Digital Interoperability: Measurement, Incentives, and Welfare*

Ao Wang† Jian Xie‡ Han Yuan§

November 21, 2026

Abstract

We develop a structural framework that connects digital interoperability measurement to firms' incentives to provide it and apply it to evaluate a non-discriminatory interoperability policy in the Chinese mobile app market. Our proposed measure, query, is constructed from app code and records whether one app requests visibility of another. Query links are associated with greater joint usage, directed disproportionately toward larger apps, and more prevalent under common ownership. Our structural estimates reveal heterogeneous interoperability costs and higher interoperability between affiliated apps. We simulate equilibrium interoperability provision under a mandate requiring uniform outgoing interoperability across recipients within app categories. The mandate expands connectivity but reduces consumer surplus and profits. Losses among frequently used app pairs outweigh gains elsewhere, highlighting how firms' reallocation of interoperability shapes the welfare consequences of regulation.

Keywords: Interoperability, Digital Markets, Common Ownership, Complementarity, Mobile Applications
JEL codes: L13, L41, L86, M15

*We thank seminar participants at the University of Warwick, the SUSTech College of Business, Renmin University of China, Zhongnan University of Economics and Law, China IO conference 2026 (SUFE), and are grateful to James Fenske, Mike Waterson, and Mo Xiao for helpful conversations. Han's research is supported by National Natural Science Foundation of China (grant number: 7260030592).

†Department of Economics and CAGE, University of Warwick. Email: ao.wang@warwick.ac.uk.
‡College of Business, Southern University of Science and Technology. Email: xiej@sustech.edu.cn.
§College of Business, Shanghai University of Finance and Economics. Email: yuanhan1992@gmail.com.

1 Introduction

In digital markets, firms compete not only on the standalone quality of their products and services, but also on how easily users can use them together. Interoperability can make combinations of services more valuable to consumers, yet firms choose which connections to support according to their private returns and implementation costs. These incentives may favor connections with affiliated services over those with rivals, raising concerns about exclusion and foreclosure. For example, in 2019, Spotify accused Apple of disadvantaging competing streaming services, including by restricting integration with services and devices such as Siri, HomePod, and Apple Watch.[1] Such concerns have motivated regulatory efforts to expand interoperability. The European Union's Digital Markets Act imposes interoperability obligations on designated gatekeepers in specified settings, while the United Kingdom's Digital Markets, Competition and Consumers Act 2024 allows the Competition and Markets Authority to impose requirements preventing designated firms from restricting interoperability.[2] Assessing these policies requires measuring interoperability, understanding firms' strategic incentives to provide it, and evaluating the consequences of their choices for consumer welfare and competition.

In this paper, we develop a structural framework that connects the measurement of digital interoperability to firms' incentives and the welfare consequences of interoperability policy, and investigate these issues in the Chinese mobile application market. We begin by proposing a directional measure based on *query* declarations, which record explicit requests by one digital product to recognize another within the operating system. Such recognition can help a product tailor its functionality to the presence and capabilities of another, supporting interactions such as opening linked content or handing off tasks. In Android, for example, developers can explicitly name other apps in their query declarations to request visibility of those apps. We extract these declarations from application files to construct a panel of directional query networks among mobile apps. Although our implementation focuses on Android, the measurement approach can extend to other operating systems with analogous mechanisms for declaring visibility requests, including iOS.

Combining query declarations with app usage and ownership data, we document three empirical patterns that support the economic relevance of our mea-

[1] See Spotify's announcement and Apple's response, which disputed these claims.

[2] For an application of the DMA to Apple's interoperability obligations, see the European Commission's March 2025 decisions. For the UK provisions, see Section 20 of the Act.

sure and motivate the structural model. First, query links are positively associated with joint usage, conditional on app-category-pair and app-week fixed effects. App pairs with reciprocal queries have approximately 6.4% more common users than pairs without query links. This relationship is consistent with interoperability facilitating the joint use of apps. Second, query links are distributed asymmetrically: leading apps receive more incoming queries than they send outgoing queries. Within an app, a 1% increase in weekly active users is associated with approximately 0.24% more incoming queries, without a corresponding increase in outgoing queries. Third, the query network is denser within ownership groups than across them. Common ownership is associated with a 7.4–13.8-percentage-point higher probability of a query link.

These patterns suggest that query captures an economically relevant dimension of interoperability and that incentives to establish connections may vary with app popularity and ownership. We develop a demand-and-supply framework for digital interoperability that captures these channels. On the demand side, users choose the app bundle that delivers the highest utility. Bundle utility combines standalone app qualities with pair-specific synergies, which reflect both directional interoperability and other functional complementarities. On the supply side, firms choose latent directional interoperability levels to maximize revenue net of interoperability costs, taking into account the effects on usage across their app portfolios. We introduce a thresholding technology that links these choices to observed query declarations: a query is observed when latent interoperability exceeds an implementation threshold.

Together, these steps connect app usage, ownership, and query network data, allowing us to estimate the demand, cost, and query-technology parameters via a likelihood approach. Concretely, our strategy combines a novel demand inverse, app owner' first-order conditions, and the query technology. For a given set of demand parameters, we invert observed marginal and pairwise usage shares to recover standalone app utility indexes and total pair-specific synergies. Plugging these recovered utility indexes and demand synergies into firms' first-order conditions of latent interoperability, we can further invert these conditions under convex interoperability cost and back out the latent directional interoperability levels given cost parameters. Finally, the query technology maps these interoperability levels into probabilities of observing query links, forming the desired likelihood.

Demand parameters are identified through exogenous variation in the marginal revenue from interoperability in firms' first-order conditions. The response of marginal revenue to this variation depends on demand parameters, while the vari-

ation is excluded from costs and the query technology. The likelihood alone, however, may not separately identify cost and query-technology parameters. Intuitively, conditional on demand, a weak response of query propensity to a positive marginal-revenue shock can reflect either steeply rising marginal costs, so that app owners barely adjust interoperability, or substantial adjustments whose effects on observed links are muted by widely dispersed query thresholds.

To resolve this issue, we use a post-MLE instrumental variables regression of recovered pair-specific synergy on the component of directional interoperability levels normalized by query-threshold dispersion. The likelihood approach identifies this normalized regressor, whose coefficient in the synergy equation is the threshold dispersion. However, the regressor may be correlated with residual pair-specific demand synergies because of firms' strategic choices of interoperability. We instrument for it using the number of non-exported activities, which shifts standalone app utility and hence equilibrium interoperability. These activities support internal functionality and are unavailable for direct launching by other apps, motivating their exclusion from residual pair-specific synergies. The IV step then identifies threshold dispersion, allowing to recover the remaining query-technology and cost parameters.

We apply the model to the mobile apps in the communication and online video categories. We find substantial heterogeneity in consumer preferences and firms' costs of providing interoperability. Allowing for heterogeneous tastes substantially improves model fit, with particularly pronounced heterogeneity among communication apps. Estimated demand synergies are also strongest between communication apps. On the supply side, cost curvature varies considerably across owners, while query thresholds differ across directed category pairs. Separating strategic interoperability from residual functional complementarity reveals that affiliated apps generally enjoy higher interoperability, whereas interoperability across owners is more dispersed and exhibits a pronounced lower tail. These estimates indicate that ownership-related differences in connections extend beyond residual functional complementarity.

We use the estimated model to simulate an interoperability mandate requiring each app to provide the same level of outgoing interoperability to all recipients within a given category (and thus a uniform probability of establishing a query), regardless of ownership, while allowing firms to choose that common level. The mandate expands connectivity across owners but reduces it within ownership groups, increasing expected query-network density by 3.82 percentage points overall. Nevertheless, the median changes in consumer surplus and aggregate owner profit are

-1.94% and -4.41%, respectively. The consumer surplus losses arise because the mandate reallocates interoperability away from some of the app combinations consumers use most frequently: gains from pairs in the bottom nine deciles of baseline joint usage are more than offset by losses among pairs in the highest decile. Thus, an interoperability mandate can generate a denser query network while reducing welfare, highlighting the importance of how interoperability is reallocated across app pairs. Evaluating its welfare implications requires accounting for firms' incentives and costs of providing interoperability, their resulting adjustments across connections, and consumers' preferences for the affected app combinations.

Related Literature. This paper contributes to the empirical literature on compatibility and interoperability by introducing a code-based, directional measure of interoperability in digital markets. Existing empirical studies primarily examine industries in which compatibility is governed by explicit technical standards, formal exclusivity agreements, or observable access prices, including modems (Augereau *et al.*, 2006), VCR standards (Park, 2004), inter-bank ATM access (Knittel & Stango, 2009, 2011), mobile money (Brunnermeier *et al.*, 2023), video game exclusivity (Lee, 2013), railway gauge (Gross, 2020), electric vehicle charging standards (Li, 2023), and electronic health record systems (Dix *et al.*, 2025). By contrast, interoperability in digital markets often takes more complex technical forms, including application interfaces, software compatibility, and data exchange, which are difficult to observe directly. We propose a technically grounded measure, *query*, that captures a common economic role of interoperability: whether one digital product recognizes and directs users toward another. Our measure is directional and therefore relates to the literature on one-way compatibility, e.g., Farrell & Saloner (1992); Manenti & Somma (2008); Adner *et al.* (2020), and differs from two-way compatibility that requires bilateral agreement to implement interoperability, e.g., Katz & Shapiro (1985); Crémer *et al.* (2000); Chen *et al.* (2009a). In particular, we provide micro-level evidence that directional interoperability is highly asymmetric with respect to market position: larger apps receive disproportionately more incoming query links than they initiate outgoing links. Although our implementation is specific to Android apps, the broader measurement logic—that recognition is essential for sophisticated technical interactions—apply more widely in digital markets.

This paper also contributes to the literature on the strategic role of interoperability. Existing theoretical and empirical work shows that firms may use interoperability strategically to soften competition (Augereau *et al.*, 2006; Doganoglu & Wright, 2006; Peitz & Sato, 2025; Ekmekci *et al.*, 2025; Chen & Rey, 2026), facilitate collu-

sion (Armstrong, 1998; Gross, 2020), or foreclose rivals (Whinston, 1990; Chen *et al.*, 2009b; Knittel & Stango, 2009, 2011; Lee, 2013; Genakos *et al.*, 2018; Heidhues *et al.*, 2024). We show that ownership can affect interoperability for two reasons: first, the benefit of interoperability can be fully internalized among affiliated apps; second, by degrading interoperability with rival apps, tech giants can further increase usage of affiliated apps. Our structural model incorporates both incentives. Our counterfactual analysis also informs the active policy debate on interoperability mandates in digital markets (Kades & Morton, 2020; Brown, 2020; OECD, 2021; Bourreau & Kraemer, 2022; Crawford *et al.*, 2023). We provide quantitative evidence that a non-discriminatory mandate requiring equal outgoing interoperability within an app category expands overall. Forcing uniform provision across all destination apps in a category induces owners to scale back interoperability on the app combinations most heavily used by consumers. This finding highlights that the welfare effects of interoperability regulation depend crucially on the reallocation of interoperability across high-value connections, not merely on aggregate network density.

Finally, our methodological framework contributes to the literature on mobile app demand (Han *et al.*, 2016; Aridor, 2025; Yuan, 2026; Frahan *et al.*, 2025) and the structural estimation of bundle demand and product complementarity (Gentzkow, 2007; Ershov *et al.*, 2025; Iaria & Wang, 2026). Existing bundle-demand models infer demand synergies from observed joint consumption or purchase shares (Iaria & Wang, 2020; Wang, 2021), typically treating pairwise synergies as exogenous primitives. We extend this literature in two respects. First, we establish the invertibility of marginal and pairwise joint-usage shares together, allowing us to recover both app-level mean utilities and total pairwise demand synergies in random-utility model of demand for bundles. Second, we distinguish the strategic component of demand synergy arising from firms' interoperability choices from residual non-strategic component in the synergy. By modeling interoperability as an endogenous choice and linking it to observed query declarations, our framework connects the value consumers derive from using products together to firms' incentives to facilitate those interactions.[3] This framework connects code-based measurement of interoperability with structural demand-and-supply estimation, providing a tractable empirical tool for evaluating interoperability policies and their welfare consequences.

[3] As shown in Iaria & Wang (2026), bundle discounts can also enter demand synergy as a strategic component. Unlike observed bundle discounts, digital interoperability is often latent and not directly observed in the data.

