER Working Paper No. 26934, 2020.

Marcus R. Keogh-Brown, Simon Wren-Lewis, W. John Edmunds, Philippe Beutels, and Richard D. Smith. The possible macroeconomic impact on the UK of an influenza pandemic. *Health Economics*, 19(11):1345–1360, 2010.

Robert J. Barro, José F. Ursúa, and Joanna Weng. The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity. NBER Working Paper No. 26866, 2020.

Abi Adams-Prassl, Teodora Boneva, Marta Golin, and Christopher Rauh. The impact of the coronavirus lockdown on mental health: Evidence from the US. Human Capital and Economic Opportunity Global Working Group Working Paper 2020-030, 2020.

Abel Brodeur, Andrew Clark, Sarah Fleche, and Nattavudh Powdthavee. COVID-19, lockdowns and well-being: Evidence from Google trends. IZA Discussion Paper No. 13204, 2020.

Pablo de Pedraza, Martin Guzi, and Kea Tijdens. Life dissatisfaction and anxiety in COVID-19 pandemic. GLO Discussion Paper 544, 2020.

Colleen P Kirk and Laura S Rifkin. I'll trade you diamonds for toilet paper: Consumer reacting, coping and adapting behaviors in the covid-19 pandemic. *Journal of Business Research*, 117: 124–131, 2020a.

Guido Lorenzoni. A theory of demand shocks. *American Economic Review*, 99(5):2050–2084, 2009.

Sangmin Aum, Sang Yoon (Tim) Lee, and Yongseok Shin. COVID-19 doesn't need lockdowns to destroy jobs: The effect of local outbreaks in Korea. NBER Working Paper No. 27264, 2020.

Truc Thi Mai Bui, Patrick Button, and Elyce G. Picciotti. Early evidence on the impact of COVID-19 and the recession on older workers. NBER Working Paper No. 27448, 2020.

Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. Labor markets during the COVID-19 crisis: A preliminary view. NBER Working Paper No. 27017, 2020b.

Eliza Forsythe, Lisa B. Kahn, Fabian Lange, and David G. Wiczer. Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims. *Journal of Public Economics*, 189: 104238, 2020.

Houseman Susan N. Autor, David H. and Sari Pekkala Kerr. The effect of work first job placements on the distribution of earnings: An instrumental variable quantile regression approach. *Journal of Labor Economics*, 35(1):149–190, 2017.

David Powell and Dana Goldman. Disentangling moral hazard and adverse selection in private health insurance. *Journal of Econometrics*, 222(1):141–160, 2021.

David Powell. A new framework for estimation of quantile treatment effects: Nonseparable disturbance in the presence of covariates. RAND Working Paper Series WR-824-1, 2013.

Nathan Kallus, Xiaojie Mao, and Masatoshi Uehara. Localized debiased machine learning: Efficient inference on quantile treatment effects and beyond, 2020.

Lars A Stole. Price discrimination and competition. *Handbook of industrial organization*, 3:2221–2299, 2007.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016a.

Brian McManus. Nonlinear pricing in an oligopoly market: The case of specialty coffee. *The RAND Journal of Economics*, 38(2):512–532, 2007.

Michaela Draganska, Michael Mazzeo, and Katja Seim. Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *QME*, 7(2):105–146, 2009.

Andrew Sweeting. The effects of mergers on product positioning: evidence from the music radio industry. *The RAND Journal of Economics*, 41(2):372–397, 2010.

Andrew Sweeting. Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica*, 81(5): 1763–1803, 2013.

Ying Fan. Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628, 2013.

Gastón Illanes and Sarah Moshary. Market structure and product assortment: Evidence from a natural experiment in liquor licensure. Technical report, National Bureau of Economic Research, 2020.

Steven T Berry and Joel Waldfogel. Do mergers increase product variety? evidence from radio broadcasting. *The Quarterly Journal of Economics*, 116(3):1009–1025, 2001.

Brian Adams and Kevin R Williams. Zone pricing in retail oligopoly. *American Economic Journal: Microeconomics*, 11(1):124–56, 2019.

Stefano DellaVigna and Matthew Gentzkow. Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, 134(4):2011–2084, 2019.

Günter J Hitsch, Ali Hortacsu, and Xiliang Lin. Prices and promotions in us retail markets: Evidence from big data. Technical report, National Bureau of Economic Research, 2019.

Francisco JR Ruiz, Susan Athey, and David M Blei. Shopper: A probabilistic model of consumer choice with substitutes and complements. *The Annals of Applied Statistics*, 14(1):1–27, 2020.

Fanglin Chen, Xiao Liu, Davide Proserpio, Isamar Troncoso, and Feiyu Xiong. Studying product competition using representation learning. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1261–1268, 2020.

Sebastian Gabel, Daniel Guhl, and Daniel Klapper. P2v-map: Mapping market structures for large retail assortments. *Journal of Marketing Research*, 56(4):557–580, 2019.

Rianne van den Berg, Thomas N Kipf, and Max Welling. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263*, 2017.

Thomas N Kipf and Max Welling. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*, 2016b.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pages 593–607. Springer, 2018.

Alex Kendall, Yarin Gal, and Roberto Cipolla. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7482–7491, 2018.

Ernst R. Berndt, Linda Bui, David H. Lucking-Reiley, and Glen L. Urban. The roles of marketing, product quality, and price competition in the growth and composition of the U.S. antiulcer drug industry. In Timothy F. Bresnahan and Robert J. Gordon, editors, *The Economics of New Goods*, pages 277–328. University of Chicago Press, Chicago, 1996.

Fei Gao, Yu Jie Zhou, Da Yi Hu, Ying Xin Zhao, Yu Yang Liu, Zhi Jian Wang, Shi Wei Yang, and Xiao Li Liu. Contemporary management and attainment of cholesterol targets for patients with dyslipidemia in China. *PLoS One*, 8(4):e47681, 2013.

National Health Commission. Circular on fully carrying out the work of promoting the comprehensive reform of public hospitals [In Chinese]. http://www.gov.cn/xinwen/2017-05/01/content_5189918.htm#allContent, April 2017.

John A. Rizzo. Advertising and competition in the ethical pharmaceutical industry: The case of antihypertensive drugs. *Journal of Law and Economics*, 42(1):89–116, 1999.

Ernst R. Berndt, Margaret Kyle, and Davina Ling. The long shadow of patent expiration: Generic entry and Rx-to-OTC switches. In Robert C. Feenstra and Matthew D. Shapiro, editors, *Scanner Data and Price Indexes*. University of Chicago Press, Chicago, 2003b.

John A. Rizzo. Do pharmaceutical sales respond to scientific evidence? *Journal of Economics and Management Strategy*, 11(4):551–594, 2002.

Gregory S. Crawford and Matthew Shum. Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173, 2005.

Julie M. Donohue and Ernst R. Berndt. Effects of direct-to-consumer advertising on medication choice: The case of antidepressants. *Journal of Public Policy and Marketing*, 23(2):115–127, 2004.

Zu-dong Shi. Institutional change and prediction of drugs group procurement [In Chinese]. *Chinese Journal of Medical Management Sciences*, 4(1):25–27, 2014.

Zongfu Mao, Xiao Shen, and Quan Wang. Retrospective study on drug's centralized purchasing system of healthcare institutions in China [In Chinese]. *Chinese Journal of Health Policy*, 7(10): 5–10, 2014.

Patrick Kline. Oaxaca-blinder as a reweighting estimator. *American Economic Review*, 101(3): 532–37, 2011.

