Blocking the Giants: Theory and Evidence from App Bans

Rachel Ruiqi Sun*†

This Version: 31 May 2026

Abstract

Bans on foreign digital products have become a prominent policy tool, yet their welfare consequences remain poorly understood. This paper develops a framework to evaluate granular trade policies in highly differentiated digital markets. I first show, using an event-study design, that product-pair elasticities of substitution are systematically related to pairwise similarity in product descriptions. Motivated by this evidence, I build a general equilibrium model in which products differ in attribute mix and quality and compete under oligopoly. I apply the framework to mobile app bans, focusing on China's Great Firewall Project. The quantitative results uncover welfare costs missed by standard models. These costs arise from the distributional effects of bans across nearby substitutes and from the "free-good" nature of digital services, under which consumers pay in attention rather than money.

Keywords: granular trade policies; generalized translog demand; digital service trade; LLM embeddings

JEL codes: F12, F13, F14, F52

*I am deeply indebted to my supervisors Daniel Trefler and Xiaodong Zhu, for their continued guidance and support. I also benefit from the comments and discussions with Costas Arkolakis, Heski Bar-Isaac, Andrew Bernard, Emily Blanchard, Eric Bond, Bernardo Blum, Ariel Burstein, Laurent Cavenile, Meredith Crowley, Maggie Chen, Zhang Chen, Jingting Fan, Jason Garred, Sungwan Hong, Ruixue Jia, Gaurav Khanna, Munseob Lee, Kevin Lim, Wenlan Luo, Matthew Mitchell, Mitsuru Igami, Kyle Handley, Keith Head, Federico Huneeus, Wen-Tai Hsu, Peter Morrow, Gianmarco Ottaviano, Scott Orr, Bruno Pellegrino, John Ries, Bob Rijkers, Joel Rodrigue, Regina Seibel, Jeffrey Sun, Kai-Jie Wu and Kai Zhu. I thank Lingpeng Kong for technical discussion on the use of Large Language Models. In memory of Jojo Zhang. All errors are my own.

†Faculty of Business and Economics, the University of Hong Kong. Email: rqsun@hku.hk.

1. Introduction

Trade policies targeting specific products or firms, rather than entire industries or countries, have become increasingly prevalent worldwide. These regulations often arise from geopolitical tensions at the intersection of technology, trade, innovation, national security, and global power dynamics (Bradford, 2023). For example, China's Great Firewall (GFW), which began in the early 2000s, effectively excluded three of the U.S. Big Four – Amazon, Meta, and Google – from the Chinese market. The ban was followed by the rise of China's own technology giants; the major products of these firms closely mirror the banned U.S. products, with Baidu serving as a substitute for Google, Tencent for Meta, and Alibaba for Amazon. Do trade policies targeted at specific digital products causally foster the growth of local substitutes? What are the welfare implications of these policies? Proposed and implemented restrictions on platforms such as TikTok and Twitter (X) make these questions increasingly important.

Answering these questions is not straightforward. Banned digital products typically have large user bases and a diverse mix of hedonic attributes, generating complex substitution patterns with 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 the analysis around a general equilibrium model in which granular, single-product firms with heterogeneous hedonic attributes and productivities compete under Bertrand oligopoly. I estimate millions of elasticity parameters using similarities between product descriptions, where similarity is measured by the cosine similarity of textual embeddings generated with large language models (LLMs). This methodology provides a practical way to capture rich demand-side heterogeneity, through hedonic attributes, alongside supply-side heterogeneity, through productivity, and highlights why this heterogeneity matters for evaluating policy interventions targeted at granular products.

In Section 2, I begin with an event study that documents substitution patterns after an unexpected ban. Following India's 2020 ban on 59 Chinese apps, apps other than the banned products expanded in response to reduced competition. For each app, I construct a measure of its similarity to the banned apps. Apps with the highest similarity to banned apps experienced substantial increases in market share, while those with the lowest similarity exhibited little change. This pattern provides evidence of heterogeneous responses to product-level trade policy interventions, a force absent from workhorse models such as Melitz (2003). The evidence also suggests that similarity measures constructed from textual product descriptions and LLMs are informative about pairwise

elasticities of substitution. I use the event-study estimates to discipline the structural estimation of the quantitative model.

Motivated by the empirical finding, I develop a general equilibrium model in Section 3 that captures rich substitution patterns while maintaining 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 Pellegrino (2025) in industrial organization, I introduce a translog expenditure function, as in Diewert (1976), in which the key parameters emerge endogenously from a microfoundation 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 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 decide whether to enter the market; in the second stage, entrants compete in prices under Bertrand oligopoly.[1]

The large number of demand parameters can be estimated under 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 are rarely available in international trade settings. The challenge is even more pronounced here because nesting would occur at the product level. The demand system in this paper instead 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 these hedonic attributes. This approach transforms the complex estimation of pairwise elasticities into a series of product-specific LLM tasks together with the identification of one aggregate elasticity parameter. More broadly, it provides a structural interpretation for language-model output, a tool increasingly used in empirical economics (Acemoglu and Restrepo, 2022, Kelly, Papanikolaou, Seru, and Taddy, 2021, Argente, Baslandze, Hanley, and Moreira, 2020).

To validate the model and the estimation strategy, I conduct a quantitative analysis using data on mobile apps. This sector and dataset are well suited for four reasons. First, the mobile app sector exhibits rich and intuitive substitution patterns for consumers. For instance, Chrome is likely to be highly substitutable with Baidu because both provide search services, whereas Chrome is only weakly substitutable with Weibo, whose primary function is social media akin to X/Twitter, even though it also includes

[1]While the baseline model applies broadly across sectors, I tailor it into a quantitative model for mobile app services in Section 4.1, where prices correspond to attention costs in this specific application.

limited search functionality. Second, textual descriptions and product-level trade data are available, making the sector ideal for quantitative analysis.[2] Third, natural experiments such as India's app ban and China's GFW policy provide exogenous variation that helps identify key parameters and validate the model's predictions. Finally, the sector is dominated by superstar firms central to current trade disputes. Among the ten largest companies globally, eight specialize in software services or hardware support, and all are involved in foreign bans, underscoring the policy relevance of the setting.

Section 4 applies the framework to mobile app bans. 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 LLMs to generate embeddings that represent standardized app characteristics. These embeddings serve as empirical counterparts to hedonic attributes in the model. The second component comprises trade shares by app and country, 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 blacklisted by China's GFW. These components are linked using unique app identifiers from the iOS App Store. Equipped with these data, I isolate the impact of the GFW from other trade frictions and provide reduced-form evidence consistent with the model's predictions regarding bans.

To quantify the welfare implications of the GFW, I estimate the quantitative model in Section 4.2. The model captures utility from both physical goods consumption and app services. The representative consumer allocates a fraction of time to app usage, and app producers monetize this engagement by selling advertisements while competing as in the theoretical model. I calibrate the quantitative model to a 2021 cross-section. The model parameters fall into two groups. The first group corresponds directly to observable data moments or externally identified parameters. In particular, the key elasticity parameter for the app sector is inferred from the event study in Section 2. The second group contains parameters that cannot be separately identified one by one. I estimate these jointly using simulated method of moments (SMM), which incorporates the general equilibrium effects of the quantitative model. The estimated model reproduces the observed patterns of trade, time allocation, and income, capturing the key empirical patterns in the data.

In Section 5, I use the estimated model to evaluate two policies. First, I assess the welfare implications of the GFW by removing the trade barriers imposed by the GFW in a counterfactual exercise. The analysis reveals two distinct effects on welfare for Chinese consumers. On the one hand, the GFW increased real incomes by shifting profits from

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

banned products to Chinese alternatives, resulting in a 1.05% welfare gain. On the other hand, this benefit was outweighed by a substantial reduction in the utility from app services. While overall time spent on apps decreased only slightly, utility per unit of app time fell by 29%. Consequently, the net welfare loss from the GFW is 7.48%. These findings suggest that policymakers may have incentives to impose such bans to boost real income, but the associated loss in app-service utility is often overlooked in policy discussions.

The model also introduces a novel measure of product diversity, defined as the dispersion of hedonic attributes among apps entering a market. The counterfactual analysis shows 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 may therefore stifle product innovation by reducing competitive pressure. This mechanism aligns with Hsieh, Klenow, and Shimizu (2021), but instead of assuming exogenous innovation efficiency, the results here arise endogenously from the trade policy. The finding provides an example of 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, such as Threads and Instagram, experience revenue increases that are 11 times greater than the average app. This result aligns with current political debates surrounding TikTok's ban, highlighting its potential to disproportionately benefit certain competitors such as Meta. This distributional outcome is absent in standard models that assume uniform substitution patterns across product pairs.

Related Literature This paper connects to several strands of literature. First, it contributes to work on granular firms in international trade. The existing literature highlights the importance of granular firms in explaining trade flows and shaping comparative advantage (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 underexplored. Relatedly, this paper 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 this literature, the paper presents a practical approach to capturing rich demand-side heterogeneity alongside standard supply-side heterogeneity with granular firms and products.

In doing so, this paper adds to the extensive body of work on demand-system estimation beyond CES in international trade. The analysis builds on the translog functional form proposed by Christensen, Jorgenson, and Lau (1975) and Diewert (1976). 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 departs from the symmetry assumption. Relatedly, papers such as Adao, Costinot, and Donaldson (2017), Lashkari and Mestieri (2019), Lind and Ramondo (2023) use cross-nested CES structures to capture flexible substitution patterns. This paper shares a similar flavor by deriving substitution from underlying hedonic attributes, but differs in two key ways: the demand system is designed to exploit LLM representations, and it yields finite choke prices that accommodate entry and exit decisions.

This paper also connects to the literature on trade shocks, competition, and welfare. A large empirical literature documents that trade liberalization reallocates resources toward 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 pro-competitive effects on welfare are quantitatively small. More recently, Baqaee, Farhi, and Sangani (2023) highlight the importance of demand-elasticity heterogeneity for fully capturing the gains from trade. This paper 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, 2025). I build on recent progress in this field, especially Pellegrino (2025), who characterizes a product network and explains the sources of market power for public corporations in the United States. This paper differs in several ways. Theoretically, I adopt a translog functional form that yields closed-form elasticities of substitution and introduce two random processes to characterize endogenous entry. Empirically, while Pellegrino (2025) uses data from Hoberg and Phillips (2016) with similarity measured by keyword frequency, the model in this paper is calibrated at the product-variety level using similarities constructed from prompted LLMs. Substantively, 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. Semantic similarity between two pieces of text can be quantified using the cosine similarity

between their corresponding embeddings, offering 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).[3] 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 the digital sector. The most closely related paper is Bernard, Bøler, Chor, Gurund, and Lu (2025), which examines the impact of the GFW on knowledge spillovers from foreign scholars to Chinese scholars. In contrast, my focus is on product competition rather than technology spillovers. Sun and Trefler (2023) find that the impact of AI deployment on app exports is halved in countries with strong cross-border data-flow restrictions. This paper differs by focusing on substitution patterns in trade under 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 focuses on international trade and competition.

Outline The remainder of the paper is structured as follows. Section 2 presents the event study, which motivates the model structure and estimation strategy. Section 3 establishes the theoretical model and demonstrates the role of pairwise elasticities of substitution. Section 4 applies the model to mobile app bans, Section 4.2 calibrates the quantitative model to the data, and Section 5 quantifies the impact of the GFW policy in China and a TikTok ban in the U.S. Section 6 concludes.

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

2. Motivating Facts: An Event Study

I begin with an event study that provides reduced-form evidence on how demand reallocates after a ban targeted at specific products. The empirical patterns documented in this section motivate the theoretical model in Section 3, where product-pair similarity governs substitution, and the estimates obtained here are later used to discipline key elasticity parameters in the quantitative model in Section 4.

2.1. Institutional background and data

In late June 2020, in response to escalating border tensions between India and China, and particularly after a violent clash in the Galwan Valley, India's Ministry of Electronics and Information Technology suddenly announced a ban on major Chinese apps, including ByteDance's video-sharing app TikTok, Lenovo's file-sharing app SHAREit, and Tencent's messaging app WeChat.[4] According to the order, the ban arose from "increasing concerns on aspects relating to data security and safeguarding the privacy," and the banned apps were described as "prejudicial to sovereignty and integrity of India, defense of India, security of state and public order."

This event provides a useful setting for studying heterogeneous responses to a product-level ban. The policy was largely unexpected, the list of banned products is known without ambiguity, and the banned apps varied substantially in both market size and primary function. These features make it possible to examine whether the apps that gained users after the ban were those that were closer substitutes for the banned products.

To document substitution after the ban, I use app-usage data from Sensor Tower. Sensor Tower is a leading data vendor specializing in mobile apps and tracks app-level usage by destination country for apps available on the iOS App Store.[5] It also provides metadata displayed on the App Store download page, including app descriptions, release dates, publishers, publisher countries, and app categories. For this exercise, I use monthly active users by app, month, and country in 2020 to construct a high-frequency market-share measure, and I use app descriptions to measure similarity across apps.[6]

[4] See the official announcement at https://www.pib.gov.in/PressReleseDetail.aspx?PRID=1635206.