2 The Chinese mobile application market

We study interoperability between apps and its economic consequences for app usage and competition in the Chinese mobile application market. Mobile applications play a central role in communication, shopping, payment, navigation, entertainment, and information access in China. Consumers use combinations of these services, so the value they obtain depends not only on the quality of individual apps but also on how easily they can use those apps together. Interoperability can reduce frictions in moving between services, sharing content, and completing tasks across apps. Its economic consequences therefore depend on which apps are connected and how consumers value them.

The scale of the Chinese market makes these questions economically significant. According to QuestMobile, China's mobile internet market reached 1.267 billion monthly active users in June 2025.[4] Large technology groups have broad and overlapping user bases: Tencent, Alibaba, ByteDance, and Baidu each reached more than one billion monthly active users across their app portfolios. These group-level figures count users once within each group, but the same user may use apps belonging to several groups. Multi-app ownership is also prevalent. Leading firms control or invest in apps across communication, shopping, payment, video, maps, browsers, and utilities. For example, among the top 200 apps in our sample, Tencent directly owns 37 apps, including WeChat and QQ.

Although the identities of the leading firms differ across countries, the underlying features—widespread use of multiple apps and firms operating broad ecosystems across services—are shared with other major app markets, including the United States (see, e.g., Bresnahan *et al.*, 2014; Aridor, 2025; Jacobides *et al.*, 2024). These features make the Chinese market a useful setting for examining how firms allocate interoperability within and across ownership boundaries, and how interoperability regulation affects connection, app usage and, consumer surplus.

2.1 Data

App usage data. Our first dataset is a proprietary weekly panel of mobile app usage from iResearch, a Chinese consulting firm specializing in the mobile Internet industry. Similar app usage data has been used in prior work on competition between mobile applications (Yuan, 2026). The dataset covers the top 200 apps

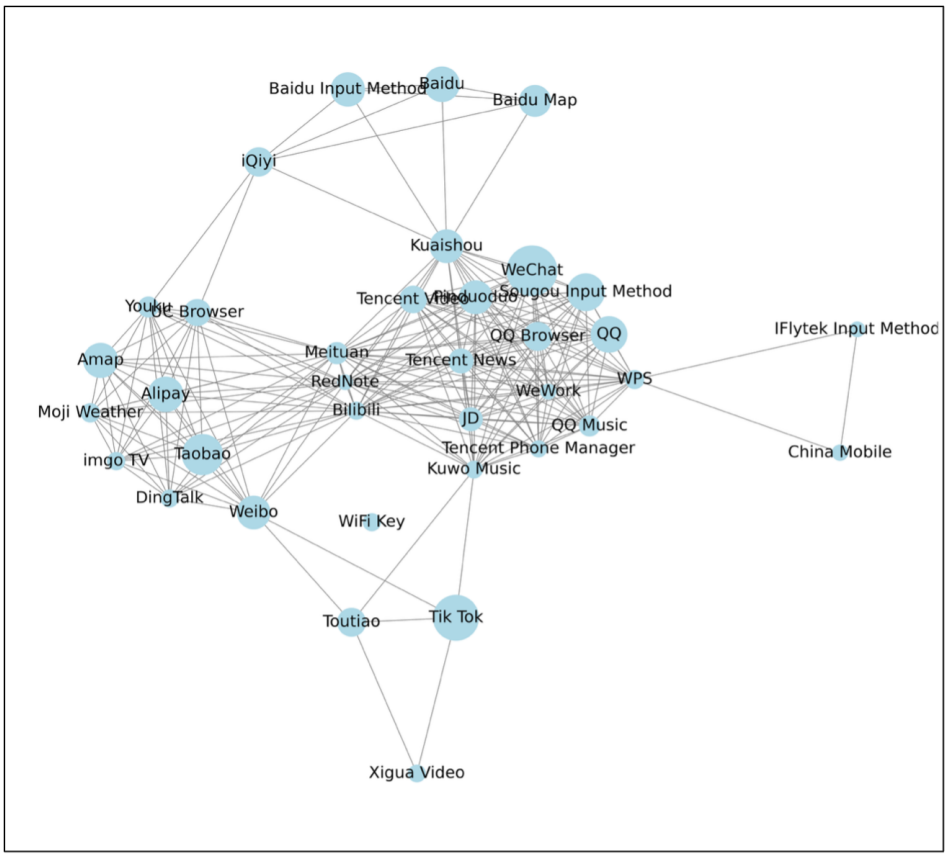
[4]See QuestMobile's *2025 China Mobile Internet Half-Year Report*, available here. These figures provide market context and refer to June 2025; our empirical sample covers 2023.

and reports, for each app, weekly active users and, for each pair of apps, weekly common users. These apps span categories ranging from communication to entertainment. Figure A.1 plots the number of active Android users, in millions, for the top 10 apps in the sample at the end of 2023. WeChat stands out as the largest app in the sample, with more than 808 million active users at the end of 2023, while AMap ranks tenth, with around 361 million active users. Over the year, the total number of mobile app users is broadly stable, increasing by only 0.3%. By contrast, WeChat's active users increased by 3.35%, faster than the overall trend.

Ownership data. The second dataset contains ownership information from Tianyancha, a business intelligence platform that provides comprehensive ownership data for registered firms in China. For each of the top 300 apps in our sample, we identify the relevant parent company and collect information on its ownership structure. In addition to common ownership, we also construct broader ownership links through investment relationships. Specifically, starting from each app's parent company, we trace major shareholders using Tianyancha's look-through ownership analysis and identify whether the app has received investment from another app's parent company. We also hand-collect the dates of these investment events so that we can identify when a cross-ownership link between two apps was established.

Common- and cross-ownership links are prevalent in the data. Among the top 300 apps in our sample, Tencent is linked to 93 apps and Alibaba to 51. Figure 1 illustrates the ownership network among the top 50 apps in 2023. The network reveals important ownership clusters, including the large Tencent cluster (in the middle), the Alibaba cluster (on the left), the Baidu cluster (at the top), and the ByteDance cluster (at the bottom).

Figure 1: Common and cross ownership network among the top 50 apps



Notes: The figure displays common- and cross-ownership relationships among the top 50 apps in the sample. Nodes correspond to apps, and the node size is proportional to their active users in 2023. Links indicate either common ownership or cross ownership through investment. The network reveals four ownership clusters, including the large Tencent cluster in the middle, the Alibaba cluster on the left, the Baidu cluster at the top, and the ByteDance cluster at the bottom.

2.2 Query: A measure of interoperability

We measure interoperability between apps through the lens of user transitions across mobile applications. Users frequently move between apps to communicate, share content, and complete tasks involving multiple services. The ease of these interactions depends on how apps support connections with one another. Higher interoperability can facilitate user flows across apps, while limited interoperability may require additional steps or less convenient alternatives. These frictions can affect app usage, complementarity, and competition.

We construct an empirical measure of interoperability based on the operating system's *query* mechanism, through which app A requests visibility of app B. This visibility enables app A to build tailored integrations and pass customized data to

app B—for instance, allowing Google Search to route users directly to Google Maps when a search query involves a location. A defining feature of this mechanism is its directionality: app A may request visibility of app B without B making a reciprocal request. Consequently, for any pair of apps, three query configurations are possible: a two-way (mutual) query, a one-way query, or no query.

Figure 2: Shared book excerpts as seen by the recipient

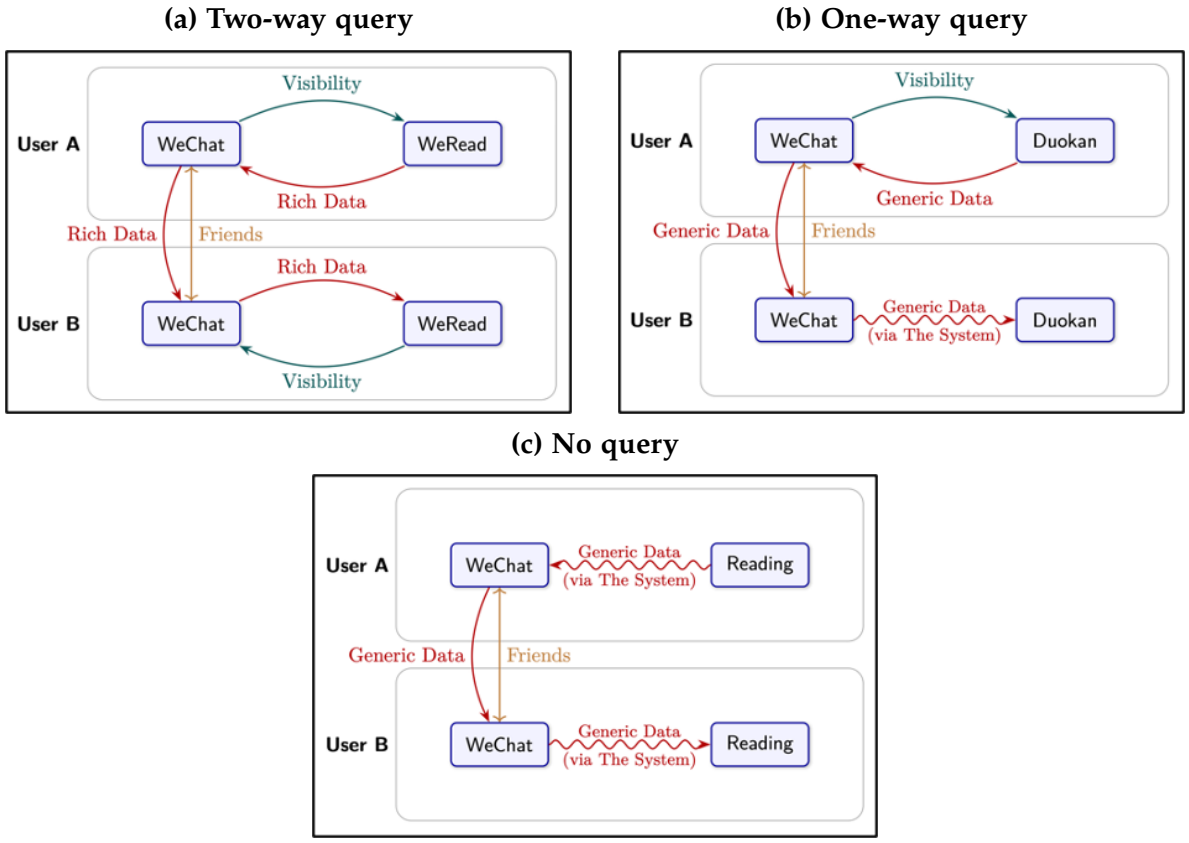
(a) WeRead **(b) Duokan**



Notes: The interactive WeRead page offers an in-app button to open WeRead. The recipient can go to the exact paragraph by tapping this button. The Duokan picture offers a QR code. After long-pressing this code, the recipient will get a link to be copied into a browser and then be directed to Duokan if this app is installed. After this long process, the recipient will see the cover of the book rather than the exact paragraph being shared.

We now present an example that illustrate how different configurations matter for sophisticated interactions and seamless transitions between apps. Consider the case of WeChat, WeRead, and Duokan. WeChat is the universally adopted superapp in China. WeRead and Duokan are two reading apps. It is natural for them to query WeChat and leverage its social network to expand usage. Besides, Tencent owns both WeChat and WeRead, but not Duokan; WeChat queries WeRead, but does not do so to Duokan. Suppose a reader shares a book excerpt with a friend on WeChat from either WeRead or from Duokan. Due the difference in their query configurations with WeChat, two experiences may substantially differ. As shown in Figure 2, a WeRead share opens an interactive page inside WeChat with a blue button to continue in the WeRead app. As a result, the recipient of a WeRead share can continue reading with a single click. In contrast, the process with Duokan is much longer and frustrating. A Duokan share displays the excerpt on a card and prompts the recipient to long-press a QR code to continue; the recipient has to copy the link from the QR code and open it in a browser and then be directed to Duokan if this app is installed. Besides, the recipient will only see the first page of the book

Figure 3: Three Query Configurations



Notes: User A is the sender and User B is the recipient. A teal arrow points from the app that is visible to the app that can discover it. Red arrows show the flow of content. Wavy arrows indicate a handoff through a generic system mechanism.

rather than the exact paragraph being shared.

More formally, the three configurations are summarized in Figure 3.

Two-way query As shown in Figure 3(a), WeRead queries WeChat on the sender's side, enabling a tailored, app-specific share. On the recipient's side, WeChat reciprocally queries WeRead. Together with the code that opens the shared passage, this reverse query enables a one-tap transition from the WeChat message directly to the exact paragraph in WeRead. Such an interactive landing experience and paragraph-level routing cannot be achieved through generic Android sharing mechanisms alone; they are only possible when a two-way query is in place.