Nicole Fortin, Thomas Lemieux, and Sergio Firpo. Decomposition methods in economics. In *Handbook of Labor Economics*, volume 4, pages 1–102. Elsevier, 2011.

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*, 33:1127–1172, 2020.

Sergio Firpo, Nicole M. Fortin, and Thomas Lemieux. Unconditional quantile regressions. *Econometrica*, 77(3):953–973, 2009.

Markus Fröolich and Blaise Melly. Estimation of quantile treatment effects with stata. *The Stata Journal*, 10(3):423–457, 2010.

Lisa Garbe, Richard Rau, and Theo Toppe. Influence of perceived threat of Covid-19 and HEXACO personality traits on toilet paper stockpiling. *PLOS ONE*, 15(6):e0234232, 2020.

Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.

Ben Etheridge and Lisa Spantig. The gender gap in mental well-being during the COVID-19 outbreak: evidence from the UK. Institute for Social and Economic Research Working Paper No. 2020-08, 2020.

Veronica Guerrieri, Guido Lorenzoni, Ludwig Straub, and Iván Werning. Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages? NBER Working Paper No. 26918, 2020.

Colleen P. Kirk and Laura S. Rifkin. I'll trade you diamonds for toilet paper: Consumer reacting, coping and adapting behaviors in the COVID-19 pandemic. *Journal of Business Research*, 117: 124–131, 2020b.

Gladys Lopez-Acevedo and Angel Salinas. How Mexico's financial crisis affected income distribution. World Bank Policy Research Working Paper 2406, 2000.

Hung-Hao Chang and Chad Meyerhoefer. COVID-19 and the demand for online food shopping services: Empirical evidence from Taiwan. NBER Working Paper No. 27427, 2020.

Jean-Charles Rochet and Jean Tirole. Two-sided markets: A progress report. *RAND Journal of Economics*, 37(3):645–667, 2006b.

Jizhou Huang, Haifeng Wang, Miao Fan, An Zhuo, Yibo Sun, and Ying Li. Understanding the impact of the covid-19 pandemic on transportation-related behaviors with human mobility data. KDD '20, page 3443–3450, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379984. doi: 10.1145/3394486.3412856. URL https://doi.org/10.1145/3394486.3412856.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.

Alon Eizenberg. Upstream innovation and product variety in the us home pc market. *Review of Economic Studies*, 81(3):1003–1045, 2014.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

Appendix A

EXAMINING THE ZERO-MARKUP DRUG POLICY IN CHINA: A STRUCTURAL APPROACH (CHAPTER 1)

A.1 Standard error

To calculate the standard errors of our estimated demand parameters, we need the derivatives of the unobserved drug quality with respect to the parameters, $\partial\xi_t/\partial\theta_2$. According to the implicit function theorem, we have

$$\frac{\partial\xi_t}{\partial\theta_2} = -\left(\frac{\partial s_t}{\partial\xi_t}\right)^{-1}\frac{\partial s_t}{\partial\theta_2}.$$

Let's denote

$$\sigma^r_{j|g}\left(P_t, M_t, \delta_t, \theta_2\right) = \frac{\exp\left\{\frac{\delta_{jt}-\alpha^r P_{jt}+\gamma^r M_{jt}}{1-\lambda}\right\}}{\sum_{j\in g}\exp\left\{\frac{\delta_{jt}-\alpha^r P_{jt}+\gamma^r M_{jt}}{1-\lambda}\right\}},$$

$$\sigma^r_g\left(p^R_t, M_t, \delta_t, \theta_2\right) = \frac{\left(\sum_{j\in g}\exp\left\{\frac{\delta_{jt}-\alpha^r P_{jt}+\gamma^r M_{jt}}{1-\lambda}\right\}\right)^{1-\lambda}}{\sum_{g\in G}\left(\sum_{j\in g}\exp\left\{\frac{\delta_{jt}-\alpha^r P_{jt}+\gamma^r M_{jt}}{1-\lambda}\right\}\right)^{1-\lambda}},$$

and

$$\kappa_t = \begin{cases} 1 & \text{before 2015} \\ \phi & \text{between 2015 and 2017Q3} \\ 0 & \text{after 2017Q3} \end{cases} \cdot$$

If we denote $\kappa^1_t = \kappa_t$ and $\kappa^2_t = 1-\kappa_t$, then

$$\frac{\partial s_{jt}}{\partial\xi_{jt}} = \sum_{r=1}^{2}\kappa^r_t\left(\frac{1-\sigma^r_{j|g}}{1-\lambda} + \sigma^r_{j|g}(1-\sigma^r_g)\right)\sigma^r_{j|g}\sigma^r_g,$$

$$\frac{\partial s_{jt}}{\partial\xi_{j't}} = \begin{cases} \sum_{r=1}^{2}\kappa^r_t\left(-\frac{\sigma^r_{j'|g}}{1-\lambda} + \sigma^r_{j'|g}(1-\sigma^r_g)\right)\sigma^r_{j|g}\sigma^r_g & \text{if } j'\in g(j) \\ -\sum_{r=1}^{2}\kappa^r_t\sigma^r_{j'|g(j')}\sigma^r_{g(j')}\sigma^r_{j|g}\sigma^r_g & \text{if } j'\notin g(j) \end{cases},$$

$$\frac{\partial s_{jt}}{\partial \theta_{1,k}} = \sum_{r=1}^{2} \kappa_t^r \left(\frac{1-\sigma_{j|g}^r}{1-\lambda} + \sigma_{j|g}^r (1-\sigma_g^r) \right) \sigma_{j|g}^r \sigma_g^r x_t^k \text{ for } k = 1, ..., 22,$$

$$\begin{aligned} \frac{\partial s_{jt}}{\partial \alpha} =& \kappa_t \left(\frac{1-\sigma_{j|g}^1}{1-\lambda} + \sigma_{j|g}^1 (1-\sigma_g^1) \right) \sigma_{j|g}^1 \sigma_g^1 (-p_{jpt}^R) \\ &+ (1-\kappa_t) \left(\frac{1-\sigma_{j|g}^2}{1-\lambda} + \sigma_{j|g}^2 (1-\sigma_g^2) \right) \sigma_{j|g}^2 \sigma_g^2 (-p_{jpt}^W), \end{aligned}$$

$$\frac{\partial s_{jt}}{\partial \gamma} = \kappa_t \left(\frac{1-\sigma_{j|g}^1}{1-\lambda} + \sigma_{j|g}^1 (1-\sigma_g^1) \right) \sigma_{j|g}^1 \sigma_g^1 m_{jpt},$$

$$\frac{\partial s_{jt}}{\partial \lambda} = \sum_{r=1}^{2} \kappa_t^r \left(\frac{1-\sigma_{j|g}^r}{1-\lambda} + \sigma_{j|g}^r (1-\sigma_g^r) \right) \sigma_{j|g}^r \sigma_g^r \frac{\delta_{jt} - \alpha^r P_{jt} + \gamma^r m_{jt}}{(1-\lambda)^2},$$

and

$$\frac{\partial s_{jt}}{\partial \phi} = \frac{\partial \kappa_t}{\partial \phi} \left(\sigma_{j|g}^1 \sigma_g^1 - \sigma_{j|g}^2 \sigma_g^2 \right)$$

where

$$\frac{\partial \kappa_t}{\partial \phi} = \begin{cases} 1 & 2015 \leq t \leq 2017\text{Q3} \\ 0 & \text{otherwise} \end{cases}.$$

Given $\partial \xi_t / \partial \theta_2$, the standard errors of our parameters are

$$\text{Std. Err.}(\theta_2) = \sqrt{\frac{1}{n} \left(\hat{Q}' \hat{W} \hat{Q} \right)^{-1} \hat{Q}' \hat{W} \hat{\Omega} \hat{W}' \hat{Q} \left(\hat{Q}' \hat{W} \hat{Q} \right)^{-1\prime}}$$

where

$$\hat{Q} = \frac{1}{n} \sum_{j,p,t} h(z_{jpt}^d) \left. \frac{\partial \xi_{jpt}}{\partial \theta_2} \right|_{\hat{\theta}_d},$$

$$\hat{\Omega} = \frac{1}{n} \sum_{j,p,t} \left(h(z_{jpt}^d) \xi_{jpt} - \bar{g} \right) \left(h(z_{jpt}^d) \xi_{jpt} - \bar{g} \right)'$$

in which $\bar{g} = \frac{1}{n} \sum_{j,p,t} h(z_{jpt}^d) \xi_{jpt}$; note that, $\hat{W} = \mathbb{I}$, and n is the full sample size.

