Blocking the Giants: Theory and Evidence from the Great Firewall

Ruiqi (Rachel) Sun*†

Job Market Paper

[Click here for latest version]

19 January 2025

Abstract

Bans on digital products from foreign tech giants have become a policy staple. Yet the welfare implications of these bans are unknown. Policy-targeted products are typically large and highly differentiated, leading to rich substitution patterns that are essential for welfare analysis. I empirically document these substitution patterns through an event study of India's ban on Chinese apps: pairwise elasticities of substitution are correlated with pairwise similarities between product descriptions. Motivated by this finding, I develop a general equilibrium model in which granular products with heterogeneous hedonic attributes and productivities engage in Bertrand oligopolistic competition. The model features pairwise elasticities of substitution with a model-consistent estimation strategy powered by large language models (LLMs). Using the estimated model, I evaluate the impact of China's Great Firewall (GFW) policy. I find that while the GFW increased Chinese real incomes, the loss in app service utility from using inferior apps overwhelms the benefits, resulting in a net welfare loss of 7.5%. Furthermore, the TikTok ban would increase Facebook's revenue by a factor of 11 compared to the average app in the U.S. This framework is easily implementable and widely applicable to research questions where pairwise substitutions are essential and rich text is available.

Keywords: digital bans; granular trade; oligopoly competition; LLM

JEL codes: F12, F13, F14, F52

*I am deeply indebted to my supervisors Daniel Trefler and Xiaodong Zhu, for their continued guidance and support, as well as Heski Bar-Isaac, Kevin Lim, Matthew Mitchell, and Jeffrey Sun for detailed and insightful comments. I also benefit from the discussions with Bernardo Blum, Laurent Cavenile, Mitsuru Igami, Guangbin Hong, Myeongwan Kim, Sijie Lin, Siyuan Liu, Peter Morrow, Regina Seibel, and Irisa Zhou. All errors are my own.

†Rotman School of Management, University of Toronto. Email: ruiqi.sun@rotman.utoronto.ca.

1. Introduction

Trade policies targeting specific products or firms, rather than entire industries or countries, have become increasingly prevalent worldwide. These regulations are often a response to the escalation of geopolitical tensions, where technology, trade, and innovation intersect with national security concerns and global power dynamics (Bradford, 2023). For example, China's Great Firewall (GFW) policy, implemented in 2009, effectively excluded three of the U.S. Big Four – Amazon, Meta, and Google – from its market. The ban was followed by the rise of China's own tech giants and I will argue that not coincidentally, each giant's major product mirrors a corresponding banned U.S. product: Baidu serves as a substitute for Google, Tencent for Meta, and Alibaba for Amazon. Do such trade policies targeted at specific digital products causally foster the growth of local substitutes in China? What are the welfare implications of these policies? The recent implementation of similar bans by countries like the U.S., Canada, India, and Brazil makes finding answers to these questions increasingly important.[1]

However, answering these questions is not a trivial task. Banned digital products typically have a vast user base and a diverse mix of hedonic attributes, leading to complex substitution patterns with the competitors. This complexity makes detailed knowledge of pairwise elasticities of substitution essential for welfare analysis. I document a novel fact: pairwise elasticities of substitution are correlated with pairwise similarities in product descriptions. Building on this finding, I anchor my analysis around a general equilibrium model where granular, single-product firms with heterogeneous hedonic attributes and productivities engage in Bertrand oligopolistic competition. Millions of elasticity parameters are estimated based on the similarity of product descriptions. Similarity here means the cosine similarity between textual embeddings estimated using large language models (LLMs). This methodology, both theoretically and empirically, provides a practical approach to capture rich demand-side heterogeneity (hedonic attributes) alongside supply-side heterogeneity (Melitz productivity) and highlights its relevance for the welfare analysis of policy interventions targeted at granular products.

In Section 2, I begin with an event study to empirically document the substitution patterns that follow an unexpected ban. After India's 2020 ban on 59 Chinese apps, I find that apps other than the banned ones expanded in response to reduced competition. For each app, I construct a measure of its similarity to the banned apps. Apps with the highest similarities to banned apps experienced a substantial increase in market share, while those with the lowest similarities exhibited almost no increase. This finding provides strong evidence of heterogenous responses to trade policy interventions, which

[1] Appendix Table A1 lists the ten most banned apps and demonstrates the worldwide prevalence of bans as a new trade tool.

are typically absent in workhorse models such as Melitz (2003). Furthermore, the similarity measures – constructed using textual product description and LLMs – effectively capture the pairwise elasticities of substitution. These estimates from the event study are subsequently utilized for the structural estimation of the quantitative model.

Motivated by the empirical finding, I develop a novel general equilibrium model in Section 3, designed to capture rich substitution patterns while maintaining high tractability. In this model, each country is endowed with a discrete number of firms that are heterogeneous along two key dimensions: hedonic attributes on the demand side and productivity on the supply side. Inspired by the work of Pellegrino (2023) in industrial organization, I introduce a Translog expenditure function as proposed by Diewert (1976), where the key parameters emerge endogenously from a micro-foundation of hedonic demand. I integrate the demand system into a granular trade model, building on Eaton, Kortum, and Sotelo (2012) and Gaubert and Itskhoki (2021), and solve for the equilibrium. In equilibrium, firms – each characterized by a unique mix of hedonic attributes and a productivity – observe the realization of all products and make sequential decisions by market. In the first stage, they determine whether to enter the market; in the second stage, those that enter engage in Bertrand oligopolistic competition to set prices.[2]

The vast number of demand parameters can be estimated with a parsimonious set of restrictions. Standard procedures for estimating elasticities within nested CES frameworks, such as Berry, Levinsohn, and Pakes (1995), require extensive market-level data, which is rarely available in the context of international trade. The challenge is even more pronounced in this model as the nesting occurs at the product level. However, the demand system here yields pairwise elasticities of substitution as a function of similarities in hedonic attributes. I propose an estimation strategy that uses LLM embeddings of textual product descriptions as proxies for the hedonic attributes. This approach transforms the complex estimation of pairwise elasticities into a series of product-specific LLM tasks, alongside the identification of one single elasticity parameter. More importantly, this approach provides a structural interpretation for the output of language models, a tool increasingly used in recent empirical studies (Acemoglu and Restrepo, 2022, Kelly, Papanikolaou, Seru, and Taddy, 2021, Argente, Baslandze, Hanley, and Moreira, 2020).

To validate the model as well as the estimation strategy, I conduct a quantitative analysis using a dataset on mobile apps. This sector and dataset are particularly well-suited for four reasons. First, the mobile app sector demonstrates rich and intuitive substitution patterns for consumers. For instance, Chrome is highly substitutable with Baidu as both are search engines, whereas Chrome is only weakly substitutable with Weibo, which functions primarily as a social media platform akin to X/Twitter but also includes limited

[2]While the baseline model applies broadly across sectors, I tailor it into a quantitative model for mobile app services in Section 5.1, where prices correspond to ad rates in this specific application.

search capabilities. Second, the availability of textual descriptions and product-level trade data makes this sector ideal for quantitative analysis.[3] Third, natural experiments such as India's app ban and China's GFW policy provide exogenous variation to identify key parameters and validate the model's predictions. Finally, the sector is dominated by superstar firms that are central to current trade disputes. Notably, among the ten largest companies globally, eight specialize in software services or hardware support, and all are involved in foreign bans, highlighting the policy relevance of this research.[4]

Section 4 details the dataset and presents the reduced-form evidence for the impact of the GFW. The dataset encompasses nearly all non-gaming apps available on the iOS App Store and consists of three key components. Using the detailed descriptions displayed in the App Store, I employ advanced LLMs to generate embeddings that represent the standardized characteristics of each app. These embeddings serve as the estimates for the hedonic attributes in the model. The second component of the dataset comprises trade shares by app and country. Trade shares are measured as the fraction of total screen time spent on a given app relative to all apps. The third component is a manually compiled list of apps that are blacklisted by China's GFW. These three components are linked using unique app ids from the iOS App Store. Equipped with this data, I isolate the impact of the GFW from other trade frictions and provide reduced-form evidence that supports the model's predictions regarding bans.

To further quantify the welfare implication of the GFW, I estimate a quantitative model in Section 5. I begin by developing a quantitative model that captures both utility from physical goods consumption and app services. The representative consumer allocates a fraction of her time as screen time on apps, and app producers monetize this engagement by selling advertisements, competing as outlined in the theoretical model. I calibrate the quantitative model to a snapshot of data in 2021. The model parameters are classified into two groups. The first group includes parameters that correspond directly to observable data moments and can be separately identified. Notably, the elasticity parameter for app sector is inferred from the event study presented in Section 2. The second group comprises parameters that cannot be separately identified individually. These are estimated jointly using a simulated method of moments (SMM) approach, which fully incorporates the general equilibrium effects of the quantitative model. The estimated model successfully reproduces the observed patterns of trade, time allocation, and income, effectively capturing the key empirical patterns in the data.

[3] A comparable dataset is NielsenIQ for manufactured goods, as used in Jaccard (2023). However, NielsenIQ offers limited product-level textual data. Instead, Argente, Baslandze, Hanley, and Moreira (2023) address this by clustering barcode-level products into 400 groups and supplementing descriptions from Wikipedia. For this paper, however, I explore variation within a sector, making NielsenIQ data less suitable.

[4] Appendix Table A2 lists the top 10 largest companies measured by market capitalization in 2024.

In Section 6, I use the estimated model to evaluate two policies. First, I assess the welfare implications of the GFW through a counterfactual analysis by removing the trade barriers imposed by the GFW. The analysis reveals two distinct effects on welfare for Chinese consumers. On the one hand, the GFW increased real incomes by shifting profits from banned products to Chinese alternatives, resulting in a 1.05% welfare gain. On the other hand, this benefit was outweighed by a significant reduction in the utility from app services. While the overall time spent on apps decreased only slightly, the utility per unit of time-spent dropped by 29%. Consequently, the net welfare loss from the GFW amounts to 7.48%. These findings suggest that policymakers may have incentives to impose such bans to boost real income. However, the loss in utility due to lower-quality screen time is often overlooked in policy considerations.

In addition, the model introduces a novel measure of product diversity, defined as the dispersion of hedonic attributes among apps entering the market. The counterfactual analysis reveals that the GFW reduced product diversity among Chinese apps by 4.71%. This finding suggests that the GFW not only replaced foreign products with inferior domestic alternatives but also led to the survival of products with less distinctive hedonic attributes. The policy intervention stifles product innovation as a result of reduced competition. It aligns with the finding in Hsieh, Klenow, and Shimizu (2021), and instead of assuming exogenous innovation efficiency, the results here are an endogenous outcome of the trade policy. The finding serves as a typical example of the hidden costs associated with policy interventions

Finally, I simulate a ban on TikTok in the U.S. and examine the distributional effects on competing apps. I find that apps with the highest similarity to TikTok and the highest productivity (e.g., Threads and Instagram) will experience revenue increases that are 11 times greater than the average app. This result aligns with the current political debates surrounding TikTok's ban, highlighting its potential to disproportionately benefit certain competitors such as Meta. Notably, this outcome is absent in standard models that assume uniform substitution patterns across all product pairs.

Related Literature This paper connects to several strands of literature. First, it contributes to the role of granular firms in international trade. The existing literature highlights the importance of granular firms in explaining trade flows and shaping comparative advantages (Hanson, Lind, and Muendler, 2015, Bernard, Moxnes, and Ulltveit-Moe, 2018, Bernard, Moxnes, and Saito, 2019). In particular, Bernard *et al.* (2018) emphasize that demand-side heterogeneity is as important as supply-side heterogeneity, yet remains under-explored. Relatedly, this work connects to empirical studies on country-level differences in exported products (Hausmann, Hwang, and Rodrik, 2007, Hummels and Klenow, 2005, Baldwin and Harrigan, 2011, Hallak and Schott, 2011,

Sutton and Trefler, 2016). Building on insights from the literature, this paper presents a practical approach to capturing rich demand-side heterogeneity alongside the standard supply-side heterogeneity with granular firms/products.

In doing so, this paper adds to the extensive body of work on demand system estimation beyond CES in international trade. This paper is built on the Translog functional form proposed by Christensen, Jorgenson, and Lau (1975) and Diewert (1976). The recent papers, such as Feenstra and Weinstein (2017) and Diewert (2023), employ special cases of the (Generalized) Symmetric Translog for identification. In contrast, this paper microfounds a demand system through hedonic demand and departures from the symmetry assumption. Relatedly, papers such as Adao, Costinot, and Donaldson (2017), Lashkari and Mestieri (2019), Lind and Ramondo (2023) use a cross-nested CES structure to capture flexible substitution patterns. This paper shares a similar flavour by deriving substitution from underlying hedonic attributes but differs in two key ways: the demand system (1) is designed to exploit the LLM representations, and (2) yields finite choke prices, accommodating entry and exit decisions.

This paper connects to the literature on trade shocks, competition and welfare. For example, vast of empirical works document that trade liberalization reallocation resources to high-performing firms and products (Pavcnik, 2002, Trefler, 2004, Mayer, Melitz, and Ottaviano, 2014). Under the CES framework, as discussed in Arkolakis, Costinot, and Rodríguez-Clare (2012), this reallocation is efficient and welfare changes depend on a single parameter. Moving beyond CES, Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2019) find that the pro-competition effect on welfare is quantitatively small. More recently, Baqaee, Farhi, and Sangani (2023) highlights the importance of demand elasticity heterogeneity in fully capturing the gains from trade. This paper also addresses demand-side heterogeneity and extends the literature by accounting for the granularity of trade shocks – where pairwise elasticities play a critical role.

This paper is related to product differentiation in industrial organization (Lancaster, 1966, Rosen, 1974, Petrin, 2002, Pellegrino, 2023). I borrow from the recent progress in this field, in particular from Pellegrino (2023). Pellegrino (2023) characterizes a product network and explains the source of market power for the public corporations in the United States. This paper differs in several key ways. Theoretically, I adopt a Translog functional form, which yields a closed-form solution for elasticities of substitution. I also introduce two random processes to characterize the endogenous entry. Empirically, while Pellegrino (2023) uses data from Hoberg and Phillips (2016) with similarity measured by keywords frequency, the model in this paper is calibrated to product-variety level with similarities constructed from prompted LLMs. Lastly, regarding the research question, this paper centers on the interaction between trade policy intervention and competition.

This paper is part of the rapidly growing literature on textual analysis in economics.

Embeddings form the backbone of recent advances in Natural Language Processing. The semantic similarity between two pieces of text can be quantified using the cosine similarity between their corresponding embeddings, which offers a straightforward and efficient way for economists to analyze textual data (Acemoglu and Restrepo, 2022, Acemoglu, Autor, Hazell, and Restrepo, 2022, Kelly *et al.*, 2021, Argente *et al.*, 2020, Sun and Trefler, 2023, Bartik, Gupta, and Milo, 2023, Boeing, Brandt, Dai, Lim, and Peters, 2024).[5] In this paper, rather than simply using similarity as a measure, I develop a model that provides an economic interpretation for LLM embeddings within an estimation context.

Lastly, this paper contributes to the literature on policy interventions in digital sector. The most closely related paper is Bernard, Boler, Chor, Gurund, and Lu (2024), which examines the impact of GFW on knowledge spillovers from foreign to Chinese scholars. In contrast, my focus is on product competition rather than technology spillovers. Sun and Trefler (2023) find the impact of AI deployment on app exports is halved in countries with strong cross-border data flow restriction. This paper differs by focusing on the substitution patterns in trade with a different policy. In addition, recent papers use similar mobile app data to examine the impacts of privacy regulations (Janssen, Kesler, Kummer, and Waldfogel, 2022, Bian, Ma, and Tang, 2021, Aridor, 2022), but none of them is about international trade and competition.

Outline The remainder of the paper is structured as follows. I start with the event study in Section 2, which motivates the model structure and estimation strategy. Section 3 establishes the theoretical model and demonstrates the role of pairwise elasticities of substitution. In Section 4, Section 5 and Section 6, I apply the model to a dataset of mobile apps. Section 4 describes the dataset, LLM-based estimation and presents reduced-form evidences supporting the main model predictions. Section 5 calibrates the quantitative model to data. Section 6 uses the estimated model to quantify the impact of the GFW policy in China and TikTok's ban in U.S. Section 7 concludes.

2. Motivating Fact: An Event Study

Before delving into the theoretical model, I begin with an event study that provides empirical evidence of substitution patterns resulting from a ban targeted at specific products. The novel empirical findings highlight the need for a richer framework beyond standard models. The estimates obtained here are later used in the estimation of Section 5 to identify a key parameter, as well as in the quantitative analysis of Section 6.

[5]Embedding models are integral to advanced language models such as ChatGPT, BERT and Cohere, as well as traditional models like Word2Vec and GloVe.

Under a Melitz framework, consider the impact of a trade shock on a specific product. The predicted change in the market share of any given product can be expressed as $\sigma\,(\mathrm{d}\log P)$, where σ is the constant elasticity of substitution (CES) and $\mathrm{d}\log P$ denotes the change in the aggregate market price induced by the shock. There are two implications: (1) the source of the trade shock becomes irrelevant, and (2) the responses of all competing products exhibit identical magnitudes. If the CES assumption holds broadly, the need for a more intricate demand system is obviated.

To test this hypothesis, I estimate the market share responses of mobile apps to India's 2020 ban on Chinese apps. On June 29 2020, in response to escalating border tensions between India and China, particularly after a violent clash in the Galwan Valley, the Indian government unexpectedly announced a ban on 59 Chinese apps, including TikTok and WeChat. I manually identify 29 of these apps available on iOS App Store and match them with the monthly-active-user data segmented by app, month and market.[6] The event study framework takes the following form:

$$s_{it,IN} = \sum_{\tau} \beta_\tau \, \mathbb{1}_{t=\tau} + \delta_i + \varepsilon_{it} \tag{1}$$

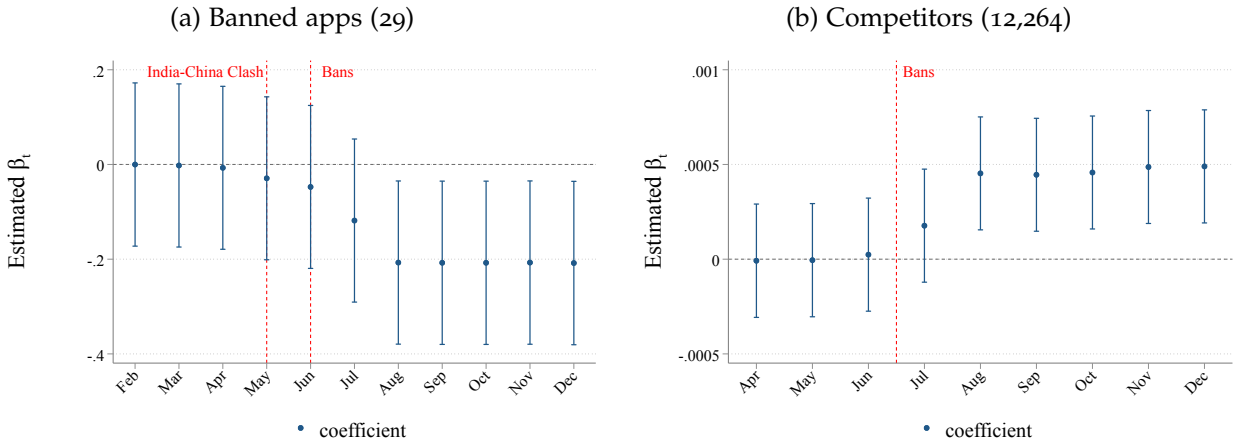
where $s_{it,IN}$ is the market share of app i in month t in India, given by the fraction of iOS monthly-active-users in percentage points, $\mathbb{1}_{t=\tau}$ are indicators for month, δ_i are fixed effects at the app level and ε_{it} is the error term that is clustered at the app level. The coefficients measure the difference in market shares between month τ and February 2020 in India. Given that the ban happened in late June 2020, the coefficients on indicator for month τ before June are expected to be zero and the coefficients on indicator for month τ after June capture the effects from the ban.