[5] A potential concern is the absence of usage data from Android systems. I focus on iOS because comprehensive Android market information for the Chinese market is limited. However, market shares are highly correlated between iOS and Android.

[6] These data are also used for the quantitative analysis in Section 4. More details are discussed later.

2.2. *Measuring exposure to the ban*

The goal of this exercise is to measure heterogeneity in apps' exposure to the common policy shock. The maintained hypothesis is that an app is more exposed when it is more similar to banned apps that had larger pre-ban market shares.

I manually identify 29 banned apps that are available on the iOS App Store, and 12,264 other apps that are active in India in at least one month before or after the ban.[7] Figure 1a plots the number of users for the matched banned apps. The most popular app on the list is TikTok, which had about 8 million users in India before the ban, while the average banned app had about 0.4 million users. Usage of the banned apps fell sharply and was close to zero by August.[8] Because post-ban usage of banned apps is mechanically bounded near zero, the policy displaced more demand from banned apps with larger pre-ban market shares.

I then use rich textual information from app descriptions to measure app similarity. App descriptions are self-reported by publishers and outline the main functionalities, benefits, and use cases of each product. Publishers also use descriptions to clarify permissions, subscriptions, and privacy disclosures. These descriptions therefore contain valuable, though noisy, information about product attributes.

To standardize these descriptions, I use large language models and a workflow motivated by recent advances in natural language processing (Josifoski, Sakota, Peyrard, and West, 2023).[9] I first use a chat model to transform each raw description into synthetic text. Specifically, I prompt the model to act as an iOS app analyst and identify the app's key features, target audiences, and unique selling points. The output is a concise paragraph summary followed by a bulleted list of core functionalities, with consistent language and tone across apps. Elements such as subscription options, privacy policies, and terms of service are excluded.

Next, I feed the synthetic text into an embedding model and estimate the corresponding vector representations. An embedding model is transformer-based: it extracts features from submitted text and encodes these features as numeric vectors, known as

[7] The remaining banned apps were available only on Android. All major apps that readers would recognize were successfully matched with the app-usage data.

[8] TikTok users declined slightly before the ban amid rising national-security concerns, but the sharp drop began after the policy announcement. Some users could still access banned apps using a Virtual Private Network (VPN), so usage need not be exactly zero; the remaining usage is negligible in the data.

[9] Recent advances in natural language processing increasingly rely on synthetic data, particularly data generated by large language 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 *et al.* (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.

"embeddings." Each embedding is normalized to be a unit vector in a fixed-dimensional space, conceptually analogous to the attribute vectors in the model. Apps with nearby vectors have similar extracted features. Since the vectors have unit length, similarity is measured by cosine similarity, which is the inner product of two vectors. The two-step procedure therefore converts app-description text into unit embedding vectors that can be used to measure product proximity.

I then construct an app-level exposure measure. Let ω_{ij} be the cosine similarity between a competing app i and a banned app j, and let $s_{jJun,IN}$ be the market share of banned app j in India in June 2020. Market share is measured as app j's users divided by users summed across all active apps in the Indian market. Exposure intensity is

$$\lambda_i = \sum_j \omega_{ij} s_{jJun,IN}. \tag{1}$$

Thus, an app receives a higher exposure value when it is similar to banned apps that accounted for a larger share of pre-ban usage.

2.3. *Specification and estimation*

The event study framework takes the following form:

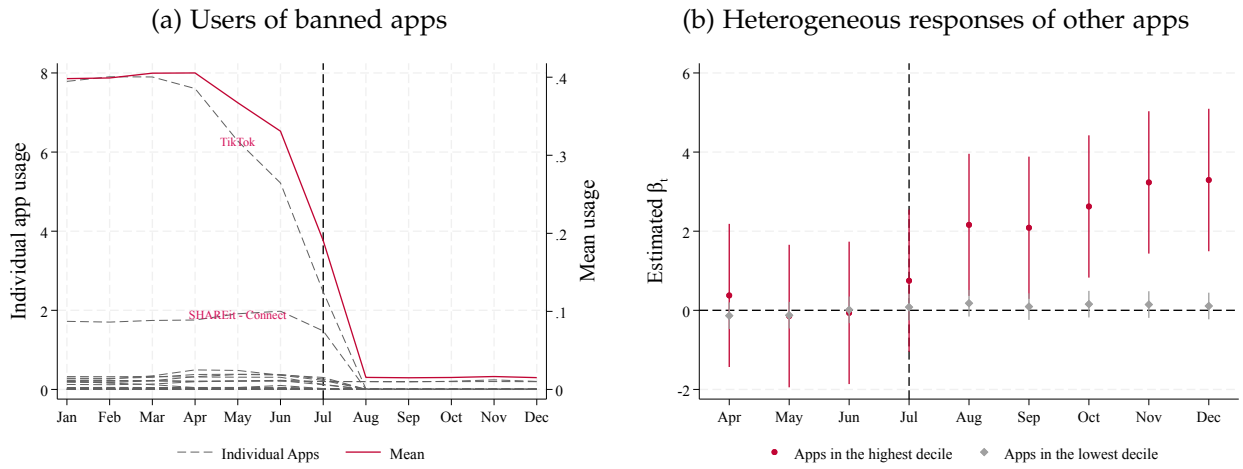
$$\Delta s_{it,IN} = \sum_\tau \beta_\tau \, \mathbb{1}_{t=\tau} \; + \; \varepsilon_{it} \tag{2}$$

where $\Delta s_{it,IN} = s_{it,IN} - s_{iJun,IN}$ is the change in app i's market share in month t relative to June, $\mathbb{1}_{t=\tau}$ are month indicators, and ε_{it} is an error term clustered at the app level. I normalize $\Delta s_{it,IN}$ by one interquartile range of $s_{iJun,IN}$, so the estimated coefficients can be interpreted in units of the cross-app dispersion in market shares at the time of the ban. If the ban primarily reallocates users toward close substitutes, coefficients for highly exposed apps should be near zero before the ban and positive afterward, while coefficients for weakly exposed apps should be much smaller.

Figure 1b presents the estimated coefficients for apps in the highest and lowest deciles of λ_i. The pre-ban coefficients are small, consistent with the absence of differential pre-trends. After the ban, apps in the highest exposure decile gain market share, reaching about 3.5 interquartile ranges by the end of 2020. In contrast, the least exposed apps show changes that are not statistically distinguishable from zero. The pattern suggests that the demand displaced by the ban was absorbed disproportionately by apps that were closer substitutes for the banned products.

One potential identification concern is that the observed patterns might be driven by COVID-related shocks or other unobserved factors affecting app usage globally. To

Figure 1: Effect of India's ban on Chinese apps



Notes: The figure displays the effect of India's ban on Chinese apps implemented at the end of June 2020. The event is indicated by the black dashed line. Figure 1a plots the number of users by banned app and the mean. Figure 1b plots the estimated event-study coefficient β_τ from equation 2 for all other non-Chinese apps. The solid dots are coefficients and the vertical lines show 95% confidence intervals. The plotted coefficients are shown in Table A1 in Appendix Appendix A.

address this concern, I repeat the exercise using market shares in the rest of the world, excluding India and China. The estimated coefficients are not significantly different from zero, suggesting that the response observed in India is not mirrored in other markets. There are also no significant differences across deciles of λ_i outside India. The results are shown in Figure A1 and Appendix Table A1.

I also estimate a more parsimonious specification using a post-ban indicator, $\mathbb{1}_{\tau \geq Jun}$, and report the results in Table 1. Column 1 uses the full sample of competing apps. Columns 2 to 5 restrict the sample to apps in the top 20, 10, 5, and 1 percent of the exposure distribution, respectively. The estimated gains rise monotonically with exposure: market shares increase by 1.6, 2.3, 2.9, and 5.7 interquartile ranges, respectively. These magnitudes indicate that the average response masks substantial heterogeneity across products. Apps that are closer to large banned apps captured much larger gains, while distant apps gained little. In a prominent example, Moj and Josh, Indian short-video apps built around TikTok-like functionality, expanded rapidly from a negligible initial presence after TikTok was removed. The evidence therefore supports the view that text-based similarity contains economically meaningful information about pairwise substitution, without requiring predefined product nests.

A natural alternative interpretation is that the response reflects broad app categories, as in a nested CES structure (Atkeson and Burstein, 2008). Such a structure can generate stronger responses for apps that share a nest with banned products. However, it is less

Table 1: Estimated responses to India's ban

	Dependent Variable: $s_{it,IN}$				
	All	Top 20%	Top 10%	Top 5%	Top 1%
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}_{\tau \geq Jun}$	0.471***	1.566***	2.310***	2.895***	5.720***
	(0.07)	(0.23)	(0.42)	(0.43)	(1.10)
Observations	97,932	19,603	9,749	4,889	961
Within-R^2	0.000	0.003	0.003	0.010	0.031
R^2	1.00	1.00	1.00	1.00	1.00
App FEs	✓	✓	✓	✓	✓

Notes: The dependent variable $s_{it,IN}$ is scaled by one interquartile range of pre-ban market shares. Columns 2–5 restrict the sample to apps in the top part of the exposure distribution λ_i. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$.

well suited to the continuous heterogeneity documented here: exposure varies at the app-pair level, and the largest gains are concentrated among apps that are particularly close to the banned products rather than among all apps in a broad category. Appendix Appendix A derives the nested-CES prediction and clarifies this distinction. The event-study evidence therefore motivates the model in Section 3, in which substitution strength varies continuously with product similarity.

3. Theoretical Model

In this section, I develop a granular trade model with generalized translog demand in which pairwise elasticities of substitution are determined by the geometry of the product network. The model is designed to quantify the distributional consequences of granular product bans. The economy consists of N countries and a single sector. Each origin country o is endowed with a finite set of potential entrants. Each firm produces a single final good and sells to destination countries indexed by d.

I characterize a static equilibrium as the outcome of a two-stage game, building on Eaton *et al.* (2012). In the first stage, firms observe the global product set and choose whether to enter each destination d. In the second stage, entrants compete in prices under Bertrand oligopoly; non-entrants face a reservation price and earn zero revenue in that destination. Prices and quantities clear goods markets in each destination. While the framework is general and can be applied across sectors, Section 4.1 specializes the model to digital services trade as embodied in mobile apps.

3.1. Model Setup

Products Let $\mathcal{M}^o$ denote the finite set of potential entrants from origin o, with cardinality $M^o \equiv \|\mathcal{M}^o\|$. Each potential entrant $i \in \mathcal{M}^o$ is endowed with (i) a product blueprint and (ii) a quality shifter. The blueprint describes the product as a bundle of hedonic attributes, in the spirit of differentiated-products models. The attribute space has K dimensions indexed by $k \in \{1, 2, \ldots, K\}$. The blueprint of product i is represented by a unit vector $\boldsymbol{a}_i \in \mathbb{R}^K$:

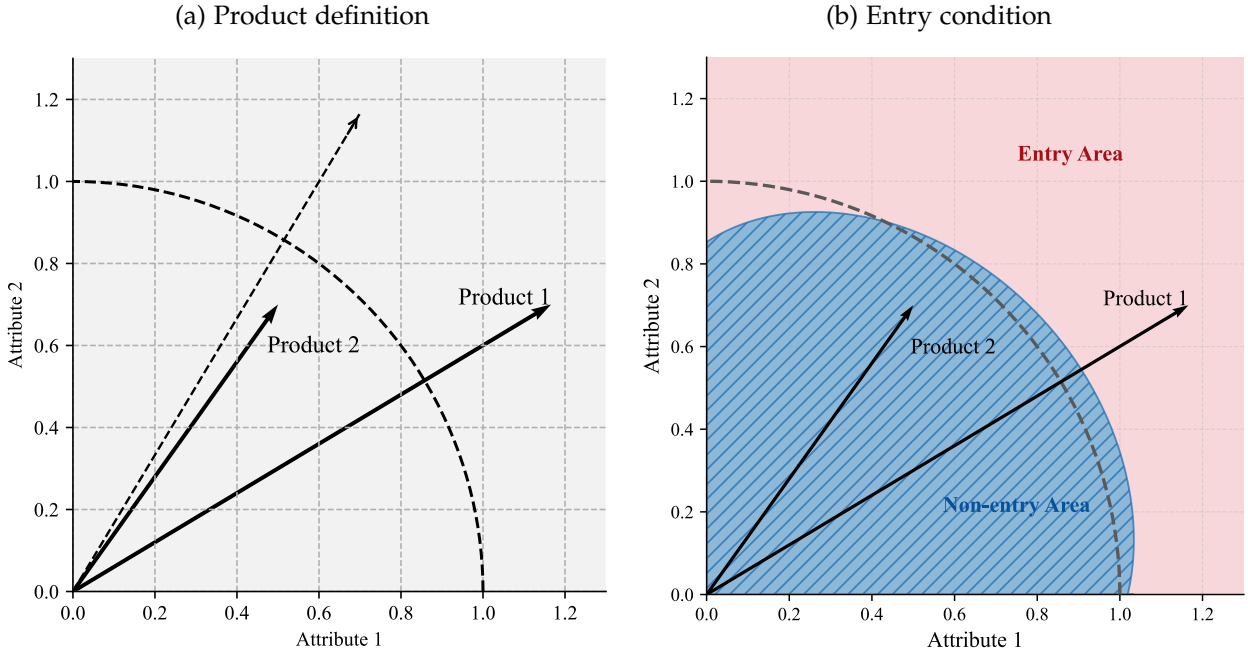
$$\boldsymbol{a}_i = \left[a_{i1}, a_{i2}, \ldots, a_{iK}\right]^\top, \qquad \sum_{k=1}^{K} a_{ik}^2 = 1. \tag{3}$$

The component a_{ik} is the loading of attribute k in product i. Each product is uniquely supplied by its blueprint owner and has quality $b_i \in \mathbb{R}_+$.[10] A product i is therefore characterized by the state $\chi_i \equiv (\boldsymbol{a}_i, b_i) \in \mathbb{R}^K \times \mathbb{R}_+$. Throughout the model, the product set and product states are taken as given.