A.2 Elasticity

Demand semi-elasticity is given by

$$\begin{aligned}
\frac{\partial \ln s_{jt}\left(\mathbf{p}_t^W\right)}{\partial p_{jt}^W} &= \frac{\partial \ln\left(\kappa_t s_{jt}^1 + (1-\kappa_t)s_{jt}^2\right)}{\partial p_{jt}^W} \\
&= \frac{1}{s_{jt}}\sum_{r=1}^{2}\kappa_t^r \frac{\partial s_{jt}^r}{\partial p_{jt}^W} \\
&= \frac{1}{s_{jt}}\sum_{r=1}^{2}\kappa_t^r \left(\frac{\partial \sigma_{j|g}^r}{\partial p_{jt}^W}\sigma_g^r + \sigma_{j|g}^r \frac{\partial \sigma_g^r}{\partial p_{jt}^W}\right) \\
&= \frac{1}{s_{jt}}\sum_{r=1}^{2}\kappa_t^r \left(\frac{\sigma_{j|g}^r\left(1-\sigma_{j|g}^r\right)}{1-\lambda}\sigma_g^r + \sigma_g^r(1-\sigma_g^r)(\sigma_{j|g}^r)^2\right)\eta^r \\
&= \frac{1}{s_{jt}}\sum_{r=1}^{2}\kappa_t^r \left(\frac{1-\sigma_{j|g}^r}{1-\lambda} + \sigma_{j|g}^r(1-\sigma_g^r)\right)\sigma_{j|g}^r\sigma_g^r\eta^r
\end{aligned}$$

where

$$\eta^1 = \begin{cases} -1.15\alpha + 0.15\gamma & p_{jt}^{Highest} \geq 1.15 p_{jt}^W \\ -\gamma & p_{jt}^{Highest} < 1.15 p_{jt}^W \end{cases}, \; \eta^2 = \begin{cases} -\alpha & p_{jt}^{Highest} \geq p_{jt}^W \\ 0 & p_{jt}^{Highest} < p_{jt}^W \end{cases}$$

and welfare semi-elasticity is

$$\begin{aligned}
\frac{\partial \ln \Delta_j w_t\left(\mathbf{p}_t^W\right)}{\partial p_{jt}^W} &= \frac{1}{\Delta_j w_t\left(\mathbf{p}_t^W\right)}\frac{\partial w_t\left(p_{jt}^W, \mathbf{p}_{-jt}^W\right)}{\partial p_{jt}^W} \\
&= \frac{1}{\Delta_j w_t\left(\mathbf{p}_t^W\right)}\textstyle\sum_{r=1}^{2}\kappa_t^r \frac{\partial\left[\delta_{jt} - \alpha^r P_{jt} + \gamma^r m_{jt} - (1-\lambda)\ln\sigma_{j|g}^r - \ln\sigma_g^r\right]}{\partial p_{jt}^W} \\
&= \frac{1}{\Delta_j w_t\left(\mathbf{p}_t^W\right)}\sum_{r=1}^{2}\kappa_t^r\eta^r\left[1 - \left(1-\sigma_{j|g}^r\right) - \sigma_{j|g}^r(1-\sigma_g^r)\right] \\
&= \frac{1}{\Delta_j w_t\left(\mathbf{p}_t^W\right)}\sum_{r=1}^{2}\kappa_t^r\eta^r\sigma_{j|g}^r\sigma_g^r
\end{aligned}$$

Note that, the (wholesale) price elasticity of demand is

$$p_{jt}^W \frac{\partial \ln s_{jt}\left(\mathbf{p}_t^W\right)}{\partial p_{jt}^W},$$

and the cross-price elasticity of demand is

$$
\begin{aligned}
& p_{j't}^{W} \frac{\partial \ln s_{jt}\left(\mathbf{p}_t^W\right)}{\partial p_{j't}^{W}} \\
&= \frac{\partial \ln\left(\kappa_t s_{jt}^1 + (1-\kappa_t) s_{jt}^2\right)}{\partial p_{j't}^{W}} \\
&= \frac{p_{j't}^{W}}{s_{jt}} \sum_{r=1}^{2} \kappa_t^r \frac{\partial s_{jt}^r}{\partial p_{j't}^{W}} \\
&= \frac{p_{j't}^{W}}{s_{jt}} \sum_{r=1}^{2} \kappa_t^r \left(\frac{\partial \sigma_{j|g}^r}{\partial p_{j't}^{W}} \sigma_g^r + \sigma_{j|g}^r \frac{\partial \sigma_g^r}{\partial p_{j't}^{W}} \right) \\
&= \begin{cases} \frac{p_{j't}^{W}}{s_{jt}} \sum_{r=1}^{2} \kappa_t^r \left(-\frac{\sigma_{j'|g}^r}{1-\lambda} + \sigma_{j'|g}^r (1-\sigma_g^r) \right) \sigma_{j|g}^r \sigma_g^r \eta^r & \text{if } j' \in g(j) \\ -\frac{p_{j't}^{W}}{s_{jt}} \sum_{r=1}^{2} \kappa_t^r \sigma_{j'|g(j')}^r \sigma_{g(j')}^r \sigma_{j|g}^r \sigma_g^r \eta^r & \text{if } j' \notin g(j) \end{cases}
\end{aligned}
$$

A.3 Dealing with unobserved package sizes

The highest price regulation varies by package size (i.e, the number of units in each package, such as 12 tablets versus 14 tablets per package), which are not observed from our data. To deal with this issue, for each aggregated drug product (a molecule-firm pair), we calculate the price caps based on several package sizes, and use the highest per unit price cap to calculate the highest possible retail price (upper bound), and the lowest to calculate the lower bound. Our calculations suggest that they are very close to each other, compared to the magnitude of the prices themselves.