One-way query In panel (b), Duokan queries WeChat, allowing Duokan to target WeChat directly when sharing an excerpt card; however, WeChat makes no reciprocal query for Duokan. On the recipient's side, the link embedded in the QR code still provides a pathway back to Duokan, but this transition must be intermediated

by a web browser—which retains default system visibility of all installed apps in Android. As illustrated in this workflow, browser intermediation introduces extra friction and the position of the shared paragraph is lost.

No query Panel (c) illustrates a generic reading app with no query in either direction. On the sending side, the user can still rely on Android's generic sharing mechanism to pass an image or link to other apps. However, because the reading app has no visibility into whether WeChat is installed on the device, the sender is prompted to manually select the target app (here, WeChat) from a system disambiguation dialog. On the receiving side, the recipient must likewise fall back on a web browser to follow the shared link. In this configuration, both handoffs rely entirely on general-purpose system mechanisms rather than app-specific integrations.

Taken together, these comparisons show that a two-way query facilitates seamless transitions and the exchange of customized data—both of which are essential for a high-quality user experience. Queries are also important when we are interested in foreclosure: to preferentially favor one app or discriminate against a rival, a focal app must first establish visibility of the target app using a query.

Empirically, for any ordered app pair (A, B) and week t, we define

$$q_{A\to B,t} = \begin{cases} 1 & \text{if app } A \text{ lists app } B \text{ in its query declarations at time } t, \\ 0 & \text{otherwise.} \end{cases}$$

The pair $(q_{A\to B,t}, q_{B\to A,t})$ distinguishes no query in either direction, a one-way query from A to B, a two-way query from B to A, and a two-way query. In the empirical analysis, we use both the directional indicators and their pair-level aggregation, $q_{A\to B,t} + q_{B\to A,t} \in \{0, 1, 2\}$, which counts the number of query links between two apps.

We construct these query variables from the manifests of Android application packages (APKs). To protect user privacy, Android 11 restricted the information apps can obtain about other installed apps. For apps targeting Android 11 or higher, query declarations provide the mechanism to request visibility beyond default system access, leaving a verifiable record of explicit, app-specific visibility requests. We extract these declarations from available versions of the top 200 apps in our sample to construct a weekly panel of directional query links. Because developers migrated to the Android 11 standard gradually, the absence of a declaration in earlier periods is difficult to interpret consistently across apps. Our empirical analysis therefore focuses on 2023, by which time adoption was widespread and coverage

of query declarations was substantially more complete. Although our application focuses on Android apps in China, this measurement approach generalizes to other ecosystems where developers must declare visibility requests or inter-app interaction capabilities.[5]

This measure provides an empirical basis for studying the economic consequences of interoperability and related regulations. Firms may have incentives to facilitate some connections while restricting others, potentially favoring affiliated services or disadvantaging competitors (Bourreau & Kraemer, 2022; European Commission. Joint Research Centre., 2021). Whether these choices benefit consumers depends on how connections affect their use of different app combinations. We next document the relationships between query patterns, app usage, and ownership. These relationships provide preliminary evidence on interoperability incentives and help motivate our structural framework for evaluating their implications for competition and welfare.

2.3 Interoperability and joint usage

We first examine whether interoperability is associated with greater joint usage of apps. Specifically, for each app pair (A, B) and week t, we estimate

$$\log N_{ABt} = \beta_q Query_{AB,t} + FE_{\text{cat}(A),\text{cat}(B)} + FE_{A,t} + FE_{B,t} + \varepsilon_{ABt}, \tag{2.1}$$

where N_{ABt} denotes the number of users who jointly use apps A and B, and $Query_{AB,t} = q_{A\to B,t} + q_{B\to A,t}$ counts the directional query links between them. We also estimate a specification replacing this count with separate indicators for one-way and reciprocal queries, with no queries as the omitted category. The app-category-pair fixed effects, $FE_{\text{cat}(A),\text{cat}(B)}$, absorb persistent differences in joint usage across category combinations, including differences in functional complementarity. The app-week fixed effects, $FE_{A,t}$ and $FE_{B,t}$, absorb time-varying factors specific to each app, such as changes in its popularity. The coefficient of interest, β_q, captures the association between query links and joint usage, conditional on these fixed effects.

Table 1 presents the regression results. Column (1) uses the number of directional query links between two apps, taking values zero, one, or two. Column (2)

[5] An analogous visibility mechanism exists on iOS: the `LSApplicationQueriesSchemes` key in an app's `Info.plist` specifies the URL schemes an app can query via `canOpenURL:`. These declarations could support a similar measure, subject to mapping URL schemes to specific apps and accounting for platform-specific rules.

replaces this count with separate indicators for one-way and reciprocal queries, with no queries as the omitted category. The estimates show a positive association between query links and joint usage, conditional on category-pair and app-week fixed effects. In Column (1), each additional query link is associated with approximately 1.4% higher joint usage. Column (2) reveals a nonlinear relationship: relative to pairs without queries, pairs with a one-way query have approximately 0.9% higher joint usage, while pairs with reciprocal queries have approximately 6.4% higher joint usage.

These patterns support the empirical relevance of query declarations as a measure of interoperability: apps connected by query links tend to be used together more frequently, with a substantially stronger association for reciprocal links. This association is consistent with the role of interoperability in facilitating joint usage, but may also reflect firms responding to demand for using apps together or other pair-specific factors affecting both outcomes. We will distinguish these channels in our structural framework.

Table 1: Interoperability and joint usage

	(1)	(2)
Number of queries	0.014	
	(0.002)	
Number of queries $= 1$		0.009
		(0.002)
Number of queries $= 2$		0.062
		(0.004)
R^2	0.984	0.984
Unique Apps	123	123
Observations	285,242	285,242
App-category-pair FE	Yes	Yes
App-week FE	Yes	Yes

Notes: The table reports regressions with $\log N_{ABt}$ in (2.1) as the dependent variable. Columns (1) and (2) use all app pairs. All specifications include app-category-pair fixed effects and app-week fixed effects. Robust standard errors in parentheses.

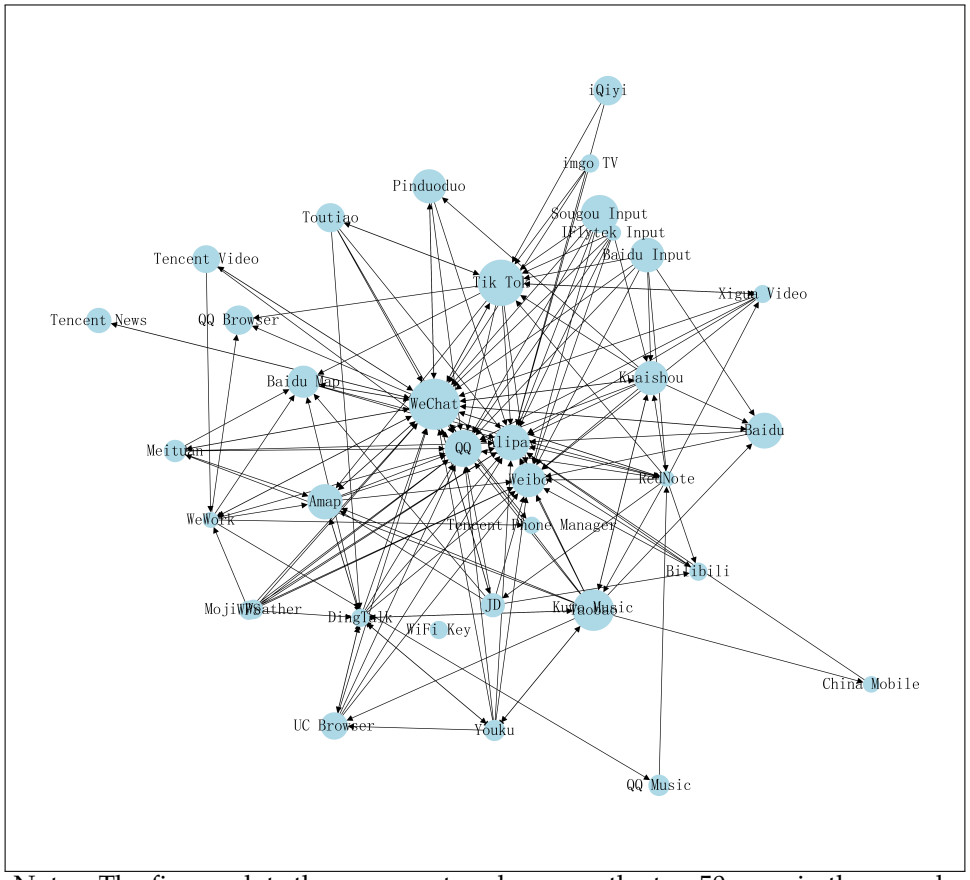
2.4 Asymmetric interoperability

Figure 4 illustrates the query network among the top 50 apps in our sample. The figure shows a clear asymmetry: leading apps, such as WeChat, receive many more incoming queries than they send outgoing queries. Smaller apps frequently request visibility of leading apps without receiving a reciprocal declaration. This pattern

suggests that directional interoperability is associated not only with functional relationships between apps but also with app popularity and market position.

To document this asymmetry more quantitatively, we use weekly active users as a proxy for an app's popularity and examine how incoming and outgoing query links depend on it. Concretely, we estimate both linear regressions of log query counts and Poisson pseudo-maximum-likelihood (PPML) specifications for query counts. The former does not include observations with zero query counts, while the latter accommodates such zeros. We also estimate a linear regression of the log incoming-to-outgoing query ratio. For each app for which we observe query declarations, we identify its outgoing query links, including those directed toward apps outside our sample. Incoming links, however, can be identified only from the declarations of other apps covered by our data. Accordingly, Columns (1)–(3) of Table 2 use links whose sender and receiver are both in our sample. Column (4) uses all observed outgoing query links for each app, regardless of the receiver.

Figure 4: Query network of top 50 apps



Notes: The figure plots the query network among the top 50 apps in the sample for the last week of 2023. Each directed edge points from the app declaring a query to the app named in that declaration. The size of each node is proportional to the number of active users of the corresponding app.

Table 2 summarizes the results. All specifications include app and week fixed effects, which absorb persistent app-specific characteristics and common weekly shocks. The PPML estimates (lower panel) show that, within an app, greater popularity is associated with more incoming queries but fewer outgoing queries to sampled apps, with estimated elasticities of approximately 0.24 and -0.21, respectively. When all observed destinations are included, the outgoing-query coefficient remains negative but is smaller in magnitude and statistically insignificant. The log-linear regressions (upper panel) exclude zero counts and therefore describe a selected sample. These estimates also show a positive association between incoming queries and popularity, while the association with outgoing queries is statistically insignificant. The log incoming-to-outgoing query ratio also increases with popularity. Together, these estimates support the asymmetry illustrated in the figure: as apps become more popular, they attract more incoming queries without a corresponding increase in outgoing queries.

These patterns suggest potentially asymmetric incentives to establish query links and facilitate interoperability between prominent apps and other apps. As an app attracts more users, facilitating interaction with it may become more valuable to other apps. The reverse connection may be less appealing to the popular app, whose incremental return from facilitating that interaction may be smaller.

Table 2: Asymmetric interoperability

	Queries between sampled apps			Queries to any app
	Incoming	Outgoing	Incoming/outgoing	Outgoing
Linear Models	(1)	(2)	(3)	(4)
Weekly active users (log)	0.2754	0.0504	0.3591	0.047
	(0.0383)	(0.0389)	(0.1170)	(0.0616)
Observations	4,527	4,379	2,727	5,502
Unique apps	136	113	103	124
R^2	0.96	0.89	0.92	0.89
Poisson Pseudo Maximum Likelihood				
Weekly active users (log)	0.2367	-0.2079	-	-0.0987
	(0.0494)	(0.0971)		(0.0754)
Observations	7,187	6,022	-	6,592
Unique apps	136	114	-	125
App FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes

Notes: The upper panel reports log-linear regressions. The dependent variables in Columns (1), (2), and (4) are the log numbers of query links. In Column (3), the dependent variable is the log number of incoming links minus the log number of outgoing links, equivalently the log incoming-to-outgoing query ratio. The lower panel reports Poisson pseudo-maximum-likelihood (PPML) estimates using query counts in levels as the dependent variables, allowing zero counts to enter the estimation. Weekly active users enter in logs in both panels. Columns (1)–(3) restrict both the sender and the receiver of each link to sampled apps. Column (4) includes all observed outgoing links from sampled apps, regardless of the receiver. Estimation samples differ across columns and panels. All specifications include app and week fixed effects. Robust standard errors are reported in parentheses.