Figure 1 presents the estimation results. First, Figure 1a plots the estimated coefficients on indicator for month τ for the banned apps. Compared to a typical month before the ban, the market shares of these apps experienced a slight, statistically insignificant decline on May and June 2020 amid the military conflicts. After the ban, the estimated coefficients $\widehat{\beta}_\tau = 0.2$ indicate that the market shares of banned apps dropped by 0.2 percentage points – corresponding to approximately 85% of the market share for a medium banned app.[7] More importantly, Figure 1b plots the estimated coefficients on indicator for month τ for the remaining 12,264 apps that were not directly restricted

[6]The remaining 30 apps were only available on Android devices. All major apps that readers would recognize were successfully matched with the app usage data. Details on the data will be discussed in Section 4.2.1.

[7]The estimated coefficient may appear small in absolute terms. However, two contextual factors are crucial for its interpretation. First, the app market is highly skewed: in India, the largest app, YouTube, holds a market share of approximately 6 percentage points, while a median app captures just 0.001 percentage points. Second, TikTok, the largest banned app, had an average market share of about 3 percentage points before the ban, which dropped to 0.005 percentage points afterward. Thus, the impact of the ban is substantial when considered within this context.

Figure 1: Estimated effect of India's ban on market share



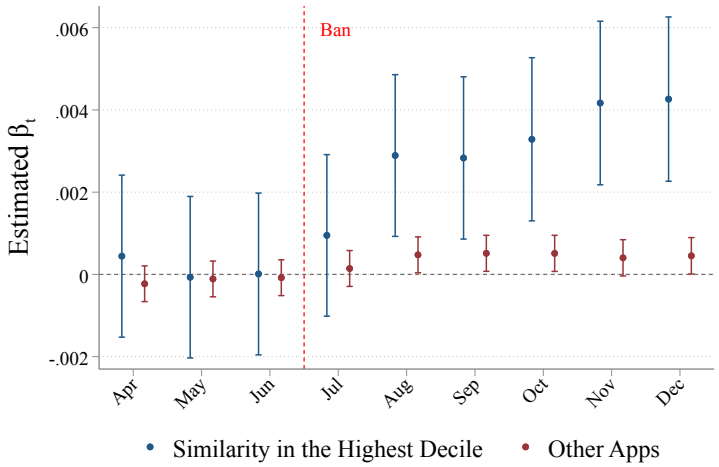
Notes: Estimated event-study coefficient β_τ of equation 1. Panel (a) plots the estimations with the observations on 29 banned apps, and Panel (b) plots the estimations with the observations on 12,264 competing apps not banned. The solid dots are coefficients and the vertical lines reflect the corresponding 95% confidence intervals. The event is defined as the bans implemented on June 2020 and indicated by the red dash line. The plotted coefficients are shown in Table A3 in Appendix B.

by the ban, reflecting the effects of competition. With the ban decreasing competition in India, I expect positive coefficients on post-June indicators. Indeed, the estimated coefficients $\widehat{\beta}_\tau = 0.0005$ are significantly positive, albeit small in magnitude. On average, competing apps gained an incremental market share of 0.0005 percentage points.

Next, I investigate the heterogenous responses across different app groups. Leaving measurement details for later discussion, I use detailed app descriptions and advanced LLMs to construct a product similarity measure for app pairs: for each app i and a banned app j, I calculate a pairwise product similarity $\omega_{ij} \in [-1, 1]$. I then construct a treatment intensity measure for each of the 12,264 competing apps as $\lambda_i = \sum_j \omega_{ij}$, which reflects the overall similarity between app i and banned apps. I then categorize the apps into groups based on deciles of the λ_i distribution. Figure 2 presents the estimated coefficients for apps in the highest decile and for those in the remaining nine deciles. The estimated coefficients for the two groups are 0.004 and 0.0005, respectively, indicating a statistically significant difference. These results demonstrate that the most similar apps experienced a significant increase in market share, whereas the dissimilar apps were virtually unaffected. This finding contradicts two key predictions of a CES framework: (1) the source of the trade shock matters, as it introduces pairwise competition effects, and (2) responses to a common shock are heterogeneous across products, driven by product similarity.

Moreover, these rich substitution patterns can be effectively identified using LLM-based similarity measures. I employ a similar specification where the explanatory vari-

Figure 2: The heterogenous estimated effects by deciles in app similarity



Notes: Estimated event-study coefficient β_τ of equation 1. The solid dots are coefficients and the vertical lines reflect the corresponding 95% confidence intervals. The event is defined as the bans issued on June 2020 and indicated by the red dash line. The plotted coefficients are shown in Table A3 in Appendix B. The robustness check with dependent variable as $s_{it,US}$ and $s_{it,RoW}$ are shown in Table A2 in Appendix B.

able $\mathbb{1}_{\tau \geq Jun}$ indicates month τ is post-June, and I show the estimation results in Table 1. In this table, each column corresponds to estimates for a specific group of apps. Column 1 and 2 repeat the estimations shown in Figure 1a and Figure 1b. Columns 3 to 5 focus on competing apps with λ_i in the top 20^{th}, 10^{th} and 5^{th} percentiles, respectively. The results show that market share gains increase significantly with greater similarity to the banned apps. In an extreme case, Moj and Josh – Indian knock-off apps replicating TikTok's core functions – rapidly expanded their market shares as replacements following the ban. This empirical evidence demonstrates that LLM-based similarity measures effectively capture pairwise substitution patterns without relying on predefined industry or product nests.

Since the event occurred during the COVID-19 pandemic, one potential identification concern is that the observed patterns might be driven by COVID-related shocks or other unobserved factors. To address this, I perform a robustness check by replacing the dependent variable with market shares in the U.S. and global markets, excluding India and China. I find that the estimated coefficients are not significantly different from 0, suggesting that the responses observed in India are not present in other markets. Furthermore, there are no significant differences in market share changes among apps in different deciles of λ_i. This robustness check helps rule out alternative channels that might explain the observed patterns and reinforces the findings presented above. The

Table 1: Estimated effect of India's ban on market share

	Banned apps	Competitors: $s_{it,IN}(\times 100)$			
		all	20^{th} percentile	10^{th} percentile	5^{th} percentile
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}_{\tau \geq Jun\text{-}20}$	-0.70***	0.03***	0.22***	0.31***	0.57***
	(0.07)	(0.01)	(0.03)	(0.05)	(0.09)
R^2	0.59	0.99	1.00	1.00	1.00
Apps	29	12,264	2,404	1,270	497
Observations	406	163,377	30,241	15,901	6,258
Fixed Effects	i	i	i	i	i

Notes: Estimated coefficient on $\mathbb{1}_{\tau \geq Jun}$. The unit of dependent variable is percentage point for banned apps, and percentage point $\times 100$ for competitors. Standard errors are clustered at app level. $^{*}\ p < 0.1$, $^{*}\ p < 0.05$, $^{*}\ p < 0.001$.

estimation results are shown in Table A2 in Appendix B.

3. Theoretical Model

In this section, I set up a granular trade model and define the equilibrium. The framework relates closely to Eaton, Kortum, and Sotelo (2012) and Gaubert and Itskhoki (2021) with a key difference: a micro-founded Translog utility function that features pairwise elasticities of substitution, as demonstrated in Section 2. In the model, there are N countries and one sector. Each country is endowed with a discrete number of firms, indexed by M. Each firm produces a single final good and can sell to multiple markets. Let i represent a firm/product, $\mathcal{M}$ the set of products, o the exporting country, and d the importing country. While this framework is broadly applicable across sectors, I will present a quantitative model specifically tailored to digital service trade embedded in mobile apps in Section 6.

Competition is modelled as a two-stage game. In the first stage, firms observe the realization of all product sets, perfectly anticipate the outcome of the second stage, and decide whether to enter each market d. In the second stage, firms that choose to enter engage in Bertrand oligopoly competition and set prices, while those that do not enter charge a reservation price and capture zero revenue. In general equilibrium, prices and wages adjust to clear both goods and labor markets in all countries.

Products Let $\mathcal{M}^o$ represent the set of potential entrants originating from country o. The number of products from country o, $M^o = \|\mathcal{M}^o\|$, is a realization of a Poisson random

process with the scale parameter $\overline{M}^o$ such that $\mathbb{E}(M^o) = \overline{M}^o$. Each of the potential entrants is born with (1) a product blueprint, and (2) a production technology. Both are drawn simultaneously from random processes. First, following the tradition of modelling differentiated products in industrial organization, a product blueprint describes a product variety as a mix of hedonic attributes. The space of attributes is fixed and indexed by $k \in \{1, 2 \ \ldots, \ K\}$, and a blueprint is a unit vector $\mathbf{a}_i$ defined in the K-dimensional attribute space:

$$\mathbf{a}_i = [a_{i1}, \ a_{i2}, \ \ldots, \ a_{iK}]^T \text{ and } \sum_s a_{is}^2 = 1 \tag{2}$$

where a_{is} is the contribution of attribute s in product i. Each potential entrant takes an *i.i.d.* draw from a von Mises-Fisher (hereforth vMF) distribution, which is a multivariate normal distribution whose support is the surface of a unit sphere.[8] Formally, the probability density function of $\mathbf{a}$ is given by:

$$f_K(\mathbf{a}, \ \boldsymbol{\mu}_o, \ \kappa) = C_K(\kappa) \exp(\kappa \boldsymbol{\mu}_o^T \mathbf{a}) \tag{3}$$

where $\boldsymbol{\mu}_o$ is a K-dimensional unit vector, $\kappa > 0$ is a scalar and $C_K(\kappa)$ is a normalization constant.[9] The distributions are independent across countries. The country-specific parameter $\boldsymbol{\mu}_o$ denotes the central direction around which product attributes $\mathbf{a}$ are concentrated. The parameter κ common to all countries, controls the dispersion of the distribution: the greater the value of κ, the more concentrated are product attributes $\mathbf{a}$ around $\boldsymbol{\mu}_o$. In the context of differentiated products, κ reflects the costliness of innovation, as unique products located further from the distribution center are rarer. By assuming the distribution, I am able to track high-dimensional product differentiation by low-dimensional directional statistics [10].

Second, each product is supplied by a unique blueprint owner with a productivity φ_i. The owner uses local labor to produce with a linear production technology, formally represented by the production function $q_i = \varphi_i L_i$. Consistent with trade literature (e.g., Melitz (2003)), I assume the productivity draws follow a Pareto distribution with shape parameter ν and scale parameter $\overline{\varphi}_o$. The distribution is defined as

$$h(\varphi_i, \ \overline{\varphi}_o, \ \nu) = 1 - (\varphi_i / \bar{\varphi}_o)^{-\nu}, \text{ where } \varphi \geq \overline{\varphi}_o \tag{4}$$

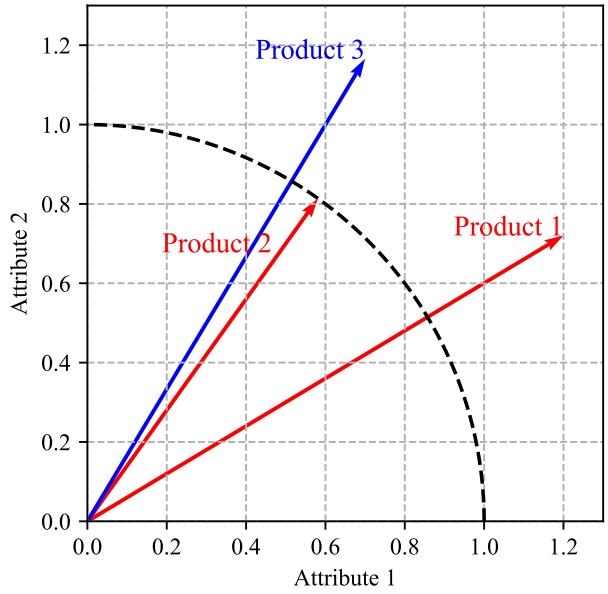
To summarize, products are heterogenous over states $\chi = (\mathbf{a}, \varphi)$, where $\mathbf{a}$ and φ are demand-side characteristic and supply-side characteristic, respectively. Figure 3 provides

[8] vMF may not be familiar to economists but is used heavily in text mining literature. Its ability to model directional data makes it highly suitable for various NLP tasks, providing a robust framework for handling the unique characteristics of text data.

[9] $C_K(\kappa) = \frac{\kappa^{K/2-1}}{(2\pi) I_{K/2-1}(\kappa)}$ where $I_s(x)$ denotes the modified Bessel function.
$I_s(x) = \sum_m^{+\infty} \frac{1}{m! \Gamma(m+s+1)} \left(\frac{x}{2}\right)^{2m+s}$.

[10] More details are provided in Appendix A.1.

Figure 3: Illustrative example of three products in 2-dimensional space



Notes: A demonstration of products when $K = 2$. There are three products depicted by three vectors. For each vector, $\mathbf{a}_i$ determines its angle and φ_i determines its length. The black dash line denotes a unit circle.

an illustrative example with three products defined in two-dimensional attribute space. Each product is depicted as a vector, where $\mathbf{a}$ determines the angle and φ determines the length. In this example, product 1 is biased towards attribute 1 and highly productive, while product 2 is biased towards attribute 2 and is less productive than product 1. Now consider a third product, represented by the blue vector, entering the market. Product 3 is highly productive, posing a competitive threat to both products 1 and 2. However, due to its similarity in attribute mix with product 2, it poses a greater competitive threat to product 2. This figures illustrates two points. Pairwise substitution patterns are essential for understanding competitive effects of entry (in this case, entry of product 3). Furthermore, there is a tradeoff between productivity and product uniqueness: product 3 can compete with the dissimilar product 1 if product 3 is productive enough. These patterns become even more complex when considering thousands of products and millions of pairwise substitution patterns.

Preferences Let us consider the demand in a given market and I drop the the market index d for simplicity. A representative consumer can observe the all product originating from all countries, such that her choice set is $\mathcal{M} = \mathcal{M}^1 \cup \mathcal{M}^2 \cup \cdots \cup \mathcal{M}^N$. To capture the distribution of products in the attribute space, I define a product matrix $\mathbf{A}$ by stacking

all hedonic attribute vectors together:

$$\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_M] = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1K} \\ a_{21} & a_{22} & \ldots & a_{2K} \\ \vdots & \ldots & \ldots & \vdots \\ a_{M1} & a_{M2} & \ldots & a_{MK} \end{pmatrix}^T \tag{5}$$

where each row $\mathbf{a}_i$ represents the hedonic attribute vector of product i, and $M = \|\mathcal{M}\|$ denotes the total number of products in the choice set. The matrix $\mathbf{A}$ summarizes the location of products in the attribute space. To further describe the properties of $\mathbf{A}$, I introduce three concepts from directional statistics (Mardia and Jupp, 2009), which are connected to the demand system discussed later.

Definition 1 *The resultant vector $\overline{\mathbf{a}}$ of matrix $\mathbf{A}$ is the sum of individual vectors and given by*

$$\overline{\mathbf{a}} = \left[\sum_i^M a_{i1}, \ \sum_i^M a_{i2}, \ \ldots, \ \sum_i^M a_{iK} \right]. \tag{6}$$

The length of the resultant vector, denoted by $\|\overline{\mathbf{a}}\|$, captures the intensity of $\mathbf{A}$; the normalized resultant vector, denoted by $\overline{\mathbf{a}}/\|\overline{\mathbf{a}}\|$, captures the average direction of $\mathbf{A}$ as a unit vector.

The resultant vector $\overline{\mathbf{a}}$ is a single vector that provides a useful summary of the overall positioning within a network. First, $\|\overline{\mathbf{a}}\|/M \propto \left(\sqrt{var(\mathbf{A})}\right)^{-1}$, ranging from 0 to 1, captures the dispersion of $\mathbf{a}$ in attribute space. A greater value indicates that the individual vectors are more aligned, suggesting that the products are more homogenous in terms of the mix of attributes. Conversely, a smaller value suggests greater dispersion, implying that the products are more differentiated. When $\|\overline{\mathbf{a}}\|/M = 1$, all products locate in the same place and the underlying $\kappa \to \infty$. In addition, resultant vector $\overline{\mathbf{a}}$ locates at the central direction of $\mathbf{A}$.[11] I further define the centrality of each product accordingly.

Definition 2 *Ω_i is the centrality of product i and given by the inner product of vector $\mathbf{a}_i$ and the resultant vector:*

$$\Omega_i = \mathbf{a}_i^T \cdot \overline{\mathbf{a}} = \sum_{j=1}^M \omega_{ij} \tag{7}$$

Ω_i measures the angular distance between product i and resultant vector $\overline{\mathbf{a}}$ and indicates how closely the product aligns with the average orientation of all vectors. When Ω_i is larger, product i is positioned closer to the center, indicating it shares a similar mix of attributes with the majority of products. In this case, product i is perceived as a

[11] When all $\mathbf{a}_i \in \mathcal{M}$ are drawn from the same distribution $f_K(\mathbf{a}, \ \boldsymbol{\mu}, \ \kappa)$, the expectation of the sample mean equals the population mean, i.e., $\mathbb{E}[\|\overline{\mathbf{a}}\|/M] = \boldsymbol{\mu}$

homogeneous good by consumers. When Ω_i is small, product i is located farther from the center, indicating that product i offers a distinctive combination of attributes. Therefore, product i is perceived as a differentiated good by consumers.

Definition 3 *The cosine similarity ω_{ij} is the inner product of vector $\mathbf{a}_i$ and vector $\mathbf{a}_j$:*

$$\omega_{ij} = \mathbf{a}_i^T \cdot \mathbf{a}_j = \sum_{s=1}^{K} a_{is} a_{js} \tag{8}$$

which indicates the product similarity between product i and j.

ω_{ij} measures the angular distance between two vectors and its value ranges from -1 to 1. In the context of attribute space, the similarity feeds directly into the patterns of product substitution. A cosine similarity closer to 1 suggests that two products share a highly overlapping mix of attributes, leading to greater substitutability between them. Conversely, lower values (closer to 0) would indicate that products are dissimilar, potentially even complementary in nature.[12] These concepts provide a foundation for interpreting the key parameters in the demand system.

A representative consumer faces a budget constraint and optimally chooses consumption levels for each product to maximize utility defined on attributes. To facilitate convenient derivations, I turn to the dual problem and derive the expenditure required to achieve a fixed level of utility. This expenditure is represented as a quadratic function of attributes and product prices, where the functional form is a specific case of the generalized hedonic demand system, as defined in equation B.1 of Pellegrino (2023), under certain homotheticity assumptions. The economic intuition for the functional form is as follows: consumers value products as bundles of attributes and fundamentally care about the price index of each attribute. However, the composition of these bundles – how attributes are "packed" together – also influences consumer preferences. This demand system yields a Translog expenditure function defined on products (Christensen *et al.*, 1975, Feenstra and Weinstein, 2017, Diewert, 2023), with parameters corresponding to the statistics described above. Formally, the expenditure required to obtain a fixed level of utility is given by:

$$\ln(e) = \alpha_0 + \sum_{i}^{M} \alpha_i \ln p_i + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta_{ij} \ln p_i \ln p_j \tag{9}$$

[12] While I allow for the possibility that products are complementary , I restrict the analysis to cases where κ is sufficiently large such that $\omega_{ij} \geq 0$ for most j and all i.

where p_i is the price of product i, α_i and the θ_{ij} are demand parameters. Notably, θ_{ij} captures the sensitivity of demand for product i to the price of product j and is given by

$$\theta_{ij} = \begin{cases} \widetilde{\theta}\omega_{ij} & j \neq i \\ -\widetilde{\theta}(\Omega_i - 1) & j = i \end{cases} \qquad \text{where } \widetilde{\theta} = \theta / \|\overline{\mathbf{a}}\| \tag{10}$$

See Appendix A.2 for detailed derivations.[13] There are three components in θ_{ij}. θ is a elasticity parameter and governs sensitivity to prices. The length of resultant vector $\|\overline{\mathbf{a}}\|$ captures a generalized love-of-variety effect – incorporating both the number of products and their dispersion in attribute space. The competition effect is amplified when dispersion increases. For the case $i \neq j$, θ_{ij} is proportional to their product similarity ω_{ij}, indicating that the demand for product i is more sensitive to price changes in product j if they locate closer to each other. For the case $i = j$, θ_{ii} is a function of $\Omega_i - 1 = \sum_{j \neq i} \omega_{ij} > 0$, indicating that the demand for product i is more responsive to its own price if it occupies a central position in $\mathbf{A}$. These three components together determine the substitution patterns. The expenditure share of product i can be solved from the *F.O.C.* with a budget constraint:

$$s_i = \alpha_i + \sum_j^M \theta_{ij} \ln p_j = \alpha_i - \widetilde{\theta}\Omega_i \left[\ln p_i - \sum_j \left(\frac{\omega_{ij}}{\sum_{j'} \omega_{ij'}} \right) \ln p_j \right] \tag{11}$$

Discussion: It is worth pointing out how the demand equation given in equation 11 differs from other recent studies. First, the Translog utility function has been used in other papers, which are the special cases of this paper. Feenstra and Weinstein (2017) examine a Symmetric Translog function where $\theta_{ij} = \theta/M$, which is a special case of my model where all products are symmetric and $\omega_{ij} = 1$. Diewert (2023) considers a case where firms are differentiated along one dimension (i.e., quality-ladder model) such that $\frac{\theta}{\sum_{i'j'} a_{i'} a_{j'}} a_i a_j$. These two cases fall within the class discussed in Matsuyama and Ushchev (2017), resulting in a single and separable price index. Therefore, they cannot generate the pairwise substitution patterns, which are the core of this paper.