Although the lower bound and upper bound are fairly close to each other, it creates a small problem in defining a "binding constraint" because in some cases $1.15p^W$ falls in the middle of the range created by the lower bound and upper bound of $p^{Highest}$. Table A.1 illustrates this potential issue by calculating the portion of observations that encounter a binding highest price (i.e., $p^{Highest} \leq 1.15p^W$) between 2012 and 2014. The binding rates calculated based on the lower bounds and upper bounds are different but in general they are fairly close to each other. Our main results will be based on the lower bound (highest per unit price cap and thus highest retail price or markup). The analysis based on the upper bound is similar. One more takeaway from Table A.1 is that the binding rates are fairly high and vary a lot by molecule, which suggests that there should

Table A.1: Rate at which retail price equals highest price (2012–2014)

	Lower bound (%)			Upper bound (%)		
	2012	2013	2014	2012	2013	2014
Molecule:						
Ezetimibe	82.61	98.78	92.77	82.61	98.78	92.77
Atorvastatin	85.23	76.33	75.31	86.91	81.00	78.75
Inositol Nicotinate	66.67	66.67	66.67	66.67	66.67	66.67
Probucol	73.81	70.68	67.16	73.81	70.68	67.16
Fluvastatin	69.83	73.41	72.02	80.45	80.35	79.17
Fenofibrate	56.83	53.19	51.84	60.43	57.60	55.81
Pravastatin	66.06	59.57	57.30	82.48	80.14	73.03
Simvastatin	56.33	55.20	60.51	58.77	58.28	64.23
Bezafibrate	45.55	46.11	40.31	63.35	62.69	54.08
Acipimox	61.38	49.60	45.38	66.90	59.20	58.46
Lovastatin	48.91	49.58	53.23	77.37	75.63	72.58
Rosuvastatin	56.82	67.95	56.59	59.09	81.79	72.20
Gemfibrozil	50.57	38.27	37.93	75.86	67.90	65.52
Pitavastatin	65.17	56.44	55.65	96.63	87.13	86.96
Jiaogulan	49.40	58.90	58.70	77.11	84.93	71.74
Zhibituo	50.00	35.71	50.00	53.85	53.57	50.00
Xuezhikang	48.96	52.17	48.94	84.38	83.70	89.36
Overall	60.66	59.86	59.35	70.20	70.62	68.81

Notes: (1) These rates are conditional frequencies calculated using the sub-sample with highest price regulations. (2) The highest price regulations can vary by drug form and size, and in our data only about 0.1% of the cases aggregate the forms and/or sizes without highest prices and the ones with highest prices, and we assume that those standardized drugs are under highest price regulations. (3) The highest price regulation also varies by package size, which is not observed from the data, and so we use the highest per unit price cap to calculate the lower bounds of the binding rates and the lowest per unit price cap to calculate the upper bounds of the binding rates.

be a good variation in retail price and hospital markup, although the retail price and markup are likely correlated.

A.4 Market structure at the national level

Table A.2 illustrates the structure of the lipid-lowering prescription drug market in China between 2012 and 2019, where we regard the whole country as a market and weight each province by the number of actual hospitals versus that of sampling hospitals. The market shares are relative to the total potential sales of lipid-lowering drugs and we assume that the treatment rate is 39 percent, based on Gao et al. [2013], for every year (so outside goods account for 69 percent of the market shares). We leave the question of the consequences of having time-varying (and later market-varying) treatment rates for future research. We also do not assume any variation in how each drug is used in each hospital.

As we can see from Table A.2, the average wholesale price of each standard unit (weighted based on a fixed "basket" of firms in 2012) is declining over time. We include the firms that appear in each year at least once in the "basket" and weight the simple average wholesale price across different quarters and provinces for each firm by its corresponding yearly market share. While the decline was modest before 2016, it became quite significant after 2017 as the ZMDP kicked in and the CDP was enhanced. Entry was observed in Atorvastatin and Rosuvastatin, the top two most popular molecules in recent years.

Table A.2: Lipid-lowering prescription drug market structure in China (2012–2019)

	2012	2013	2014	2015	2016	2017	2018	2019
Wholesale price (CNY)/std unit	3.25	3.19	3.16	3.13	3.08	2.96	2.85	2.63
Market shares of top brands								
Lipitor (Pfizer)	11.17%	11.99%	12.47%	11.79%	11.95%	12.77%	13.40%	13.21%
Zocor (MSD)	5.73%	4.31%	3.62%	3.00%	2.70%	2.16%	1.46%	0.90%
Crestor (AstraZeneca)	4.00%	4.97%	6.43%	7.03%	7.36%	8.14%	8.07%	5.90%
A Le (Jialin)	2.56%	2.81%	3.16%	3.72%	4.02%	3.99%	3.96%	5.39%
Lescol (Novartis)	2.30%	2.21%	1.96%	2.04%	1.77%	1.12%	1.03%	0.90%
Jing Bi Shu Xin (Jingxin)	1.36%	0.94%	0.69%	0.46%	0.44%	0.27%	0.21%	0.19%
You Jia (Topfond/Sinopharm)	1.08%	1.20%	1.11%	1.10%	0.81%	0.68%	0.48%	0.30%
Market shares of selected molecules								
Atorvastatin	14.82%	16.09%	16.86%	16.82%	17.39%	18.13%	18.85%	20.22%
Simvastatin	11.06%	8.86%	7.21%	6.29%	5.29%	4.47%	3.24%	2.44%
Rosuvastatin	5.59%	7.35%	9.01%	9.95%	10.77%	11.85%	12.44%	11.46%
Fluvastatin	2.66%	2.51%	2.18%	2.23%	2.01%	1.24%	1.10%	0.96%
Xuezhikang	0.53%	0.46%	0.35%	0.37%	0.34%	0.29%	0.24%	0.24%
Acipimox	0.32%	0.28%	0.18%	0.18%	0.17%	0.18%	0.17%	0.25%
Probucol	0.22%	0.18%	0.17%	0.18%	0.20%	0.19%	0.16%	0.17%
# firms of selected molecules								
Atorvastatin	4	6*	6	6	6	6	6	7
Simvastatin	42	39	42	39	35	38	37	36
Rosuvastatin	5	5	5	7	7	7	7	7
Fluvastatin	2	2	2	2	2	2	2	2
Xuezhikang	1	1	1	1	1	1	1	1
Acipimox	3	3	3	4	4	4	4	3
Probucol	2	2	2	2	2	2	2	2

Notes: (1) Wholesale price per standard unit is calculated using the "basket" of firms that appear in every year, weighted by their market shares in 2012. (2) Market shares are relative to the total potential sales of lipid-lowering drugs (the treatment rate is 39%). (3) The reported mean values of prices and market shares are weighted according to the hospital distribution in China versus in the sample. * In 2003, Topfond merged into Sinopharm, so we merge them into one firm here.

Appendix B

TRAINING, PRODUCTIVITY, AND WAGES: AN INVESTIGATION OFCHINAS MANUFACTURING ENTERPRISES IN A PRIVATIZATION ERA (CHAPTER 2)

Table B1: Types of Large-Scale Chinese Manufacturing Enterprises (1998–2013)

Year	State-owned		Collectively-owned		Private		Foreign		Total
	Subtotal	(%)	Subtotal	(%)	Subtotal	(%)	Subtotal	(%)	
1998	48,329	32.29	54,384	36.33	20,878	13.95	26,093	17.43	149,684
1999	43,409	29.51	50,071	34.04	27,202	18.49	26,434	17.97	147,116
2000	35,822	24.16	46,153	31.13	38,294	25.83	28,010	18.89	148,279
2001	28,608	18.24	39,511	25.20	57,733	36.82	30,964	19.75	156,816
2002	23,731	14.22	35,257	21.13	73.912	44.29	33,968	20.36	166,868
2003	18,209	10.05	29,439	16.25	95,490	52.70	38,048	21.00	181,186
2004	20,356	7.85	23,383	9.01	159,145	61.35	56,528	21.79	259,412
2005	11,556	4.62	20,350	8.14	162,636	65.04	55,495	22.19	250,037
2006	10,237	3.67	17,494	6.26	191,535	68.58	60,016	21.49	279,282
2007	7,197	2.30	16,079	5.14	223,280	71.32	66,490	21.24	313,046
2008	6,201	1.61	14,333	3.72	290,251	75.27	74,809	19.40	385,594
2009	5,402	1.58	12,202	3.58	258,039	75.65	65,444	19.19	341,087
2010	5,720	1.38	15,743	3.79	323,218	77.79	70,833	17.05	415,514
2011	3,962	1.41	5,619	2.00	219,075	78.15	51,654	18.43	280,310
2012	4,098	1.36	5,537	1.84	238,299	79.11	53,290	17.69	301,224
2013	3,031	0.94	3,805	1.18	261,236	81.36	53,028	16.51	321,100

Notes: (1) Numbers are calculated by the authors based on the Chinese Industrial Enterprises Database. Some other authors may have different data sources and small differences can occur. (2) Large-scale means the main business income of an enterprise was larger than 5 million RMB, and this standard was revised to 20 million RMB in 2011.