2.5 Ownership and interoperability

We next examine how interoperability is associated with ownership links between apps. Figure 5 plots the query network using the same node positions as the ownership network in Figure 1. Apps belonging to the same firm are more likely to be connected by queries, particularly reciprocal queries. The same relationship is reflected in the comparison between within-firm and overall query densities. The overall density of the query network is 0.014, whereas the within-firm query density reaches 0.31 for Alibaba and 0.14 for ByteDance.[6] Thus, query links are about 22 times denser within Alibaba and 10 times denser within ByteDance than in the network as a whole. These patterns suggest that interoperability is concentrated not only around leading apps but also within ownership groups.

We now turn to a more systematic examination of these patterns. For each ordered app pair (A, B) and week t, we estimate

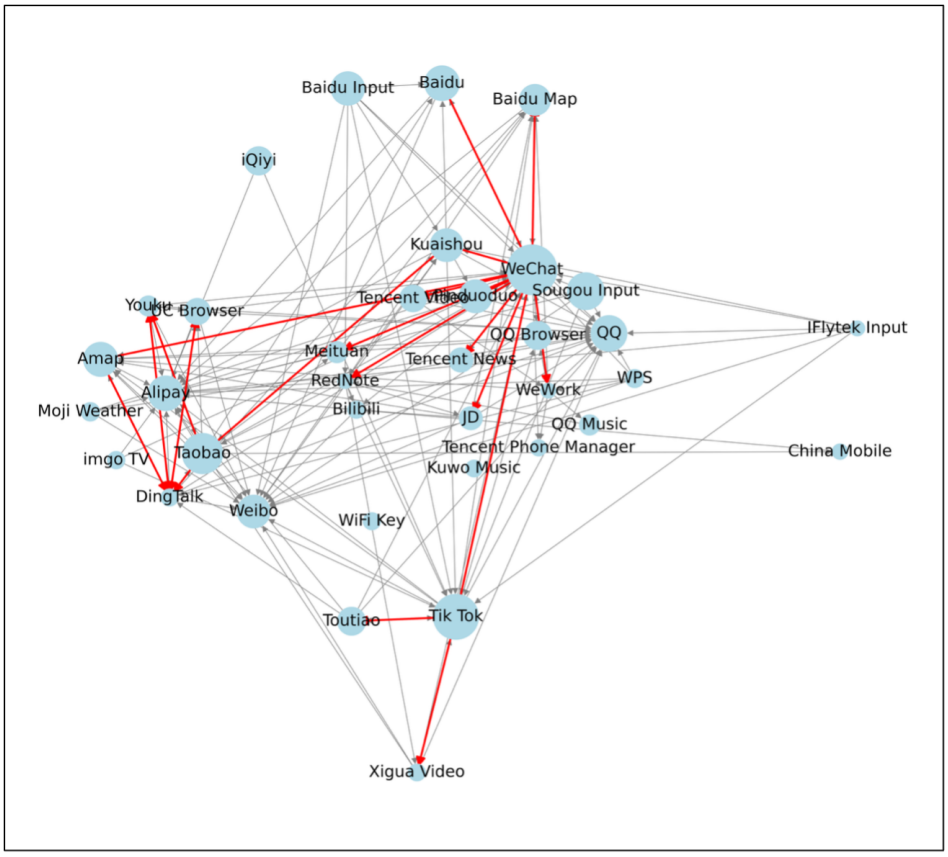
$$\begin{aligned}\mathbf{1}\{A \text{ queries } B \text{ at } t\} = {} & \gamma_A ActiveUser_{A,t} + \gamma_B ActiveUser_{B,t} \\ & + \gamma_c \mathbf{1}\{A \text{ and } B \text{ are owned by the same firm at } t\} \\ & + \alpha_A^{(o)} + \alpha_B^{(d)} + \xi_t + \Gamma_{\text{cat}(A),\text{cat}(B)} + \varepsilon_{ABt}.\end{aligned} \tag{2.2}$$

The coefficients γ_A and γ_B capture the associations between the user bases of the sending and receiving apps and query links. The coefficient of interest, γ_c, captures the association between common ownership and the probability of a query link, conditional on the controls. The sender and receiver fixed effects, $\alpha_A^{(o)}$ and $\alpha_B^{(d)}$, absorb persistent differences in apps' propensities to send and receive queries, while ξ_t captures common weekly shocks. We also include app-category-pair fixed effects, $\Gamma_{\text{cat}(A),\text{cat}(B)}$, to account for broad differences in functional complementarity across category combinations.[7] These controls are relevant because firms may assemble portfolios of functionally complementary apps, which may be more likely to establish query links independently of ownership. The specification therefore compares commonly owned and other app pairs within the same category combination.

[6] The query density within a firm is the ratio of the number of observed directional query links among apps owned by the firm to the total number of potential directional query links among those apps.

[7] App-pair fixed effects would absorb all time-invariant pair-specific determinants of query links. However, common ownership varies little within app pairs during our 2023 sample, leaving limited variation for identifying its association with query links under such a specification.

Figure 5: Queries Network



Notes: This plot displays the corresponding directional query network, where a gray edge represents a unidirectional query between apps, and a red edge denotes a bidirectional relationship. The position of each node is determined by the ownership network in Figure 1.

Table 3 confirms the visual pattern in Figure 5. The association between common ownership and query links is positive and statistically significant across specifications. Apps with a common owner are substantially more likely to be connected by a query link, even after controlling for sender, receiver, week, and app-category-pair fixed effects. The magnitude is economically large: common ownership is associated with a 7.4-percentage-point higher probability of a query link in the linear specification, with larger positive marginal effects in the Probit and Logit specifications, at 12.8 and 13.8 percentage points, respectively.

As a robustness check, we broaden the ownership-link indicator to include both common ownership and minority-shareholding links. The broader indicator equals one if the apps share a common owner or if app A's owner is a minority shareholder of app B's owner, or vice versa, at time t. Column (4) reports the results using this indicator in a linear specification. Although it is weaker than that for common ownership alone, the association remains positive and statistically significant, at 2.8

percentage points.

These patterns are relevant to regulatory concerns that firms may use interoperability to favor affiliated apps or disadvantage competitors. The concentration of query links within ownership groups is consistent with preferential connections among affiliated apps, but does not by itself establish exclusion or foreclosure. It may also reflect lower costs or greater functional complementarity among these apps. Assessing the implications for regulation therefore requires understanding both firms' incentives to provide interoperability and the value of the resulting connections to users. We next develop a structural framework that brings these elements together to evaluate how policies governing interoperability affect app usage, competition, and welfare.

Table 3: Dyadic regressions of query links on ownership

	Linear (1)	Probit (2)	Logit (3)	Linear (4)
Active User (sender)	−0.002	−0.017	−0.015	−0.002
	(0.001)	(0.004)	(0.004)	(0.001)
Active User (receiver)	0.001	0.014	0.011	0.001
	(0.001)	(0.004)	(0.004)	(0.001)
Common Ownership	0.074	0.128	0.138	
	(0.002)	(0.003)	(0.003)	
Common/Cross-ownership				0.028
				(0.001)
Observations	806,916	270,426	270,426	806,916
R^2	0.37	-	-	0.37
Sender FE	Yes	Yes	Yes	Yes
Receiver FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
App-Category-Pair FE	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy indicating if app A queries app B in week t. All specifications include sender fixed effects, receiver fixed effects, week fixed effects, and app-category-pair fixed effects. The Probit and Logit models will drop observations whose query links can be perfectly predicted by fixed effects. As a result, they have fewer observations. For Probit and Logit, we report marginal effects. Robust standard errors are reported in parentheses.

3 A Model of Digital Interoperability

Our structural framework incorporates user's demand for apps, app owners' strategic decisions on directional latent inter-app interoperability, and the query technology that maps latent interoperability into directed query links observed in app code.

3.1 Demand for Apps

Let $\mathbf{J}$ denote the set of available apps, $\mathbf{B}$ the set of app bundles, and 0 the outside option. A user can choose a single app in $\mathbf{J}$, a bundle of apps in $\mathbf{B}$, or the outside option 0. The set $\mathbf{B}$ can differ from the set of all possible combinations of apps from $\mathbf{J}$ and be its subset, e.g., $\mathbf{B}$ is the collection of bundles whose sizes are at most L. For example, if one studies interoperability between communication and short-video apps, one can write $\mathbf{J} = \mathbf{J}_c \cup \mathbf{J}_v$, where $\mathbf{J}_c$ and $\mathbf{J}_v$ denote communication and short-video apps, respectively. If $L = 4$, i.e., a user can choose a bundle of at most four different apps, then $\mathbf{B} = \cup_{l=2}^{4}\{\mathbf{b} = \{j_1, ..., j_l\} : j_k \in \mathbf{J},\ 1 \leq k \leq l;\ j_k \neq j_r \text{ for } k \neq r\}$. The outside option is interpreted as not choosing any app from the communication and short-video categories.

User i's indirect utility from using app $j \in \mathbf{J}$ at time t is

$$u_{itj} = \delta_{tj} + \beta_i x_{tj} + \varepsilon_{itj}, \tag{3.1}$$

where δ_{tj} is the population-average utility of app j at time t, β_i is a random coefficient with zero mean that captures user's potential heterogeneous preference for app characteristics x_{tj}, and ε_{itj} is an idiosyncratic error term.

Following the literature of demand for bundles, we specify user i's indirect utility from using bundle $\mathbf{b} \in \mathbf{B}$ as

$$u_{it\mathbf{b}} = \sum_{j \in \mathbf{b}} (\delta_{tj} + \beta_i x_{tj}) + \Gamma_{t\mathbf{b}} + \varepsilon_{it\mathbf{b}}, \tag{3.2}$$

where $\Gamma_{t\mathbf{b}}$ captures the additional utility from jointly using the apps in bundle $\mathbf{b}$. The indirect utility of the outside option is normalized as $u_{it0} = \varepsilon_{it0}$.

We impose a pairwise structure on bundle-level demand synergy:

$$\Gamma_{t\mathbf{b}} = \frac{1}{2} \sum_{j,j' \in \mathbf{b},\, j \neq j'} \Gamma_{tjj'}, \qquad \Gamma_{tjj'} = \Gamma_{tj'j}. \tag{3.3}$$

Thus, the additional utility from jointly using several apps is the sum of the synergies associated with each app pair contained in the bundle. This restriction serves two purposes. First, it substantially reduces the dimensionality of an otherwise unrestricted bundle-level synergy term and follows the logic of second-order specifications of bundle synergy (Gentzkow, 2007). Second, it connects the demand model directly to the structure of our usage data. As described in Section 2.1, we observe marginal usage shares for individual apps as well as pairwise joint-usage

shares. As we will see in Proposition 4.1, the pairwise specification allows to recover an app-pair-specific synergy term from these marginal and joint-usage shares.

For each app pair $\{j, j'\}$, we decompose pairwise demand synergy as

$$\Gamma_{tjj'} = \gamma_{tjj'} + \alpha_j a_{t,j\to j'} + \alpha_{j'} a_{t,j'\to j}.$$

The first component, $\gamma_{tjj'}$, captures the part of pairwise demand synergy that is unrelated to the app owners' strategic interoperability choices. It can reflect functional complementarity or substitutability between apps, app-pair-specific demand shocks, or other determinants of joint use. We allow $\gamma_{tjj'}$ to vary flexibly across app pairs and over time.

The second and third components capture the utility effect of directional interoperability. The variable $a_{t,j\to j'}$ denotes the interoperability chosen by the owner of app j toward app j', and $a_{t,j'\to j}$ denotes interoperability in the opposite direction. These choices need not be symmetric. The parameter $\alpha_j > 0$ converts the interoperability choice of app j into units of user utility. Because interoperability is latent, its scale cannot be separately identified from α_j. We therefore normalize $\alpha_j = 1$ for all $j \in \mathbf{J}$ and interpret $a_{t,j\to j'}$ directly as the utility-relevant directional interoperability index. Under this normalization, $a_{t,j\to j'} + a_{t,j'\to j}$ is the total contribution of interoperability to the demand synergy of the pair, and

$$\Gamma_{tjj'} = \gamma_{tjj'} + a_{t,j\to j'} + a_{t,j'\to j}. \tag{3.4}$$

An important feature of this decomposition is that $\gamma_{tjj'}$ is unrestricted and may be correlated with app owners' interoperability choices $a_{t,j\to j'}$ and $a_{t,j'\to j}$: they choose $a_{t,j\to j'}$ strategically after accounting for the demand, including the collection of pair-specific synergies $(\gamma_{tjj'})_{j\neq j'}$. Thus, the interoperability components in (3.4) depend on $(\gamma_{tjj'})_{j\neq j'}$ and are therefore endogenous. This is analogous to endogenous prices in standard differentiated-products demand models in which prices depend on unobserved demand shocks.