Figure 2a illustrates the product space in a two-attribute space. Each product i is represented by a vector: the blueprint $\boldsymbol{a}_i$ determines its direction, and quality b_i determines its length. In the figure, product 1 loads relatively more on attribute 1, whereas product 2 loads relatively more on attribute 2. Product 1 has higher quality than product 2, as indicated by its longer vector. In what follows, I formally map this geometry in attribute space into competitive interactions.

[10] I normalize quality shifters within the given product set so that $\sum_{i=1}^{M} b_i = 1$.

Figure 2: Illustrative example of two products in a two-attribute space



Notes: This figure illustrates the case $K = 2$ with two products represented by vectors. For each product, the direction (angle) is determined by the blueprint $\boldsymbol{a}$, and the length (radius) is determined by quality b. The dashed circle denotes the unit circle. In the right panel, the red shaded area and the blue shaded area indicate entry and non-entry regions, respectively, given that products 1 and 2 are already active.

Following Pellegrino (2025), I define the product network matrix

$$\boldsymbol{A} \equiv \begin{pmatrix} \boldsymbol{a}_1^\top \\ \boldsymbol{a}_2^\top \\ \vdots \\ \boldsymbol{a}_M^\top \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{MK} \end{pmatrix} \in \mathbb{R}^{M \times K}, \tag{4}$$

where $\mathcal{M} \equiv \cup_o \mathcal{M}^o$ is the global set of potential products and $M \equiv \|\mathcal{M}\|$. The matrix $\boldsymbol{A}$ records the location of every product in attribute space. I next introduce three sufficient statistics from the directional statistics literature (Mardia and Jupp, 2009). They summarize the geometry of $\boldsymbol{A}$ and enter the demand system directly.

Definition 1 (Resultant vector) *The resultant vector $\overline{\boldsymbol{a}}$ of matrix $\boldsymbol{A}$ is the sum of individual attribute vectors:*

$$\overline{\boldsymbol{a}} \equiv \sum_{i=1}^{M} \boldsymbol{a}_i = \left[\sum_{i=1}^{M} a_{i1}, \sum_{i=1}^{M} a_{i2}, \ldots, \sum_{i=1}^{M} a_{iK} \right]^\top. \tag{5}$$

Its length $\|\overline{\boldsymbol{a}}\|$ measures the degree of concentration of $\{\boldsymbol{a}_i\}$, and its direction is the mean direction of $\{\boldsymbol{a}_i\}$ in attribute space.

The resultant vector provides a compact summary of the configuration of products. Its normalized resultant length

$$\overline{R} \equiv \frac{\|\overline{\boldsymbol{a}}\|}{M} \in [0,1] \tag{6}$$

is an inverse measure of product differentiation: $\overline{R} = 1$ when all products share the same direction, while smaller values indicate greater dispersion and hence stronger product differentiation in attribute space.

Definition 2 (Centrality) *The centrality of product i is its alignment with the mean direction:*

$$\Omega_i \equiv \boldsymbol{a}_i^\top \overline{\boldsymbol{a}}. \tag{7}$$

Ω_i is a network measure that reflects how "typical" product i is within the product set: Ω_i is large when i is similar to many other products and small when i is relatively isolated in attribute space.

Definition 3 (Cosine similarity) *For products i and j, cosine similarity*

$$\omega_{ij} \equiv \boldsymbol{a}_i^\top \boldsymbol{a}_j, \tag{8}$$

measures product similarity in attribute space. A larger ω_{ij} corresponds to greater attribute similarity. Graphically, it corresponds to a smaller angular distance between i and j.[11]

Demand I now introduce a microfoundation for demand in attribute space. In this environment, angular distance governs substitution patterns, while quality serves as a taste primitive under *generalized translog demand*. Consumers choose consumption bundles to maximize utility, where aggregate utility is a Cobb–Douglas aggregator of attribute-level utilities:

$$\ln U = \frac{1}{K} \sum_{s=1}^{K} \ln U_s. \tag{9}$$

To characterize demand for attribute s, I start with a generalized symmetric translog price index (Diewert, 2023) and parameterize it with attribute loadings. The price index for attribute s is

$$\ln P_s = \sum_{i=1}^{M} b_i \ln p_i + \frac{\tilde{\theta}}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \left(a_{is} a_{js} - \mathbb{1}_{\{i=j\}} a_{is} \overline{a}_s \right) \ln p_i \ln p_j, \tag{10}$$

where p_i is the price; a_{is} is the s entry of $\boldsymbol{a}_i$ in equation (3); $\overline{a}_s$ is the s entry of $\overline{\boldsymbol{a}}$ in equation (5); and $\tilde{\theta} > 0$ governs the strength of substitution. Aggregating across attributes

[11] While ω_{ij} can be negative in principle, I focus on environments in which the attribute configuration is sufficiently concentrated that $\omega_{ij} \geq 0$ for most pairs, consistent with interpreting ω_{ij} as a similarity measure.

yields the market price index

$$\ln P = \frac{1}{K}\sum_{s=1}^{K} \ln P_s = \sum_{i=1}^{M} b_i \ln p_i + \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \theta_{ij} \ln p_i \ln p_j \tag{11}$$

where

$$\theta_{ij} = \begin{cases} \theta\omega_{ij} & j \neq i \\ -\theta(\Omega_i - 1) & j = i \end{cases} \qquad \text{with } \theta = \tilde{\theta}/K. \tag{12}$$

Equations (11)–(12) define a generalized translog price index parameterized by cosine similarity ω_{ij}, centrality Ω_i, the quality shifter b_i, and the curvature parameter θ.[12] Under the restrictions in Appendix B.1, the price index corresponds to a homogeneous utility function. The specification nests several demand systems used in the literature, including Cobb–Douglas, symmetric translog (Feenstra and Weinstein, 2017), and generalized symmetric translog demand (Diewert, 2023). Appendix B.2 discusses these special cases.

By Shephard's lemma, the expenditure share of product i is

$$s_i = b_i + \sum_{j=1}^{M} \theta_{ij} \ln p_j = b_i - \Theta_i \left(\ln p_i - \ln \mathcal{P}_i\right) \tag{13}$$

where $\Theta_i = \theta(\Omega_i - 1)$ and

$$\ln \mathcal{P}_i = \sum_{j \neq i} \left(\frac{\omega_{ij}}{\Omega_i - 1}\right) \ln p_j. \tag{14}$$

The formulas below focus on products with $\Omega_i > 1$; an isolated product with no similar competitors is obtained as a limiting case. Equation (13) is central to the analysis. It shows that the competitive environment faced by product i depends on competitors' prices, i's position in the product network, and i's quality.

First, the similarity-weighted price aggregator $\mathcal{P}_i$ breaks the symmetric-substitution structure and distinguishes this demand system from the three classes discussed in Matsuyama and Ushchev (2017). Products that are closer to i in attribute space receive larger weights in $\mathcal{P}_i$ and therefore exert stronger competitive pressure on product i. Moreover, the overall sensitivity of demand to relative prices is governed by $\Theta_i = \theta(\Omega_i - 1)$. As a result, products located in denser regions of the attribute space are more exposed to competitive pressure, whereas niche products are less responsive to competitors' price changes. This feature is absent from models with symmetric demand systems.

The following proposition characterizes the implied price elasticities and Allen–Uzawa elasticities of substitution:

[12] The coefficients on the linear terms capture baseline demand weights before relative-price effects are taken into account, while the coefficients on the quadratic terms govern substitution and thus determine cross-price demand interactions. This interpretation applies both at the attribute level and, after aggregation, at the product level.

Proposition 1 *For any two products i and j, the compensated price elasticity of demand is*

$$\eta_{ij} \equiv \left.\frac{\partial \ln q_i}{\partial \ln p_j}\right|_{U \text{ const.}} = \begin{cases} s_j + \dfrac{\theta_{ij}}{s_i}, & j \neq i, \\ -1 + s_i + \dfrac{\theta_{ii}}{s_i}, & j = i, \end{cases} \tag{15}$$

and the Allen–Uzawa partial elasticity of substitution is

$$\sigma_{ij} \equiv \frac{\eta_{ij}}{s_j} = \begin{cases} 1 + \dfrac{\theta_{ij}}{s_i s_j}, & j \neq i, \\ 1 + \dfrac{\theta_{ii}}{s_i^2} - \dfrac{1}{s_i}, & j = i. \end{cases} \tag{16}$$

Appendix B.3 provides the derivation. Proposition 1 shows that, for $j \neq i$, the pairwise elasticity of substitution σ_{ij} is increasing in product similarity ω_{ij} and decreasing in the market shares of the two products. Products that are closer in attribute space are therefore more substitutable, while larger products are less sensitive to a given cross-price shock. This result is the key mechanism through which product geometry shapes competition in the model.

To illustrate, reconsider Figure 2a and suppose a third product, represented by the dashed vector, becomes available. Product 3 has high quality, as reflected in its longer length, and therefore competes strongly with both products 1 and 2. However, because product 3 is closer to product 2 in attribute space, it exerts greater competitive pressure on product 2 than on product 1. The demand system captures this asymmetry through ω_{ij}: similarity in attribute mix strengthens substitution and amplifies the competition effect.

Second, unlike CES demand, equation (13) expresses expenditure shares in levels rather than logarithms. This distinction implies that a sufficiently high price can drive demand to zero, which makes the entry margin easily defined. Following Feenstra and Weinstein (2017), define the reservation price $\overline{p}(\boldsymbol{a}_i \mid \mathcal{M})$ as the price threshold above which a product with blueprint $\boldsymbol{a}_i$ earns zero sales, holding fixed the global product set $\mathcal{M}$. Because substitutability varies across the attribute space, the reservation price also varies across products. The next proposition gives the implied expression.

Proposition 2 *Given a product set $\mathcal{M}$, the reservation price at $\boldsymbol{a}_i$ is the solution to $s_i = 0$ in equation (13), and is given by*

$$\ln \overline{p}_i(\boldsymbol{a}_i \mid \mathcal{M}) = \frac{b_i}{\theta(\Omega_i - 1)} + \sum_{j \neq i} \left(\frac{\omega_{ij}}{\Omega_i - 1} \right) \ln p_j. \tag{17}$$

The entry condition is harder to satisfy when $\overline{p}_i$ is lower. This occurs when: (i) $\boldsymbol{a}_i$ is closer to low-price products, implying stronger substitution toward more attractive competitors; or (ii) $\boldsymbol{a}_i$

is more central in the product space, implying stronger competitive pressure from overlapping attributes.

Figure 2b illustrates Proposition 2. Suppose there are initially two products in the market, products 1 and 2, and consider an additional potential entrant. The boundary between the hatched blue and red regions indicates the minimum quality required for the entrant to charge a price below its reservation price. Only products located in the red region can profitably enter and survive in the market.

Firms In the absence of overhead entry costs, firm decisions are separable across markets. I solve the model by backward induction and begin with the second-stage pricing problem. Firm i produces with marginal cost ϕ_i. Trade costs take the standard iceberg form: delivering one unit of product i from origin o to destination d requires shipping $\tau_{i,od} \geq 1$ units, with $\tau_{i,oo} = 1$ for all i. The effective marginal cost of serving market d is therefore $\phi_{i,od} = \tau_{i,od}\phi_i$. Conditional on entering destination d in the first stage, firm i chooses its price $p_{i,od}$ to maximize operating profits in that market, taking destination expenditure and competitors' prices as given:

$$\Pi_{i,od} = s_{i,od}E_d - \phi_{i,od}q_{i,od}, \tag{18}$$

where E_d denotes total expenditure in destination d, and $s_{i,od}$ is given by equation (13). For a fixed active set, the Bertrand first-order condition implies the markup rule

$$p_{i,od} = \left(1 + \frac{s_{i,od}}{\theta(\Omega_i - 1)}\right)\phi_{i,od}. \tag{19}$$

The markup is increasing in the firm's expenditure share and decreasing in local competitive density $\theta(\Omega_i - 1)$. Operating profits are a fraction $\frac{s_{i,od}}{s_{i,od}+\theta(\Omega_i-1)}$ of revenue. Hence, larger firms charge higher markups and earn higher profits in equilibrium, consistent with Marshall's second law of demand. In addition, firms with more differentiated products, as captured by a lower centrality measure Ω_i, face weaker competitive pressure, charge higher markups, and earn higher profits.[13]

In each market d, firms make entry decisions following a given sequence.[14] Given the active set induced by that sequence, firms play the second-stage Bertrand game,

[13]This pricing rule is nest-like, but not exactly a nested demand system. When products cluster into a small number of regions in attribute space, the similarity matrix becomes approximately block-diagonal, and the model delivers substitution patterns resembling nested demand systems. The key difference is that the present model allows substitution strength to vary continuously across product pairs. I discuss a special case that is closely related to nested logit demand in Appendix B.4.