Table B2: Results of Manufacturing Sectors

	OLS			ACF			
Sector	ϕ_T	λ_T	$\phi_T - \lambda_T$	ϕ_T	λ_T	$\phi_T - \lambda_T$	Observations
1 Non-staple Food Processing	0.181***	0.064***	0.116***	0.102***	0.049***	0.053***	28,796
	(0.024)	(0.008)	(0.020)	(0.020)	(0.007)	(0.020)	
2 Food Products	0.157***	0.053***	0.104***	0.077	0.020	0.058	10,962
	(0.029)	(0.015)	(0.020)	(0.058)	(0.014)	(0.051)	
3 Beverage Products	0.317***	0.175***	0.141***	0.193***	0.118***	0.074*	7,155
	(0.035)	(0.014)	(0.029)	(0.050)	(0.013)	(0.044)	
4 Tobacco Processing	0.244	0.184	0.059	0.638	0.072	0.565	346
	(0.167)	(0.120)	(0.055)	(6.44)	(0.108)	(0.565)	
5 Textile Products	0.252***	0.074***	0.177***	0.136***	0.062***	0.074*	49,527
	(0.031)	(0.012)	(0.023)	(0.043)	(0.011)	(0.041)	
6 Wearing Apparel, Shoes and Caps	0.267***	0.070***	0.196***	0.134*	0.045**	0.089	25,329
	(0.072)	(0.023)	(0.051)	(0.080)	(0.018)	(0.074)	
7 Leather, Fur and Feather Products	0.265***	0.019	0.246***	0.109**	-0.007	0.116**	12,939
	(0.028)	(0.016)	(0.028)	(0.052)	(0.015)	(0.052)	
8 Timber Processing and Wood Products	0.298***	0.060**	0.237***	0.161*	0.056***	0.104	9,135
	(0.091)	(0.029)	(0.069)	(0.088)	(0.021)	(0.084)	
9 Furniture	0.248***	0.093***	0.154**	0.168*	0.071***	0.096	6,282
	(0.071)	(0.026)	(0.069)	(0.087)	(0.026)	(0.082)	
10 Papermaking and Paper Products	0.118*	0.054**	0.064	0.061	0.024	0.036	13,458
	(0.063)	(0.024)	(0.041)	(0.072)	(0.019)	(0.059)	
11 Printing and Record Medium Reproduction	0.143***	0.063***	0.079**	0.057	0.026**	0.030	9,578
	(0.030)	(0.013)	(0.031)	(0.034)	(0.012)	(0.035)	
12 Cultural, Educational and Sporting Goods	0.282***	0.061**	0.220***	0.136*	0.033	0.102	6,871
	(0.050)	(0.025)	(0.048)	(0.081)	(0.025)	(0.077)	
13 Petroleum and Nuclear Fuel Processing	0.500	0.156**	0.344	0.044	0.073	-0.028	3,622
	(0.313)	(0.077)	(0.244)	(0.229)	(0.061)	(0.192)	
14 Chemical Products	0.193***	0.096***	0.096***	0.088**	0.061***	0.026	38,192
	(0.023)	(0.011)	(0.017)	(0.038)	(0.008)	(0.034)	
15 Medical and Pharmaceutical Products	0.108***	0.060	0.048***	0.066**	0.063*	0.002	8,116
	(0.038)	(0.040)	(0.014)	(0.028)	(0.033)	(0.033)	
16 Chemical Fiber Products	0.289***	0.148***	0.141***	0.174	0.114***	0.060	3,014
	(0.089)	(0.037)	(0.073)	(0.121)	(0.030)	(0.121)	
17 Rubber	0.270***	0.090***	0.179***	0.129	0.059**	0.070	6,547
	(0.066)	(0.025)	(0.068)	(0.089)	(0.028)	(0.084)	
18 Plastic	0.170***	0.032***	0.138***	0.081***	0.010	0.071***	25,184
	(0.021)	(0.010)	(0.015)	(0.021)	(0.008)	(0.020)	
19 Non-metallic Mineral Products	0.299***	0.083***	0.215***	0.141***	0.048***	0.092***	42,693

	(0.035)	(0.009)	(0.028)	(0.043)	(0.007)	(0.041)	
20 Ferrous Metal Smelting and Processing	0.138***	0.099***	0.038**	0.067***	0.059***	0.008	13,085
	(0.031)	(0.028)	(0.017)	(0.027)	(0.023)	(0.030)	
21 Non-ferrous Metal Smelting and Processing	0.145***	0.032**	0.112***	0.049	0.018	0.031	9,808
	(0.036)	(0.015)	(0.026)	(0.036)	(0.012)	(0.035)	
22 Metal Products	0.185***	0.048***	0.136***	0.111***	0.025***	0.086***	23,787
	(0.026)	(0.014)	(0.019)	(0.027)	(0.012)	(0.024)	
23 General Machinery	0.147***	0.060***	0.087***	0.067**	0.046***	0.020	42,481
	(0.027)	(0.013)	(0.016)	(0.027)	(0.010)	(0.024)	
24 Special Equipment	0.152**	0.084**	0.067**	0.071	0.059*	0.011	20,441
	(0.067)	(0.042)	(0.028)	(0.064)	(0.035)	(0.045)	
25 Transportation Equipment	0.199***	0.122***	0.076***	0.116***	0.067***	0.048**	23,757
	(0.026)	(0.019)	(0.017)	(0.021)	(0.014)	(0.020)	
26 Electrical Machinery and Equipment	0.170***	0.064***	0.105***	0.088***	0.044***	0.044*	32,524
	(0.025)	(0.013)	(0.017)	(0.026)	(0.011)	(0.024)	
27 Communication, Computers and Electronic Equipment	0.236***	0.072***	0.164***	0.104***	0.032***	0.071***	14,289
	(0.038)	(0.030)	(0.019)	(0.020)	(0.022)	(0.021)	
28 Instruments and Office Machinery	0.081	0.018	0.062*	0.027	0.001	0.026	8,190
	(0.048)	(0.013)	(0.036)	(0.038)	(0.009)	(0.031)	
29 Artware and Others	0.509***	0.098***	0.411***	0.258***	0.063***	0.195***	10,356
	(0.075)	(0.023)	(0.061)	(0.085)	(0.020)	(0.081)	

Note: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) In ACF, we control the endogeneity of input and training. (3) Standard errors in parentheses and ϕ_T are computed as in Table 2.2. .