We suppose that the idiosyncratic errors $(\varepsilon_{it\mathbf{b}})_{\mathbf{b}\in\mathbf{J}\cup\mathbf{B}\cup\{0\}}$ are independently distributed with respect to β_i. Denoted by θ_e the parameters that govern the distribution of $(\varepsilon_{it\mathbf{b}})_{\mathbf{b}\in\mathbf{J}\cup\mathbf{B}\cup\{0\}}$ and the corresponding individual-level choice probability function of $\mathbf{b}$ with utility indexes $(v_{i\mathbf{b}})_{\mathbf{b}\in\mathbf{J}\cup\mathbf{B}}$ by $\sigma_{i\mathbf{b}}\left((v_{i\mathbf{b}})_{\mathbf{b}\in\mathbf{J}\cup\mathbf{B}}; \theta_e\right)$. Let $F(\cdot;\theta_F)$ denote the distribution of β_i with parameters θ_F. The demand for bundle $\mathbf{b}$ can then

be written as

$$
\begin{aligned}
s_{t\mathbf{b}} &= \Pr\left(u_{it\mathbf{b}} > u_{it\mathbf{b}'},\ \forall \mathbf{b}' \in \mathbf{J} \cup \mathbf{B} \cup \{0\},\ \mathbf{b}' \neq \mathbf{b}\right) \\
&= \int \sigma_{i\mathbf{b}} \left(\left(\sum_{j \in \mathbf{b}'} (\delta_{tj} + \beta_i x_{tj}) + \Gamma_{t\mathbf{b}'} \right)_{\mathbf{b}' \in \mathbf{J} \cup \mathbf{B}} ; \theta_e \right) dF(\beta_i; \theta_F) \\
&= \sigma_{\mathbf{b}}(\delta_t, \Gamma_t; (x_{tj})_{j \in \mathbf{J}}, \theta_D),
\end{aligned}
\tag{3.5}
$$

where $\delta_t = (\delta_{tj})_{j \in \mathbf{J}}$ collects app-level mean utilities, $\Gamma_t = (\Gamma_{tjj'})_{j<j'}$ collects pairwise demand synergies, and $\theta_D = (\theta_e, \theta_F)$. We use the convention $\Gamma_{tj} = 0$ for $j \in \mathbf{J}$ in (3.5).

The specification of θ_D depends on the assumed distribution of consumer heterogeneity. For example, one may take the distribution of $(\varepsilon_{it\mathbf{b}})_{\mathbf{b} \in \mathbf{J} \cup \mathbf{B} \cup \{0\}}$ as known. For example, these shocks are i.i.d. Type-I extreme value, integrating the resulting logit bundle-choice probabilities over the distribution of β_i yields a mixed-logit demand for bundles. Alternatively, one can impose $\beta_i = \beta$ and parameterize the distribution of the idiosyncratic shocks, such as in a nested-logit specification, with the corresponding parameters included in θ_e. For notational simplicity, we suppress the dependence of σ_t^m on $(x_{tj})_{j \in \mathbf{J}}$ below.

We do not observe the market share of every possible bundle. Instead, our usage data provide the marginal usage share of each app and the pairwise joint-usage share of each pair of apps. These objects are obtained by aggregating the bundle-level choice probabilities in (3.5). The marginal usage share of app j in market m and week t is

$$
s_{tj}^m = \sigma_{tj}^m(\delta_t, \Gamma_t; \theta_D) = \sum_{\mathbf{b} \in \mathbf{J} \cup \mathbf{B} : j \in \mathbf{b}} s_{t\mathbf{b}}. \tag{3.6}
$$

Similarly, for any pair $j < j'$, the pairwise joint-usage share is

$$
s_{tjj'}^m = \sigma_{tjj'}^m(\delta_t, \Gamma_t; \theta_D) = \sum_{\mathbf{b} \in \mathbf{B} : j, j' \in \mathbf{b}} s_{t\mathbf{b}}. \tag{3.7}
$$

Thus, s_{tj}^m includes the probability of every bundle containing app j, while $s_{tjj'}^m$ includes the probability of every bundle containing both j and j'.

3.2 Supply of Latent Interoperability

Let $f(j)$ denote the owner of app j, and let $\mathcal{J}_f$ denote the set of apps owned by f. Owner f chooses the collection of directional interoperability lev-

els $\left(a_{t,j\to j'}\right)_{j\in\mathcal{J}_f,\, j'\neq j}$ to maximize the revenue generated by its portfolio of apps net of the implementation costs of interoperability. The profit of owner f is

$$\Pi_{ft} = R_{ft}(\delta_t, \Gamma_t; \theta_D) - \sum_{j\in\mathcal{J}_f} C_j((a_{t,j\to j'})_{j'\neq j}; \theta_C), \tag{3.8}$$

where $R_{ft}(\delta_t, \Gamma_t; \theta_D)$ denotes the revenue generated by the apps owned by firm f, and $C_j((a_{t,j\to j'})_{j'\neq j}; \theta_C)$ is the implementation cost associated with interoperability supplied from app j to other apps. Vector θ_C collects the cost parameters. The total interoperability implementation cost for owner f is additive across apps. In contrast, within the same app $j \in \mathcal{J}_f$, the cost can be interdependent across app pairs. A leading example of revenue specification is advertising-based income: revenue is proportional to the total usage of the apps in the firm's portfolio. Normalizing revenue per unit of usage to one, we write

$$R_{ft}(\delta_t, \Gamma_t; \theta_D) = \sum_{k\in\mathcal{J}_f} \sigma^m_{tk}(\delta_t, \Gamma_t; \theta_D). \tag{3.9}$$

An app owner values interoperability insofar as it changes the total usage of the apps that it owns. Importantly, the owner internalizes the effect of interoperability on all apps in its portfolio, rather than only on the sender app. We discuss other revenue specifications in Appendix F, e.g., subscription-based revenue.

Given the interoperability choices of other app owners, owner f solves

$$\left(a^*_{t,j\to j'}\right)_{j\in\mathcal{J}_f,\, j'\neq j} = \arg \max_{\left(a_{t,j\to j'}\right)_{j\in\mathcal{J}_f,\, j'\neq j}} \Pi_{ft}. \tag{3.10}$$

Define the marginal revenue of a directional interoperability choice from app j to app j' as

$$MR_{t,j\to j'} \equiv \frac{\partial R_{f(j)t}(\delta_t, \Gamma_t; \theta_D)}{\partial \Gamma_{tjj'}}. \tag{3.11}$$

From (3.4), we have $\partial \Gamma_{tjj'} / \partial a_{t,j\to j'} = 1$. Therefore, at an interior optimum, the first-order condition of (3.10) for the directional choice $a_{t,j\to j'}$ is

$$\frac{\partial R_{f(j)t}(\delta_t, \Gamma_t; \theta_D)}{\partial \Gamma_{tjj'}} = \partial_{a_{j\to j'}} C_j\left((a^*_{t,j\to r})_{r\neq j}; \theta_C\right). \tag{3.12}$$

Equation (3.12) summarizes the basic trade-off governing the supply of interoperability. Higher marginal revenue from increasing pairwise demand synergy induces the owner to supply more interoperability, while implementation costs discipline its

level. Because $MR_{t,j\to j'}$ is computed from the revenue generated by the entire portfolio owned by $f(j)$, ownership affects interoperability incentives through the extent to which the firm internalizes usage gains and losses across its apps.

3.3 Query Technology

Latent interoperability $a^*_{t,j\to j'}$ is not directly observed in the data. Instead, we observe whether app j declares a query toward app j'. As discussed in Section 2.2, such declarations provide an imperfect measure of underlying directional interoperability. Conceptually, supporting richer interactions with another app can make explicit visibility of that app more valuable, while more limited interactions may be implemented without a query declaration. This motivates a positive relationship between the level of interoperability and the likelihood of observing a query. We capture this relationship through a threshold-crossing rule: a query is declared when latent interoperability exceeds a threshold, which reflects differences in the technical requirements of supporting interactions across apps.

Specifically, an observed query from app j to app j' is generated according to

$$q_{t,j\to j'} = \mathbf{1}\left\{a^*_{t,j\to j'} > \tau_{t,j\to j'}\right\}, \tag{3.13}$$

where $a^*_{t,j\to j'}$ is the equilibrium level of latent interoperability chosen by the owner of app j, and $\tau_{t,j\to j'}$ is a threshold associated with implementing interoperability from app j to app j' (unobserved by the econometrician). Thus, a higher level of latent interoperability increases the probability that a query link is observed.

As an example, consider the setting with communication and short-video apps. One can allow the mean query threshold to vary across directed category pairs:

$$\tau_{t,j\to j'} = \mu_{\text{cat}(j),\text{cat}(j')} + \nu_{t,j\to j'}, \qquad \nu_{t,j\to j'} \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma^2_\tau). \tag{3.14}$$

The parameter $\mu_{\text{cat}(j),\text{cat}(j')}$ captures systematic differences in the threshold for generating a query across directed category pairs, while $\nu_{t,j\to j'}$ captures unobserved heterogeneity in the threshold across app pairs and markets. In addition, suppose that the threshold shocks $(\nu_{t,j\to j'})_{j,j'}$ are independent of the demand and supply shocks that determine the equilibrium interoperability choices $(a^*_{t,j\to j'})_{j,j'}$. Combining (3.13) and (3.14), the probability of observing a query is

$$\Pr\left(q_{t,j\to j'} = 1 \mid a^*_{t,j\to j'}\right) = \Phi\left(\frac{a^*_{t,j\to j'} - \mu_{\text{cat}(j),\text{cat}(j')}}{\sigma_\tau}\right), \tag{3.15}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. Query declarations thus provide an imperfect measure of latent interoperability: higher interoperability makes a query more likely, while heterogeneous thresholds allow the same interoperability level to generate different query outcomes across app pairs and markets.

4 Identification and Estimation

We propose an identification and estimation strategy for the demand, cost, and query technology parameters. This strategy relies on a demand invertibility result of the observed marginal and pairwise joint-usage shares, and combines it with the first-order conditions of app owners' optimal supply of interoperability and the query technology to construct a likelihood approach. The estimates of the parameters of interest are then obtained via the maximum likelihood approach.

4.1 Identification

Demand inverse. Let $\sigma_t^m(\cdot;\theta_D)$ denote the demand mapping from the app-level mean utilities and pairwise synergies into the marginal and pairwise usage shares:

$$\left((s_{tj}^m)_{j\in\mathbf{J}}, (s_{tjj'}^m)_{j<j'}\right) = \sigma_t^m(\delta_t, \Gamma_t; (x_{tj})_{j\in\mathbf{J}}, \theta_D). \tag{4.1}$$

Recall that $\delta_t = (\delta_{tj})_{j\in\mathbf{J}}$ and $\Gamma_t = (\Gamma_{tjj'})_{j<j'}$. The next proposition characterizes conditions under which one can invert σ_t^m to recover (δ_t, Γ_t) from the corresponding marginal and pairwise usage shares. The proof is included in Appendix B.

Proposition 4.1 (Demand inverse)**.** *Suppose that* $\mathbf{B}$ *contains every two-app bundle* $\{j, j'\}$ *with* $j < j'$ *and* $j, j' \in \mathbf{J}$*. Moreover, the following conditions hold:*

1. *The domain of* (δ_t, Γ_t) *is convex.*
2. *For every admissible* (δ_t, Γ_t)*, the Jacobian* $\frac{\partial\left(\sigma_{i\mathbf{b}}\left((v_{i\mathbf{b}'})_{\mathbf{b}'\in\mathbf{J}\cup\mathbf{B}};\theta_e\right)\right)_{\mathbf{b}\in\mathbf{J}\cup\mathbf{B}}}{\partial\left(v_{i\mathbf{b}'}\right)_{\mathbf{b}'\in\mathbf{J}\cup\mathbf{B}}}$ *is positive definite for a set of* β_i *with positive probability under* $F(\cdot;\theta_F)$ *where* $v_{i\mathbf{b}'} = \sum_{j\in\mathbf{b}'}\left(\delta_{tj} + x_{tj}\beta_i\right) + \Gamma_{t\mathbf{b}'}$.