Second, equation 11 has a similar expression to CES-like demand functions, such as in Baqaee *et al.* (2023), but with two key differences. First, s_i on the left-hand side is expressed in levels instead of in logarithms. This difference means that while the choke price in the CES case is infinite, in my case, a sufficiently high price can lead to

[13] This function satisfies two restrictions on θ_{ij}: (1) *symmetric restrictions* where $\theta_{ij} = \theta_{ji}$; and (2) *homotheticity restrictions* where $\sum_j \theta_{ij} = 0$ for any i and j. Additionally, by imposing a homotheticity restriction on α such that $\sum_i \alpha_i = 1$ and $\alpha_i \in (0, 1)$, the function becomes homogeneous of degree one, as discussed in Diewert (1976) and Christensen *et al.* (1975). This paper focuses on the impacts of competition, and the main conclusions do not rely on the specific parameterization of α_i. I will discuss α_i further in the quantitative exercise of Section 5.

zero sales for a product. Second, the price aggregator faced by product i is given by $\sum_j \left(\frac{\omega_{ij}}{\sum_{j'} \omega_{ij'}}\right) \ln p_j$, a weighted average of competitor prices where weights correspond to product similarity with product i. This competition effect depends on the joint distribution of $\mathbf{A}$ and $\{p_j\}_{j=1}^M$, a feature absent in other models.

As a consequence, none of the papers mentioned above can capture the pairwise competition that is central to my paper. I formalize the concept in Proposition 1 by the expressions of price elasticities and elasticities of substitution:

Proposition 1 *For any two products i and j, the cross-price elasticities of demand – measuring the responsiveness of the quantity demanded of product i to a change in the price of product j – are defined as:*

$$\eta_{ij} \equiv \left.\frac{\partial \ln q_i}{\partial \ln p_j}\right|_{U \text{ const.}} = \begin{cases} s_j + \frac{\theta_{ij}}{s_i} & j \neq i \\ -1 + s_i + \frac{\theta_{ii}}{s_i} & j = i \end{cases} \tag{12}$$

and the Allen-Uzawa partial elasticities of substitution between i and j are defined as:

$$\sigma_{ij} \equiv \frac{\eta_{ij}}{s_j} = \begin{cases} 1 + \frac{\theta_{ij}}{s_i s_j} & j \neq i \\ 1 + \frac{\theta_{ii}}{s_i^2} - \frac{1}{s_i} & j = i \end{cases} \tag{13}$$

See Appendix A.5 for derivations. When $\widetilde{\theta} = 0$, the demand function simplifies to Cobb-Douglas functional form, where all elasticities of substitution equal unity. In contrast, when $\widetilde{\theta} > 0$, the pairwise elasticities of substitution σ_{ij} between product i and product $j \neq i$ is proportional to the pairwise product similarity ω_{ij}, scaled by the market shares of the two products. This indicates that products closer in the attribute space are more substitutable for one another, while larger products – those with higher market shares – are relatively less substitutable. There are $M(M+1)/2$ pairwise elasticities of substitution among the M products. However, the model structure reduces the number of parameters to estimate by representing the attribute space with $M \times K$-dimensional vectors and a single elasticity parameter θ. This structure not only captures the rich heterogeneity of substitution patterns but also ensures tractability by significantly reducing the dimensionality of the estimation problem.

Similarly, pairwise competition also affects the extensive margin. Due to the Translog functional form, a product can achieve zero sales if its price is sufficiently high or its productivity is sufficiently low. Following the literature, I define the choke price $\overline{p}(\mathbf{a}_i)$ as the price threshold above which the product at $\mathbf{a}_i$ captures exactly zero sales. Notably, the choke price varies across the attribute space, reflecting differences in substitutability with other competitors. I formalize this concept in Proposition 2, providing an explicit expression for the choke price:

Proposition 2 *Given a product set $\mathcal{M}$, the choke price at $\mathbf{a}_i$ is the solution to $s_i = 0$ in equation 11, and is expressed as:*

$$\ln \overline{p}(\mathbf{a}_i) = \frac{\alpha_i}{\widetilde{\theta}\Omega_i} + \sum_j \left(\frac{\omega_{ij}}{\sum_{j'} \omega_{ij'}} \right) \ln p_j \tag{14}$$

The entry probability at $\mathbf{a}_i$ is then given by $\mathbb{P}[p_i(\varphi_i, \mathbf{a}_i|\mathcal{M}) < \bar{p}(\mathbf{a}_i|\mathcal{M})]$. The entry probability decreases when $\bar{p}(\mathbf{a}_i|\mathcal{M})$ is lower, which occurs under the following conditions:

1. *$\mathbf{a}_i$ is relatively closer to products with lower prices p_j, reflecting stronger substitution effects.*
2. *$\mathbf{a}_i$ is more central among all products, indicating higher competition due to overlapping attributes.*

See Appendix A.6 for derivations. Figure 4 illustrates Proposition 2. Suppose there are two products in the market: product 1 and product 2. Now consider an additional potential entrant, represented by a vector pointing to a position in the figure and determined probabilistically by random processes. The edge of the blue-shaded area indicates the minimum productivity required to charge a price lower than the choke price. Products falling within the shaded area cannot survive in the market. Notably, the edge is not a circular arc, reflecting the fact that the choke productivity varies with the angle ($\mathbf{a}$) in the attribute space. The minimum entering productivity is relatively lower in directions far from the central region, and it is also lower when the angle is closer to product 2 than to product 1. Compared to CES demand, this model derives choke prices without the need for fixed costs, avoiding the fixed-cost non-convexity typically encountered when solving the general equilibrium model.

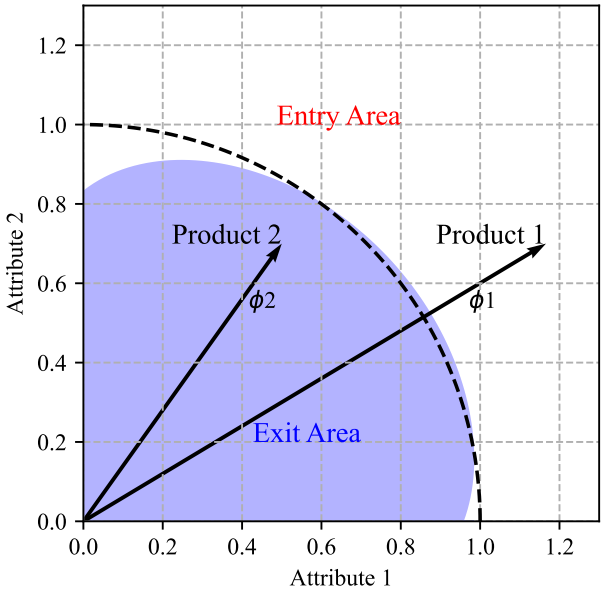
Market Structure Let us consider trade and competition. The equilibrium is solved backward, so I start with the second stage. I introduce impediments to trade in a standard way: selling one unit of product i to market d from source o requires exporting $\tau_{i,od} \geq 1$, where domestic costs are set to $\tau_{i,oo} = 1$ for all i. The effective marginal cost of product i from country o in market d is

$$c_{i,od} = \tau_{i,od} c_i = \tau_{i,od} \left(\frac{w_o}{\varphi_i} \right) \tag{15}$$

Due to the linearity of the production function, each firm's profit maximization problem is separable across markets and hence can be considered one market at a time. Given that firm i enters market d in the first stage, it choose its price $p_{i,od}$, taking as given the wage rate and the prices of its competitors, to maximize its profit from serving the market:

$$\Pi_{i,od} = s_{i,od} E_d - c_{i,od} q_{i,od} \tag{16}$$

Figure 4: Illustrative example of choke prices



Notes: A demonstration of choke prices when $K = 2$. There are two existing products depicted by product 1 and product 2. I let the price be $1/\varphi$. The edge of blue-shaded areas represents the productivity that captures zero market share at angle $\mathbf{a}$ following equation 14. The black dash line denote a unit circle.

where the expression for market share $s_{i,od}$ is given by equation 11. The solution for this Bertrand-Nash competition game is a markup price setting rule:

$$p_{i,od} = \frac{\theta_{ii} - s_{i,od}}{\theta_{ii}} c_{i,od} = \left(1 + \frac{s_{i,od}}{\widetilde{\theta}(\Omega_i - 1)}\right) c_{i,od} \tag{17}$$

Operating profits are a fraction $\frac{s_{i,od}}{s_{i,od}+\widetilde{\theta}(\Omega_i-1)}$ of revenue. Firms that are larger in size charge a higher markup and earn greater profits in equilibrium. This is commonly referred as *Marshall's second law of demand* in the literature. In addition, firms whose products are more differentiated from others, as indicated by a smaller Ω_i, also charge higher markups and gain more profits in equilibrium.

Entry In the first stage, an equilibrium of the entry game is achieved when the equilibrium profits of a subset of firms, as given by equation 16, are non-negative, while for any additional entrant profits upon entry would be negative. With a discrete number of potential entrants, multiple equilibria may exist in the entry game. To address this, we consider a sequential entry game in each market separately.[14] Specifically, consumers in

[14]The entry decision is separable by market in the absence of overhead costs. However, when overhead costs are incurred at entry, the equilibrium of the entry game is achieved when a firm's overall profits, aggregated across all markets, exceed these costs.

market d can observe all products, so the initial choice set is $\mathcal{M}$. In the first hypothetical round, all firms are assumed to enter and engage in the Bertrand price-setting game as described earlier. The profits in this hypothetical equilibrium are derived from substituting equation 17 into equation 16.

In the first round, I reassign indexes i such that $\Pi_{1^{(1)d}} \geq \Pi_{2^{(1)d}} \geq \cdots \geq \Pi_{M^{(1)d}}$ where $^{(1)}$ indicates the round number. If the least profitable firm earns on-negative profits, $\Pi_{M^{(1)d}} \geq 0$, then this constitutes an equilibrium where all potential entrants enter market d. If not, at least one firm must exit market d. I let the least profitable firm, indexed by $M^{(1)d}$, exits, and the remaining $M-1$ firms, $i \in \mathcal{M} \setminus \{M^{(1)d}\}$, repeat the Bertrand price-setting game. An equilibrium is then defined as the sequential equilibrium following definition 4:

Definition 4 *A sequential equilibrium of the entry game in market d is reached when:*

1. $\Pi_{M_d+1^{(M-M_d-1)d}} < 0$ *in the* $(M-M_d-1)^{th}$ *round*
2. $\Pi_{M_d{}^{(M-M_d)d}} \geq 0$ *in the* $(M-M_d)^{th}$ *round*

$\mathcal{M}_d = \mathcal{M} \setminus \{M^{(1)d}, M-1^{(2)d}, \ldots, M_d+1^{(M-M_d-1)d}\}$ *is the set of active firms. Firms choose their prices as follows:*

$$p_{i,od} = \begin{cases} \left(1 + \frac{Ms_{i,od}}{\theta(\Omega_i - 1)}\right) c_{i,od} & i \in \mathcal{M}_d \\ \\ \bar{p}_{i,d} & i \in \mathcal{M} \setminus \mathcal{M}_d \end{cases} \tag{18}$$

where $\bar{p}_{i,d}$ is the reservation price from equation 14, at which the firm captures zero sales.

The sequential entry equilibrium is defined in a way reminiscent of Gaubert and Itskhoki (2021). In their model, firms are heterogeneous in productivity only, allowing for an ex-ante ranking by productivity. Starting with the most productive firms and proceeding to the least productive ensures that the resulting equilibrium is the most efficient among all possible equilibria. In contrast, such a straightforward ranking is not feasible in this model due to the complex interplay between hedonic attributes and productivity, which precludes a single ex-ante indicator for firm ranking. To address this, I frame the problem on the exit side, using the dual approach. In the appendix, I explore alternative sequences to test the robustness of the welfare outcomes. Moreover, consistent with Feenstra and Weinstein (2017), the equilibrium expenditure fully accounts for the introduction of new goods and the disappearance of existing ones, with finite reservation prices set by inactive firms.

General Equilibrium is defined by vectors of wages and prices such that goods and labor markets clear in all countries. Specifically, the aggregate incomes equal aggregate expenditures in any market d

$$\sum_{i \in \mathcal{M}} p_{i,od} q_{i,od} = E_d = w_d L_d + \Pi_d \tag{19}$$

where E_d denotes the aggregate expenditures, w_d is the wage rate, L_d is the labor supply, and Π_d represents the aggregate profits. In addition, labor market clearing requires that the aggregate labor supply equals the total labor demand from all domestic firms in each country:

$$\sum_{i \in \mathcal{M}^o} \sum_{d}^{N} \frac{q_{i,od}}{\varphi_i} = L_o \tag{20}$$

For any realization of the product sets, a unique equilibrium exists following the sequential entry process described above. The equilibrium is defined in Proposition 3:

Proposition 3 *Given the set of parameters* $\{L_n,\ \overline{M}_n,\ \overline{\varphi}_n,\ \nu,\ \boldsymbol{\mu}_n,\ \kappa,\ \theta\}_n$ *and trade barriers* $\{\tau_{od}\}_{o,d}$*, there exists a unique equilibrium defined by* $\{\mathcal{M}_n,\ w_n,\ \{p_{i,od},\ s_{i,od}\}_{i \in \mathcal{M}_n}\}_n$ *for any realization of* $\{\chi_i\}_{i \in \mathcal{M}}$*, following the entry sequence described above.*

See Appendix A.7 for a proof.

4. Model Implications and Empirical Evidence

4.1. Granular bans in the model

Bans targeting specific products have become increasingly common. For instance, TikTok faces restrictions in multiple countries and is banned outright in India; X/Twitter has faced legal issues with Brazil's Supreme Court; and Huawei and ZTE are restricted in the U.S. and Canada. Understanding how these bans shape competition is critical. This model offers novel predictions that are absent from standard trade models. I represent bans as non-tariff trade barriers. Specifically, the iceberg costs defined in equation 15 consist of two components:

$$\ln \tau_{i,od} = \ln \tau_{od} + \ln \overline{\tau}\, \mathbb{1}_{i,\text{policy}_d} \tag{21}$$

where $\ln \tau_{od}$ is the standard dyadic gravity costs and $\ln \overline{\tau}$ is the costs from policy intervention. Specifically, a ban on product i by country d increase i's marginal costs by $\overline{\tau}$ in market d. In the context of digital services, modelling a ban as an iceberg cost reflects the fact that the policy interventions often result in slowdown in response times and limited

local services. Even when an app is fully banned, a non-negligible number of local users may still access it via virtual private networks (VPNs).[15]

I evaluate the impact of bans by a comparative statics between the initial equilibrium without the bans to the new equilibrium with the bans. Following the convention of trade literature, I use hat algebra notation to denote the changes from the initial equilibrium to the new equilibrium – formally, $\hat{x} = x_1 - x_0$.[16] The policy intervention is a ban on a set of products $j \in \{j \mid \mathbb{1}_{j,\text{policy}} = 1\}$. First, the changes in choke prices are given by:

$$\widehat{\ln \bar{p}}_{i,d} = \left(\sum_{j|\mathbb{1}_{j,\text{policy}}=1} \frac{\omega_{ij}}{\Omega_i} \right) \ln \overline{\tau} \tag{22}$$

The ban induces the entry of new products by raising the choke prices, and more importantly, more entries occur in locations relatively closer to the banned products in the attribute space. Figure 5 illustrates the changes in choke prices and their impact on firm entry. Consider a ban on product i, which reduce the effective productivity of product 1 from φ_1 to $\frac{2}{3}\varphi_1$. This policy intervention diminishes competition, shifting the exit area from the blue-shaded region to the red-shaded region. The changes in choke prices, $\widehat{\ln \bar{p}}_{i,d}$, are depicted as the difference between the edges of these two regions.

Second, the changes of the market shares for continuing products are given by:

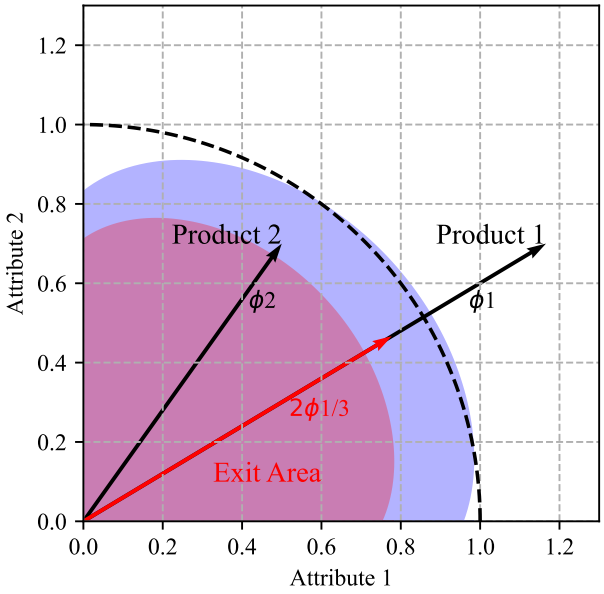
$$\widehat{s}_{i,od} = \widetilde{\theta} \left(\sum_{j|\mathbb{1}_{j,\text{policy}}=1} \omega_{ij} - \Omega_i \mathbb{1}_{i,\text{policy}} \right) \ln \overline{\tau} \tag{23}$$

For banned products, the changes in market shares include both the direct effect of the ban and the competition effect from other banned products. The overall changes are always negative as $\omega_{ij} \leq 1$, indicating that the banned products shrink in market shares. For products that are not banned, the changes are solely due to the competition component and are strictly positive. As a result, competing products benefit from reduced competition, with those that are more similar to the banned products – and thus more substitutable – expanding their market shares to a greater extent. Notably, equation 23 corresponds to the empirical specifications presented in Section 2. To fully account for the long-term responses, I turn to China's GFW implemented decade ago and quantify its impact from a full quantitative model.

[15] Prior to 2022, Apple almost never removed banned products from the market. However, amid escalating political tensions, Apple began removing apps like WhatsApp and Threads from the Chinese App Store in response to government orders, starting in April 2024.

[16] Compared with Dekle, Eaton, and Kortum (2007), the changes as defined in levels not in ratios for simplicity purpose. This difference is due to the Translog utility function. I will also use ratios as the robustness check for empirical evidences.

Figure 5: Illustrative example of policy intervention



Notes: A demonstration of the impact of a granular ban on product 1. The black vectors denote the products and the red vector denote the effective productivity after product 1 is banned. The blue-shaded region denote original exit area and the red-shaded region denote new exit area after the ban. The black dash line denote a unit circle.