Appendix C

UNCONDITIONAL QUANTILE TREATMENT EFFECT OF COVID-19 IN A TWO-SIDED MARKET: EVIDENCE FROM A LIVE-STREAMING PLATFORM (CHAPTER 3)

C.1 The Proof of Proposition 1

Proof: Without generality, we only show θ^F. To obtain $\frac{dp^F}{\theta^F}$ and $\frac{dp^A}{\theta^F}$, we use the implicit function theorem:

$$\frac{dp}{d\theta^F} = -\left(\frac{D\varphi}{Dp}\right)^{-1}\frac{\partial\varphi}{\partial\theta^F}$$

where

$$\left(\frac{D\varphi}{Dp}\right)^{-1} = \frac{1}{\det\frac{D\varphi}{Dp}}\begin{bmatrix} \frac{\partial\lambda^A}{\partial p^A} + \frac{1}{(p^A+p^F-c)^2} & -\frac{1}{(p^A+p^F-c)^2} \\ -\frac{1}{(p^A+p^F-c)^2} & \frac{\partial\lambda^F}{\partial p^F} + \frac{1}{(p^A+p^F-c)^2} \end{bmatrix}$$

and

$$\frac{\partial\varphi}{\partial\theta^F} = \begin{pmatrix} \frac{\partial\lambda^F}{\partial\theta^F} \\ \frac{\partial\lambda^A}{\partial\theta^F} \end{pmatrix}$$

It's straightforward that $\frac{\partial\lambda^A}{\partial\theta^F} = 0$. Due to the log-concavity assumption, we also have $\frac{\partial\lambda^F}{\partial\theta^F} < 0$ and $\frac{\partial\lambda^A}{\partial p^A} > 0$. Again, due to the log-concavity, we have

$$\det\frac{D\varphi}{Dp} = \frac{\partial\lambda^A}{\partial p^A}\frac{\partial\lambda^F}{\partial p^F} + \frac{1}{(p^A+p^F-c)^2}\left(\frac{\partial\lambda^A}{\partial p^A} + \frac{\partial\lambda^F}{\partial p^F}\right) > 0$$

Therefore,

$$\begin{pmatrix} \frac{dp^F}{d\theta^F} \\ \frac{dp^A}{d\theta^F} \end{pmatrix} = \frac{-\frac{\partial\lambda^F}{\partial\theta^F}}{\det\frac{D\varphi}{Dp}}\begin{bmatrix} \frac{\partial\lambda^A}{\partial p^A} + \frac{1}{(p^A+p^F-c)^2} \\ -\frac{1}{(p^A+p^F-c)^2} \end{bmatrix} \begin{matrix} > 0 \\ < 0 \end{matrix}$$

C.2 The Proof of Proposition 2

Proof: Without generality, we only prove θ_F. From the equilibrium condition we could get $p^A = \frac{b_2+c+2\theta^A-\theta^F}{3}$ and $p^F = \frac{b_2+c+2\theta^F-\theta^A}{3}$. So

$$
\begin{aligned}
W^A &= \mathbb{E}\left(\max\{U^A, 0\}\right) \\
&= \int_{b^A \geq p^A - \theta^A} \left(b^A + \theta^A - p^A\right) N^F \frac{1}{b_2 - b_1} db^A \\
&= \frac{N^F b^A}{b_2 - b_1} \left(\frac{b^A}{2} + \theta^A - p^A\right) |_{p^A - \theta^A}^{b_2} \\
&= \frac{N^F}{b_2 - b_1} \left(\frac{b_2^2}{2} + b_2(\theta^A - p^A) + \frac{(\theta^A - p^A)^2}{2}\right) \\
&= \frac{b_2 + \theta^F - p^F}{(b_2 - b_1)^2} \frac{(b_2 + \theta^A - p^A)^2}{2} \\
&= \frac{(2b_2 + \theta^A + \theta^F - c)^3}{54(b_2 - b_1)^2}
\end{aligned}
$$

And similarly

$$
W^F = \mathbb{E}\left(\max\{U^F, 0\}\right) = \frac{(2b_2 + \theta^A + \theta^F - c)^3}{54(b_2 - b_1)^2}
$$

So it's obvious that $\frac{W^A}{\partial \theta^F} > 0$ and $\frac{W^F}{\partial \theta^F} > 0$

C.3 Summary Statistics by Month

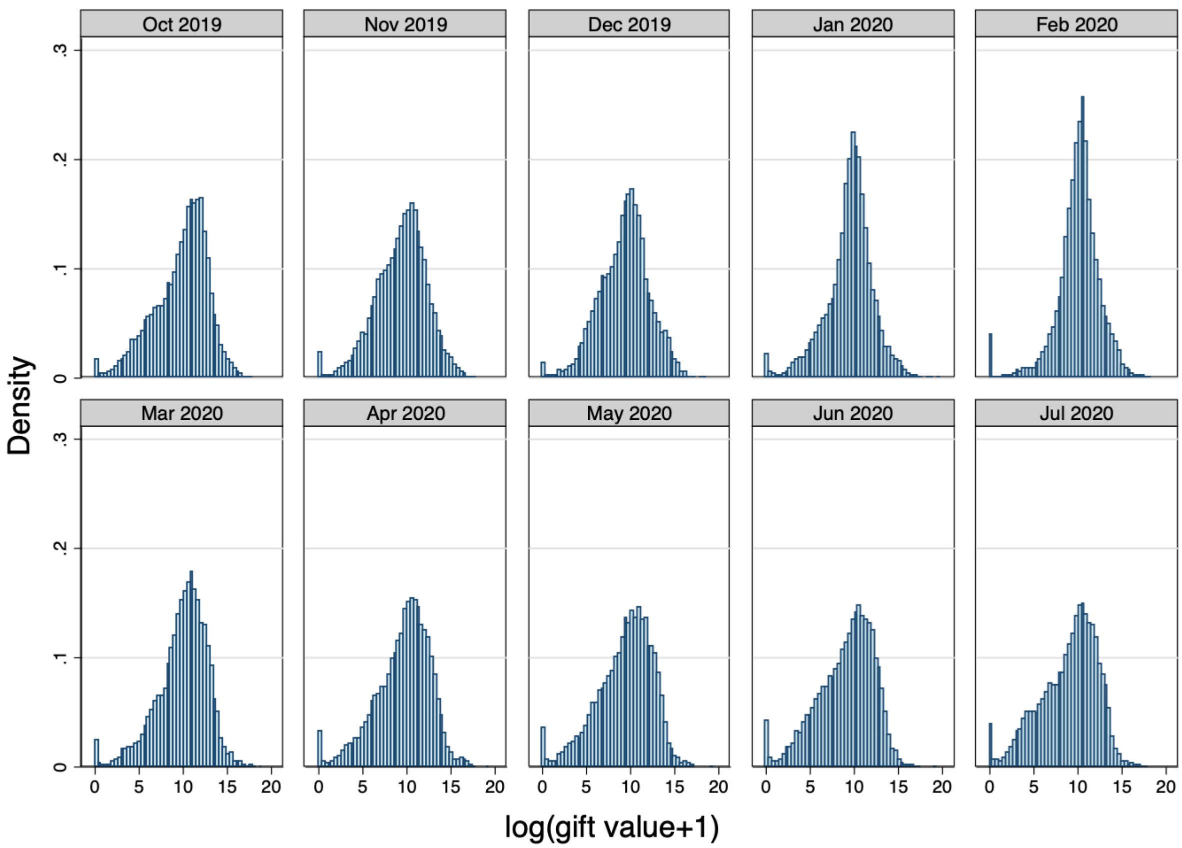

Figure C1: Histograms of gift values per live streaming video in each month

Table C1: Summary Statistics by Month (Oct 2019–Jul 2020)