[14]Because the set of potential entrants is discrete, the entry game may admit multiple equilibria, as in Atkeson and Burstein (2008) and Gaubert and Itskhoki (2021). I therefore characterize the equilibrium selected by any given entry sequence.

which yields prices, market shares, and operating profits. In equilibrium, all active firms earn non-negative operating profits, while any additional entrant would earn negative operating profits after the Bertrand game is resolved.

In equilibrium, active firms choose prices according to the Bertrand markup rule, and inactive firms charge their reservation prices:

$$p_{i,od} = \begin{cases} \left(1 + \frac{s_{i,od}}{\theta(\Omega_i - 1)}\right) \phi_{i,od} & i \in \mathcal{M}_d, \\ \overline{p}_{i,od} & i \in \mathcal{M} \setminus \mathcal{M}_d, \end{cases} \tag{20}$$

where $\mathcal{M}_d$ is the active set in destination d, and $\overline{p}_{i,od}$ is the reservation price from equation (17).

Equilibrium I formalize the equilibrium in the following proposition:

Proposition 3 *For a given set of preference parameter θ, expenditures $\{E_d\}_{d=1}^N$, firm's production technology $\{\phi_i\}_{i\in\mathcal{M}}$, product states $\{\chi_i = (\boldsymbol{a}_i, b_i)\}_{i\in\mathcal{M}}$, trade costs $\{\tau_{i,od}\}_{i,o,d}$, and an entry sequence in each destination d, an equilibrium is a collection*

$$\{\mathcal{M}_d, p_{i,od}, s_{i,od}, q_{i,od}\}_{i\in\mathcal{M},\, d=1,\ldots,N} \tag{21}$$

such that, in each destination d, (i) active firms earn non-negative operating profits, $\Pi_{i,od} \geq 0$ for all $i \in \mathcal{M}_d$; (ii) any inactive firm would earn negative operating profits if it were added to $\mathcal{M}_d$ and the Bertrand game were re-solved; (iii) prices satisfy equation (20); (iv) market shares satisfy equation (13); and (v) quantities satisfy $q_{i,od} = s_{i,od}E_d/p_{i,od}$ for active firms and $q_{i,od} = 0$ for inactive firms. Following any given entry sequence, the equilibrium is unique.

Appendix C proves uniqueness.

3.2. Implications of Granular Product Bans

The model delivers novel insights into the incidence of granular product bans. In particular, it characterizes not only how displaced expenditure is reallocated across incumbent products, but also how the entry margin shifts across the attribute space. To make these mechanisms transparent analytically, I first consider a partial-equilibrium benchmark that isolates the direct substitution and entry channels, holding fixed the prices of all non-banned products. The full general-equilibrium effects are then quantified in the structurally estimated model in Section 4.

Consider a ban imposed by country d on product i. In the short-run benchmark, the prices of all products other than i are held fixed at their pre-ban levels. The ban is

implemented by increasing the trade costs faced by product i in country d so that its expenditure share falls to zero. To simplify notation, I suppress the destination subscript d. To implement the ban, raise p_i until product i reaches its reservation price and captures zero sales. Since

$$\frac{\partial s_i}{\partial \ln p_i} = \theta_{ii} = -\theta(\Omega_i - 1), \tag{22}$$

the change in the log price of product i required to drive its expenditure share from s_i to zero is

$$\Delta \ln p_i = -\frac{s_i}{\theta_{ii}} = \frac{s_i}{\theta(\Omega_i - 1)}. \tag{23}$$

Holding fixed all prices except p_i, the effect of a change in the price of product i on the expenditure share of product $j \neq i$ is

$$\frac{\partial s_j}{\partial \ln p_i} = \theta_{ji} = \theta\omega_{ji}. \tag{24}$$

Substituting equation (23) into equation (24) yields the short-run change in the expenditure share of any surviving product $j \neq i$:

$$\Delta s_j = \theta\omega_{ji}\Delta \ln p_i = \frac{\omega_{ji}}{\Omega_i - 1} s_i. \tag{25}$$

Equation (25) provides a sharp characterization of the immediate incidence of the ban. The banned product loses share $\Delta s_i = -s_i$, and its entire expenditure share is reallocated across the remaining products. Importantly, the reallocation is local in attribute space: *products with higher similarity to the banned product gain more, while distant products gain less.*[15]

The same logic also holds for the entry margin. Consider the reservation price at location $\boldsymbol{a}_j$. Holding fixed all prices except p_i, the effect of a change in the banned product's price on the reservation price at $\boldsymbol{a}_j$ is

$$\frac{\partial \ln \overline{p}(\boldsymbol{a}_j \mid \mathcal{M})}{\partial \ln p_i} = \frac{\omega_{ji}}{\Omega_j - 1}. \tag{26}$$

Substituting equation (23) into equation (26) yields

$$\Delta \ln \overline{p}(\boldsymbol{a}_j \mid \mathcal{M}) = \frac{\omega_{ji}}{\Omega_j - 1} \cdot \frac{s_i}{\theta(\Omega_i - 1)}. \tag{27}$$

Equation (27) shows that a ban raises reservation prices most strongly in locations close to the banned product. Intuitively, removing product i relaxes competition in its neighborhood of the attribute space, making entry easier for nearby products. Thus, the

[15] The formula satisfies adding up across surviving products: $\sum_{j\neq i} \Delta s_j = \sum_{j\neq i} \frac{\omega_{ji}}{\Omega_i - 1} s_i = s_i$ because $\sum_{j\neq i} \omega_{ji} = \Omega_i - 1$. Hence $\Delta s_i + \sum_{j\neq i} \Delta s_j = 0$.

model delivers a sharp prediction for the entry margin: *a granular ban disproportionately improves entry prospects for products that are close substitutes to the banned good*.

Taken together, equations (25) and (27) highlight the central mechanism of the model. A granular ban does not induce a proportional reallocation of expenditure across surviving products, nor does it shift entry incentives uniformly. Instead, both incumbent gains and entry responses are governed by the geometry of the product space. Products located near the banned variety absorb a disproportionate share of displaced demand and experience the largest increases in reservation prices, while products in distant regions of the product network are affected much less. These mechanisms correspond to the empirical patterns documented in Section 2.

4. An Application to App Bans

I apply the framework to study app bans. This setting is especially well suited to the question for four reasons. First, the mobile app sector exhibits rich and intuitive substitution patterns. For example, Chrome is likely to be a close substitute for Baidu because both provide search services, but a much weaker substitute for Weibo, whose primary function is social media rather than search. Second, the sector offers an unusual combination of detailed textual descriptions and product-level cross-country usage data, making it particularly suitable for quantitative analysis.[16] Third, policy episodes provide plausibly exogenous shocks that are useful both for identification and for assessing the model's empirical implications. Fourth, the sector is dominated by superstar technology firms that sit at the center of current trade disputes. App bans have become an increasingly common policy instrument in digital markets, especially in software and AI-related services, yet their economic effects remain poorly understood. These features make the app sector both empirically tractable and highly policy relevant.

This section proceeds as follows. I first present a quantitative model of the app sector built on the general framework. I then describe the data and estimation strategy. Finally, I use the estimated model to conduct counterfactual policy exercises.

[16] A related dataset for manufacturing goods is NielsenIQ. However, NielsenIQ provides limited product-level textual information. Argente *et al.* (2023), for example, address this limitation by clustering barcode-level products into broader groups and supplementing product descriptions with external sources such as Wikipedia. By contrast, my analysis focuses on substitution within a narrowly defined sector, making the app setting especially attractive.

4.1. Quantitative Model

I now develop a quantitative version of the model to study the effects of app bans. The key extension is to embed the app sector in a broader economy in which apps are monetized through advertising. Following the literature on leisure and advertising (Greenwood, Ma, and Yorukoglu, 2024, Kopytov, Roussanov, and Taschereau-Dumouchel, 2023, Rachel, 2021), I assume that app producers provide services to users for free and monetize attention by selling advertising slots. Advertisers are producers of physical goods that invest in advertising to build brand equity. Both physical goods and app services are tradable across countries, whereas advertising is local.

Households A representative consumer in country d derives utility from physical consumption and app-based leisure:

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

where Q_d denotes consumption of physical goods and l_d denotes leisure-time consumption of app services.[17] Each consumer is endowed with one unit of time, which is allocated between work and app usage. Let h_d denote time spent using apps. The budget constraint is

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

where P_d^p is the price index for physical goods in country d (the p superscript stands for "physical" hereafter), w_d is the wage, and π_d denotes profit income rebated to households.

App Services Apps compete as in Section 3. Each country o and app category c is endowed with a given set of potential apps. The full set of apps from country o is

$$\mathcal{M}^o = \mathcal{M}^{o1} \cup \mathcal{M}^{o2} \cup \cdots \cup \mathcal{M}^{oC}. \tag{30}$$

Let P_d^a denote the app-service price index in country d, measured in units of time per unit of utility. App-based leisure is then

$$l_d = \frac{h_d}{P_d^a}, \tag{31}$$

where h_d is total time spent using apps in country d. The price index is induced by the generalized translog demand system derived above:

$$\ln P_d^a = \sum_{i=1}^{M} b_i \ln p_{i,od} + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta_{ij} \ln p_{i,od} \ln p_{j,od}. \tag{32}$$

[17] It is straightforward to add non-app leisure as an additional Cobb–Douglas component.

Under this time-based interpretation, $p_{i,od}$ is the time required for a user in destination d to obtain one unit of utility from app i, and $c_{i,od}$ is the underlying time cost of delivering that unit of utility. The markup $\mu_{i,od}$ therefore captures the wedge between the underlying time requirement of providing the service and the time ultimately borne by users. In the app context, this wedge can be interpreted as reflecting design choices that increase user attention and thereby expand the supply of advertising slots.[18]

I further decompose marginal cost as

$$\ln \phi_{i,od} = \ln \phi_i + \tau_{od} + \tau_{i,d}, \tag{33}$$

where τ_{od} is a dyadic barrier that shifts the time cost of supplying apps from country o to country d, and $\tau_{i,d}$ captures destination-specific policies targeted at product i, including bans. App producers provide services to users at a zero monetary price and sell user attention in market d to the local ad exchange at price P_d^e. Since $s_{i,od}$ is the share of total attention time allocated to app i, total app revenue in destination d is[19]

$$s_{i,od} P_d^e h_d. \tag{34}$$

Physical Goods I model the physical goods sector parsimoniously using a Krugman-style model. There is a continuum of physical varieties, and consumers in country d aggregate them according to

$$Q_d = \left(\sum_{o=1}^{N} \int b_{i,od}^{\psi} \left(q_{i,od}^P \right)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \tag{35}$$

where $q_{i,od}^P$ is the quantity of physical variety i from country o consumed in country d, σ is the elasticity of substitution, $\psi \in (0,1)$ governs the effectiveness of advertising, and $b_{i,od}$ denotes brand equity created through advertising expenditures.

Physical goods producers choose prices and advertising expenditures to maximize profits in each destination market. Under CES demand, optimal pricing implies the standard constant markup $\sigma/(\sigma-1)$. Optimal advertising expenditure is proportional to revenue:

$$b_{od} = \frac{\psi}{\sigma - 1} x_{od}^P, \tag{36}$$

[18]The interpretation of prices, marginal costs, and markups in the app sector is from the user's perspective. On the producer side, I abstract from the detailed engineering margin and assume that supplying app services entails zero marginal production cost, consistent with the view that the marginal cost of digital services is close to zero. Appendix D discusses an extension in which app producers hire software engineers and incur maintenance and development fixed costs.

[19]Accordingly, the expenditure share $s_{i,od}$ can be interpreted as app i's share of total app-usage time in destination d.

where x_{od}^P denotes exports from country o to country d in the physical goods sector. Producers use labor to produce with a linear production technology. Combining optimal pricing and demand yields

$$x_{od}^P = \frac{b_{od}^{\psi} \left(w_o \tau_{od}^P / T_o^P\right)^{1-\sigma}}{\sum_{o'} b_{o'd}^{\psi} \left(w_{o'} \tau_{o'd}^P / T_{o'}^P\right)^{1-\sigma}} E_d, \tag{37}$$

where w_o is the wage rate of labor, T_o^P is productivity in physical goods and τ_{od}^P is the iceberg trade cost.