Then $\sigma_t^m(\cdot; (x_{tj})_{j\in\mathbf{J}}, \theta_D)$ *is one-to-one in* (δ_t, Γ_t).

Proposition 4.1 is related to the invertibility argument in Iaria & Wang (2020) and Wang (2021) that uses product-level marginal shares to back out the mean utilities of products. Differently, the invertibility argument in Proposition 4.1 additionally uses the app-pair-level joint usage shares that enable to recover the corresponding demand synergies. Technically, Condition 1 is the counterpart of the convex-domain condition in Wang (2021)'s Theorem 1, applied to (δ_t, Γ_t). Condition 2 imposes the same positive-definiteness requirement as his Theorem 1 on the individual bundle-demand Jacobian for a set of random coefficients with positive probability. Under random utility maximization, this Jacobian is symmetric positive semidefinite. Condition 2 ensures that its population average is positive definite, which allows to apply the invertibility arguments in Gale & Nikaido (1965) and accommodate complementarity among products.

Using Proposition 4.1, we can express $\left(\delta_t, \left(\Gamma_{tjj'}\right)_{j<j'}\right)$ as

$$\left(\delta_t, \left(\Gamma_{tjj'}\right)_{j<j'}\right) = \left(\sigma_t^m\right)^{-1}\left(\left(s_{tj}^m\right)_{j\in\mathbf{J}}, \left(s_{tjj'}^m\right)_{j<j'}; \left(x_{tj}\right)_{j\in\mathbf{J}}, \theta_D\right). \tag{4.2}$$

In other words, we can write (δ_t, Γ_t) as functions of the demand parameters θ_D, given the observed app-level and app-pair-level usage shares and app characteristics.

Cost assumption and optimal supply of interoperability. Write the first-order conditions for the optimal supply of interoperability in (3.12), for $\left(a_{t,j\to j'}\right)_{j'\neq j}$, as

$$\left(\frac{\partial R_{f(j)t}(\delta_t, \Gamma_t; \theta_D)}{\partial \Gamma_{tjj'}}\right)_{j'\neq j} = \left(\partial_{a_{j\to j'}} C_j\left(\left(a^*_{t,j\to r}\right)_{r\neq j}; \theta_C\right)\right)_{j'\neq j}. \tag{4.3}$$

Because of (4.2), the marginal revenues on the left-hand side of (4.3) are known up to θ_D. The following assumption on the cost function allows us to recover the equilibrium interoperability choices $\left(a^*_{t,j\to r}\right)_{r\neq j}$ from these marginal revenues given the cost parameters θ_C.

Assumption 1 (Convex cost). *The cost $C_j\left(\left(a_{j\to r}\right)_{r\neq j}; \theta_C\right)$ is strictly increasing in $a_{j\to r}$ for each (j, r) with $r \neq j$, and strictly convex in $\left(a_{j\to r}\right)_{r\neq j}$.*

Assumption 1 is motivated by the institutional feature that building and sustaining richer interactions between apps entails increasing marginal costs (Leung & Fan, 2002). Appendix E illustrates this mechanism using a piecewise-linear cost function linked to the threshold-based query technology in (3.13). In this specification, higher levels of interoperability, associated with a query declaration and richer

inter-app functionality, entail a higher constant marginal cost than lower levels that can be implemented without a query declaration. In addition, we show in Appendix D that under identical, strictly convex link-specific costs, a lower marginal revenue from interoperability in one direction implies a lower equilibrium level of interoperability in that direction. When dominant apps face lower marginal returns from connecting to other apps than those apps face from connecting to them, this mechanism can rationalize the directional query asymmetry documented in Section 2.4. This provides a theoretical motivation for Assumption 1.

A leading example of a convex cost specification satisfying Assumption 1 is

$$C_j\left((a_{t,j\to j'})_{j'\neq j};\theta_C\right)=\sum_{j'\neq j} r_{jj'}c_j(a_{t,j\to j'};\theta_C), \tag{4.4}$$

where $c_j(\cdot)$ is strictly increasing and convex in latent interoperability a and $r_{jj'}>0$ is a constant. Under this specification, the cost of implementing interoperability is separably additive across directional app pairs.

Assumption 1 implies that the Jacobian of the left-hand side of (4.3) is positive definite. According to Gale & Nikaido (1965) (Theorem 6), $\left(\partial_{a_{j\to j'}}C_j\left((a_{t,j\to j'})_{j'\neq j};\theta_C\right)\right)_{j'\neq j}$ is invertible in $(a_{t,j\to r})_{r\neq j}$ in the relevant domain. Consequently, from (4.3) and (4.2), we recover $(a^*_{t,j\to r})_{r\neq j}$ as:

$$\begin{aligned}(a^*_{t,j\to r})_{r\neq j}&=(\partial_{a_{j\to j'}}C_j)^{-1}_{j'\neq j}\left(\left(\frac{\partial R_{f(j)t}(\delta_t,\Gamma_t;\theta_D)}{\partial\Gamma_{tjj'}}\right)_{j'\neq j};\theta_C\right)\\&=(\partial_{a_{j\to j'}}C_j)^{-1}_{j'\neq j}\left(\left(\frac{\partial R_{f(j)t}\left((\sigma^m_t)^{-1}\left((s^m_{tj})_{j\in\mathbf{J}},(s^m_{tjj'})_{j<j'};(x_{tj})_{j\in\mathbf{J}},\theta_D\right);\theta_D\right)}{\partial\Gamma_{tjj'}}\right)_{j'\neq j};\theta_C\right)\end{aligned} \tag{4.5}$$

where $(\partial_{a_{j\to j'}}C_j)^{-1}_{j'\neq j}$ is the inverse function of $(\partial_{a_{j\to j'}}C_j)_{j'\neq j}$.

Likelihood construction. The optimality conditions above recover latent interoperability $(a^*_{t,j\to j'})_{j\neq j'}$ as a function of the demand and cost parameters. If $a^*_{t,j\to j'}$ were directly observed, estimation could proceed by directly matching the observed interoperability choices to their model-implied counterparts on the right-hand side of (4.5). In our setting, however, latent interoperability is not directly observed. Instead, we observe the query indicator $q_{t,j\to j'}$ that provides an imperfect measure of $a^*_{t,j\to j'}$ described in (3.13). We next impose conditions on the distribution of the query threshold $\tau_{t,j\to j'}$ in (3.13) that, together with (4.5), give rise to the likelihood

used for estimation.

Assumption 2 (Query technology). *Conditional on* $(x_{tj})_{j\in\mathbf{J}}$, $\tau_{t,j\to j'}$ *is independent of* $(\delta_t, (\gamma_{tjj'})_{j\neq j'})$ *and independent across directed app pairs.*

Assumption 2 imposes conditional independence between the unobserved threshold for generating a query link and the demand shocks that determine equilibrium latent interoperability. It also imposes the independence across directed app pairs, which allows the likelihood to be constructed from the individual pair-level query-link probabilities.

Denote by $G(\cdot; (x_{tj})_{j\in\mathbf{J}}, \theta_\tau)$ the conditional distribution function of $\tau_{t,j\to j'}$ in (3.13), where θ_τ collects the parameters governing this distribution. Combining (3.13), (4.5), and Assumption 2, we obtain

$$
\begin{aligned}
&\Pr\left(q_{t,j\to j'}=1 \,\middle|\, (x_{tj})_{j\in\mathbf{J}}, (s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; \theta_D, \theta_C, \theta_\tau\right) \\
=&G\left(a^*_{t,j\to j'}; (x_{tj})_{j\in\mathbf{J}}, \theta_\tau\right) \\
=&G\left(\left[\left(\partial_{a_{j\to r}} C_j\right)^{-1}_{r\neq j}\left(\left(\frac{\partial R_{f(j)t}\left((\sigma^m_t)^{-1}\left((s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; x_t, \theta_D\right); \theta_D\right)}{\partial \Gamma_{tjr}}\right)_{r\neq j}; \theta_C\right)\right]_{j'}; (x_{tj})_{j\in\mathbf{J}}, \theta_\tau\right),
\end{aligned}
\tag{4.6}
$$

where $[\cdot]_{j'}$ denotes the component corresponding to the directional choice from app j to app j'.

We now shed light on the identification intuition based on (4.6).

Example 1 (Identification). *Suppose that the pair-specific implementation cost in* (4.4) *takes the exponential form*

$$
c_j(a_{jj'}; \kappa, \rho) = \frac{\kappa_{jj'}}{\rho}\exp(\rho a_{jj'}), \qquad \kappa_{jj'} > 0, \quad \rho > 0,
$$

so that $\partial_a c_j(a; \kappa, \rho) = \kappa_{jj'}\exp(\rho a)$. *The parameter* ρ *controls the common curvature of the implementation cost. The scale parameter* $\kappa_{jj'}$ *captures potential heterogeneity in cost across app pairs, e.g., the cost of implementing interoperability between two apps of the same owner can be lower than that between a pair of apps of different owners. Note that* $\kappa_{jj'}$ *is not separately identified from the level of latent interoperability and the residual pairwise demand synergy. To see this, the first-order condition implies* $a^*_{t,j\to j'} = \left[\ln MR_{t,j\to j'} - \ln\kappa_{jj'}\right]/\rho$. *Hence, normalizing* $\kappa_{jj'} = 1$ *amounts to redefining* $\widetilde{a}^*_{t,j\to j'} = a^*_{t,j\to j'} + \ln\kappa_{jj'}/\rho$. *The corresponding residual pairwise synergy is correspondingly redefined as* $\widetilde{\gamma}_{tjj'} = \gamma_{tjj'} - 2\ln\kappa_{jj'}/\rho$, *so that the total pairwise synergy* $\Gamma_{tjj'}$ *and the*

first-order conditions remain unchanged. We therefore normalize $\kappa_{jj'} = 1$ below.

Suppose further that the query threshold is normally distributed, $\tau_{t,j\to j'} \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma_\tau^2)$. With $\kappa_{jj'} = 1$, the supply-side first-order condition gives

$$a^*_{t,j\to j'} = \frac{1}{\rho} \ln \left(\frac{\partial R_{f(j)t}((\sigma_t^m)^{-1} \left((s_{tj}^m)_{j\in \mathbf{J}}, (s_{tjj'}^m)_{j<j'}; (x_{tj})_{j\in \mathbf{J}}, \theta_D \right); \theta_D)}{\partial \Gamma_{tjj'}} \right). \tag{4.7}$$

Combining (4.7) with the query technology yields

$$\begin{aligned} &\Pr \left(q_{t,j\to j'} = 1 \,\middle|\, (x_{tj})_{j\in \mathbf{J}}, (s_{tj}^m)_{j\in \mathbf{J}}, (s_{tjj'}^m)_{j<j'} \right) \\ =&\Phi \left[\frac{1}{\sigma_\tau \rho} \ln \left(\frac{\partial R_{f(j)t}((\sigma_t^m)^{-1} \left((s_{tj}^m)_{j\in \mathbf{J}}, (s_{tjj'}^m)_{j<j'}; (x_{tj})_{j\in \mathbf{J}}, \theta_D \right); \theta_D)}{\partial \Gamma_{tjj'}} \right) - \frac{\mu}{\sigma_\tau} \right]. \end{aligned} \tag{4.8}$$

The identification intuition behind (4.8) is akin to that of nonlinear least squares. Concretely, the propensity to observe a query link depends on the app-pair-specific index

$$\frac{1}{\sigma_\tau \rho} \ln \left(\frac{\partial R_{f(j)t}((\sigma_t^m)^{-1} \left((s_{tj}^m)_{j\in \mathbf{J}}, (s_{tjj'}^m)_{j<j'}; (x_{tj})_{j\in \mathbf{J}}, \theta_D \right); \theta_D)}{\partial \Gamma_{tjj'}} \right) - \frac{\mu}{\sigma_\tau}.$$

The demand parameters θ_D enter the index only through the demand inverse and hence through the marginal revenue from interoperability. Consequently, the variation in the marginal revenue provides useful information for identifying θ_D. Such variation could originate from the cross-sectional dimension, i.e., across app pair (j, j'), for a given t, or from the time dimension, i.e., the change in observed marginal and pairwise usage shares for a given (j, j').