4.2. Data and institutional background

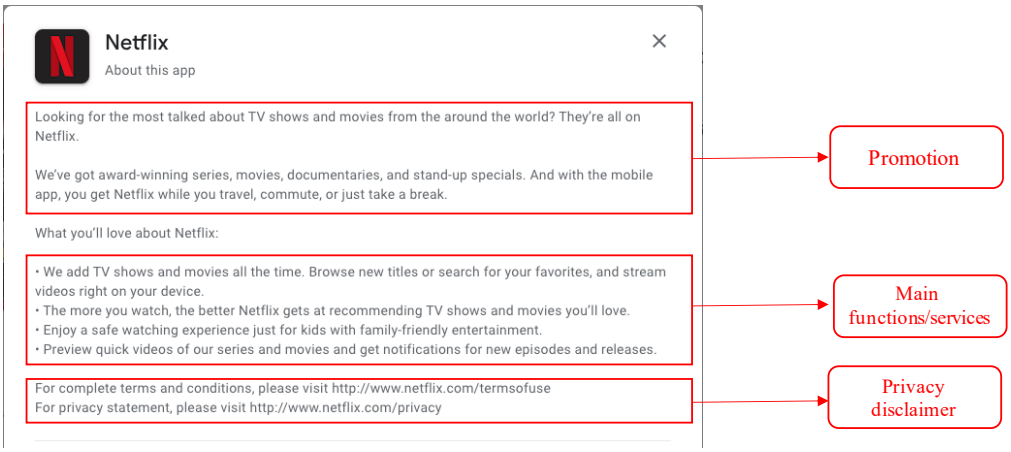
4.2.1. Mobile app data

To validate the mechanisms of the model and quantify the welfare implications, I apply the model on a dataset for mobile apps. The primary database is on app usage data purchased from SensorTower. Sensor Tower is the largest and most reliable company providing app-level metadata.[17] For each app, I have detailed metadata scraped from their App Store pages, including app descriptions, release dates, publishers, publisher countries, and app categories. I use the publisher's country as the exporting country of the app and determine the production costs and trade costs faced by each app. More importantly, I leverage rich textual descriptions at the product variety level to estimate the product attributes and pairwise similarities.

The data also track app-level usage by destination country the apps available on the iOS App Store. I have monthly data on the number of monthly active users by destination country for a substantial period, from 2015 to 2021. Using this data, I construct a high-

[17] A potential concern may be the absence of usage data from Android systems. First, as Google Play is unavailable in China, comprehensive Android market information for the Chinese market is limited. Second, within each platform, an app's market share is highly correlated between iOS and Android.

Figure 6: Example of app description for Netflix



Notes: A screenshot of Netflix's iOS app store page: https://apps.apple.com/us/app/netflix/id363590051

frequency market share measure, calculated as the fraction of users for an app relative to the total users across all iOS apps. This variable is used as the dependent variable in Section 2. In addition, I have a snapshot of the time spent on each app by destination country in 2021, representing the total number of hours from all users on an app in that year. This variable captures equilibrium time allocation and is used as a dependent variable in Section 4.3 as well as the trade flow in Section 5 and Section 6.

4.2.2. Measuring app attributes

One key reason for utilizing this dataset is its rich textual data at individual app level. App descriptions, provided by publishers, outline the main functionalities, benefits, and use cases, helping users understand what the app does and how it can enhance their experience. A well-crafted description not only explains features but also addresses app regulations, encouraging downloads and long-term engagement. Developers can use descriptions to clarify permissions, subscriptions, or in-app purchases, fostering transparency with potential users. Thus, the descriptions contain valuable information about product attributes, though they also include unrelated noises. Figure 6 displays an example for Netflix. In this example, there are rich information about the main function and service for Netflix along with the promotion and privacy disclaimer. The goal of the estimation is to get a numerical measure with a focus on product attributes, which is achieved from a two-step procedure with LLMs.

To handle complex tasks and improve accuracy, high-quality labeled data is often required, though it is scarce and costly. As a result, recent advances in natural language processing increasingly rely on synthetic data, particularly data generated by large lan-

guage models (LLMs) (Bubeck, Chandrasekaran, Eldan, Gehrke, Horvitz, Kamar, Lee, Lee, Li, Lundberg *et al.*, 2023, Liu, Wei, Liu, Si, Zhang, Rao, Zheng, Peng, Yang, Zhou *et al.*, 2024). Josifoski, Sakota, Peyrard, and West (2023) show that when certain tasks cannot be directly solved by LLMs, a more efficient approach is to reverse the task by prompting the LLM to generate input text (which LLMs excel at) from predefined output structures. In this paper, I adopt a similar methodology to that proposed by Josifoski *et al.* (2023) to estimate the product network for universal mobile apps.

First, I utilize the "`GPT-4o-mini`" model, one of the most advanced and efficient LLMs available, to construct a zero-shot prompt for generating synthetic data. The knowledge base for this task is derived from the descriptions of all iOS apps. Specifically, I prompt the LLM to act as an iOS app analyst, identifying key features, target audiences, and unique selling points to evaluate how closely apps compete within the same market segment. The output is structured as a concise paragraph, followed by a bulleted list of core functionalities. The language and tone remain consistent across outputs. To minimize noise from subjective or inconsistent information, elements such as subscription options, privacy policies, and terms of service are excluded. Potential biases in the data and further details on prompt construction are discussed in Appendix C.1.

Next, I use the synthetic data generated from this process to feed an embedding model, "`text-embedding-3-large`", which estimates the corresponding vector representations. A transformer-based model extracts features from user-submitted text and encodes these features as numeric vectors, known as embeddings. Each embedding is a unit vector in a fixed-dimensional space, conceptually similar to the attribute vectors, as demonstrated in Figure 3, where each vector represents a distinct subject area within the LLM's corpus. The synthetic data generated in the first step effectively bridges the gap between product space and corpus space, ensuring that the vectors are meaningful in both contexts. The outputs are unit embedding vectors with 3,072 dimensions.

In summary, the two-step process outlined above allows me to generate unit embedding vectors from synthetic data, based on the app description knowledge base. In the subsequent quantitative analysis, these embedding vectors serve as the measurements of **a** in the model. Figure 7 displays a visualization of estimated embeddings on the surface of a three-dimensional ball. Each colour block represents model predicted app categories within which apps are very similar, such as search engines or photo sharing platform.

4.2.3. The Great Firewall policy

This paper uses the data to examine the impact of the Great Firewall (GFW) policy in China, a set of bans implemented during the early internet era between the two largest economies. The GFW is a combination of legislative actions and technologies enforced by

Figure 7: Estimation of **A** for mobile apps and app categories



Notes: This is a visualization of **A** from textual app descriptions. I use UMAP algorithm to reduce 3072-dimensional vectors to 3-dimensional vectors. Each node represents an app in my dataset and all apps are distributed on a unit ball. Each colour represents an app category predicted from aggolomerative clustering algorithm based on the 3-dimensional vectors. Notice that relative locations of the nodes cannot perfect reflect the ω_{ij} in quantitative exercise due to the dimensionality reduction.

the Chinese government to regulate the domestic internet market. Concerned about the spread of illegal and harmful content – such as pornography, violence, fake news, and politically sensitive information that could threaten social stability and national security – the Chinese government initiated a project to monitor sensitive words and filter content, which became known as the "Great Firewall" project. This initiative began in 1998 and was completed in 2008.

A key function of the GFW is to selectively block access to content using methods like DNS spoofing, URL filtering, and IP range bans. In 2009, China's Ministry of Industry and Information Technology (MIIT) issued the "Notice on Establishing a Blacklist Management System for Illegal Domestic Websites."[18] In the same year, YouTube and Facebook were added to the blacklist. In 2010, Google exited the Chinese market, and all its domains and products were subsequently blacklisted. Today, many mobile apps and

[18] *Notice on Establishing a Blacklist Management System for Illegal Domestic Websites [2009] No.371*

websites are blocked, particularly those of companies unwilling to comply with China's strict regulations on data collection and privacy, including Google and Meta.

Although the existence of the blacklist is well known, the standards and timeline for banning sites have been opaque, leaving users uncertain about when a foreign app or website might be banned. Chinese internet users could continue using foreign services while being aware of the risk of sudden, unannounced bans. Even after official bans, some Chinese users bypass the firewall using VPNs. Certain universities and foreign-owned companies provide VPN access for their students and employees under special permissions. However, this method carries increased costs and potential political risks. For instance, in October 2023, a software engineer had 1 million RMB of illegal income confiscated. According to the disclosed verdict, he was penalized for illegally "using a VPN to access Google Chrome, log into GitHub, and set up meetings on Zoom."[19]

I manually collect the information of blacklist apps from public information and crosscheck with the results from China Firewall Test (chinafirewalltest.co), which is a real-time check on banning status on domains.[20] I found 268 banned apps that can be matched with my dataset. 131 are main banned apps (e.g., Instagram), and the rest are supplementary apps (e.g., Layout from Instagram). The majority of bans happened around 2010. Therefore, I treat the snapshot in 2021 as the equilibrium with the bans. Although I do not observe the pre-treated equilibrium, I will construct a valid control group for each of the banned apps and address the identification with the cross-sectional data variation.

4.2.4. Final dataset

The final dataset includes app attribute vectors, app usage by country, and a list of apps banned in China, merged using unique app ids from the App Store. Figure 8 illustrates app usage patterns across countries, with monthly active users in December 2020 and time spent in January 2021. On average, a smartphone user actively uses 15 apps on her phone and spends 6 hours per month on each app,[21] amounting to approximately 850 total hours spent on apps. This is a non-trivial utility component for consumers. Table 2 lists the top 40 apps by global time-spent on apps. First, the distribution is highly skewed: total hours spent on the most popular app, YouTube, reaches 390 billion hours, while the 40th most popular app, Reddit, are 3 billion hours. Second, the market is predominantly dominated by apps from the United States and China. Third, among the top 40 apps, China has banned nearly all foreign apps, with the exceptions of Microsoft

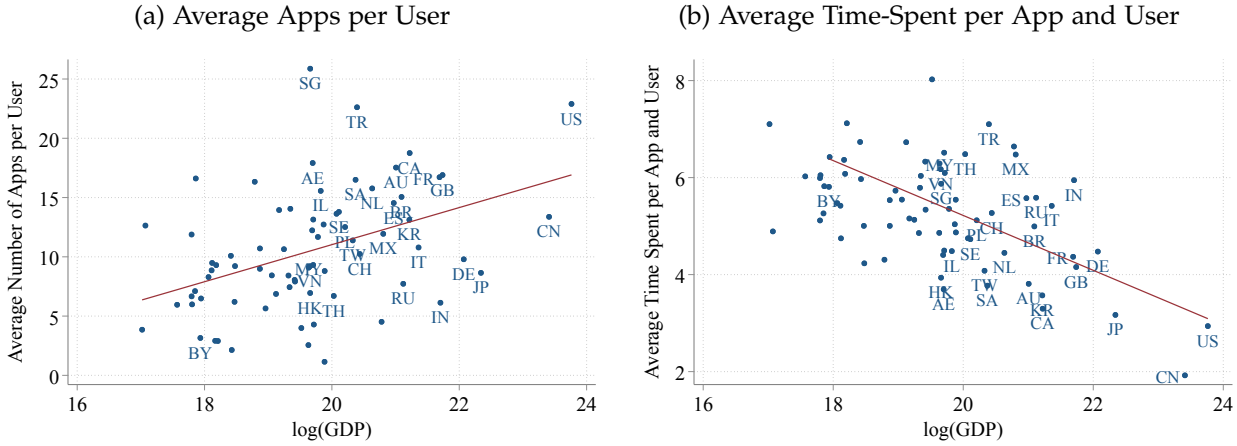
[19]See news sources from https://www.sohu.com/a/732098677_121124373

[20]See details in appendix Appendix D

[21]The average time is calculated using market share as weights. The unweighted average time spent per app per user is 1.2 hours per month, reflecting intensive use on popular apps.

and VK. These features create a natural experimental setting to examine the impact of bans under the GFW policy.

Figure 8: App usage by country



Notes: This figure plot the average apps per user and average time-spent per app and user by country in 2020. Each dot represent a country. The GDP data is from CEPII. There are 30 countries labelled with 2 digital ISO code. They have the largest market share in app market.

4.3. Reduced-form evidences

Lacking in information in the initial equilibrium before the policy intervention, I cannot directly test the predictions in equation 22 and equation 23. Instead, I construct control groups to isolate the impact of the GWF. First, consistent with the literature, the trade barriers can be inferred from the difference between domestic market share and foreign market share – formally:

$$\begin{aligned} \Delta s_{i,od} &= s_{i,oo} - s_{i,od} \\ &= \frac{\theta \Omega_i}{M} (\ln \tau_{od} + \ln \overline{\tau}\, \mathbb{1}_{i,\text{policy}} - \Delta \ln P_{i,od}) \end{aligned} \tag{24}$$

where $\Delta \ln P_{i,od}$ is the difference in the multilateral resistant term. Consider two products i and i' developed from the same country o, occupying the same position in attribute space, while product i being banned in market d and product i' not. The net effect of ban is simply given by

$$\ln \overline{\tau} = \frac{M}{\theta \Omega_i} \left(\Delta s_{i,od} - \Delta s_{i',od} \right) \tag{25}$$

The differences eliminate the potential omitted variables including culture barriers, IPRs, productivity and etc. Therefore, for each banned app, I find their control group as the

Table 2: Top 40 apps by worldwide time spent in 2021

	Name	Exporter	Worldwide time spent (bn hours)		Name	Exporter	Worldwide time spent (bn hours)
1	YouTube*	US	390	21	VK	RU	8
2	TikTok^{+}	CN	188	22	Google Maps*	US	8
3	Facebook*	US	142	23	Spotify*	SE	6
4	WhatsApp*	US	114	24	Discord*	US	6
5	Instagram*	US	109	25	Pinterest*	US	6
6	WeChat^{+}	CN	60	26	Waze*	US	6
7	Messenger*	US	29	27	Wattpad*	CA	6
8	Snapchat*	US	29	28	Redbook	CN	5
9	Twitter*	US	26	29	Zalo	CN	4
10	QQ^{+}	CN	21	30	Tencent Video	CN	4
11	Chrome*	US	19	31	Gmail*	US	4
12	iQiyi	CN	19	32	Bilibili	CN	4
13	Netflix*	US	14	33	KakaoTalk^{+}	KR	4
14	LINE*	JP	14	34	Microsoft Teams	US	4
15	Baidu	CN	11	35	Kugou Music	CN	3
16	Kwai	CN	11	36	ZOOM^{+}	US	3
17	Youku	CN	11	37	Yahoo! JAPAN*	JP	3
18	Twitch*	US	9	38	YouTube Kids*	US	3
19	Taobao	CN	9	39	Alipay	CN	3
20	Telegram*	AE	9	40	Reddit*	US	3

Notes: This Table lists the top 40 apps ranked by their worldwide time spent in 2021. $*$ indicates apps banned in China and $+$ indicates apps not banned but under restrictions with the services. Almost all foreign big apps are banned or partially banned in China, while two exceptions are VK from Russia and Microsoft Teams from the U.S.

Table 3: Summary statistics for market shares (percentage points)

	Observations	mean	25^{th} percentile	medium	75^{th} percentile
$s_{i,oo}$					
Banned Group	78	16.386	0.325	3.245	18.816
Control Group	781	1.146	0.005	0.033	0.282
$s_{i,o-CN}$					
Banned Group	78	3.552	0.001	0.132	1.735
Control Group	781	0.272	0	0	0.009
$s_{i,o-HK}$					
Banned Group	78	13.677	0.102	1.605	15.609
Control Group	781	0.328	0	0	0.009

apps from the same exporting country and in the same model predicted category (same colour block in Figure 7). For example, the control group for Chrome is Bing and Firefox. The empirical specification is

$$\Delta s_{i,oCN} = Ban_i + \lambda_c + \lambda_o + \varepsilon_{i,oCN} \tag{26}$$

where Ban_i is an indicator for whether the app is banned in China, λ_c is the app cluster fixed effects, λ_o is the exporting country fixed effects and $\varepsilon_{i,od}$ is the error term. Table 3 the market shares in the domestic country, mainland China, and Hong Kong for both the ban and control groups respectively. The banned apps are significantly larger than those in the control group, suggesting that the GFW policy is not solely about information censorship; major players with similar functionalities are more likely to be banned. Moreover, both groups have smaller market shares in mainland China and Hong Kong compared to their domestic markets, indicating that gravity factors such as language barriers, cultural differences, and localization still play a role. However, the banned apps experience a relatively larger market share loss in mainland China.

Table 4 presents the estimation results. In column 1 to 3, I include different fixed effects, and the results are robust across these specifications. Compared to the control group, banned apps lost an additional 11 percentage points in market share in China. Column 4 replicates the specification from column 3 but includes all supplementary apps. While the estimated effect is smaller, it remains both statistically and economically significant. In column 6, I replace the dependent variable with $\Delta s_{i,o-HK}$ and find that the market share of banned apps is not significantly different from that of apps in the control group in Hong Kong. This suggests that the identified coefficient captures the causal impact of the GFW policy. The magnitude of this effect is substantial.

Table 4: The effect of bans on foreign apps in China (mainland) and Hong Kong

	$\Delta s_{i,o-CN}$				$\Delta s_{i,o-HK}$
	(1)	(2)	(3)	(4)	(6)
Ban	11.72***	11.51***	11.06***	7.46***	1.09
	(0.41)	(2.47)	(2.56)	(1.54)	(1.02)
Observations	859	859	859	2,335	848
R^2	0.138	0.163	0.172	0.116	0.064
Fixed Effects	-	*c*	*c*, *o*	*c*, *o*	*c*, *o*

Next, I further investigate the impacts on Chinese apps competing with the banned foreign apps. I categorize apps into approximately 300 narrowly defined categories based on their similarities and classify these categories into two groups: those with banned foreign apps and those without. I expect that more Chinese apps enter and that these apps are larger in size in categories with banned apps. Referring to the illustrative example in Figure 5, the differences between app categories resemble the distinction between regions near banned products and regions near other products, reflecting variations in choke prices. The empirical model is specified as follows:

$$y_{cCN} = Ban_i + \lambda_{\bar{c}} + \varepsilon_{cCN} \tag{27}$$

where $y_{ct,CN}$ denotes the number of Chinese apps and their market share in category c, $\lambda_{\bar{c}}$ denotes the fixed effects of categories defined in the App Store (e.g., social networking and productivity) and ε_c is the error term. Table 5 presents the estimation results. Columns 1 and 3 show results for app categories with at least one main banned app, while columns 2 and 4 include categories with at least one supplementary app as well. Categories with banned foreign apps have around 10 additional Chinese apps, and their market shares are 15 percentage points higher compared to categories without bans. The median size of an app category is 102 products with 0.63 million users, indicating that the effect is substantial.

To summarize, these reduced-form evidences support the main predictions on the impacts of the bans, as presented in equation 22 and equation 23. To understand the welfare implications, I will build a quantitative GE model and calibrate the model to data in the next section.

Table 5: The effect of the GFW on the entry and market share of Chinese apps

	Number of Product		Market share	
	(1) Main	(2) All	(3) Main	(4) All
ban	10.79***	5.83***	14.27***	7.02***
	(1.29)	(0.70)	(3.71)	(1.93)
Observations	222	293	222	293
R^2	0.051	0.038	0.032	0.029

5. Quantification

5.1. Quantitative Model

This section introduces a quantitative trade model incorporating both physical goods and app services, where app service producers engage in Bertrand competition à la the theoretical model. I rationalize the monetization of app sector through advertisement sector (Greenwood, Ma, and Yorukoglu, 2024, Kopytov, Roussanov, and Taschereau-Dumouchel, 2023, Rachel, 2021).[22] In this model, app producers provide services to consumers for free and sell user attention as ad slots to advertisers, with a market-specific ad price. The advertisers are physical goods producers who endogenously determine their ad expenditure. Both physical goods and app services are tradable across countries, while ads are non-tradable.