Ad Exchange In each destination market d, a passive ad exchange sets a uniform local price P_d^e for attention. Advertising is non-tradable: app producers sell local user attention to advertisers serving the same destination market. I assume that the ad exchange operates competitively and earns zero profits. Market clearing in the advertising sector therefore requires total ad expenditure to equal total payments to the app sector:

$$P_d^e h_d = \sum_{o=1}^{N} b_{od} = \frac{\psi}{\sigma - 1} E_d. \tag{38}$$

Thus, ad revenue is proportional to destination expenditure and is distributed across apps according to their attention shares.

General Equilibrium A general equilibrium is a collection of wages, prices, time allocation, and incomes such that: (i) households choose consumption and app-usage time optimally; (ii) app producers choose entry and prices optimally; (iii) physical-goods producers choose prices and advertising optimally; (iv) the advertising market clears in every destination; and (v) labor markets clear in every country. Labor is used in the physical goods sector, while households receive labor income together with profit rebates from both the physical goods and app sectors.

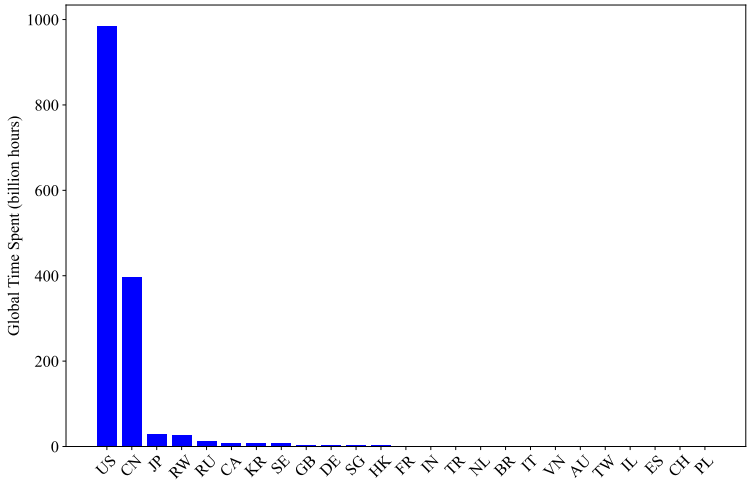
4.2. Model Estimation

4.2.1. Additional assumptions and data

Table 2 summarizes the parameters in the quantitative model by sector. I use the vMF and Pareto distributions to characterize 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 idea that consumers value product uniqueness and assign greater utility weight to less central products. When $\eta = 0$, so that $\alpha_i = 1/M$, differences in market shares arise only from

the quadratic terms, which capture competition effects. These assumptions allow me to translate high-dimensional estimation at the product level into estimation on lower-dimensional statistics.

Figure 3: Global time spent on apps by exporting country



Notes: Global time spent is the sum of time spent across apps produced by exporting countries and across all markets in 2021.

I estimate the parameters using cross-sectional data from 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 3 illustrates global time spent on apps by exporting country, revealing a substantial 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), where RoW 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. 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, 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-sector employment share is drawn from the OECD Going Digital Toolkit, supplemented by national statistical agencies for economies not covered by the OECD.

Table 2: Summary of Parameters

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

4.2.2. Estimation strategy

I estimate the model parameters in two steps. First, I calibrate parameters that can be directly inferred from observed data moments or identified externally. Second, I use simulated method of moments (SMM) to estimate the remaining five parameters. Because the model is granular, the simulation relies on realizations of the random processes. I fix the random seeds throughout the estimation and conduct robustness checks using alternative random seeds. I begin by introducing the separately identified parameters and then discuss the SMM procedure.

Estimating $\widetilde{\theta}$ The parameter $\widetilde{\theta}$ is a key elasticity parameter in the pairwise elasticities of substitution. I infer $\widetilde{\theta}$ from the event study in Section 2. The empirical specification corresponds to changes in market shares, which gives a structural interpretation to the estimated coefficient. In particular, the model-predicted change in market share is

$$s_{i\text{-}post,IN} - s_{i\text{-}pre,IN} = \widetilde{\theta} \lambda_i \ln \overline{\tau}, \tag{39}$$

where λ_i is the exposure measure defined in Section 2.2. The corresponding empirical specification is

$$s_{it,IN} = \beta \times \lambda_i \times \mathbb{1}_{t \geq Jun\text{-}20} + \delta_i + \varepsilon_{it}. \tag{40}$$

The estimated coefficient $\widehat{\beta}$ therefore reflects the value of $\widetilde{\theta} \times \ln \overline{\tau}$ in the model. Table 3 presents the estimation results. Column 1 includes 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 statistically insignificant. In addition, the explanatory power of the empirical model, measured by within R^2, increases when I restrict the sample to larger apps.

The nature of digital bans provides a useful interpretation of $\ln \tau$. After a ban, local consumers may still access banned services through a VPN. However, VPN usage can slow internet speeds because of encryption, server distance, and network congestion. Average speed reductions range from roughly 30% for paid VPNs to 50% for free VPNs,

Table 3: Dependent Variable: $s_{it,IN}(\times 10^{-4})$

	(1)	(2)	(3)	(4)
$Post_t \times \lambda_i$		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	$s \geq 10^{-4}$

although the exact effect varies by service and server location. I set the shock magnitude to $\ln \tau = 0.69$, corresponding to a 50% speed reduction for a user relying on a free VPN to access the banned apps. Consequently, the estimated value is $\widetilde{\theta} = \hat{\beta} / \ln \tau = 1.7 \times 10^{-5} / 0.69 = 2.46 \times 10^{-5}$. Figure 4 plots the kernel density of the pairwise elasticities of substitution in India before the ban, with a market-share-weighted average elasticity of substitution of 6.05.

Estimating vMF mean I assume that apps from 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

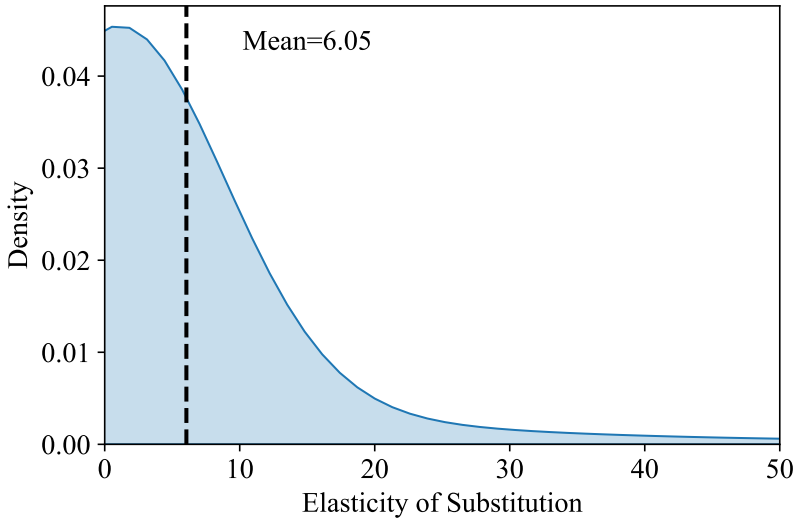
$$E\left[\frac{\overline{\mathbf{a}}^{o,c}}{\|\overline{\mathbf{a}}^{o,c}\|}\right] = \boldsymbol{\mu}^{o,c} \tag{41}$$

with a sufficiently large number of realizations. I directly calibrate $\boldsymbol{\mu}_o^c$ by calculating the resultant vectors in the data. The dimension of the embedding vectors is fixed by the embedding model, which yields 3,027 dimensions in this exercise. For computational efficiency, I reduce dimensionality using the Truncated SVD Algorithm while retaining most of the information on pairwise cosine similarity.[20] Figure 5 shows the mean squared 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 substantially 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, the 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

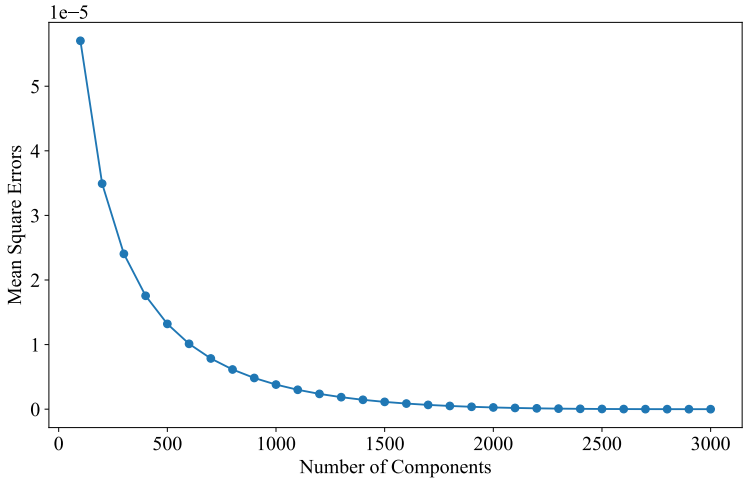
[20] A more familiar approach is Principal Component Analysis (PCA), but the Truncated SVD Algorithm performs better in predicting cosine similarities.

Figure 4: The density of pairwise elasticity of substitution



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

Figure 5: Mean squared error with lower-dimensional embeddings



distribution can be rationalized by different combinations 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 competitive effects of inactive firms. I set all $\overline{M}^{o,c}$ to the mean of $\overline{M}^{o,c}$ in the data, focusing sectoral differences on productivity without introducing an excess number of shadow products into the economy.

As in many trade models, this model faces an identification challenge between productivity and unilateral trade costs. Because the counterfactual analysis focuses on the additional costs from direct bans on U.S. apps in the Chinese market, the conclusions do not hinge on separating baseline productivity from unilateral trade costs. 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 one. Table 4 presents the estimation results. The additional trade cost due to the GFW is estimated as $\overline{\tau} = 10$, indicating that banned U.S. apps face trade costs about 10 times higher than otherwise comparable apps.

Table 4: Trade costs in the app sector

	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 I identify σ and ψ by matching two data moments. The gravity equation derived from 42 is

$$\frac{\pi^P_{od}}{\pi^P_{dd}} = \left(\tau^P_{od} \frac{w^P_o / T^P_o}{w^P_d / T^P_d} \right)^{\frac{1-\sigma}{1-\psi}}. \tag{42}$$

This equation takes the same form as the standard gravity equation, with the trade elasticity for physical goods given by $(1-\sigma)/(1-\psi)$. I use an externally identified value from the literature, such as Caliendo and Parro (2015), and set the physical-goods trade elasticity to 4. Second, the expenditure share on advertising is given by $\psi/(\sigma-1)$ from equation 38. I choose this ratio to match the average expenditure share in the data and set $\psi/(\sigma-1)$ to 0.05. These two moments pin down σ and ψ.

Estimating T^P_o and τ^P_{od} I estimate T^P_o and τ^P_{od} from bilateral trade flows in the physical-goods sector. I follow Head and Ries (2001) to estimate bilateral trade costs. Assuming $\tau^P_{od} = \tau^P_{do}$ and taking the trade elasticity as given, I recover symmetric trade barriers using the Head-Ries index:

$$\tau^P_{od} = \left(\frac{\pi^P_{od}}{\pi^P_{dd}} \frac{\pi^P_{do}}{\pi^P_{oo}} \right)^{\frac{1-\psi}{2(1-\sigma)}}. \tag{43}$$

Taking the auxiliary value of the relative wage as given, I estimate T^P_o as the residual in equation 42, with the normalization $T^P_{CN} = 1$. The relative productivity between the two normalizations, $\varphi^{CN,1}/T^P_{CN}$, is calibrated in the SMM estimation.

Table 5 summarizes the first group of parameters. These parameters are separately identified through direct matches to observable data moments or external estimates.

SMM Estimation I use simulated method of moments (SMM) to estimate the remaining five parameters of the model, $\Theta = \{\kappa,\ \nu,\ \eta,\ \beta,\ \varphi^{CN,1}/T^P_{CN}\}$. For a given parameter

Table 5: Summary of separately identified parameters

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	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^P_{od}	country-pair iceberg costs	Head-Ries Index	$\left(\frac{\pi^P_{od}}{\pi^P_{dd}}\frac{\pi^P_{do}}{\pi^P_{oo}}\right)$

vector Θ, I simulate the model, compute a list of cross-sectoral moments $\mathbb{M}(\Theta)$, and compare them with the corresponding moments in the data $\widetilde{\mathbf{m}}$. I search for the 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, I search for the best-fitting point on the grid. The details of the SMM procedure are described in the appendix.