	2019			2020						
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Anchors										
Female	0.36	0.36	0.34	0.33	0.33	0.34	0.33	0.32	0.32	0.31
	(0.48)	(0.48)	(0.47)	(0.47)	(0.47)	(0.48)	(0.47)	(0.47)	(0.47)	(0.46)
Total fans (1M)	5.41	5.41	5.44	5.55	5.96	6.18	6.63	6.74	6.95	7.11
	(4.74)	(4.61)	(4.71)	(4.49)	(5.24)	(5.32)	(5.80)	(5.46)	(5.82)	(6.06)
City characteristics:										
–*New COVID-19 cases*	0.00	0.00	0.00	7.38	235.46	27.99	8.28	1.00	10.10	4.49
(past 2 weeks)	(0.00)	(0.00)	(0.00)	(74.75)	(2225.8)	(303.27)	(27.72)	(4.17)	(42.50)	(19.57)
–*Lockdown policies*	0.00	0.00	0.03	7.10	20.32	23.03	46.38	29.91	26.82	25.78
(past month)	(0.00)	(0.00)	(0.25)	(7.03)	(4.15)	(6.81)	(11.92)	(13.23)	(11.69)	(14.92)
–*Reopening policies*	0.00	0.00	0.00	0.00	1.26	17.38	18.31	22.91	13.43	9.39
(past month)	(0.00)	(0.00)	(0.00)	(0.00)	(1.23)	(6.16)	(4.16)	(7.09)	(3.50)	(4.35)
–*Population (100K)*	83.94	84.98	84.61	86.59	88.61	87.49	87.29	88.52	87.67	90.43
	(58.04)	(58.16)	(58.68)	(59.79)	(62.50)	(60.56)	(60.03)	(62.13)	(62.87)	(63.02)
–*Area (10K km^2)*	1.83	1.87	1.85	1.77	1.87	1.82	1.74	1.78	1.74	1.72
	(2.28)	(2.55)	(2.40)	(2.19)	(2.70)	(2.61)	(1.98)	(2.11)	(2.04)	(1.90)
–*Quarterly GDP (1T yuan)*	0.24	0.25	0.25	0.20	0.20	0.20	0.23	0.24	0.23	0.24
	(0.28)	(0.29)	(0.30)	(0.36)	(0.25)	(0.29)	(0.25)	(0.26)	(0.26)	(0.26)
Fans										
Female ratio	0.46	0.48	0.47	0.45	0.44	0.46	0.47	0.42	0.43	0.43
	(0.25)	(0.25)	(0.25)	(0.24)	(0.23)	(0.25)	(0.25)	(0.22)	(0.21)	(0.21)
City characteristics:										
–*New COVID-19 cases*	0.00	0.00	0.00	5.50	133.75	12.76	4.72	0.73	6.82	3.09
(past 2 weeks)	(0.00)	(0.00)	(0.00)	(14.65)	(192.58)	(24.61)	(3.44)	(1.05)	(9.71)	(4.45)
–*Lockdown policies*	0.00	0.00	0.03	7.14	20.31	22.40	44.81	28.75	26.08	25.15
(past month)	(0.00)	(0.00)	(0.18)	(6.97)	(3.52)	(4.86)	(7.19)	(10.19)	(9.20)	(12.23)
–*Reopening policies*	0.00	0.00	0.00	0.00	1.26	17.41	18.17	22.36	13.19	9.01
(past month)	(0.00)	(0.00)	(0.00)	(0.00)	(1.19)	(5.90)	(2.91)	(2.66)	(2.11)	(1.83)
–*Population (100K)*	74.07	74.28	74.17	74.12	74.48	74.06	73.96	79.74	79.29	78.47
	(10.33)	(10.53)	(10.17)	(10.04)	(10.39)	(9.60)	(9.78)	(17.02)	(14.70)	(13.98)
–*Area (10K km^2)*	2.50	2.49	2.49	2.60	2.54	2.47	2.51	2.91	2.97	3.79

	(1.42)	(1.45)	(1.48)	(1.63)	(1.42)	(1.23)	(1.55)	(10.16)	(10.57)	(16.36)
–*Quarterly GDP (1T yuan)*	0.21	0.21	0.21	0.14	0.14	0.14	0.22	0.26	0.25	0.24
	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)	(0.11)	(0.23)	(0.19)	(0.18)
Live Streaming										
Length (hours)	2.01	2.06	2.21	2.16	2.30	2.18	2.11	1.93	1.99	1.92
	(1.47)	(1.76)	(1.64)	(1.74)	(1.72)	(1.52)	(1.59)	(1.47)	(1.47)	(1.40)
Average attendance (1K)	4.74	4.21	4.62	5.29	7.54	6.88	7.00	5.58	8.44	7.85
	(10.93)	(10.26)	(12.58)	(12.07)	(16.17)	(13.95)	(14.51)	(11.99)	(17.80)	(16.95)
Log of gift value+1 (yuan)	9.83	9.47	9.41	9.56	10.14	10.02	9.68	9.45	9.13	9.06
	(3.02)	(2.87)	(2.76)	(2.63)	(2.49)	(2.79)	(3.05)	(3.09)	(3.14)	(3.21)
Observations	19,757	20,284	19,781	15,143	17,997	21,981	18,922	25,300	18,503	16,837

Notes: (1) Standard deviations are in the parentheses under each mean value. (2) In late December 2019, COVID-19 cases were still not officially confirmed in our data set, but the government already started assigning quarantine areas near the Huanan Seafood Wholesale Market and within hospitals in Wuhan for suspected cases.

C.4 Additional Results

Table C2: OLS Regressions for Average Treatment Effects

	(1) Unadjusted	(2) Adjusted
COVID-19 Intensity (Treatment)		
Anchor side (cases/100K people, past 2 weeks)	0.000	0.003
	(0.003)	(0.004)
Fans side (cases/100K people, past 2 weeks)	0.133**	0.121**
	(0.053)	(0.057)
Proneness Variables		
Lockdown policies in the anchor's city (past month)	–	-0.032*
		(0.018)
Reopening policies in the anchor's city (past month)	–	-0.008
		(0.017)
Lockdown policies in a fan' city (past month)	–	0.035**
		(0.018)
Reopening policies in a fan's city (past month)	–	0.003
		(0.018)
Quarterly GDP of the anchor's city (1T yuan)	–	0.257
		(0.354)
Quarterly GDP of a fan's city (1T yuan)	–	0.535
		(0.503)
Area of the anchor's city (10K km^2)	–	-0.041
		(0.031)
Area of a fan's city (10K km^2)	–	-0.040***
		(0.010)
Constant	9.561***	9.577***
	(0.120)	(0.211)
Observations	194,505	194,505

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) Robust standard errors, clustered on anchor, are in parentheses.

Table C3: QTEs of Single Binary Treatment: Fans Side > Anchor Side

	(1) Firpo [2007]	(2) LDML	(3) GQR
QTE at Percentile			
10	1.480***	0.740***	1.481
	(0.085)	(0.068)	(1.621)
20	1.606***	0.571***	1.415
	(0.052)	(0.046)	(4.403)
30	1.684***	0.758***	1.124***
	(0.054)	(0.061)	(0.299)
40	1.297***	0.596***	0.776***
	(0.047)	(0.035)	(0.219)
50	0.795***	0.438***	0.524**
	(0.046)	(0.061)	(0.253)
60	0.435***	0.263***	0.354
	(0.037)	(0.051)	(0.317)
70	0.166***	0.109	0.245
	(0.034)	(0.085)	(0.540)
80	-0.022	0.077	0.156
	(0.035)	(0.096)	(0.681)
90	-0.103***	-0.180	0.029
	(0.033)	(0.197)	(1.262)
Observations		194,505	

Notes: (1) QTE = quantile treatment effect; LDML = localized debiased machine learning; GQR = generalized quantile regression. (2) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (3) The LDML part is based on Kallus et al. [2020] using random forests to fit the conditional expectation functions and spliting the data into 5 folds. (4) The GQR part is based on Powell [2019] using the "linear" model to predict the probability the outcome is below the quantile function conditional on the proneness variables in Table 3.1, and the numerical optimization proceeds via a Nelder-Mead algorithm. (5) Conventional standard errors are in parentheses.