Demand The utility function of a representative consumer in country d is given by:

$$\ln u_d = \beta \ln Q_d + (1-\beta) \ln l_d \tag{28}$$

where Q_d represents physical goods consumption, and l_d represents digital service consumption, specifically from mobile apps. The consumer has one unit of time that can be allocated either to work or to using apps. All income is spent on physical goods consumption, subject to the following budget constraint:

$$P_d^p Q_d = w_d (1 - h_d) + \pi_d \tag{29}$$

where P_d^p is the price index for physical goods, h_d is the time allocated to app usage, w_d is the wage rate and π_d is profit rebate distributed to local workers.

[22]This is a common modelling approach for leisure enhancing technologies and digital services, including television, web browsing, and mobile apps.

App Services Potential app entrants are realized from exogenous random processes based on exporting country o and category c. Specifically, the set of potential entrants $\mathcal{M}^{o,c}$ are generated as follows: (1) the number of potential entrants $M^{o,c} = \|\mathcal{M}^{o,c}\|$ is drawn from Poisson process with the scale parameter $\overline{M}^{o,c}$; (2) hedonic attributes $\mathbf{a}_i$ is drawn from vMF distribution $f_K(\mathbf{a}, \boldsymbol{\mu}^{o,c}, \kappa)$; and (3) productivity φ_i is drawn from Pareto distribution $h(\varphi, \overline{\varphi}^{o,c}, \nu)$. The set of potential entrants from country o is given by $\mathcal{M}^o = \mathcal{M}^{o,1} \cup \mathcal{M}^{o,2} \cup \cdots \cup \mathcal{M}^{o,C}$. I allow for endogenous entry across exporting country $o \times$ category c blocks.

The digital service consumptin derived from using apps in market d is given by $l_d = h_d / P_d^a$, where P_d^a represents the effective cost of the app usage and follows the Translog functional form specified in equation 9:

$$\ln(P_d^a) = \alpha_0 + \sum_i^M \alpha_i \ln p_{i,od} + \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \theta_{ij} \ln p_{i,od} \ln p_{j,od}. \tag{30}$$

The price of an app $p_{i,od}$ is defined as the amount of time required to derive one unit of utility. For example, if a consumer's goal is to find a piece of news, the price of using Chrome is measured by the time spent searching and downloading information. The markup $\mu_{i,od}$ reflects strategy of app producers, who intentionally design user interactions to slow down processes, thereby maximizing the time spent on their app. This increases "eyeballs" available as ad slots for advertisers.

App producers employ software engineers, denoted by H_o^E, to produce and maintain services that are free in monetary terms but require time from consumers. Taking the wage rate w_o and total screen-time h_d in each country as given, app producers engage in the two-stage game described in the theoretical model and equilibrium is solved by vectors of $\{\mathcal{M}_d, \{p_{i,od}, q_{i,od}\}_{i \in \mathcal{M}}\}_{d=1}^N$. In addition, I explicitly consider the trade costs from the GFW such that $\mathbb{1}_{i,\text{GFW}} = 1$ if app i is banned in China.

Physical Goods I introduce the physical goods in the simplest way possible, namely, using a Krugman (1980) model. There exists a continuum of goods, with one unit originating from each country o. The consumption of physical goods in market d can be expressed as:

$$Q_d = \left(\sum_o^N \int b_{od}^{\psi}(i) q_{od}^P(i)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}} \tag{31}$$

where $q_{od}^P(i)$ is the quantity of product i from country o, σ denotes the elasticity of substitution in physical goods sector, $\psi \in (0, 1)$ is a parameter that governs the efficiency of advertisement, and $b_{od}(i)$ is the brand equity built from advertisement investment

associated with product i (Cavenaile and Roldan-Blanco, 2021, Rachel, 2021).[23] All producers from country o use the same production technology and choose prices and advertisement expenditures to maximize their profit in each market:

$$\Pi_{od}^{P} = \max_{b_{od}, p_{od}} \left\{ x_{od}^{P} - c_{od}^{P} q_{od}^{P} - b_{od} \right\} \tag{32}$$

where x_{od}^{P} represents the trade value of physical goods from country o to country d. q_{od}^{P} is the trade quantity such that $q_{od}^{P} = T_{o}^{P} L_{od}^{P}$ with T_{o}^{P} as the productivity. The unit cost is given by $c_{od}^{P} = w_{o} \tau_{od}^{P}$, where τ_{od}^{P} denotes iceberg trade costs in the physical goods sector. The optimal monopolistic constant markup is solved as $\sigma/(\sigma-1)$. The optimal advertisement expenditure is derived as a fraction of trade value from exporting country o in importing country d, as follows:

$$b_{od} = \frac{\psi}{\sigma - 1} x_{od}^{P} \quad \text{and} \quad x_{od}^{P} = \frac{b_{od}^{\psi} (w_{o}^{P} \tau_{od}^{P} / T_{o}^{P})^{1-\sigma}}{\sum_{o'} b_{o'd}^{\psi} (w_{o'}^{P} \tau_{o'd}^{P} / T_{o'}^{P})^{1-\sigma}} E_{d} \tag{33}$$

where E_{d} is the aggregated expenditure in market d.

Ad Exchange There is a single Ad Exchange platform that sets a uniform price for ads in each market d (Decarolis and Rovigatti, 2021).[24] For simplicity, I assume the platform operates with zero profit. Ads can only be sold locally, meaning app producers sell consumer attention in market d to physical goods producers operating in the same market. The monetary profit of app i acquired from market d is given by

$$\Pi_{i,od} = p_{d}^{b} \left(s_{i,od} h_{d} \right) - c_{i,od} q_{i,od} \tag{34}$$

where p_{d}^{b} represents the price of ads in market d. The ad price is is determined by the condition that total ad expenditure equals the total revenue of the app sector in each market:

$$p_{d}^{b} h_{d} = \sum_{o}^{N} b_{od} = \frac{\psi}{\sigma - 1} E_{d} \tag{35}$$

General Equilibrium is achieved through vectors of wages and labor supplies that clear the labor markets, app markets, physical goods markets, and ad markets in all countries. Consumers allocate their time optimally, satisfying the first-order condition:

$$\frac{w_{d} h_{d}}{P_{d}^{p} Q_{d}} = \frac{1 - \beta}{\beta} \tag{36}$$

[23] This demand specification is a special case of Rachel (2021) where $\chi = 1$. An alternative approach is to model advertisement as reducing search costs to reach more consumers, as in Arkolakis (2010). We adopted this approach in our prior work Sun and Trefler (2023). Both approaches yield consistent conclusions, but this method is computationally simpler.

[24] In this case, the Ad Exchange is Apple Transparent Tracking for iOS apps

which determines the screen time d. Assuming manufacturing workers and software engineers are perfectly substitutable, the total labor force in each country is fixed and denoted by $\overline{L}_d$. Labor markets clear when:

$$\sum_{d}^{N} \frac{q_{od}^{P}}{T_o^P} + \sum_{i \in \mathcal{M}^{o,c}} \sum_{d}^{N} \frac{q_{i,od}}{\varphi_i} = H_d^P + H_d^E = (1 - h_d)\overline{L}_d \tag{37}$$

where H_d^P and H_d^E are labor supplied to physical goods sector and app service sector respectively.

5.2. Estimation Procesure

In this section, I discuss the estimation procedure of parameters in the quantitative model and fit the model to data.

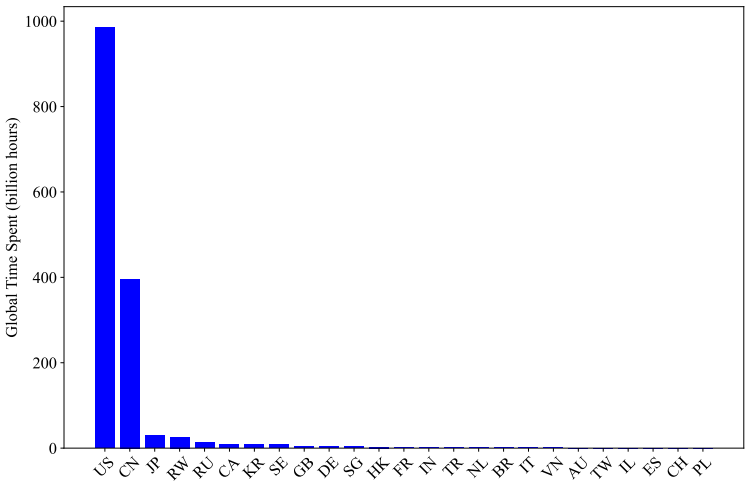
5.2.1. Additional assumptions and data

Table 6 summarizes all the parameters in the quantitative model by sectors. I use the vMF and Pareto distribution to characterize the hedonic attributes **a** and productivity φ. I further parameterize α_i as $\alpha_i = \frac{\Omega_i^{-\eta}}{\sum_j \Omega_j^{-\eta}}$ where $\eta \geq 0$. This expression captures the "tech enthusiasm" phenomenon, where consumers value product uniqueness and assign it greater weight in the utility function. When $\eta = 0$, meaning $\alpha_i = 1/M$, the differences in market share only comes from the quadratic terms, capturing the competition effect. These assumptions enable me to translate high-dimensional estimation at the product level into estimation on low-dimensional statistics.

I estimate the parameters using a cross-sectional data in 2021. The estimation focuses on 25 economies that collectively account for 98% of global app service exports, 84% of physical goods exports, 81% of world GDP, and 62% of global employment. Figure 9 illustrates global time spent on apps by exporting country, revealing a significant concentration of exports from the U.S. and China. As a result, I model a world with three economies ($N = 3$): the U.S., China and the rest of the world (RoW), which represents the remaining 23 economies excluding the U.S. and China. Product realization processes are defined at the country-category level, resulting in $N \times C = 30$ blocks. The model scales with world nominal GDP, so I normalize China's nominal GDP to 1. In addition, the labor forces of each economy are fixed, and I also normalize $\overline{L}_{CN} = 1$.

The core of my analysis relies on the dataset described in Section 4.2, which captures bilateral trade flows measured by time spent on apps, disaggregated by product and market. Complementary data on bilateral trade flows for physical goods, GDP, population, and gravity covariates are sourced from the CEPII BACI Database. Employment data and average annual working hours are obtained from the Penn World Table 10.0. The ICT

Figure 9: The density of pairwise elasticity of substitution



Notes: The global time spent is the summation of time-spent across apps produced by exporting countries and all markets in 2021

sector employment share is drawn from the OECD Going Digital Toolkit, supplemented by national statistics bureaus for economies not covered by the OECD.

Table 6: Summary of Parameters

	Mobile App	Physical Goods	Labor
Demand	$\{ \boldsymbol{\mu}^{o,c}, \kappa, \eta, \theta \}$	$\{ \sigma, \psi \}$	$\{ \beta \}$
Supply	$\{ \overline{M}^{o,c}, \overline{\varphi}^{o,c}, \nu \}$	$\{ T_o, \tau_{od}^P \}$	$\{ \overline{L}_o \}$

5.2.2. Estimation strategy

I estimate the model parameters in two steps. First, I calibrate the parameters that can be directly inferred from observed data moments or identified externally. In the second step, I use the simulated method of moments (SMM) to estimate the remaining five parameters. Due to the model's granularity, the simulation relies on realizations of the random processes. I fix the random seeds throughout the estimation and conduct robustness checks using different random seeds. I start with introducing the separately identified parameters and then discuss the SMM procedure.

Estimating $\widetilde{\theta}$ is a key elasticity parameter in the pairwise elasticities of substitution. The estimates of $\widetilde{\theta}$ is inferred from the event study presented in Section 2. The empirical specification corresponds to the changes in market shares shown in equation 23, which

Table 7: Dependent Variable: $s_{it,IN}(\times 10^{-4})$

	(1)	(2)	(3)	(4)
$Post_t \times \omega_{ij}$		0.01***	0.04***	0.17***
		(0.00)	(0.01)	(0.06)
$Post_t$	0.02***	-0.17***	-0.46***	-1.97**
	(0.01)	(0.05)	(0.09)	(0.80)
Observations	165,227	165,227	129,965	10,501
R^2	0.99	0.99	0.99	0.99
Within R^2	0.0001	0.0002	0.0004	0.0016
Fixed Effects	i	i	i	i
Sample	all	all	weighted	$s \geq 10^{-4}$

gives a structural interpretation of the estimated coefficient. Particularly, the model predicted changes in market share are:

$$s_{i\text{-}post,IN} - s_{i\text{-}pre,IN} = \widetilde{\theta} \left(\sum_{j|\mathbb{1}_{j,\text{policy}}=1} \omega_{ij} \right) \ln \overline{\tau} = \widetilde{\theta} \lambda_i \times ln\overline{\tau} \tag{38}$$

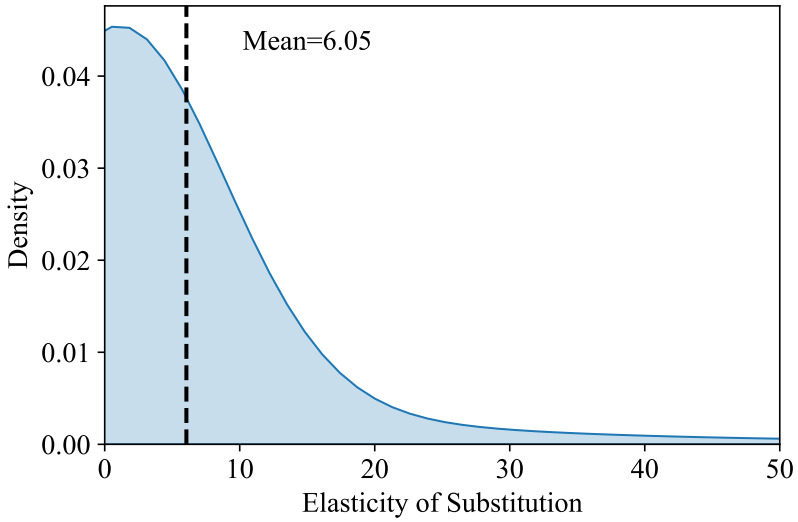
Therefore, the corresponding empirical specification is as follows:

$$s_{it,IN} = \beta \times \lambda_i \times \mathbb{1}_{t \geq Jun\text{-}20} + \delta_i + \varepsilon_{it} \tag{39}$$

where the estimated coefficient $\widehat{\beta}$ reflect the value of $\widetilde{\theta} \times ln\overline{\tau}$ in the model. Table 7 presents the estimation results. Column 1 shows results with only the post-period dummy as the explanatory variable, while columns 2 to 4 include interactions between the post-period dummy and λ_i. After controlling for the interaction term, the coefficients on the dummy become negative and insignificant. In addition, the explanatory power of the empirical model, measured by within R^2, increases when I restrict the sample to large apps.

The unique nature of digital bans provides a reasonable interpretation of $\ln \tau$. After the ban, local consumers still can get access to the banned services via VPN. However, using a VPN can slow down internet speeds due to factors like data encryption, server distance, and network congestion. On average, VPN usage can reduce speeds by 30% (for a paid VPN at $10 USD/month) to 50% (for free VPNs), although the exact impact varies based on the VPN service and server location; higher-quality VPNs experience less slowdown compared to free or poorly optimized ones. I set the shock magnitude to $\ln \tau = 0.69$ reflecting that a user now use a free VPN to get access to those banned apps and are slowed down by 50%. Consequently, the estimated $\widetilde{\theta} = \hat{\beta} / \ln \tau = 1.7 \times 10^{-5} / 0.69 = 2.46 \times 10^{-5}$. Figure 10 plots the kernel density of the pairwise elasticities of substitution in India before the ban, with a market share-weighted average elasticities of substitution of 6.05.

Figure 10: The density of pairwise elasticity of substitution



Notes: This figure is plotted by app usage data in India on May 2020. $\widetilde{\theta}$ is set to be 2.46×10^{-5}.

Estimating vMF mean I assume that apps from a country o and category c are drawn from the vMF distribution $f_K(\mathbf{a}, \boldsymbol{\mu}^{o,c}, \kappa)$. Therefore, the resultant vector for apps $i \in \mathcal{M}^{o,c}$ satisfies the convergence condition:

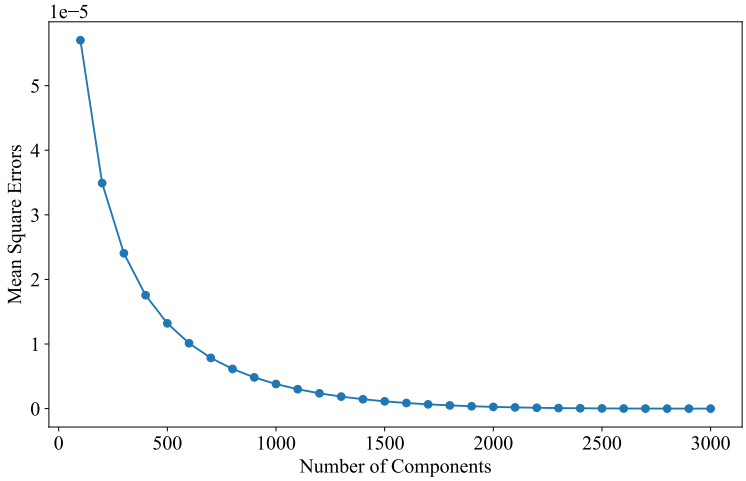
$$E\left[\frac{\overline{\mathbf{a}}^{o,c}}{\|\overline{\mathbf{a}}^{o,c}\|}\right] = \boldsymbol{\mu}^{o,c} \tag{40}$$

with a sufficient large number of realizations. I can directly calibrate $\boldsymbol{\mu}_o^c$ by calculating the resultant vectors in data. The dimension of embedding vectors is fixed with the embedding model, which is 3,027 in this exercise. For computational efficiency, I reduce the dimensionality using the Truncated SVD Algorithm, while retaining most of the information on pairwise cosine similarity.[25] Figure 11 shows the mean square error (MSE) in predicting pairwise cosine similarity with lower-dimensional vectors. The MSE is less than 1×10^{-5} when $K \geq 500$, with only minor changes beyond this point. Thus, the additional dimensions do not contribute significantly to predicting cosine similarity between embeddings, and I use 500-dimensional vectors as proxies for $\overline{\mathbf{a}}^{o,c}$.

Estimating $\overline{M}^{o,c}$, $\overline{\varphi}^{o,c}$, τ_{od} and $\overline{\tau}$ I estimate the number of potential entrants, scale parameter in productivity and iceberg trade costs to match observed trade flows in the app sector. With the Poisson-Pareto productivity structure, any observed productivity

[25] A more familiar approach is Principal Component Analysis (PCA), but the Truncated SVD Algorithm performs better in predicting cosine similarities.

Figure 11: Mean Square Error (MSE) with low dimensionality



distribution can be realized by various combination of $\overline{M}_o$ and $\overline{\varphi}^{o,c}$, making it impossible to identify these two parameters separately. However, the Translog functional form accounts for the effect of inactive firms on competition. I set all $\overline{M}^{o,c}$ to be the mean of $\overline{M}^{o,c}$ in the data, focusing sectoral differences on productivity without introducing an excess of shadow products into the economy.

As with many trade models, this model faces an identification challenge between productivity and unilateral trade costs. However, since the counterfactual analysis focuses on the extra costs from direct bans on U.S. apps in the Chinese market, the conclusions do not hinge on separating productivity from unilateral trade costs. In my approach, I use the market share in RoW as a reference point, take the auxiliary value of the relative wage as given, and normalize $\varphi^{CN,1}$ to be 1. Table 8 presents the estimation results. In addition, the extra trade costs due to the GFW are estimated as $\overline{\tau} = 10$, indicating that banned U.S. apps face trade costs 10 times higher than those not banned.

Table 8: Trade Costs in App Setor

	Importer		
	CN	US	ROW
CN	1.00	2.33	1.01
US	8.37	1.00	0.59
ROW	5.56	1.55	1.00

Estimating σ and ψ σ and ψ are identified from matching two data moments. The gravity equation derived from 41 is as follows:

$$\frac{\pi_{od}^P}{\pi_{dd}^P} = \left(\tau_{od}^P \frac{w_o^P / T_o^P}{w_d^P / T_d^P} \right)^{\frac{1-\sigma}{1-\psi}}. \tag{41}$$

This equation takes the same form as the standard gravity estimation where the trade elasticity of physical goods is now given by $(1-\sigma)/(1-\psi)$. I borrow the externally identified value from the literature (e.g., Caliendo and Parro (2015)) and let the trade elasticity for physical goods sector $(1-\sigma)/(1-\psi)$ to be 4. Second, the expenditure share on advertisement is given by $\psi/(\sigma-1)$ from equation 35. I let the parameter match the average expenditure share in data and set $\psi/(\sigma-1)$ to be 0.05. These two data moments fix the value of σ and ψ.