The effect of such variation on the index is compounded by $1/(\sigma_\tau \rho)$, the coefficient on log marginal revenue. Under suitable regularity conditions, by relying on the nonlinearity of the log marginal revenue with respect to θ_D, one can separately identify the effect of θ_D on the index from that of $\sigma_\tau \rho$, and hence their identification. Finally, once θ_D and $\sigma_\tau \rho$ are identified, the remaining μ/σ_τ can be identified as the constant term in the index.[8]

This likelihood approach based on (4.8) alone, however, does not separately identify ρ and σ_τ. We solve this problem by an additional instrumental variable step. To see this, we

[8]Note that one can also adapt this identification analysis to accommodate heterogeneity in the mean query threshold, for example by replacing μ with $\mu_{\text{cat}(j),\text{cat}(j')}$ for each directed category pair.

use $\Gamma_{tjj'} = a^*_{t,j\to j'} + a^*_{t,j\to j'} + \gamma_{tjj'}$ *and* (4.7) *and obtain the following linear regression:*

$$\begin{aligned}\Gamma_{tjj'} &= \sigma_\tau \times \frac{a^*_{t,j\to j'} + a^*_{t,j'\to j}}{\sigma_\tau} + \gamma_{tjj'} \\ &= \sigma_\tau \times \frac{1}{\sigma_\tau \rho}\left[\ln\left(\frac{\partial R_{f(j)t}((\sigma^m_t)^{-1}\left((s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; (x_{tj})_{j\in\mathbf{J}}, \theta_D\right); \theta_D)}{\partial \Gamma_{tjj'}}\right)\right. \\ &\left. + \ln\left(\frac{\partial R_{f(j')t}((\sigma^m_t)^{-1}\left((s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; (x_{tj})_{j\in\mathbf{J}}, \theta_D\right); \theta_D)}{\partial \Gamma_{tjj'}}\right)\right] + \gamma_{tjj'}.\end{aligned} \tag{4.9}$$

Given the identification of θ_D*, the left-hand side* $(\Gamma_{tjj'})_{t,j\neq j'}$ *is then identified from the demand inverse* (4.2)*. Moreover, the regressor for* σ_τ *is identified via the likelihood approach. This regressor is generally endogenous: the marginal revenues are functions of the equilibrium demand synergies and hence may be correlated with the unobserved component* $\gamma_{tjj'}$*.*

To address this endogeneity issue, one can consider excluded variables that shift app-level mean utilities δ_t *and hence marginal revenues, but are uncorrelated with* $\gamma_{tjj'}$*, conditional on the included controls. One candidate is the number of non-exported activities in apps* j *and* j' *at time* t*, potentially combined through their sum or product. An activity represents a focused task that a user can perform within an app.*[9] *The number of non-exported activities can therefore proxy for the scope of an app's internal functionality, providing a source of variation in its standalone utility and the marginal revenue from interoperability. Unlike exported activities, which can provide entry points for interactions initiated by other apps, non-exported activities are not accessible to other apps. This distinction motivates the exclusion restriction: conditional on the included controls, variation in non-exported activity counts shifts standalone app utilities without directly affecting, or being correlated with, the residual pair-specific synergy* $\gamma_{tjj'}$*.*

Once σ_τ *is identified, we recover* $\rho = \frac{\sigma_\tau \rho}{\sigma_\tau}$*, and* $\mu = \sigma_\tau\left(\frac{\mu}{\sigma_\tau}\right)$*. We can then recover directional latent interoperability from* (4.7) *and, finally, the residual pairwise demand synergy* $\gamma_{tjj'} = \Gamma_{tjj'} - a^*_{t,j\to j'} - a^*_{t,j'\to j}$*. Thus, the demand inverse, the query likelihood, and the additional exclusion restriction together recover the objects in demand, supply, and query technology that are required for the counterfactual analysis.*

4.2 Estimation

Let $\mathcal{D}_t$ denote the set of directed app pairs $j \to j'$ for which the query outcome is observed at time t. Conditional on the observed app characteristics and usage

[9] See the Android documentation.

shares, the log-likelihood contribution of market t is

$$\sum_{(j,j')\in\mathcal{D}_t} \Big\{ q_{t,j\to j'} \log \Pr\left(q_{t,j\to j'} = 1 \,\middle|\, (x_{tj})_{j\in\mathbf{J}}, (s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; \theta_D, \theta_C, \theta_\tau \right) + (1 - q_{t,j\to j'}) \log \left[1 - \Pr\left(q_{t,j\to j'} = 1 \,\middle|\, (x_{tj})_{j\in\mathbf{J}}, (s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; \theta_D, \theta_C, \theta_\tau \right) \right] \Big\}$$

The finite-sample log-likelihood is therefore

$$\ell_T(\theta_D, \theta_C, \theta_\tau) = \sum_{(j,j')\in\mathcal{D}_t} \Big\{ q_{t,j\to j'} \log \Pr\left(q_{t,j\to j'} = 1 \,\middle|\, (x_{tj})_{j\in\mathbf{J}}, (s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; \theta_D, \theta_C, \theta_\tau \right) + (1 - q_{t,j\to j'}) \log \left[1 - \Pr\left(q_{t,j\to j'} = 1 \,\middle|\, (x_{tj})_{j\in\mathbf{J}}, (s^m_{tj})_{j\in\mathbf{J}}, (s^m_{tjj'})_{j<j'}; \theta_D, \theta_C, \theta_\tau \right) \right] \Big\}.$$

The maximum likelihood estimator is defined as

$$\left(\widehat{\theta}_D, \widehat{\theta}_C, \widehat{\theta}_\tau \right) \in \arg \max_{(\theta_D, \theta_C, \theta_\tau)\in\Theta_D\times\Theta_C\times\Theta_\tau} \ell_T(\theta_D, \theta_C, \theta_\tau). \tag{4.10}$$

In some parametric specifications, e.g., Example 1, the likelihood depends on the primitive cost and query-technology parameters only through particular combinations. In that case, (4.10) should be understood as maximizing the likelihood with respect to the identified reparameterization of (θ_C, θ_τ). The remaining scale parameters can then be recovered using the additional restrictions such as the instrumental variable regression described in Example 1.

5 Empirical Application

We apply the model to the apps market in China and implement (4.10). We further implement an interoperability mandate counterfactual simulation and investigate its implications on the interoperability across mobiles apps and social welfare. We focus on 13 apps in the communication and online-video categories: 4 communication apps and 9 online-video apps. For each app and app pair, we observe weekly active users over 53 weeks in 2023, from which we construct app-level marginal usage shares and app-pair joint-usage shares. We also observe weekly directional query links between app pairs.

We let $\mathbf{J}$ contain the 13 apps in our sample and construct $\mathbf{B}$ to include all multi-app bundles containing up to L distinct apps. The outside option corresponds to using none of the communication or online-video apps in our sample. In practice, the maximum bundle size L is not directly observed. Nevertheless, using the ob-

served app-level marginal shares and app-pair joint-usage shares, we construct a diagnostic that provides a necessary condition for a given maximum bundle size to rationalize these observed shares. We choose $L = 6$, the smallest value that satisfies this necessary condition in all markets. Together with the individual-app choices, this gives 4,095 nonempty bundles. We describe the construction of this diagnostic in Appendix C.1.

We allow consumer heterogeneity in the utility of individual apps according to

$$\delta_{itj} = \delta_{tj} + \eta_{itj}, \qquad \eta_{itj} \sim \mathcal{N}\left(0, \sigma^2_{\text{cat}(j)}\right).$$

The dispersion of consumer tastes is allowed to differ between the communication and online-video categories. The demand parameters to be estimated are therefore $\theta_D = (\sigma_{\text{comm}}, \sigma_{\text{online}})$.

We adopt the separably additive cost specification in (4.4) and parameterize its pair-specific component along the lines of Example 1

$$c_j\left(a; \rho_{f(j)}\right) = \frac{1}{\rho_{f(j)}} \exp\left(\rho_{f(j)} a\right), \qquad \rho_{f(j)} > 0.$$

We allow the curvature parameter $\rho_{f(j)}$ to vary across app owners. We specify the mean query threshold to depend on directed category pairs:

$$\tau_{t,j\to j'} = \mu_{\text{cat}(j),\text{cat}(j')} + \nu_{t,j\to j'}, \qquad \nu_{t,j\to j'} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2_\tau).$$

Define $\lambda_f \equiv 1/\sigma_\tau \rho_f$ and $\mu^N_{cc'} \equiv \mu_{cc'}/\sigma_\tau$. As Example 1, the likelihood approach directly identifies the parameter vector $\left(\sigma_{\text{comm}}, \sigma_{\text{online}}, (\lambda_f)_f, (\mu^N_{cc'})_{c,c'}\right)$. In our specification, there are 8 app owners and 4 directed category-pair thresholds, so that the likelihood contains 14 parameters: 2 demand heterogeneity parameters, 8 owner-specific λ_f's, and 4 normalized query-threshold parameters. Finally, we follow the IV regression approach (4.9) to further estimate σ_τ. We use the sum of the log of the numbers of non-exported activities of apps j and j' at time t as the instrument. Besides, we include app-pair and time fixed effects to further alleviate the endogeneity in (4.9).

5.1 Results

We summarize the MLE and post-MLE IV regression results in Table 4. We compare the results from the multinomial and mixed-logit specifications of demand.

The top panel summarizes the demand estimates, while the middle panel reports the estimates of query technology. The scale parameter σ_τ is estimated using the post-MLE IV approach. Its first stage achieves an F-statistic of 201.14 and a partial R^2 of 0.1254 for the scenario of the mixed-logit demand. We also compare the IV regression with an OLS regression that does not account for the endogeneity of the regressor in (4.9). The results are summarized in Table C.1. We additionally report the estimated mean query thresholds $\mu_{cc'}$.

On the demand side, the estimates reveal substantial heterogeneity in demand for apps. The standard deviations $\sigma_{\text{Communication}}$ and $\sigma_{\text{Online Video}}$ are estimated significantly different from zero. In particular, the estimated standard deviation of heterogeneous tastes is much larger for communication apps than for online-video apps, suggesting that accounting for heterogeneity is important for communication-app demand. Relative to the multinomial-logit specification, allowing for such unobserved heterogeneity in demand also remarkably improves model fitness, increasing the log likelihood from -1454.48 to -1321.51. The recovered pairwise demand synergies also differ across categories. Under mixed logit, the mean synergy is 5.524 among communication apps, compared with 0.739 between communication and online-video apps and 0.622 among online-video apps. In contrast, the multinomial-logit specification seems to estimate considerably more modest synergies among communication apps.

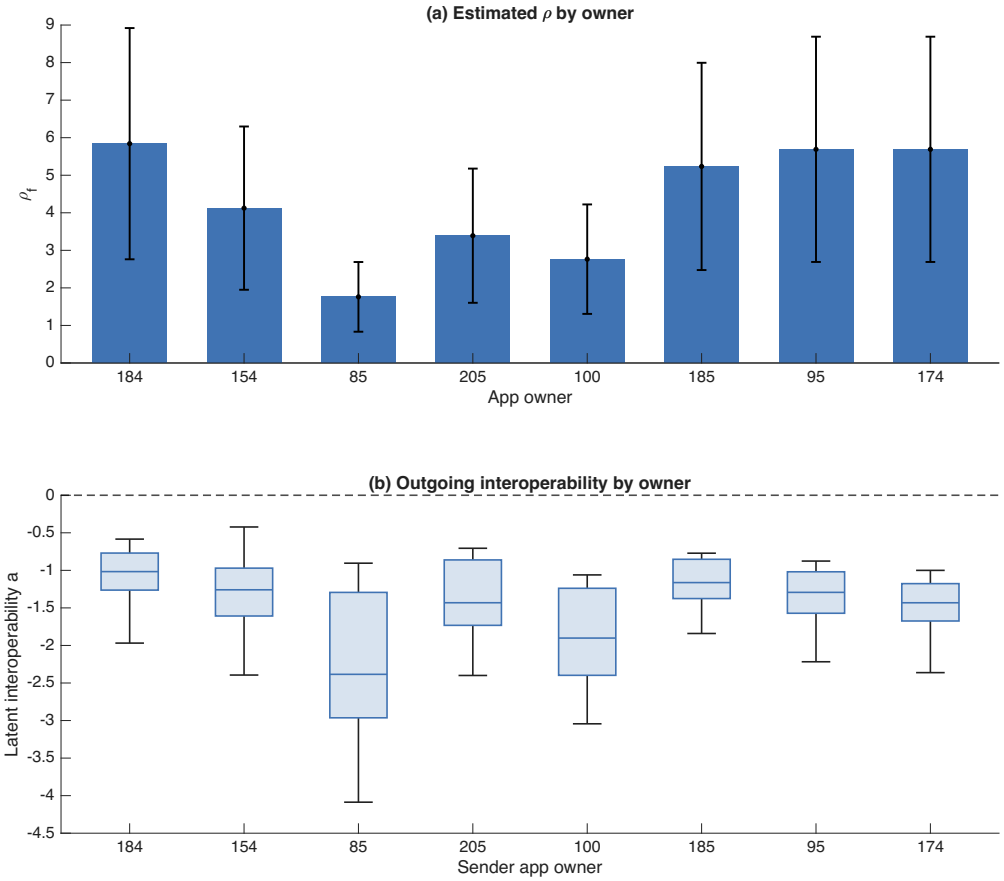
On the supply side, Figure 6(a) shows substantial differences in the estimated cost-curvature parameter ρ_f across the eight app owners. The cost of increasing interoperability therefore responds differently across owners, contributing to the heterogeneous distributions of outgoing interoperability across sender owners illustrated in Figure 6(b). Furthermore, the query-technology estimates indicate that the probability of observing a query can differ by the direction of the category pair, conditional on latent interoperability. For example, the estimated threshold mean is lower for online-video apps querying communication apps than for the reverse direction. This is consistent with the empirical pattern that non-dominant apps are more likely to query dominant ones such as WeChat that is in the communication category. Finally, the post-MLE IV estimate of σ_τ is smaller than its OLS counterpart (Table C.1), consistent with the possibility that app owners choose interoperability in response to unobserved components of pairwise demand synergy. We also give an explanation for the upward bias of the OLS estimate in Appendix C.3.