I target five groups of data moments, corresponding to sectoral outcomes in the three economies. Table 6 summarizes the parameters and the main targeted moments for each parameter. First, the concentration parameter κ is identified by the dispersion of product attributes. Theoretically, κ solves

$$\overline{a} = \frac{I_{K/2}(\kappa)}{I_{K/2-1}(\kappa)} \tag{44}$$

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

$$\hat{\kappa} \approx \frac{\overline{a}(K - \overline{a}^2)}{1 - \overline{a}^2}. \tag{45}$$

However, competition reduces density in more central locations. Thus, $\hat{\kappa}$ cannot be mapped directly into 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 demand-side valuation of uniqueness and targets the covariance between market shares and product centrality. Fourth, the utility weight on real consumption is informed by time spent on apps relative to working hours. Fifth, the relative level of the two

Table 6: 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 productivity	employment shares	1.1

productivity normalizations in China targets the employment share of the app sector relative to the manufacturing sector.

Model Fit Table 7 reports the model-based values of the moments used in the estimation and compares them with their empirical counterparts. Overall, the model provides a reasonable fit across the five groups of targeted 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 estimation closely matches hedonic-attribute dispersion for active apps in each market. The model also reproduces bilateral trade shares in both sectors. Figure 6 shows bilateral trade shares predicted by the model against those observed in the data for the physical-goods and app sectors. The points cluster around the 45-degree line, indicating a close fit. However, the model predicts less skewness in market shares than observed in the data, which may reflect the assumption of constant returns to scale and the absence of externalities. Second, the model aligns reasonably well with observed time allocation and labor allocation across countries and sectors. Third, it performs well on non-targeted moments, such as relative wages and relative nominal GDP.

5. 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 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 uniform dyadic trade costs and productivity differences, ensuring that the quantitative results do not depend on separately identifying baseline trade barriers from productivity parameters.

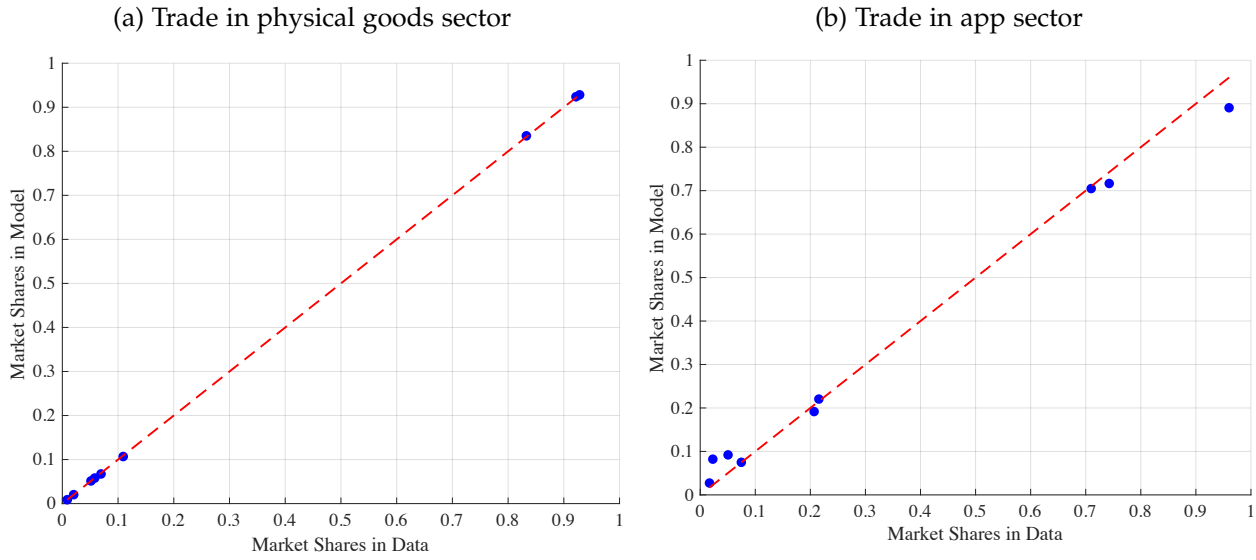
I conduct two counterfactual exercises using the estimated parameters. First, I consider the absence of China's GFW policy by removing the policy-induced trade barriers. This exercise allows me to decompose the resulting welfare changes for Chinese consumers.

Table 7: 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 data moments and model moments using the estimates in Tables 5 and 6. The current estimation searches over 20 grid points and can be further refined with a finer grid.

Figure 6: Model fit in bilateral trade flows



Notes: This figure plots bilateral trade shares predicted by the model against the corresponding shares in the data for the physical-goods and app sectors.

Second, I examine the impact of a ban on TikTok – China's most productive social-networking app – on market outcomes in the U.S. This exercise highlights distributional effects among competing apps that arise from heterogeneous substitution patterns.

5.1. The impact of the Great Firewall policy

Armed with the estimated model, I quantify the impact of the GFW policy in China. I consider a counterfactual scenario with $\overline{\tau} = 1$ and interpret the difference between the data and the counterfactual outcome as the impact of the GFW policy. Table 8 displays the changes in welfare components. The change in x is defined as $x^{\text{Data}}/x^{\text{Counterfactual}} - 1$. Quantitatively, the GFW policy has two distinct effects on the 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 substantial reduction in utility from app services. While overall time spent on apps decreased only slightly, utility per unit of app time fell 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 9 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 by only 0.34% in China, while profits rise by 7.65%, driven by head-to-head

oligopolistic competition among technology giants. This channel is consistent with the idea that domestic incumbents, such as Tencent, benefited from the absence of major foreign competitors. The loss in app services, however, reflects substitution toward lower-quality domestic alternatives, such as replacing Chrome with Baidu. While overall time spent on apps remains nearly constant, the effective cost of app usage rises sharply, by approximately 40%. This increase can be interpreted as a substantial decline in the quality-adjusted value of app time. These findings suggest that policymakers may have incentives to impose such bans to boost real income, but the associated loss in app-service utility is often overlooked in policy discussions.

Table 8: 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 product diversity among active apps. The model provides a measure of product diversity based on the dispersion of product attributes among active apps in a market. I find that 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 homogeneous goods with fewer distinctive hedonic attributes. This finding suggests that reduced competition may stifle product differentiation and innovation. It aligns with Hsieh *et al.* (2021), and rather than assuming exogenous innovation efficiency, the results here show that lower innovation can arise endogenously from reduced competition.

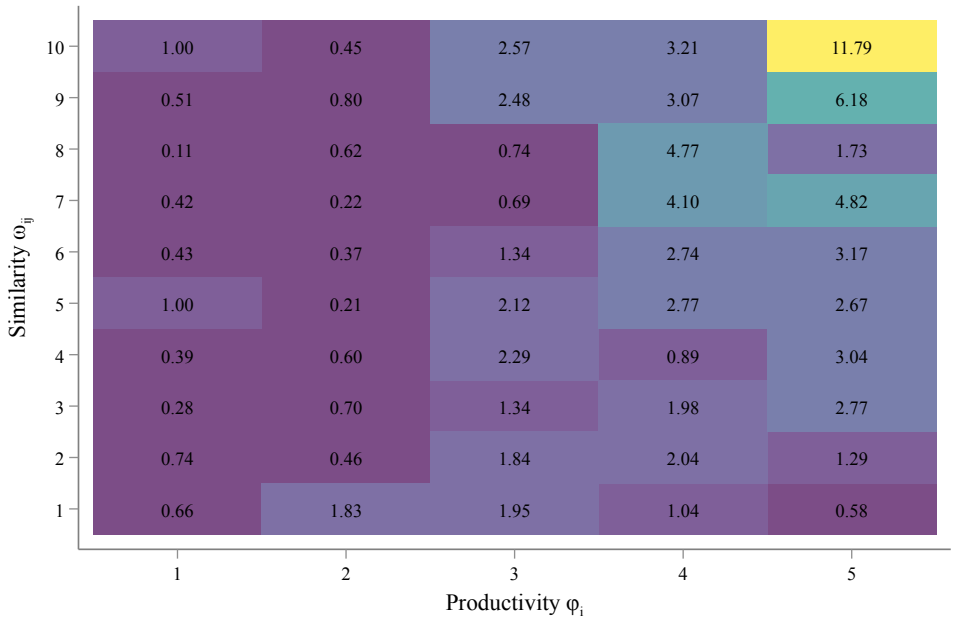
Table 9: The impact of the GFW on Chinese app sector

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

5.2. The impact of a ban on TikTok

Talk of banning TikTok began years ago as its influence grew in the U.S. In April 2024, President Biden signed legislation requiring ByteDance to divest TikTok by January 19,

Figure 7: Changes in revenue in the U.S. market



2025, or face restrictions in the U.S. market. The policy remained politically contested: 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.

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 7 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.

6. Conclusion

This paper examines the welfare implications of trade policies targeted at specific foreign products. Standard tools fall short in addressing this issue because targeted products typically have large user bases and differentiated hedonic attributes, generating 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 increases in market share after India's ban on Chinese apps. More importantly, the substitution patterns can be measured 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 emerge endogenously from a microfoundation of hedonic demand. The model is designed 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 two-stage competition.

I estimate the model 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 fostered Chinese technology 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 utility losses from the use of inferior apps, yielding a net welfare change of -7.48%. I also uncover an interaction between trade barriers and product diversity: reduced competition from the policy intervention led to a decline in product differentiation among Chinese domestic products. Finally, I find that the revenue increase for highly productive apps that are highly similar to TikTok, such as Threads and Instagram, is 11 times larger than that of the average app in the U.S. after a TikTok 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. Demand-side heterogeneity has been underexplored 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 readily implementable and broadly applicable to research areas where rich textual data are available, such as occupational substitution in labor markets 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, 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.

Atkeson, Andrew and Ariel Burstein. 2008. Pricing-to-market, trade costs, and international relative prices. *American Economic Review* 98(5):1998–2031.

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

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 Bøler, Davin Chor, Sirig Gurund, and Wei Lu. 2025. The great firewall and knowledge diffusion. Working paper, Dartmouth.

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: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.

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

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. "You Will": 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.

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. 2025. Product differentiation and oligopoly: A network approach. *American Economic Review* 115(4):1170–1225.

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. Working paper.

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

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.

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.

Appendix

Appendix A. Supplementary empirical results

Table A1: Event-study coefficients

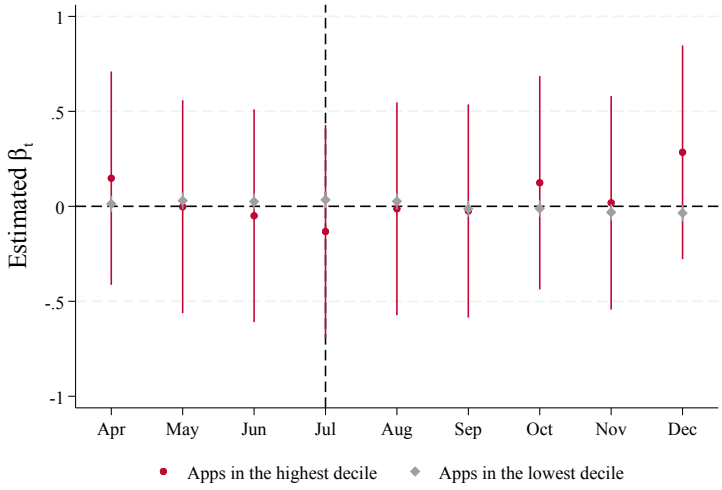
	Dependent Variable: $s_{it,IN}$			
	All	1^{st} Decile	10^{th} Decile	RoW
	(1)	(2)	(3)	(4)
$\tau = Apr,20$	-0.008	0.376	-0.136	0.148
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = May,20$	-0.005	-0.145	-0.124	-0.002
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Jun,20$	0.025	-0.065	0.013	-0.050
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Jul,20$	0.176	0.748	0.081	-0.133
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Aug,20$	0.449***	2.161**	0.179	-0.013
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Sep,20$	0.442***	2.087**	0.094	-0.024
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Oct,20$	0.454***	2.625***	0.155	0.124
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Nov,20$	0.483***	3.233***	0.146	0.018
	(0.15)	(0.92)	(0.17)	(0.29)
$\tau = Dec,20$	0.486***	3.294***	0.109	0.285
	(0.15)	(0.92)	(0.17)	(0.29)
Observations	104,974	9,749	9,759	9,961
Within-R^2				
R^2	0.99	1.00	0.97	0.15
App FEs	✓	✓	✓	✓

Under nested CES demand, the market share of product i is given by

$$s_i = \left(\frac{p_i}{P^c}\right)^{1-\sigma^c} \left(\frac{P^c}{P}\right)^{1-\eta}, \tag{46}$$

where σ^c is the elasticity of substitution within category c, η is the cross-category elasticity of substitution, $P^c = \left(\sum_{j \in M^c} p_j^{1-\sigma^c}\right)^{1/(1-\sigma^c)}$ is the category-level price index, and $P = \left(\sum_c (P^c)^{1-\eta}\right)^{1/(1-\eta)}$ is the market price index. Hence, the response of product i to

Figure A1: Market shares of apps in the rest of the world (excluding India and China)



Notes: This figure reports the robustness check corresponding to Figure 1b.

the policy shock is

$$d\ln s_i = (\sigma^c - \eta)\left(\sum_{j\in M^c_{ban}} s^c_j d\ln p_j\right) + (\eta - 1)\left(\sum_{j\in M_{ban}} s_j d\ln p_j\right), \tag{47}$$

where s^c_j represents the market share of product j within its category. Apps in the same category as banned apps have an additional response. However, conditional on category membership, nested CES predicts the same response for apps within the category. This contrasts with the continuous app-pair heterogeneity emphasized in the main text.