Estimating T_o^P and τ_{od}^P I estimate the T_o^P and τ_{od}^P from the bilateral trade flows in physical goods sectcor. I follow Head and Ries (2001) to estimate bilateral trade costs. Assume that $\tau_{od}^P = \tau_{do}^P$ and given trade elasticity, we can back out symmetric trade barriers as the symmetric Head-Ries index:

$$\tau_{od}^P = \left(\frac{\pi_{od}^P}{\pi_{dd}^P} \frac{\pi_{do}^P}{\pi_{oo}^P} \right)^{\frac{1-\psi}{2(1-\sigma)}} \tag{42}$$

Taking the auxiliary value of the relative wage as given, I can estimate T_o^P as the residual in equation 41 with a normalization $T_{CN}^P = 1$. The relative productivity between two normalizers, $\varphi^{CN,1}/T_{CN}^P$, is calibrated in SMM estimation.

Table 9: Summary of parameters separately identified

Parameter	Interpretation	Data Moments/Source	Value
App sector			
$\widetilde{\theta}$	elasticity parameter	$\hat{\beta}/\ln\overline{\tau}$	2.46×10^{-5}
$\boldsymbol{\mu}^{o,c}$	unit resultant vectors	LLM embeddings	$\overline{\mathbf{a}}^{o,c}/\|\overline{\mathbf{a}}^{o,c}\|$
$\overline{M}^{o,c}$	number of potential entrants	the average number	150
τ_{od}	dyadic trade costs	trade shares	$s_{i,od}$
$\overline{\tau}$	GFW costs	banned apps	10.49
Physical goods sector			
$(\sigma-1)/(1-\psi)$	trade elasticity	Caliendo and Parro (2015)	4
$\psi/(\sigma-1)$	revenue share of apps	expenditure share on ICT	0.05
τ_{od}^P	country-pair iceberg costs	Head-Ries Index	$\left(\frac{\pi_{od}^P}{\pi_{dd}^P}\frac{\pi_{do}^P}{\pi_{oo}^P}\right)$

Table 9 summarizes the parameters in the first group. These parameters are separately identified by a direct match with observable data moments. The targeting data moments and value of parameters are displayed in the table.

SMM Estimation We use simulated method of moments (SMM) to estimate the remaining five parameters of the model, $\Theta = \{\kappa,\ \nu,\ \eta,\ \beta,\ \varphi^{CN,1}/T_{CN}^{P}\}$. The estimation proceeds as follows: for a given parameter vector Θ, I simulate the model, compute a list of cross-sectoral moments $\mathbb{M}(\Theta)$, and contrast them with the equivalent moments in the data $\widetilde{\mathbf{m}}$. I search for parameter vector $\widehat{\Theta}$ that minimizes the distance between the model and data moments, according to the loss function $\mathcal{L}(\Theta) \equiv (\mathbb{M}(\Theta) - \widetilde{\mathbf{m}})' \mathbf{W} (\mathbb{M}(\Theta) - \widetilde{\mathbf{m}})$, where $\mathbf{W}$ is the weighting matrix. Specifically, we search for the best fitting points from the grid. The details for SMM procedure is described in appendix.

I target five groups of data moments, which corresponds the sectoral outcomes in three economies. Table 10 summarize the parameters and the main targeted moments for each parameter. First, the concentration parameter κ is identified by the dispersion of product attributes. Theoretically, the concentration parameter κ is the solution for

$$\overline{a} = \frac{I_{K/2}(\kappa)}{I_{K/2-1}(\kappa)} \tag{43}$$

where $\overline{a} = \|\overline{\mathbf{a}}\| / M$ is the resultant vector length and I_v is the modified Bessel function of the first kind and order v. For a large number of observations, κ can be estimated using the following approximation (Sra, 2012) :

$$\hat{\kappa} \approx \frac{\overline{a}(K - \overline{a}^2)}{1 - \overline{a}^2} \tag{44}$$

However, the competition reduce the density in more central locations. Thus, the estimation of $\hat{\kappa}$ cannot be directly mapped to κ. Second, following Gaubert and Itskhoki (2021), I estimate the Pareto shape parameter ν from the tail distribution of market shares. Third, η captures the love-of-innovativeness from the demand side and target the covariance between market shares and product centrality. Fourth, the utility weight on real consumption is informative on the time-spent on apps relative to working hours. Fifth, the relative level of the two normalizations on productivities in China target to the employment shares of app sector relative to the manufacturing sector.

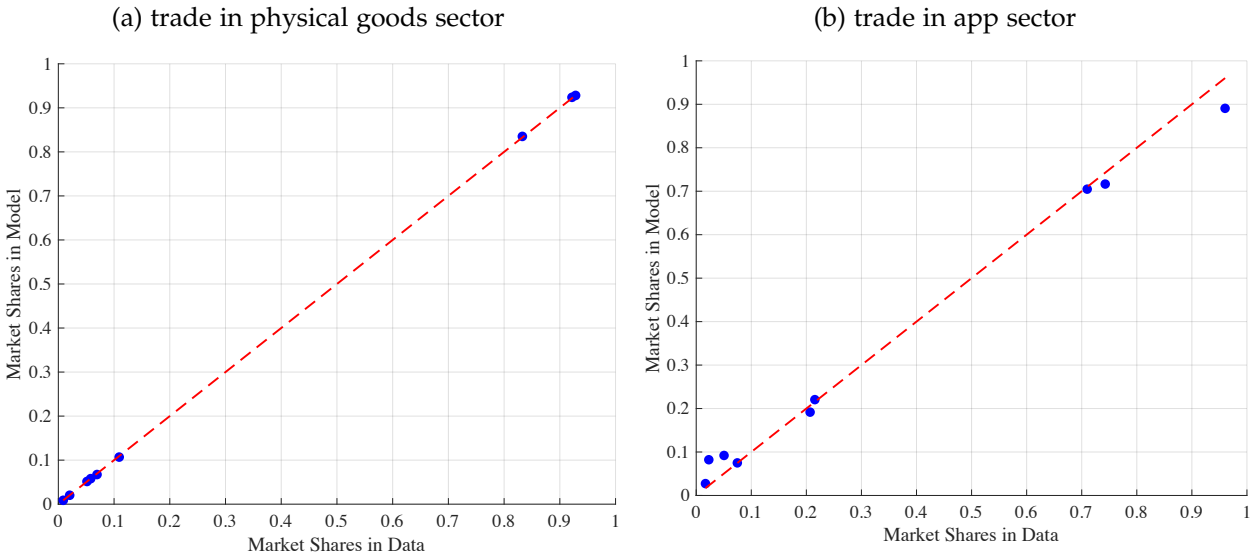
Model Fit Table 11 reports the model-based values of the moments used in the estimation, comparing them to their empirical counterparts. Overall, the model provides a reasonable fit to the data across all five groups of moments. Due to computational constraints, the current SMM estimation is performed on a parameter grid of only 20 points, so further refinements on a finer grid are likely to improve accuracy. First, the

Table 10: Summary of parameters in SMM

Parameter	Interpretation	Main targeted moments	Value
κ	spherical concentration	$var(\mathbf{a}_i)$	360
ν	Pareto shape parameter	tail distribution of s_i	1.21
η	love-of-innovativeness	$cov(s_i, \Omega_i)$	1.03
β	utility weight on real consumption	time-allocation	0.75
$\varphi^{CN,1}/T^P_{CN}$	relative producivity	employment shares	1.1

estimation closely matches the hedonic attribute dispersion for all active apps in each market. Notably, the model accurately replicates bilateral trade shares in both sectors. Figure 12 shows the bilateral trade shares predicted by the model against those observed in the data for the physical goods and app sectors. The points cluster tightly around the 45-degree line, indicating a strong match. However, the model predicts lower skewness in market shares than the data exhibit, which may stem from its assumption of constant returns to scale and the absence of externalities. Second, the model aligns well with observed time allocation and labor allocation across countries and sectors. Third, it also performs well in predicting non-targeted moments, such as relative wages and relative nominal GDP.

Figure 12: Model fit in bilateral trade flows



Notes: This figure plot the average apps per user and average time-spent per app and user by country in 2020. Each dot represent a country. The GDP data is from CEPII. There are 30 countries labelled with 2 digital ISO code. They have the largest market share in app market.

Table 11: Moments in data and estimation

	Moments		Data $\widetilde{\mathbf{m}}$	Model $\mathbb{M}(\Theta)$
Targeted Moments				
1.	Variance of hedonic attributes	$var(\mathbf{a}_{i\in\mathcal{M}})$	0.6074	0.6235
	– active apps in China	$var(\mathbf{a}_{i\in\mathcal{M}_{CN}})$	0.6733	0.6564
	– active apps in the US	$var(\mathbf{a}_{i\in\mathcal{M}_{US}})$	0.5921	0.5432
	– active apps in RoW	$var(\mathbf{a}_{i\in\mathcal{M}_{ROW}})$	0.5514	0.5259
2.	Top-product sales share	$\overline{s}_{1,od}$	0.386	0.164
	Top-3 product sales share	$\sum_{i=1}^{3}\overline{s}_{i,od}$	0.620	0.382
	Top-5 product sales share	$\sum_{i=1}^{5}\overline{s}_{i,od}$	0.687	0.454
	Top-10 product sales share	$\sum_{i=1}^{10}\overline{s}_{i,od}$	0.783	0.511
3.	Covariance of market share and centrality	$cov(s_i, \Omega_i)$	-0.05	-0.08
4.	Average time-spent on apps	h_d		
	– time-spent in China	h_{CN}	0.274	0.295
	– time-spent in the US	h_{US}	0.326	0.309
	– time-spent in RoW	h_{RoW}	0.335	0.313
5.	Employment share of app sector	$L_d^A/\overline{L}_d$		
	– employment in China	$L_{CN}^A/\overline{L}_{CN}$	0.035	0.059
	– employment in the US	$L_{US}^A/\overline{L}_{US}$	0.034	0.013
	– employment in RoW	$L_{RoW}^A/\overline{L}_{RoW}$	0.083	0.051
Non-targeted Moments				
1.	Number of active products			
2.	Relative wage	w_d		
	– US/China	w_{US}/w_{CN}	1.57	1.29
	– RoW/China	w_{RoW}/w_{CN}	6.40	6.21
3.	GDP	E_d		
	– US/China	E_{US}/E_{CN}	2.27	1.60
	– RoW/China	E_{RoW}/E_{CN}	1.42	1.28

Notes: This table displays the data moments using the estimation in Table 9 and Table 10. The current is subject to searching on 20 grids and will be improved with finer searching.

6. Counterfactuals

In this section, I use the quantitative model to evaluate the aggregate and distributional implications of bans targeted at specific foreign products. I focus on the policy-induced trade costs, represented by $\ln \overline{\tau}\ \mathbb{1}_{i,\text{policy}_d}$, to replicate the effects of actual bans. These additional costs operate over and above the uniform dyadic trade costs and any productivity differences, ensuring that the quantitative results are not tied to the identification of baseline trade barriers or productivity parameters.

I conduct two counterfactual exercises using the estimated parameters. First, I consider the absence of China's GFW policy by removing the policy-induced trade barriers. This exercise allows me to decompose the resulting welfare changes for Chinese consumers. Second, I examine the impact of a ban on TikTok – China's most productive social networking app-on market outcomes in the U.S. In this exercise, I uncover the ignored distributional effect among competing apps as a result of heterogenous substitution patterns.

6.1. The impact of the Great Firewall policy

Armed with the estimated model, I now quantify the impact of the GFW policy in China. I consider a counterfactual scenario where $\overline{\tau} = 1$ and take the difference between data and the counterfactual outcome as the impact of the GFW policy. Table 12 displays the changes in welfare components. The changes in x is defined as $x^{\text{Data}}/x^{\text{Counterfactual}} - 1$. Quantitatively, the GFW policy has two distinct effects on welfare of Chinese consumers. On the one hand, the GFW increased real incomes by shifting profits from banned products to Chinese alternatives, resulting in a 1.05% welfare gain. On the other hand, this benefit was outweighed by a significant reduction in the utility from app services. While the overall time spent on apps decreased only slightly, the utility per unit of time-spent dropped by 29%. Consequently, given the calibrated utility weight on real income (75%), the net welfare loss from the GFW policy is 7.48% for Chinese consumers.

Table 13 further decomposes the sources of welfare changes. The income effect primarily arises from rent shifting within the app sector: the number of active Chinese apps increases only by 0.34% in China, while profits rise by 7.65%, driven by head-to-head oligopolistic competition among the tech giants. For instance, Tencent would not have reached a valuation of $486.24 billion if Meta had remained in China. The loss in app services, however, stems from using inferior domestic substitutes, such as replacing Chrome with Baidu. While the overall time spent on apps remains nearly constant, the effective cost of app usage rises sharply – by approximately 40%. It indicates that a Chinese consumer receives 40% less valuable information during the same search time on Baidu

compared to a U.S. consumer using Chrome. These findings suggest that policymakers may have incentives to impose such bans to boost real income. However, the loss in utility due to lower-quality screen time is often overlooked in policy considerations.

Table 12: Welfare decomposition of the GFW by country

Country	Welfare	Real Incomes	App Services
China	-7.48%	1.05%	-29%
US	0.11%	-0.81%	2.88%
RoW	0.59%	0.00%	2.38%

In addition, I examine the impact on the level of product diversity among active apps. The model exhibits a novel measure of product diversity by the dispersion of product attributes among active apps in a market. I find that the product dispersion among Chinese apps decreased by 4.71%, indicating that the policy not only replaced foreign products with inferior domestic alternatives but also led to the survival of more homogenous goods with fewer distinctive hedonic attributes. This finding suggests that reduced competition stifles product innovation. It aligns with the finding in Hsieh *et al.* (2021), and rather than assuming exogenous innovation efficiency, the results here demonstrate that lower innovation can be an endogenous outcome of reduced competition.

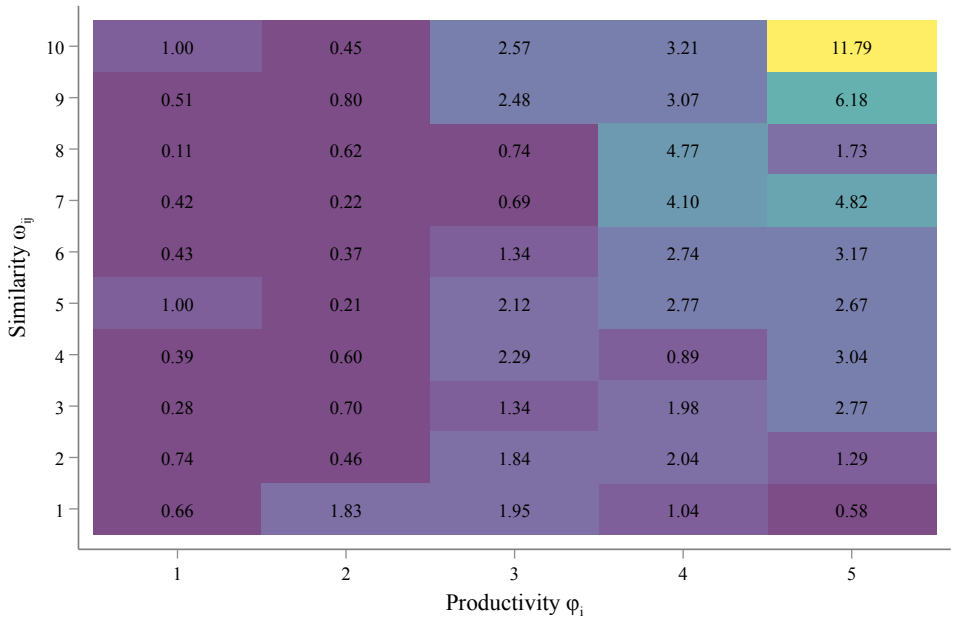
Table 13: The impact of the GFW on Chinese app sector

	App Services		Real Incomes		Product
	Screen Hours	Effective Costs	Profits	Entry	Dispersion
Changes	-0.63%	39.97%	7.65%	0.34%	-4.71%

6.2. The impact of ban on TikTok

Talk of banning TikTok began years ago as its influence grew in the U.S. In April 2024, President Biden issued a ban requiring ByteDance to sell the platform by January 19, 2025 – just one day before Inauguration Day – making it difficult for President Trump to reverse the decision. Meanwhile, Trump pledged to "never ban TikTok" and argued that "such a ban would only benefit Mark Zuckerberg and Facebook". In this counterfactual exercise, I simulate the policy by imposing a $\overline{\tau}$ barrier on the top Chinese social networking app, representing TikTok.

Figure 13: Changes in revenue on the US market



I examine the distributional effects of the ban on apps in the U.S. market. For each app, I measure its similarity to TikTok using $\omega_{i,\text{TikTok}}$. I then group apps into fifty bins, formed by intersecting deciles of similarity and quintiles of productivity. Figure 13 shows the average changes in revenue for each bin, where revenue changes are defined as $(x^{\text{Data}}/x^{\text{Counterfactual}}) - 1$, and the average change is normalized to 1. A key insight is that apps with high similarity to TikTok and high productivity gain more from the ban. In particular, apps in the highest similarity and productivity category (including Facebook and Instagram) increase their U.S. revenue by 11 times more than the average app. This result is consistent with the political debate surrounding TikTok, but would not emerge from a standard model assuming equal substitution patterns among all product pairs.

7. Conclusion

This paper examines the welfare implications of trade policies targeted at specific foreign products. The standard tools fall short in addressing this issue, as policy-targeted products typically have a vast user base and a differentiated mix of hedonic attributes, leading to complex substitutability patterns with competitors. I find causal evidence of these substitution patterns through an event study: apps with high similarity to banned apps captured greater increase in market share after the India's ban on Chinese apps. More importantly, the substitution patterns can be identified using LLM-based similarity measures.

Building on this empirical finding, I develop a model in which large, single-product firms with heterogeneous productivities and hedonic attributes engage in Bertrand oligopolistic competition. I introduce a Translog expenditure function, in which the key parameters endogenously emerge from a micro-foundation of hedonic demand. The model is purpose-built to exploit textual product descriptions and large language models: pairwise elasticities of substitution are proportional to pairwise cosine similarities between product descriptions. I integrate the demand system with a granular trade model and solve for the static equilibrium through a two-stage competition.

This model is estimated with a detailed dataset on mobile apps. Equipped with the estimated model, I quantify the welfare implications of China's Great Firewall policy: a censorship system restricting the local availability and functionality of major information-sensitive, foreign (mainly U.S.) apps. I find that the Great Firewall policy causally fostered Chinese tech giants by shifting profits away from their foreign counterparts, resulting in a 1.05% increase in Chinese real incomes. However, this gain was offset by the utility loss from the use of inferior apps, yielding a net welfare change of -7.48%. I also uncover a novel interaction between trade barriers and product diversity: reduced competition from policy intervention led to a decline in product differentiation among Chinese domestic products. Finally, I find that the revenue increase for apps highly similar to TikTok and very productive (e.g., Threads & Instagram) is by 11 times more than that of the average app in the U.S. after TikTok's ban. This finding aligns with the current political debate regarding the disproportionate benefits accruing to Meta.

This paper contributes to the growing literature on granular firms in international trade. The demand-side heterogeneity has been under-explored, despite its empirical importance alongside supply-side heterogeneity. This paper – both methodologically and empirically – provides a practical approach to capturing rich demand-side heterogeneity and demonstrates its relevance for welfare analysis. Moreover, the model and estimation strategy are easily implementable and widely applicable to research areas where rich textual data is available such as the occupational substitution in labor market or factor substitution in production.

References

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo. 2022. Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics* 40(S1):S293–S340.