Table C4: Single QTEs by LDML using Different Methods

	(1) Forest	(2) Neural Net	(3) LASSO	(4) Boosting
QTE at Percentile				
10	0.740***	1.199***	2.626***	0.932***
	(0.068)	(0.097)	(0.069)	(0.078)
20	0.571***	1.796***	2.285***	1.049***
	(0.046)	(0.080)	(0.090)	(0.060)
30	0.758***	1.364***	1.931***	0.898***
	(0.061)	(0.074)	(0.085)	(0.053)
40	0.596***	1.080***	1.419***	0.586***
	(0.035)	(0.070)	(0.084)	(0.049)
50	0.438***	0.784***	0.963***	0.386***
	(0.061)	(0.069)	(0.079)	(0.047)
60	0.263***	0.536***	0.715***	0.300***
	(0.051)	(0.071)	(0.072)	(0.050)
70	0.109	0.660***	0.548***	0.249***
	(0.085)	(0.093)	(0.082)	(0.064)
80	0.077	0.343***	0.437***	0.194**
	(0.096)	(0.126)	(0.110)	(0.090)
90	-0.180	0.737***	0.284	0.035
	(0.197)	(0.273)	(0.180)	(0.137)
Observations		194,505		

Notes: (1) QTE = quantile treatment effect; LDML = localized debiased machine learning. (2) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (3) The LDML part is based on Kallus et al. [2020] using random forests, neural networks, LASSO, and gradient boosting to fit the conditional expectation functions and spliting the data into 5 folds. (4) Conventional standard errors are in parentheses.

Table C5: Predict probability based on the "logit" model

	COVID-19 Intensity	
	Anchor Side	Fans Side
QTE at Percentile		
10	0.004	0.345***
	(0.003)	(0.078)
20	0.001	0.311***
	(0.002)	(0.042)
30	-0.001	0.266***
	(0.001)	(0.028)
40	0.000	0.180***
	(0.003)	(0.044)
50	0.001	0.110***
	(0.001)	(0.031)
60	0.001	0.049*
	(0.001)	(0.029)
70	0.000	-0.001
	(0.001)	(0.016)
80	-0.000	-0.042
	(0.004)	(0.084)
90	-0.000	-0.049
	(0.005)	(0.096)
Observations	194,505	

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) We use the "logit" model to predict the probability the outcome is below the quantile function conditional on the proneness variables in Table 3.1. (3) Numerical optimization proceeds via a Nelder-Mead algorithm. (4) Conventional standard errors are in parentheses.

Table C6: Predict probability based on the “probit” model

	COVID-19 Intensity	
	Anchor Side	Fans Side
QTE at Percentile		
10	0.003	0.344***
	(0.003)	(0.095)
20	0.001	0.311***
	(0.002)	(0.041)
30	-0.001	0.265***
	(0.001)	(0.027)
40	0.000	0.180***
	(0.002)	(0.044)
50	0.001	0.110***
	(0.001)	(0.031)
60	0.001	0.049
	(0.001)	(0.030)
70	0.000	0.004
	(0.001)	(0.031)
80	-0.000	-0.042
	(0.004)	(0.086)
90	-0.000	-0.049
	(0.005)	(0.093)
Observations	194,505	

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) We use the “probit” model to predict the probability the outcome is below the quantile function conditional on the proneness variables in Table 3.1. (3) Numerical optimization proceeds via a Nelder-Mead algorithm. (4) Conventional standard errors are in parentheses.

Table C7: Use an adaptive MCMC optimization procedure

	COVID-19 Intensity	
	Anchor Side	Fans Side
QTE at Percentile		
10	0.003***	0.322***
	(0.001)	(0.009)
20	0.001**	0.304***
	(0.000)	(0.007)
30	0.000	0.231***
	(0.000)	(0.005)
40	0.001***	0.158***
	(0.000)	(0.005)
50	0.001***	0.107***
	(0.000)	(0.003)
60	0.001***	0.049***
	(0.000)	(0.004)
70	0.001***	-0.005
	(0.000)	(0.004)
80	0.000	-0.056***
	(0.000)	(0.005)
90	0.000	-0.061***
	(0.000)	(0.007)
Observations	194,505	

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) We use the "linear" model to predict the probability the outcome is below the quantile function conditional on the proneness variables in Table 3.1. (3) Numerical optimization proceeds via an adaptive MCMC procedure with 1,000 draws where the first 30% of them are dropped as a burn-in period, the acceptance rate of the algorithm is set to be 0.5, and the initial covariance matrix is based on the estimator proposed by Koenker and Bassett Jr. [1978]. (4) Conventional standard errors are in parentheses.

Table C8: Use a grid-search optimization procedure

	COVID-19 Intensity	
	Anchor Side	Fans Side
QTE at Percentile		
10	0.004	0.325***
	(0.004)	(0.119)
20	0.001	0.310***
	(0.002)	(0.041)
30	0.000	0.235***
	(0.001)	(0.030)
40	0.001	0.160***
	(0.001)	(0.033)
50	0.001	0.105***
	(0.001)	(0.028)
60	0.001	0.045
	(0.001)	(0.031)
70	0.001	-0.010
	(0.001)	(0.018)
80	0.000	-0.050
	(0.002)	(0.046)
90	0.000	-0.050
	(0.005)	(0.097)
Observations	194,505	

Notes: (1) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (2) We use the "linear" model to predict the probability the outcome is below the quantile function conditional on the proneness variables in Table 3.1. (3) Numerical optimization proceeds via a grid-search procedure within the range $[-0.003, 0.005] \times [-0.05, 0.325]$ and with increments 0.001 and/or 0.005. (4) Conventional standard errors are in parentheses.

Table C9: Control for province fixed effects

	COVID-19 Intensity: Cases / 10K Population	
	Anchor Side	Fans Side
ATE	0.001	0.148***
	[0.003]	[0.052]
QTE at Percentile		
10	0.001	0.347***
	(0.034)	(0.067)
20	0.001	0.311***
	(0.002)	(0.042)
30	-0.001	0.266***
	(0.001)	(0.029)
40	-0.001	0.175***
	(0.003)	(0.060)
50	-0.001	0.119***
	(0.001)	(0.038)
60	-0.000	0.060
	(0.001)	(0.038)
70	-0.000	0.006
	(0.001)	(0.018)
80	-0.000	-0.034
	(0.009)	(0.130)
90	0.332***	-0.064
	(0.007)	(0.269)
Observations	194,505	

Notes: (1) ATE = average treatment effect; QTE = quantile treatment effect. (2) ***, **, * denote significance level at 1%, 5%, and 10%, respectively. (3) The ATE is obtained by linear regression adjustment (see column 2 of **??**). (4) We use the "linear" model to predict the probability the outcome is below the quantile function conditional on the following eight proneness variables: lockdown policies, reopening policies, quarterly GDP, and the area in the anchor's city and in an average fan's city, respectively (see Table 3.1); in addition, we control for province fixed effects. (5) Numerical optimization proceeds via a Nelder-Mead algorithm. (6) Conventional standard errors are in parentheses for QTEs, while robust standard errors clustered on the anchor are in the square brackets for ATEs.