Appendix B. Derivations of the Generalized Translog Demand

The translog expenditure function, originally developed by Christensen *et al.* (1975), provides a flexible second-order approximation to an arbitrary expenditure function and has been widely used in international trade and industrial organization. Existing work typically imposes structured parameterizations on the second-order terms to make the system tractable for calibration or estimation; see, for example, Feenstra and Weinstein (2017) and Diewert (2023). The microfoundation developed in Section 3 delivers a particular parameterization in which substitution patterns are linked to product similarity in attribute space. This appendix collects the parameter restrictions, derivations, and special cases discussed in the main text.

Appendix B.1. Properties and parameter restrictions

For the expenditure function in equation (11) to be homogeneous of degree one in prices, concave, and consistent with valid expenditure shares, the following restrictions are sufficient.

(R1) **Homogeneity.** The translog price index

$$\ln P = \sum_{i=1}^{M} b_i \ln p_i + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta_{ij} \ln p_i \ln p_j$$

is homogeneous of degree one in prices if

$$\sum_i b_i = 1 \qquad \text{and} \qquad \sum_j \theta_{ij} = 0 \quad \forall i.$$

The first condition is imposed by the normalization of quality shifters in the main text. The second condition follows from the geometric structure of the model. Since

$$\theta_{ij} = \begin{cases} \theta\omega_{ij}, & j \neq i, \\ -\theta(\Omega_i - 1), & j = i, \end{cases}$$

and $\Omega_i = \sum_{j=1}^{M} \omega_{ij}$ with $\omega_{ii} = 1$, we have

$$\sum_j \theta_{ij} = \theta \sum_{j \neq i} \omega_{ij} - \theta(\Omega_i - 1) = \theta(\Omega_i - 1) - \theta(\Omega_i - 1) = 0.$$

Hence equation (11) is homogeneous of degree one in prices.

(R2) **Symmetry.** The translog system requires $\theta_{ij} = \theta_{ji}$ for all $i \neq j$. This condition is automatically satisfied because

$$\theta_{ij} = \theta\omega_{ij} = \theta\omega_{ji} = \theta_{ji},$$

where $\omega_{ij} = \boldsymbol{a}_i^\top \boldsymbol{a}_j$ is symmetric by construction.

(R3) **Concavity.** A sufficient condition for concavity of the expenditure function in log prices is that the matrix $\boldsymbol{\Theta} = (\theta_{ij})$ be negative semidefinite on the relevant price domain. Since $\boldsymbol{\Theta} = \theta \boldsymbol{G}$, where

$$(\boldsymbol{G})_{ij} = \omega_{ij} \quad (i \neq j), \qquad (\boldsymbol{G})_{ii} = -(\Omega_i - 1) = -\sum_{j \neq i} \omega_{ij}, \tag{48}$$

it is sufficient to show that $\boldsymbol{G} \preceq 0$ when $\theta > 0$.

Define the signed Laplacian

$$L := -\boldsymbol{G}. \tag{49}$$

Then $\boldsymbol{G} \preceq 0$ is equivalent to $L \succeq 0$.

Case 1: all pairwise similarities are nonnegative. If $\omega_{ij} \geq 0$ for all $i \neq j$, then for any $x \in \mathbb{R}^M$,

$$x^\top \boldsymbol{G} x = -\frac{1}{2} \sum_{i \neq j} \omega_{ij} (x_i - x_j)^2 \leq 0.$$

Therefore $\boldsymbol{G} \preceq 0$, or equivalently $L \succeq 0$.

Case 2: negative similarities are localized on a small subset. To allow for the possibility of complementarity, define the positive and negative parts of ω_{ij}:

$$\omega_{ij}^+ := \max\{\omega_{ij}, 0\}, \qquad \omega_{ij}^- := \max\{-\omega_{ij}, 0\}, \qquad \omega_{ij} = \omega_{ij}^+ - \omega_{ij}^-. \tag{50}$$

Let L^+ denote the ordinary Laplacian associated with the nonnegative weights ω_{ij}^+:

$$(L^+)_{ij} = \begin{cases} -\omega_{ij}^+, & i \neq j, \\ \sum_{k \neq i} \omega_{ik}^+, & i = j. \end{cases} \tag{51}$$

By construction, $L^+ \succeq 0$. Let $B \subseteq \{1, \ldots, M\}$ denote the subset of products on which negative weights may occur, and let $P = \{1, \ldots, M\} \setminus B$.

Assumption 1 (Localized negative similarities) *There exists a subset $B \subseteq \{1, \ldots, M\}$ such that $\omega_{ij}^- = 0$ whenever $i \notin B$ or $j \notin B$. Equivalently, all negative similarities are confined to edges with both endpoints in B.*

Under Assumption 1, the signed Laplacian can be written as

$$L = L^+ - \begin{pmatrix} 0 & 0 \\ 0 & L_B^- \end{pmatrix}, \tag{52}$$

where L_B^- is the Laplacian associated with the negative weights ω_{ij}^- on the induced subgraph B.

Partition L^+ according to (P,B):

$$L^+ = \begin{pmatrix} L^+_{PP} & L^+_{PB} \\ L^+_{BP} & L^+_{BB} \end{pmatrix}. \tag{53}$$

Define the effective Laplacian on B by the Kron reduction

$$L^+_{\text{eff}} := L^+_{BB} - L^+_{BP}(L^+_{PP})^\dagger L^+_{PB}, \tag{54}$$

where $(\cdot)^\dagger$ denotes the Moore–Penrose pseudoinverse.

Lemma A1 *Under Assumption 1, if*

$$L^+_{\text{eff}} \succeq L^-_B, \tag{55}$$

then $L \succeq 0$, and therefore $\boldsymbol{G} \preceq 0$.

Proof Fix any $x \in \mathbb{R}^M$ and partition it as $x = (x_P,x_B)$. From equation (52),

$$x^\top L x = x^\top L^+ x - x_B^\top L^-_B x_B. \tag{56}$$

For fixed x_B, the quadratic form $x^\top L^+ x$ is minimized over x_P at the Kron-reduced value

$$\min_{x_P} x^\top L^+ x = x_B^\top L^+_{\text{eff}} x_B. \tag{57}$$

Combining equations (56) and (57),

$$x^\top L x \geq x_B^\top (L^+_{\text{eff}} - L^-_B) x_B.$$

If condition (55) holds, the right-hand side is nonnegative for all x_B. Hence $x^\top L x \geq 0$ for all x, so $L \succeq 0$. Therefore $\boldsymbol{G} \preceq 0$. ■

Interpretation. Case 1 corresponds to the empirically relevant case in which pairwise similarities are nonnegative, so all products are substitutes. In that case, concavity follows immediately. Case 2 allows for a limited amount of complementarity: some product pairs may have negative similarity, but these negative links must be sufficiently sparse and weak relative to the positive part of the network. In the application, estimated similarities between mobile apps are positive, so Case 1 is the relevant benchmark; Case 2 is included only to show that the framework can accommodate limited complementarity.

Appendix B.2. Comparison with other special cases

The demand system nests several familiar specifications as special cases.

Case 1: Cobb–Douglas. If $\theta = 0$, then $\theta_{ij} = 0$ for all i and j. The expenditure function reduces to

$$\ln P = \sum_i b_i \ln p_i,$$

and all Allen–Uzawa elasticities of substitution are equal to one.

Case 2: symmetric translog. In Feenstra and Weinstein (2017), the second-order terms are symmetric across products. This case is nested by the present framework when all products have the same attribute vector, so that $\boldsymbol{a}_i = \boldsymbol{a}$ for all i. Then $\omega_{ij} = 1$ for all i,j, $\Omega_i = M$, and

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

Equivalently, if one normalizes the common translog curvature parameter by the number of products, this reproduces the standard symmetric translog structure.

Case 3: generalized symmetric translog. The formulation in Diewert (2023) is nested when products differ only along a single scalar attribute dimension. Abstracting from the unit-vector normalization for this special case, let a_i denote the scalar attribute loading. The second-order terms become

$$\theta_{ij} = \begin{cases} \theta a_i a_j & j \neq i \\ \theta \left[a_i^2 - a_i \sum_j a_j \right] & j = i \end{cases}$$

which matches the generalized translog parameterization in that setting.

Appendix B.3. Derivations of market shares and elasticities

Let the homothetic expenditure function be $e(p,U) = UP(p)$, where $P(p)$ is the translog price index in equation (11). By Shephard's lemma, the expenditure share of product i is

$$s_i \equiv \frac{p_i h_i(p,U)}{e(p,U)} = \frac{\partial \ln e(p,U)}{\partial \ln p_i} = \frac{\partial \ln P(p)}{\partial \ln p_i},$$

where $h_i(p,U)$ is Hicksian demand. Because

$$\ln P = \sum_{i=1}^{M} b_i \ln p_i + \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta_{ij} \ln p_i \ln p_j,$$

and $\boldsymbol{\Theta}$ is symmetric, differentiation gives

$$s_i = b_i + \sum_{j=1}^{M} \theta_{ij} \ln p_j. \tag{58}$$

Using the definitions of θ_{ij}, ω_{ij}, and Ω_i, this can be rewritten as

$$\begin{aligned} s_i &= b_i - \theta(\Omega_i - 1)\ln p_i + \theta \sum_{j \neq i} \omega_{ij} \ln p_j \\ &= b_i - \theta(\Omega_i - 1)\left[\ln p_i - \sum_{j \neq i} \frac{\omega_{ij}}{\Omega_i - 1} \ln p_j\right], \end{aligned} \tag{59}$$

which is equation (13) in the main text.

Elasticities. The compensated price elasticity of demand is defined by

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

Since $q_i = s_i e(p,U)/p_i$ and utility is held constant, we have

$$\ln q_i = \ln s_i + \ln e(p,U) - \ln p_i.$$

Differentiating with respect to $\ln p_j$ at constant U gives

$$\eta_{ij} = \frac{1}{s_i}\frac{\partial s_i}{\partial \ln p_j} + \frac{\partial \ln e(p,U)}{\partial \ln p_j} - \mathbb{1}_{i=j}.$$

Using Shephard's lemma and equation (58),

$$\frac{\partial \ln P}{\partial \ln p_j} = s_j, \qquad \frac{\partial s_i}{\partial \ln p_j} = \theta_{ij},$$

so

$$\eta_{ij} = s_j + \frac{\theta_{ij}}{s_i} - \mathbb{1}_{i=j},$$

and the Allen–Uzawa partial elasticity of substitution is simply

$$\sigma_{ij} \equiv \frac{\eta_{ij}}{s_j} = 1 + \frac{\theta_{ij}}{s_i s_j} - \mathbb{1}_{i=j}\frac{1}{s_i},$$

which yields the expressions in Proposition 1.

Reservation price. The reservation price is defined as the price that drives product i's expenditure share to zero, holding fixed the product set and competitors' prices. Setting $s_i = 0$ in equation (13) gives

$$0 = b_i - \theta(\Omega_i - 1)\left(\ln p_i - \sum_{j \neq i} \frac{\omega_{ij}}{\Omega_i - 1} \ln p_j\right).$$

Solving for p_i yields the reservation price

$$\ln \overline{p}(\boldsymbol{a}_i \mid \mathcal{M}) = \frac{b_i}{\theta(\Omega_i - 1)} + \sum_{j \neq i}\left(\frac{\omega_{ij}}{\Omega_i - 1}\right)\ln p_j,$$

which is equation (17) in the main text.

Appendix B.4. Block-Similarity Special Case

This subsection considers a special case of the similarity matrix that generates a nested pattern of competition and clarifies the relationship between the pricing rule in the main text and nested-logit demand.

Partition the product set into nests indexed by $g \in \{1, \ldots, G\}$. Let $g(i)$ denote the nest to which product i belongs, and let M_g denote the number of products in nest g. Suppose pairwise similarities satisfy

$$\omega_{ij} = \begin{cases} 1, & i = j, \\ \rho, & i \neq j \text{ and } g(i) = g(j), \\ \lambda, & g(i) \neq g(j), \end{cases} \qquad 1 \geq \rho \geq \lambda \geq 0. \tag{60}$$

Thus, products within the same nest share a common similarity ρ, while products in different nests share a common similarity λ.

Under this structure, for any product i in nest g,

$$\Omega_i = 1 + (M_g - 1)\rho + (M - M_g)\lambda. \tag{61}$$

Substituting this expression into the pricing rule in the main text yields

$$p_i = \left(1 + \frac{s_i}{\theta\big[(M_g - 1)\rho + (M - M_g)\lambda\big]}\right) c_i. \tag{62}$$

Equation (62) implies that markups increase with a product's expenditure share and decrease with the density of competitors in its relevant neighborhood. In particular, greater within-nest similarity ρ, greater across-nest similarity λ, or a larger nest all reduce markups by intensifying substitution.