Acemoglu, Daron and Pascual Restrepo. 2022. Tasks, automation, and the rise in us wage inequality. *Econometrica* 90(5):1973–2016.

Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson. 2017. Nonparametric counterfactual predictions in neoclassical models of international trade. *American Economic Review* 107(3):633–689.

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira. 2020. Patents to products: Product innovation and firm dynamics .

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira. 2023. Patents to Products: Product Innovation and Firm Dynamics. Working paper, Federal Reserve Bank of Atlanta.

Aridor, Guy. 2022. Drivers of digital attention: Evidence from a social media experiment. *Available at SSRN 4069567* .

Arkolakis, Costas. 2010. Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy* 118(6):1151–1199.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. 2019. The elusive pro-competitive effects of trade. *The Review of Economic Studies* 86(1):46–80.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. New trade models, same old gains? *American Economic Review* 102(1):94–130.

Baldwin, Richard and James Harrigan. 2011. Zeros, quality, and space: Trade theory and trade evidence. *American Economic Journal: Microeconomics* 3(2):60–88.

Banerjee, Arindam, Inderjit S Dhillon, Joydeep Ghosh, Suvrit Sra, and Greg Ridgeway. 2005. Clustering on the unit hypersphere using von mises-fisher distributions. *Journal of Machine Learning Research* 6(9).

Baqaee, David Rezza, Emmanuel Farhi, and Kunal Sangani. 2023. The darwinian returns to scale. *The Review of Economic Studies* :rdad061.

Bartik, Alexander, Arpit Gupta, and Daniel Milo. 2023. The costs of housing regulation: Evidence from generative regulatory measurement. Working paper.

Bernard, Andrew B., Esther Boler, Davin Chor, Sirig Gurund, and Wei Lu. 2024. The great firewall and knowledge diffusion. Working paper, EITI 2024.

Bernard, Andrew B, Andreas Moxnes, and Yukiko U Saito. 2019. Production networks, geography, and firm performance. *Journal of Political Economy* 127(2):639–688.

Bernard, Andrew B, Andreas Moxnes, and Karen Helene Ulltveit-Moe. 2018. Two-sided heterogeneity and trade. *Review of Economics and Statistics* 100(3):424–439.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* 63(4):841–890.

Bian, Bo, Xinchen Ma, and Huan Tang. 2021. The supply and demand for data privacy: Evidence from mobile apps. *Available at SSRN* .

Boeing, Philipp, Loren Brandt, Ruochen Dai, Kevin Lim, and Bettina Peters. 2024. The anatomy of chinese innovation: Insights on patent quality and ownership. *ZEW-Centre for European Economic Research Discussion Paper* (24-016).

Bradford, Anu. 2023. *Digital empires: The global battle to regulate technology*. Oxford University Press.

Bubeck, Sébastien, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, *et al.* 2023. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712* .

Caliendo, Lorenzo and Fernando Parro. 2015. Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies* 82(1):1–44.

Cavenaile, Laurent and Pau Roldan-Blanco. 2021. Advertising, innovation, and economic growth. *American Economic Journal: Macroeconomics* 13(3):251–303.

Christensen, Laurits R, Dale W Jorgenson, and Lawrence J Lau. 1975. Transcendental logarithmic utility functions. *The American Economic Review* 65(3):367–383.

Decarolis, Francesco and Gabriele Rovigatti. 2021. From mad men to maths men: Concentration and buyer power in online advertising. *American Economic Review* 111(10):3299–3327.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2007. Unbalanced trade. *American Economic Review* 97(2):351–355.

Diewert, W Erwin. 1976. Exact and superlative index numbers. *Journal of econometrics* 4(2):115–145.

Diewert, W Erwin. 2023. A generalization of the symmetric translog functional form. *Journal of International Economics* :103821.

Eaton, Jonathan, Samuel S. Kortum, and Sebastian Sotelo. 2012. International trade: Linking micro and macro. working paper, National bureau of economic research.

Feenstra, Robert C and David E Weinstein. 2017. Globalization, markups, and us welfare. *Journal of Political Economy* 125(4):1040–1074.

Gaubert, Cecile and Oleg Itskhoki. 2021. Granular comparative advantage. *Journal of Political Economy* 129(3):871–939.

Greenwood, Jeremy, Yueyuan Ma, and Mehmet Yorukoglu. 2024. âĂIJyou will:âĂİ a macroeconomic analysis of digital advertising. *Review of Economic Studies* :rdae067.

Hallak, Juan Carlos and Peter K Schott. 2011. Estimating cross-country differences in product quality. *The Quarterly journal of economics* 126(1):417–474.

Hanson, Gordon H, Nelson Lind, and Marc-Andreas Muendler. 2015. The dynamics of comparative advantage. Technical report, National bureau of economic research.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. What you export matters. *Journal of economic growth* 12:1–25.

Head, Keith and John Ries. 2001. Increasing returns versus national product differentiation as an explanation for the pattern of us–canada trade. *American Economic Review* 91(4):858–876.

Hoberg, Gerard and Gordon Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124(5):1423–1465.

Hsieh, Chang-Tai, Pete Klenow, and Kazuatsu Shimizu. 2021. Romer or ricardo? Working paper, National Bureau of Economic Research.

Hummels, David and Peter J Klenow. 2005. The variety and quality of a nation's exports. *American economic review* 95(3):704–723.

Jaccard, Torsten. 2023. Import demand in differentiated products. Technical report, Tech. rep., University of Toronto.

Janssen, Rebecca, Reinhold Kesler, Michael E Kummer, and Joel Waldfogel. 2022. Gdpr and the lost generation of innovative apps. Technical report, National Bureau of Economic Research.

Josifoski, Martin, Marija Sakota, Maxime Peyrard, and Robert West. 2023. Exploiting asymmetry for synthetic training data generation: Synthie and the case of information extraction. *arXiv preprint arXiv:2303.04132* .

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. Measuring technological innovation over the long run. *American Economic Review: Insights* 3(3):303–20.

Kopytov, Alexandr, Nikolai Roussanov, and Mathieu Taschereau-Dumouchel. 2023. Cheap thrills: the price of leisure and the global decline in work hours. *Journal of Political Economy Macroeconomics* 1(1):80–118.

Krugman, Paul R. 1980. Scale economies, product differentiation, and the pattern of trade. *American Economic Review* 70(5):950–959.

Lancaster, Kelvin J. 1966. A new approach to consumer theory. *Journal of political economy* 74(2):132–157.

Lashkari, Danial and Martí Mestieri. 2019. How substitutable are the products of different countries? working paper.

Lind, Nelson and Natalia Ramondo. 2023. Trade with correlation. *American Economic Review* 113(2):317–353.

Liu, Ruibo, Jerry Wei, Fangyu Liu, Chenglei Si, Yanzhe Zhang, Jinmeng Rao, Steven Zheng, Daiyi Peng, Diyi Yang, Denny Zhou, *et al.* 2024. Best practices and lessons learned on synthetic data for language models. *arXiv preprint arXiv:2404.07503* .

Mardia, Kanti V and Peter E Jupp. 2009. *Directional statistics*. John Wiley & Sons.

Matsuyama, Kiminori and Philip Ushchev. 2017. Beyond ces: three alternative classes of flexible homothetic demand systems. Working Paper 17-109, Global Poverty Research Lab.

Mayer, Thierry, Marc J Melitz, and Gianmarco IP Ottaviano. 2014. Market size, competition, and the product mix of exporters. *American Economic Review* 104(2):495–536.

Melitz, Marc J. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica* 71(6):1695–1725.

Pavcnik, Nina. 2002. Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *The Review of economic studies* 69(1):245–276.

Pellegrino, Bruno. 2023. Product differentiation and oligopoly: a network approach. Working paper, CESifo.

Petrin, Amil. 2002. Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy* 110(4):705–729.

Rachel, Lukasz. 2021. Leisure-enhancing technological change. *Job market paper, London School of Economics* .

Rosen, Sherwin. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy* 82(1):34–55.

Salop, Steven C. 1979. Monopolistic competition with outside goods. *The Bell Journal of Economics* :141–156.

Sra, Suvrit. 2012. A short note on parameter approximation for von mises-fisher distributions: and a fast implementation of $I_s(x)$. *Computational Statistics* 27:177–190.

Sra, Suvrit and Inderjit Dhillon. 2005. Generalized nonnegative matrix approximations with bregman divergences. *Advances in neural information processing systems* 18.

Sun, Ruiqi and Daniel Trefler. 2023. The impact of AI and cross-border data regulation on international trade in digital services: A large language model approach. Working paper, National Bureau of Economic Research.

Sutton, John and Daniel Trefler. 2016. Capabilities, wealth, and trade. *Journal of Political Economy* 124(3):826–878.

Trefler, Daniel. 2004. The long and short of the canada-us free trade agreement. *American Economic Review* 94(4):870–895.

Zhong, Shi and Joydeep Ghosh. 2003. A unified framework for model-based clustering. *The Journal of Machine Learning Research* 4:1001–1037.

Appendix

Appendix A. Derivations and Proofs

Appendix A.1. von Mises-Fisher Distribution

The circular distribution is particularly useful to summarize the high-dimensional unit sphere data by low-dimensional angular statistics and is widely used in machine learning studies. In this model, I use von Mises-Fisher (vMF) distribution to model the arrival of product blueprints. It is normal distribution defined on high dimensional sphere. Starting form a normal distribution with isotropic covariance $\kappa^{-1}\mathbf{I}$ and mean $\boldsymbol{\mu}$ in $K-$dimensional space $\mathbb{R}^K$, we can get the probability density function as follows:

$$f_K(\mathbf{x}, \boldsymbol{\mu}, \kappa) = \left(\sqrt{\frac{\kappa}{2\pi}}\right)^K \exp\left(-\kappa\frac{(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})}{2}\right) \tag{45}$$

vMF distribution is simply obtained by conditioning on $\|\mathbf{x}\| = 1$ and $\|\boldsymbol{\mu}\| = 1$. By expanding

$$(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\mu}^T\boldsymbol{\mu} - 2\mathbf{x}^T\boldsymbol{\mu} = 2(1-\mathbf{x}^T\boldsymbol{\mu}), \tag{46}$$

the pdf of Von Mises-Fisher distribution is recovered by adjust the normalization constant as a integral over unit sphere – formally:

$$f_K(\mathbf{x}, \boldsymbol{\mu}, \kappa) = C_K(\kappa)\exp(\kappa\boldsymbol{\mu}^T\mathbf{x}) \tag{47}$$

Consider a case where $K = 2$, the nodes are distributed on a unit circle (e.g., Salop (1979)). The vMF distribution collapses to von Mises distribution, whose *pdf* is given by

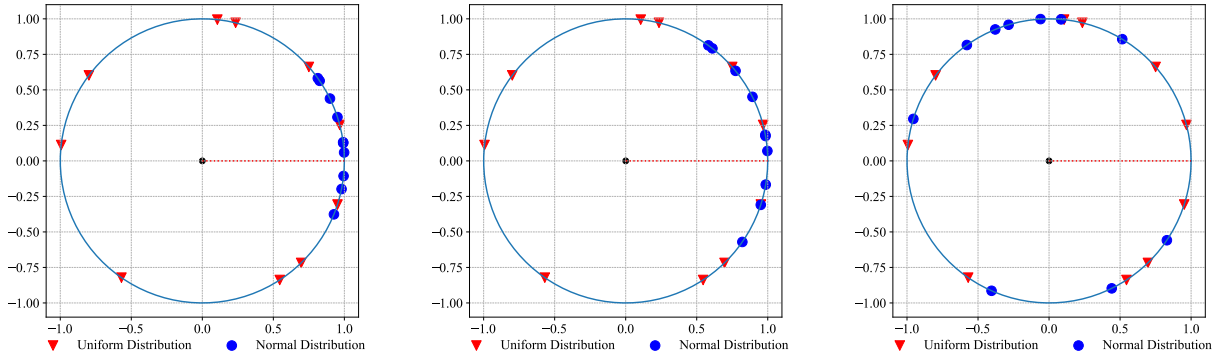
$$f_2(x, \mu, \kappa) = \frac{exp(\kappa\cos(x-\mu))}{2\pi I_0(\kappa)} \tag{48}$$

Figure A1 displays 10 random draws (red triangles) from a uniform distribution and 10 random draws (blue rounds) from von Mises distribution. Both of the distribution are centred at $(1, 0)$. In figure A1c, figure A1b and figure A1a, we set $\kappa = 0.1$, 4, and 10. When κ increases, the spherical variance of the nodes from the von Mises distribution decreases.

The vMF distribution's ability to model directional data makes it highly suitable for various tasks in text mining. Its applications range from topic modeling and clustering to improving information retrieval and NLP tasks, providing a robust framework for handling the unique characteristics of text data represented in high-dimensional spaces (Banerjee, Dhillon, Ghosh, Sra, and Ridgeway, 2005, Sra and Dhillon, 2005, Zhong and Ghosh, 2003).

Figure A1: Illustration of circular normal distribution

(a) $\kappa = 10$ (b) $\kappa = 4$ (c) $\kappa = 0.1$



Notes: These figures plot 10 random nodes drawn from two distributions. The random seed is set to 10. The red triangles are drawn from a uniform distribution. The blue rounds are drawn from a circular normal distribution.

Appendix A.2. Generalized Linear Hedonic demand

Suppose there are M products, indexed by $i \in \{1, 2, \ldots, M\}$, in the economy. These products are distributed on the K dimensional attribute space and are described by the matrix $\mathbf{A}$. The price index for a product attribute s is defined as follows

$$\left(\sum_{i=1}^{M} a_{is}\right) \ln p_s = \sum_{i=1}^{M} a_{is} \ln p_i \tag{49}$$

We can rewrite equation 49 as vectors and matrices. The vector of attribute price, indexed by $\tilde{p}$, is a function of product matrix $\mathbf{A}$ and the vector of product price $\mathbf{p}$:

$$\ln \tilde{\mathbf{p}} = diag(\overline{\mathbf{a}})^{-1} \mathbf{A} \ln \mathbf{p} \tag{50}$$

where $\overline{\mathbf{a}}$ is the resultant vector of matrix $\mathbf{A}$, $\ln \tilde{\mathbf{p}} = [\ln \tilde{p}_1, \ \ln \tilde{p}_2, \ldots, \ln \tilde{p}_K]^T$ and $\ln \mathbf{p} = [\ln p_1, \ \ln p_2, \ldots, \ln p_M]^T$. The diagonal matrix $diag(\overline{\mathbf{a}})$ is a matrix where the s^{th} diagonal entry is $\overline{a}_s$ – formally given by:

$$diag(\overline{\mathbf{a}}) = \begin{pmatrix} \sum_i^M a_{i1} & 0 & \ldots & 0 \\ 0 & \sum_i^M a_{i2} & \ldots & 0 \\ \vdots & \ldots & \ldots & \vdots \\ 0 & 0 & \ldots & \sum_i^M a_{iK} \end{pmatrix}^T \tag{51}$$

The representative consumer face an expenditure cost, which is a quadratic function of attribute prices and product prices. The function is as follows:

$$\ln e \ = \ a_0 + \ln \mathbf{p}^T \boldsymbol{\alpha} + \frac{\theta}{2} \ln \tilde{\mathbf{p}}^T \tilde{\mathbf{D}} \ln \tilde{\mathbf{p}} - \frac{\theta}{2} \ln \mathbf{p}^T \mathbf{D} \ln \mathbf{p} \tag{52}$$

where $\widetilde{\mathbf{D}}$ and $\mathbf{D}$ are weight matrices and defined as

$$\widetilde{\mathbf{D}} = \begin{pmatrix} \frac{\overline{a}_1^2}{\|\overline{\mathbf{a}}\|} & 0 & \dots & 0 \\ 0 & \frac{\overline{a}_2^2}{\|\overline{\mathbf{a}}\|} & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & 0 & \dots & \frac{\overline{a}_K^2}{\|\overline{\mathbf{a}}\|} \end{pmatrix} \text{ and } \mathbf{D} = \begin{pmatrix} \frac{\Omega_1}{\|\overline{\mathbf{a}}\|} & 0 & \dots & 0 \\ 0 & \frac{\Omega_2}{\|\overline{\mathbf{a}}\|} & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & 0 & \dots & \frac{\Omega_M}{\|\overline{\mathbf{a}}\|} \end{pmatrix}$$

where $\|\overline{a}\| = \sum_s (\sum_i a_{is})^2$ is the length of resultant vector. The weights within each matrix are sum to 1 such that $\sum_s^K \widetilde{\mathbf{D}}_{ss} = 1$ and $\sum_i^M \mathbf{D}_{ii} = 1$. The two weights measure the relative importance of each product or attribute. The intuition of the demand system is as following. Under the unbundling assumption, the consumers value product as a bundle of attributes so they fundamentally care about the price of each attributes. However, the bundle packing itself is also valuable. Hence, each product get a discount on price which is negatively correlated with its centrality – a rare bundle packing gets more credit from the consumer. The equation 52 results in the following expenditure cost function:

$$\begin{aligned} \ln(e) &= a_0 + \ln \mathbf{p}^T \boldsymbol{\alpha} + \frac{1}{2} \ln \mathbf{p}^T \Theta \ln \mathbf{p} \\ &= a_0 + \sum_{i=1}^{N} \alpha_i \ln p_i + \frac{1}{2} \sum_{i=1} \sum_{j=1} \theta_{ij} \ln p_i \ln p_j \end{aligned} \tag{53}$$

where θ_{ij} are given by:

$$\theta_{ij} = \begin{cases} \frac{\theta}{\|\overline{\mathbf{a}}\|} \omega_{ij} & j \neq i \\ -\frac{\theta}{\|\overline{\mathbf{a}}\|} (\Omega_i - 1) & j = i \end{cases} \tag{54}$$

This function satisfies the two restrictions on θ_{ij}: (1) *symmetric restrictions* $\theta_{ij} = \theta_{ji}$; (3) *homotheticity restrictions*: $\sum_j \theta_{ij} = 0$ for any i and j. Further Imposing the homotheticity restriction on $\boldsymbol{\alpha}$ such that $\sum_i \alpha_i = 1$ and $\alpha_i \in (0, 1)$, equation 53 is exact translog expenditure cost function proposed by Diewert (1976) and Christensen *et al.* (1975).

Additional Derivations for θ_{ij}:

$$\begin{aligned} \ln \mathbf{p}^T \Theta \ln \mathbf{p} &= \theta \ln \widetilde{\mathbf{p}}^T \widetilde{\mathbf{D}} \ln \widetilde{\mathbf{p}} - \theta \ln \mathbf{p}^T \mathbf{D} \ln \mathbf{p} \\ &= \theta \left(diag(\overline{\mathbf{a}})^{-1} \mathbf{A} \ln \mathbf{p} \right)^T \widetilde{\mathbf{D}} \left(diag(\overline{\mathbf{a}})^{-1} \mathbf{A} \ln \mathbf{p} \right) - \theta \ln \mathbf{p}^T \mathbf{D} \ln \mathbf{p} \\ &= \ln \mathbf{p}^T \left(\frac{\theta}{\|\overline{a}\|} \mathbf{A}^T \mathbf{A} \right) \ln \mathbf{p} - \ln \mathbf{p}^T \left(\frac{\theta}{\|\overline{a}\|} diag(\Omega_i) \right) \ln \mathbf{p} \\ &= \ln \mathbf{p}^T \left[\frac{\theta}{\|\overline{a}\|} \left(\mathbf{A}^T \mathbf{A} - diag(\Omega_i) \right) \right] \ln \mathbf{p} \end{aligned} \tag{55}$$

Appendix A.3. Comparison with other translog functional forms

This section compares the translog functional form defined in equation 53 with other popular functional forms discussed in rencet trade literature – specifically, Feenstra and Weinstein (2017) and Diewert (2023).

First, Feenstra and Weinstein (2017) consider a special case where all products are symmetric and all of the second order parameters are set to be θ/M. These assumptions yield a Symmetric Translog (ST) Function and give neat analytical solutions. The ST function is a special case in our model with zero variance in product network. Therefore, $\omega_{ij} = 1$ for any i and j and $\|\overline{a}\| = M$. We can solve θ_{ij} as

$$\theta_{ij} = \begin{cases} \frac{\theta}{M} & j \neq i \\ \theta(\frac{1-M}{M}) & j = i \end{cases} \tag{56}$$

This is exactly the same with equation (3) in Feenstra and Weinstein (2017).

Second, Diewert (2023) propose a Generalized Symmetric Translog (GST) Function and allow for product specific second order parameters. It corresponds to a special case of my model where consumer only care product differentiation along one dimension. Let consumer gives all weights on attribute 1 so that weight matrix $\widetilde{\mathbf{D}}$ is given by

$$\widetilde{\mathbf{D}} = \begin{pmatrix} \frac{\sum_{ij} a_{i1}a_{j1}}{\sum_{ij} a_{i1}a_{j1}} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \text{ and } \mathbf{D} = \begin{pmatrix} \frac{a_{11}(\sum_i a_{i1})}{\sum_{ij} a_{i1}a_{j1}} & 0 & \dots & 0 \\ 0 & \frac{a_{21}(\sum_i a_{i1})}{\sum_{ij} a_{i1}a_{j1}} & \dots & 0 \\ \vdots & \dots & \dots & \vdots \\ 0 & 0 & \dots & \frac{a_{M1}(\sum_i a_{i1})}{\sum_{ij} a_{i1}a_{j1}} \end{pmatrix}$$

We can solve θ_{ij} as

$$\theta_{ij} = \begin{cases} \frac{\theta}{\sum_{ij} a_{i1}a_{j1}} \, a_{i1}a_{j1} & j \neq i \\ \frac{\theta}{\sum_{ij} a_{i1}a_{j1}} \, \left(a_{i1}^2 - a_{i1}\sum_{j'} a_{j'1}\right) & j = i \end{cases} \tag{57}$$

where $\widetilde{\theta}$. This is exactly the same with equation (11) in Diewert (2023). Notice that both of functional forms satisfy the discussion of separable translog in Example 3a from Matsuyama and Ushchev (2017) and therefore yield single price aggregator.

Appendix A.4. Derivations for market shares

The Langrangian function is

$$\min_{q_i} \left(e - \lambda(\sum_{i=1}^{N} p_i q_i - E) \right) \tag{58}$$

F.O.C. is

$$\begin{aligned} \frac{e}{p_i}\frac{\partial \ln e}{\partial \ln p_i} &= \lambda q_i \\ \frac{e}{p_i}(\alpha_i + \sum_j \theta_{ij} \ln p_j) &= \lambda q_i \end{aligned} \tag{59}$$

Summing equation 59 over i, we can derive:

$$\lambda = \frac{e}{E} \tag{60}$$

Therefore, the real utility $U = 1/\lambda$. Taking equation 60 into 59, we have

$$s_i = \frac{p_i q_i}{E} = \alpha_i + \sum_j \theta_{ij} \ln p_j \tag{61}$$

Appendix A.5. Derivations for the elasticities

$$\eta_{ij} \equiv \left.\frac{\partial \ln q_i}{\partial \ln p_j}\right|_{U\ const.} \tag{62}$$

From equation 59, we can derive q_i is

$$q_i = \frac{e s_i}{p_i} \cdot U \tag{63}$$

We can get

$$\begin{aligned} \eta_{ij} &= \frac{\partial \ln\left(\frac{e s_i}{p_i} \cdot U\right)}{\partial \ln p_j} \\ &= \frac{\partial \ln e}{\partial p_j} + \frac{\partial s_i}{\partial p_j}\frac{1}{s_i} - \mathbb{1}_{i=j} \\ &= s_j + \frac{\theta_{ij}}{s_i} - \mathbb{1}_{i=j} \end{aligned} \tag{64}$$

Also, we can derive the Allen-Uzawa partial elasticities of substitution:

$$\begin{aligned} \sigma_{ij} &= \frac{1}{s_j}\frac{\partial \ln q_i}{\partial \ln p_j} \\ &= 1 + \frac{\theta_{ij}}{s_i s_j} - \mathbb{1}_{i=j}\frac{1}{s_i} \end{aligned} \tag{65}$$

The unconditional price elasticities and elasticities of substitution can be derived in similar way:

$$\eta_{ij} \equiv \frac{\partial \ln q_i}{\partial \ln p_j} = \begin{cases} \frac{\theta_{ij}}{s_i} & j \neq i \\ -1 + \frac{\theta_{ii}}{s_i} & j = i \end{cases} \tag{66}$$

and

$$\begin{aligned}\sigma_{ij} &\equiv \frac{\eta_{ij}}{s_j} \\ &= \begin{cases} \frac{\theta_{ij}}{s_i s_j} & j \neq i \\ -\frac{1}{s_i} + \frac{\theta_{ii}}{s_i^2} & j = i \end{cases}\end{aligned} \tag{67}$$

Appendix A.6. Derivations for the choke prices

Let $s_i = 0$ in equation 11, we can derive

$$\alpha_i + \sum_j^M \theta_{ij} \ln p_j = \alpha_i - \widetilde{\theta}\Omega_i \left[\ln p_i - \sum_j \left(\frac{\omega_{ij}}{\sum_{j'} \omega_{ij'}} \right) \ln p_j \right] = 0 \tag{68}$$

Taking the prices of other products as given, we can solve $s_i(\overline{p}_i) = 0$:

$$\ln \overline{p}_i = \frac{\alpha_i}{\widetilde{\theta}\Omega_i} + \sum_j \left(\frac{\omega_{ij}}{\sum_{j'} \omega_{ij'}} \right) \ln p_j \tag{69}$$

Appendix A.7. Uniqueness and Existence

Let $\mathcal{M}_d$ denote the set of active products in market d. The equilibrium price is given by

$$p_{i,od} = \begin{cases} \left(1 + \frac{s_{i,od}}{\widetilde{\theta}(\Omega_i - 1)}\right) c_{i,od} & i \in \mathcal{M}_d \\ \\ \bar{p}_{i,d} & i \in \mathcal{M} \setminus \mathcal{M}_d \end{cases} \tag{70}$$

Stacking the market shares for all products i together, we have

$$\mathbf{S} = \boldsymbol{\alpha} + \Theta \mathbf{P} \tag{71}$$

where $\mathbf{S} = [s_1, s_2, \ldots s_M]$, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots \alpha_M]$, and $\mathbf{P} = [\ln p_1, \ln p_2, \ldots \ln p_M]$. In the first hypothetical round, there exists a unique solution for $\mathbf{S}$ (Feenstra and Weinstein, 2017), although s_i may exceed 1 or be negative. If $0 < s_i < 1$ hold for all i, then $\mathbf{S}$ represents the equilibrium expenditure shares. However, if $s_i < 0$ for at least one product, the next round sets the minimum market share from the previous solution to zero, assigning the corresponding price to the choke price. The new functional system becomes:

$$\mathbf{S} = \boldsymbol{\alpha} + \Theta \mathbf{P} + \overline{\mathbf{P}} \tag{72}$$

where $\mathbf{S} = [s_1, s_2, \ldots s_{M-1}]$ and $\overline{\mathbf{P}}$ can be treated as a constant. This process is repeated iteratively until $0 < s_i < 1$ hold for all i, ensuring the solution represents valid equilibrium expenditure shares.

Appendix B. Additional Empirical Results

Table A1: Top ten most banned apps worldwide

	App	Category	Users (Brillion)	Banned in
1	Facebook	social media	3.0	CN, RU, IR, KP, MM, TM, UG
2	WhatsApp	messaging	3.0	CN, KP, SY, QA, AE
3	Youtube	video	2.70	CN, KP, TH, MA, DE(partially)
4	Intagram	social media	2.0	CN, VN, KP, TK
5	Google	search engine	1.5	CN
6	WeChat	social media	1.3	IN, UK, US/CA/NZ (partially)
7	TikTok	social media	1.0	IN, US/CA/NZ/TW/ID (partially)
8	X/Twitter	social media	0.4	CN, RU, BR, IR, KP, MM, TM
9	Baidu	search engine	0.7	IN
10	Spotify	music	0.6	CN, IR, ET, LY, IQ, CU, MM

Notes: This table uses data from Linkee.ai and public information.

Table A2: Top 10 public corporation by market capitalization in 2024

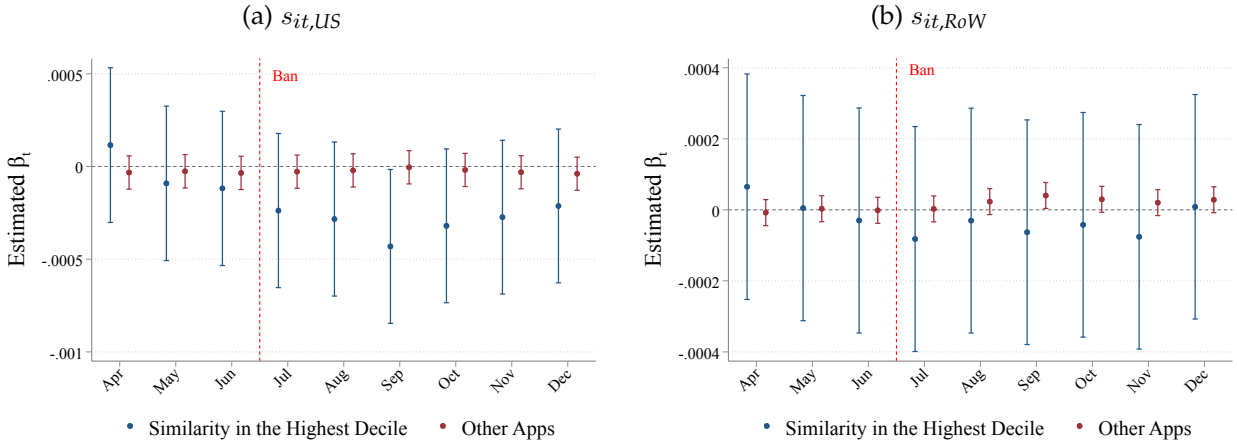
Rank	Company	Market Cap (trillion USD)	Main product
1	Microsoft	3.32	Software
2	Apple	3.23	Electronics, software
3	Nvidia	3.18	Semiconductors
4	Alphabet	2.27	Software
5	Amazon	2.01	Software
6	Meta	1.28	Software
7	TSMC	0.90	Semiconductors
8	Berkshire Hathaway	0.88	Conglomerate
9	Eli Lilly	0.82	Pharmaceutical
10	Broadcom	0.75	Semiconductors, software

Notes: The rank and market cap are up to date as of 30 June 2024. See data resources in https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization.

Table A3: Event Study Coefficients

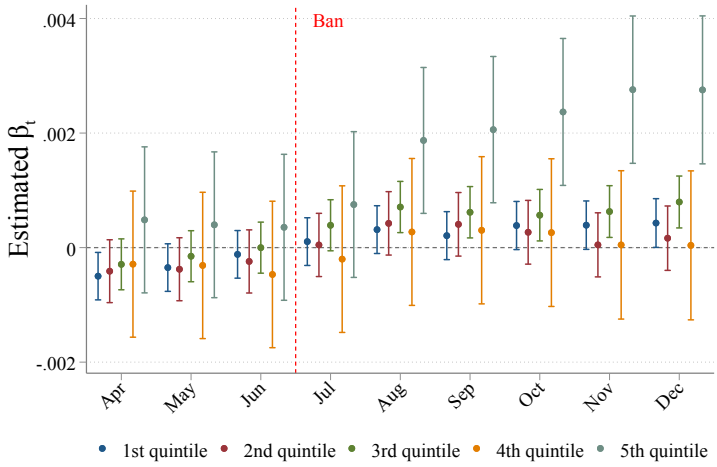
	Banned apps	Competing apps ($\times 10^{-2}$)			
		all	20^{th} percentile	10^{th} percentile	5^{th} percentile
	(1)	(2)	(3)	(4)	(5)
$\tau =$ Dec-19	0.032	-0.001	-0.021	-0.040	-0.058
	(0.19)	(0.02)	(0.08)	(0.12)	(0.23)
$\tau =$ Jan-20	0.078	-0.006	-0.033	-0.064	-0.062
	(0.19)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Feb-20	0.100	-0.008	-0.015	-0.025	0.023
	(0.19)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Mar-20	0.101	-0.008	0.068	0.101	0.243
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Apr-20	0.063	-0.008	0.119	0.148	0.229
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ May-20	-0.093	-0.008	0.112	0.098	0.151
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Jun-20	-0.208	-0.004	0.109	0.108	0.247
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Jul-20	-0.543***	0.011	0.150*	0.203*	0.430*
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Aug-20	-0.855***	0.039**	0.265***	0.402***	0.726***
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Sep-20	-0.853***	0.038*	0.286***	0.399***	0.697***
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Oct-20	-0.848***	0.039**	0.318***	0.446***	0.793***
	(0.18)	(0.02)	(0.08)	(0.12)	(0.22)
$\tau =$ Nov-20	-0.848***	0.043**	0.358***	0.536***	1.014***
	(0.18)	(0.02)	(0.08)	(0.12)	(0.23)
$\tau =$ Dec-20	-0.849***	0.043**	0.359***	0.547***	1.137***
	(0.18)	(0.02)	(0.08)	(0.12)	(0.23)
Observations	406	163,377	30,241	15,901	6,258
R^2	0.64	0.99	1.00	1.00	1.00
Fixed Effects	app	app	app	app	app
Number of apps	29	12,264	2,404	1,270	497

Figure A2: Robustness Check on US and RoW



Notes: This figure corresponds to robustness checks for Figure 2.

Figure A3: Heterogeneity by similarity quintiles



Notes: This figure corresponds to robustness checks for Figure 2.

Appendix C. Estimation Procedure with LLMs

In this paper, I use the large language model for two tasks.

Appendix C.1. Generating synthetic data

The raw data are the app description for consumers in iOS App Store. I use a prompt chat model to generate the synthetic data from raw product descriptions. The pre-trained model is the most efficient GPT chat model up to date. I use one-shot prompt to focus on the functionality and services. Here is the prompt structure:

system: You act as an iOS app analyst, specializing in analyzing market competition among apps. When provided with app descriptions, you identify core features, target audiences, and unique selling points to assess how closely apps compete within the same market segment. You maintain a professional, clear, and concise tone and always give the response in English

user: Here is a description of an iOS mobile app delimited by triple backticks. Please summarize the app in the following format: -A Overall summary of the app,: 'Overall summary: xxxx' -A list of main functions/services in bullet points: '1. xxxx 2.xxx...' Ensure your response is in English, focusing on the most important functions and omitting subscription options, privacy policies, and terms of service. App description: "'{raw_description}"'

model: `gpt-4o-mini`

temperature: 0

seed:123

Following the given prompt, the synthetic description for Netflix is given by

Netflix is a mobile app that provides access to a vast library of award-winning TV shows, movies, documentaries, and stand-up specials, allowing users to stream content anytime and anywhere.

Main functions/services:

1. Browse and stream a wide variety of TV shows and movies.
2. Personalized recommendations based on viewing habits.
3. Create up to five individual profiles for different users.
4. Family-friendly content with a dedicated kids' section.
5. Quick video previews and notifications for new episodes and releases.
6. Download titles for offline viewing.

Appendix C.2. Generating embeddings

The embedding model is `text-embedding-3-large`

Figure A4: Example of App Description for Netflix

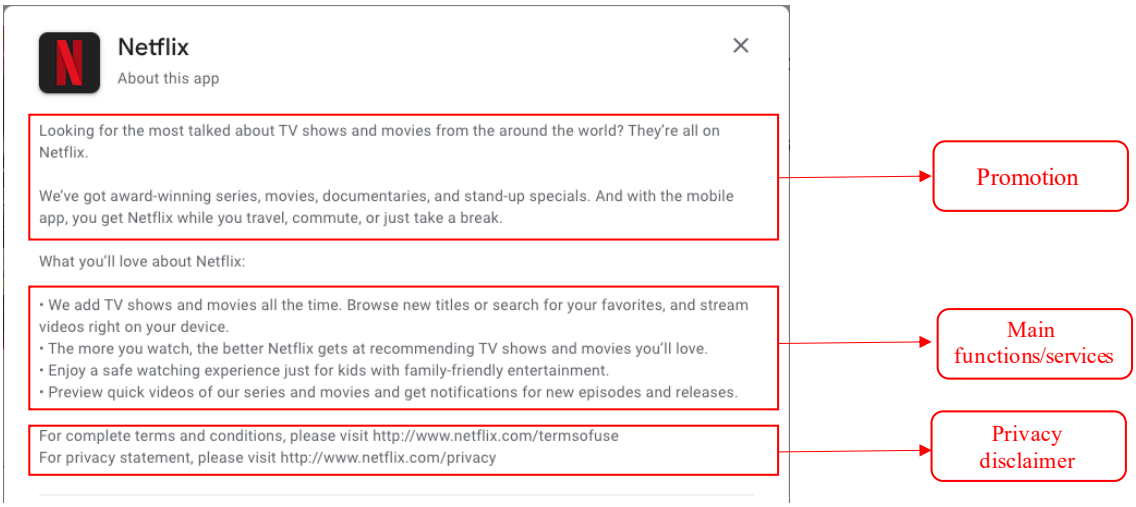
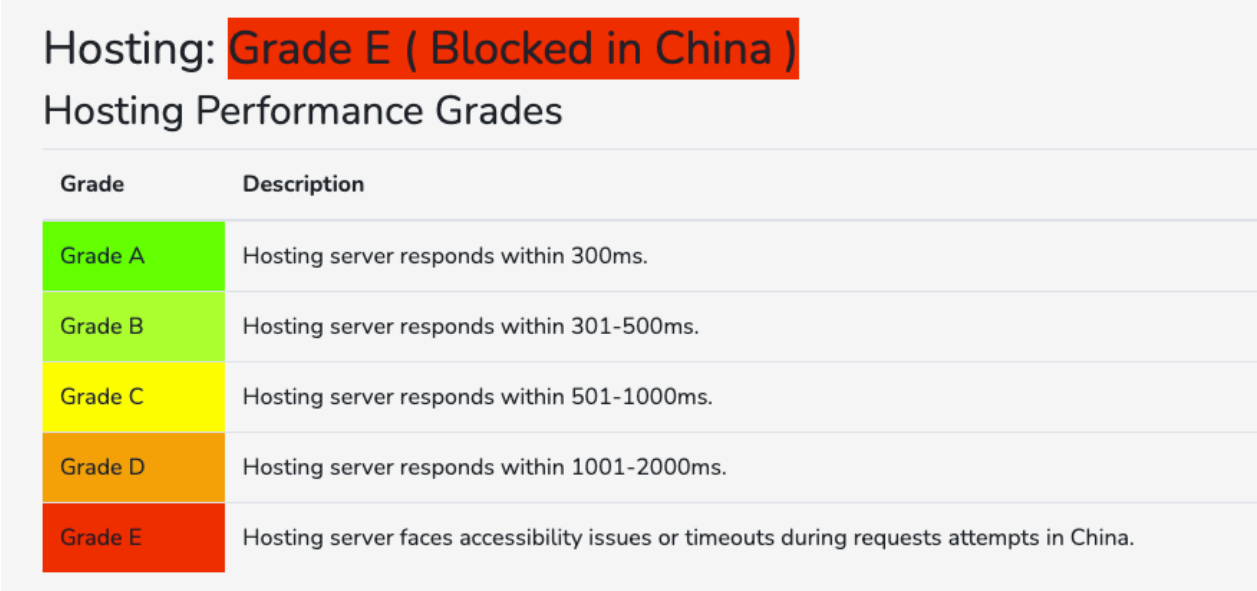


Figure A5: Online test for the GFW block



Notes: This figure is a snapshot of the test response from *China Firewall Test*

Appendix D. Firewall Test on Apps

The test of Great Firewall (GFW) block is conducted using the *China Firewall Test* (https://www.chinafirewalltest.co/), which evaluates the accessibility of any domain within mainland China. Figure A5 provides a snapshot of the test response from the *China Firewall Test*. Accessibility is measured by the response time, which corresponds to the "price" of an app in the model. If the URL associated with an app is graded as C, D, or E, I classify the app as being blocked by the Great Firewall in China.