This special case is closely related in spirit to nested logit. In nested logit, substitution is stronger within nests than across nests, so the elasticity faced by a product depends on the structure of demand within its nest. The corresponding markup rule for a single-product firm can be written as

$$p_i - c_i = \frac{1}{\lambda\left[\frac{1}{1-\rho} - \frac{\rho}{1-\rho} s_{i|g} - s_i\right]}, \tag{63}$$

where $s_{i|g} \equiv s_i / s_g$ denotes product i's share within nest g.

Equations (62) and (63) deliver the same qualitative implication: products face stronger competitive pressure from products in the same nest, and markups are lower when within-nest competition is more intense. The key difference lies in the source of that force.

In nested logit, the nest structure is primitive and substitution patterns are generated by a discrete-choice tree. In the present model, substitution patterns are generated by product geometry. The block-similarity case should therefore be interpreted as a geometric analogue of nested logit rather than as an exact special case.

Appendix C. Uniqueness of Equilibrium

This appendix establishes uniqueness of the equilibrium, conditional on the given product set and a given entry sequence in each destination. The argument has two steps. First, for any fixed active set of firms, the second-stage Bertrand equilibrium is unique. Second, because the entry sequence is fixed, the sequence evaluates a unique profit vector at each candidate active set and therefore selects a unique active set satisfying the equilibrium inequalities in Section 3.

Fix a destination market and an active set $\mathcal{A} \subseteq \mathcal{M}$. Let $x_i \equiv \ln p_i$ for $i \in \mathcal{A}$, and let c_i denote the effective marginal cost in that market. Prices of inactive products are fixed at their reservation prices, so their contribution to the active firms' expenditure shares can be absorbed into constants

$$\widetilde{b}_{i,\mathcal{A}} \equiv b_i + \sum_{j \notin \mathcal{A}} \theta_{ij} \ln \overline{p}_j, \qquad i \in \mathcal{A}.$$

Combining the market-share system in equation (13) with the Bertrand pricing rule in equation (19), equilibrium prices solve

$$\widetilde{b}_{i,\mathcal{A}} + \sum_{j \in \mathcal{A}} \theta_{ij} x_j - \theta(\Omega_i - 1)\left(e^{x_i - \ln c_i} - 1\right) = 0, \qquad i \in \mathcal{A}.$$

This system is the first-order condition of the function

$$\Psi_{\mathcal{A}}(x) = \sum_{i \in \mathcal{A}} \widetilde{b}_{i,\mathcal{A}} x_i + \frac{1}{2} x^\top \boldsymbol{\Theta}_{\mathcal{A}} x - \sum_{i \in \mathcal{A}} \theta(\Omega_i - 1)\left(e^{x_i - \ln c_i} - x_i\right), \tag{64}$$

where $\boldsymbol{\Theta}_{\mathcal{A}}$ is the restriction of the demand-curvature matrix to the active set.

The Hessian of $\Psi_{\mathcal{A}}$ is

$$\nabla^2 \Psi_{\mathcal{A}}(x) = \boldsymbol{\Theta}_{\mathcal{A}} - \operatorname{diag}\left(\theta(\Omega_i - 1) e^{x_i - \ln c_i}\right)_{i \in \mathcal{A}}. \tag{65}$$

By Appendix B.1, $\boldsymbol{\Theta}_{\mathcal{A}}$ is negative semidefinite. The second term is strictly negative definite because $\theta(\Omega_i - 1) > 0$ and $e^{x_i - \ln c_i} > 0$ for all $i \in \mathcal{A}$. Hence $\Psi_{\mathcal{A}}$ is strictly concave. It follows that the first-order conditions admit a unique solution. Therefore, for any fixed active set $\mathcal{A}$, the second-stage equilibrium in prices and shares is unique.

Now fix an entry sequence in destination d. Since the pricing equilibrium is unique for every candidate active set along that sequence, the operating profit of each potential entrant is uniquely determined whenever its entry is considered. Thus the sequence leaves no ambiguity about which firms can enter while satisfying the two equilibrium inequalities: all active firms earn non-negative operating profits, and any additional entrant would earn negative operating profits if added to the active set and the Bertrand game were re-solved. Because the number of potential entrants is finite, following the sequence yields a unique terminal active set. Combining this unique active set with the unique Bertrand equilibrium for that set gives a unique equilibrium, conditional on the given product set and entry sequence.

Appendix D. Extension of Quantitative Model

This extension adds producer-side development and maintenance costs to the app-sector environment. Let app producers hire software engineers to pay a global development cost and, when serving destination d, a market-specific maintenance cost. These costs enter the first-stage entry problem but do not change the second-stage pricing condition, because they are fixed with respect to user activity. Thus the generalized translog demand system and the Bertrand markup rule remain unchanged. The entry condition is modified so that operating profits must cover the relevant fixed costs, and the aggregate labor-market-clearing condition must include engineering labor used in development and maintenance.

Online Appendices to

"Blocking the Giants: Theory and Evidence from App Bans"

by

©Rachel Ruiqi Sun

Online Appendix A. von Mises–Fisher Distribution

This section describes the von Mises–Fisher (vMF) distribution used in the quantitative model to characterize the distribution of product blueprints on the attribute space. In the main text, each product i is associated with a blueprint vector $\mathbf{a}_i \in \mathbb{R}^K$ satisfying $\|\mathbf{a}_i\| = 1$, where $\mathbf{a}_i$ summarizes the mix of hedonic attributes embodied in the product. Because the support of $\mathbf{a}_i$ is the unit sphere, a directional distribution is a natural choice. The vMF distribution provides a parsimonious way to model both the central direction of product designs and the degree of dispersion around that direction.

In the quantitative analysis, the blueprint vector $\mathbf{a}_i$ is not directly observed. Instead, I recover it from product-description text using embedding representations generated by a pretrained language model. These embeddings map each product description into a high-dimensional semantic vector. After normalization to unit length, the resulting vectors lie on the unit sphere and can be interpreted as empirical counterparts to the blueprint vectors in the model. Under this interpretation, text embeddings serve as measurement objects for the latent blueprint rather than as economic primitives themselves. This measurement strategy makes directional statistics, and in particular the vMF distribution, a natural tool for modeling the cross-product distribution of blueprints.

Formally, let $\mathbb{S}^{K-1} \equiv \{\mathbf{x} \in \mathbb{R}^K : \|\mathbf{x}\| = 1\}$ denote the unit sphere in $\mathbb{R}^K$. A random vector $\mathbf{x} \in \mathbb{S}^{K-1}$ follows a von Mises–Fisher distribution with mean direction $\boldsymbol{\mu} \in \mathbb{S}^{K-1}$ and concentration parameter $\kappa \geq 0$ if its density is

$$f_K(\mathbf{x};\boldsymbol{\mu},\kappa) = C_K(\kappa) \exp\left(\kappa \boldsymbol{\mu}^\top \mathbf{x}\right), \tag{66}$$

where

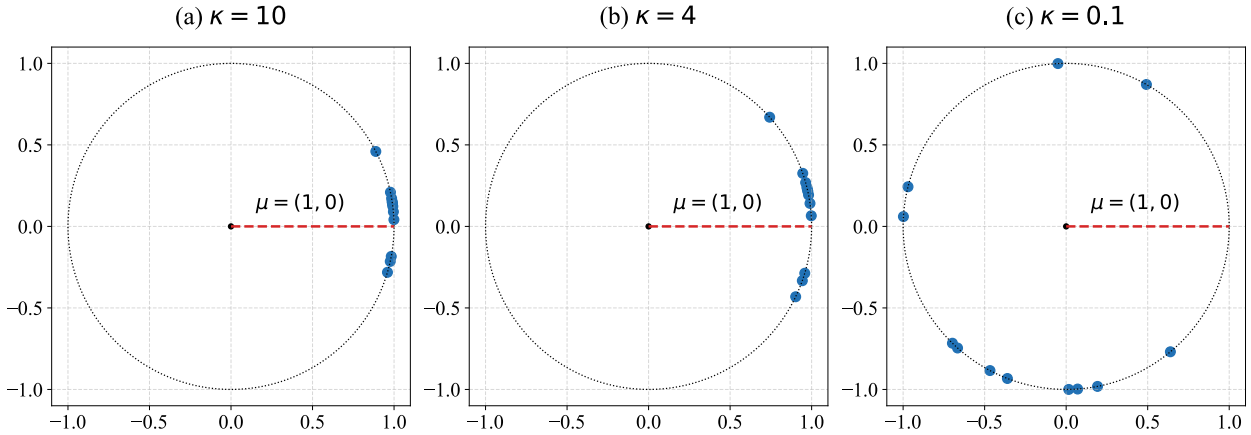
$$C_K(\kappa) = \frac{\kappa^{K/2-1}}{(2\pi)^{K/2} I_{K/2-1}(\kappa)}, \tag{67}$$

and $I_\nu(\cdot)$ is the modified Bessel function of the first kind of order ν. The parameter $\boldsymbol{\mu}$ determines the central direction of the distribution, while κ governs concentration around $\boldsymbol{\mu}$.

The density in equation (66) depends on $\mathbf{x}$ only through the inner product $\boldsymbol{\mu}^\top \mathbf{x} = \cos(\angle(\mathbf{x},\boldsymbol{\mu}))$. As a result, probability mass is increasing in directional proximity to $\boldsymbol{\mu}$. When $\kappa = 0$, the density is uniform on the sphere. As κ increases, the distribution becomes more concentrated around $\boldsymbol{\mu}$. In the limit as $\kappa \to \infty$, draws collapse onto the mean direction. In the quantitative model, this property is useful because κ maps directly into the degree of product differentiation: low κ implies that blueprints are dispersed in attribute space, while high κ implies that products are clustered around a common design direction.

The vMF distribution is the directional analogue of the multivariate normal distribution. In particular, it plays on the unit sphere the role that the Gaussian plays in Euclidean space: $\boldsymbol{\mu}$ is a location parameter and κ is an inverse-dispersion parameter. This interpretation is consistent with the discussion in the main text, where the mean direction summarizes the center of the product set and the concentration parameter governs how tightly products are clustered around that center.

Figure A1: Illustration of the von Mises distribution on the unit circle



Notes: This figure plots 10 random draws from the von Mises distribution on the unit circle. In all three panels, the mean direction is $\mu = (1,0)$. Panels (a), (b), and (c) use $\kappa = 0.1$, $\kappa = 4$, and $\kappa = 10$, respectively. Larger values of κ imply greater concentration around the mean direction. The random seed is set to 10.

The moments of the vMF distribution connect naturally to the directional-statistics objects used in the theory. If $\mathbf{a}_1, \ldots, \mathbf{a}_M$ are i.i.d. draws from $f_K(\mathbf{a}; \boldsymbol{\mu}, \kappa)$, then the sample resultant vector

$$\overline{\mathbf{a}} = \sum_{i=1}^{M} \mathbf{a}_i \tag{68}$$

summarizes the location and dispersion of the sample. Its normalized version, $\overline{\mathbf{a}}/\|\overline{\mathbf{a}}\|$, estimates the population mean direction $\boldsymbol{\mu}$, while the mean resultant length $\|\overline{\mathbf{a}}\|/M$ is increasing in κ. In the present setting, this link is useful because the same geometric objects that summarize the text-based blueprint vectors also enter the demand system through product similarity and centrality.

When $K = 2$, the support is the unit circle and the vMF distribution reduces to the von Mises distribution. Writing a point on the circle as an angle $x \in [-\pi, \pi)$ and the mean direction as angle μ, the density becomes

$$f_2(x; \mu, \kappa) = \frac{\exp\big(\kappa \cos(x - \mu)\big)}{2\pi I_0(\kappa)}. \tag{69}$$

This special case is useful for visualization because the geometry can be represented in two dimensions while preserving the economic interpretation of directional proximity.

Figure A1 illustrates the role of κ in the two-dimensional case. Each panel plots 10 draws from a von Mises distribution centered at $\mu = (1,0)$. As κ rises, the draws become more tightly concentrated around the mean direction. The figure therefore provides a visual counterpart to the product-space interpretation in the main text: larger κ corresponds to less dispersion in attribute space and, in turn, to stronger similarity across products.

In the quantitative model, I use the vMF distribution as a parsimonious statistical model for the arrival of product blueprints. This specification serves two purposes. First,

it respects the geometric structure of the blueprint vectors, whether treated as latent objects in the theory or as normalized text embeddings in the data. Second, it allows the degree of horizontal differentiation to be summarized by a small number of interpretable parameters, notably the mean direction $\boldsymbol{\mu}$ and the concentration parameter κ. These parameters can then be estimated jointly with the rest of the model and mapped into the substitution patterns implied by generalized translog demand.

Online Appendix B. App Data