Table 4: MLE and Post-MLE Estimates

	Multinomial logit	Mixed logit
Demand parameters		
$\sigma_{\text{Communication}}$	–	24.2672
		(4.1753)
$\sigma_{\text{Online Video}}$	–	0.2256
		(0.0067)
Demand synergies $\Gamma_{tjj'}$		
Mean	0.6872	1.0240
	(0.0360)	(0.2422)
Mean, within communication	0.8372	5.5241
	(0.1241)	(2.8921)
Mean, within online video	0.6215	0.6219
	(0.0511)	(0.0511)
Mean, between communication and online video	0.7328	0.7392
	(0.0499)	(0.0502)
Query technology		
σ_τ	0.2914	0.2259
	(0.0738)	(0.1191)
Threshold means $\mu_{cc'}$		
Communication → Communication	−0.3038	−0.7644
	(0.0769)	(0.4030)
Communication → Online video	−0.1693	−0.6251
	(0.0429)	(0.3296)
Online video → Communication	−1.1199	−1.1011
	(0.2835)	(0.5805)
Online video → Online video	−0.7393	−0.7661
	(0.1871)	(0.4039)
Log likelihood	−1454.48	−1321.51
Observations	4770	4770

Notes: Standard errors are in parentheses. Standard errors for demand parameters are Hessian-based. Standard errors for mean synergies are clustered by app pair and market, conditional on the recovered synergies. Post-MLE estimates are implemented with pair and market fixed effects. Their standard errors are two-way clustered and conditional on the first-step estimates.

Figure 6: Interoperability Costs and Latent Interoperability by Owner



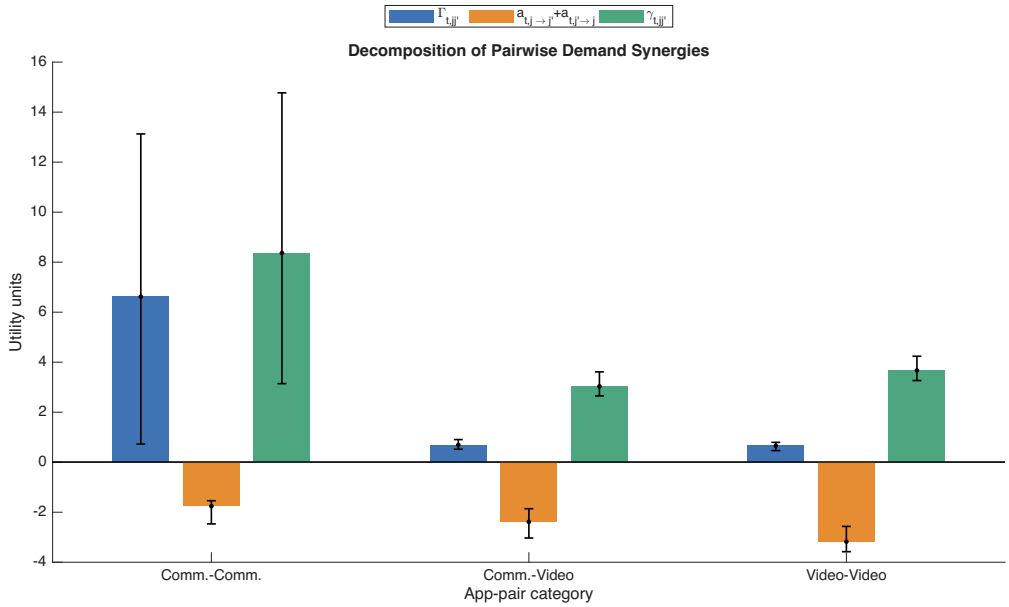
Notes: Panel (a) reports the recovered owner-specific cost-curvature parameters ρ_f, with error bars indicating one standard error. Panel (b) shows the distribution of latent interoperability $a_{t,j\to j'}$ by the owner of the sending app. Boxes represent the interquartile range, with the median indicated by the horizontal line.

Let us turn to demand synergy estimates. In Figure 7, we decompose demand synergies $\Gamma_{tjj'}$ into the strategic component $a_{t,j\to j'} + a_{t,j'\to j}$ and residual pairwise synergy $\gamma_{t,jj'}$ for category-pair groups. The median demand synergy is positive across the category-pair groups and estimated stronger among communication apps. This is mainly driven by the residual component $\gamma_{tjj'}$, whose median is estimated positive, whereas the median $a_{t,j\to j'} + a_{t,j'\to j}$ is negative.[10] Figures 8(a) and 8(b) summarize the the distributions of residual pairwise demand synergy and the strategic component ownership, respectively. Within-owner app pairs have a slightly lower median residual demand synergy than cross-owner pairs; but it also exhibit a substantially thicker right tail, indicating that some affiliated app pairs have particularly strong demand synergy even without the strategic components. By contrast, the estimated latent interoperability $a_{t,j\to j'}$ is generally higher for apps

[10]Because of the location normalizations on $a_{t,j\to j'}$ and $\gamma_{tjj'}$ (see Example 1), the negative levels of $a_{t,j\to j'} + a_{t,j'\to j}$ do not necessarily imply low incentive of establishing inter-app connectivity.

of the same owner, and is more dispersed among cross-owner directions with a more pronounced negative tail. Consistent with the empirical pattern that common ownership is associated with higher query degrees among affiliated apps, these patterns indicate app owners' greater incentive of creating better interoperability among affiliated apps by strategically establishing higher latent interoperability.

Figure 7: Decomposition of Pairwise Demand Synergies

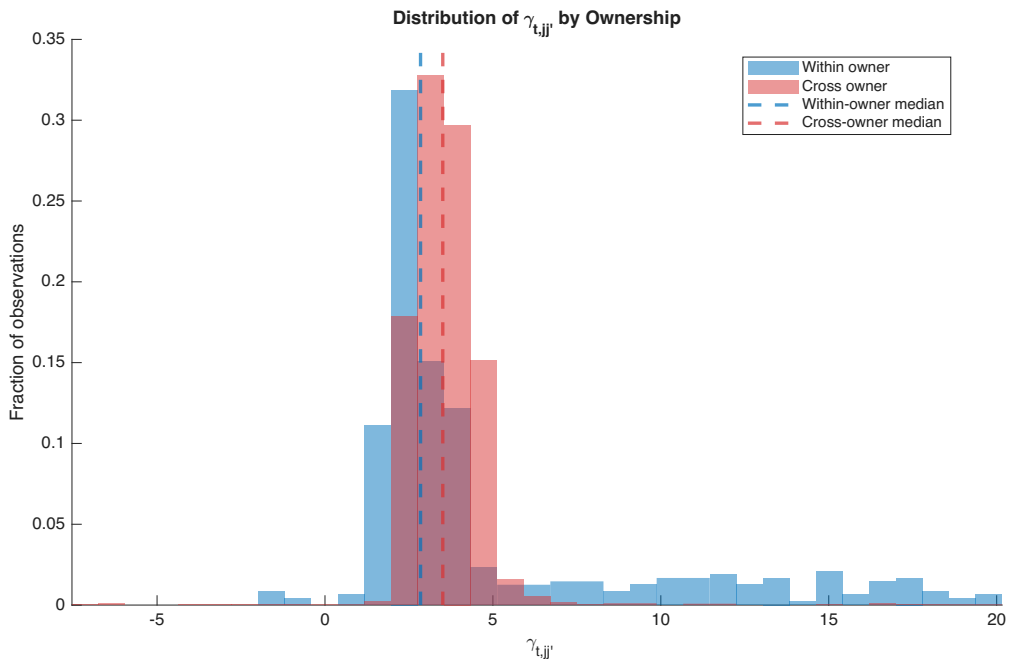


Notes: The figure decomposes the estimated pairwise demand synergy $\Gamma_{tjj'}$ into the contribution of strategic interoperability $a_{t,j\to j'} + a_{t,j'\to j}$ and residual functional complementarity $\gamma_{tjj'}$. Results are reported separately for communication–communication, communication–online video, and online video–online video app pairs. Bars indicate median values, and error bars represent the interquartile range (25th to 75th percentiles).

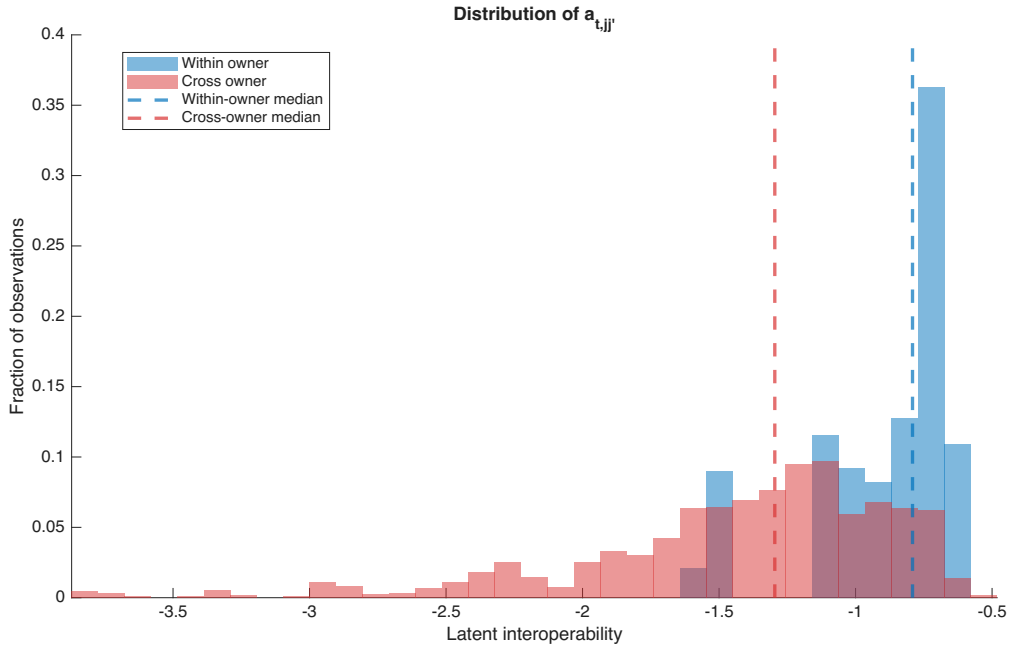
Together, the estimates of demand synergies highlight the distinction between residual demand synergies and strategic latent interoperability, and how ownership is differently associated with each component in the observed scenario. In the counterfactual interoperability mandate simulations, we will hold $\gamma_{tjj'}$'s fixed and allow app owners to change the levels of $a_{t,j\to j'}$ from their apps to others. By doing so, we investigate how app owners' incentive of establishing interoperability interacts with the mandate policy and its welfare implications.

Figure 8: Distributions of Residual Demand Synergies and Latent Interoperability by Ownership

(a) Distribution of $\gamma_{tjj'}$



(b) Distribution of Latent Interoperability



Notes: The figure compares the estimated distributions of residual pairwise demand synergy $\gamma_{tjj'}$ and directional latent interoperability $a_{t,j\to j'}$ between within-owner (blue) and cross-owner (red) app pairs. Each unordered app pair contributes one observation per market to the distribution of $\gamma_{tjj'}$, while the distribution of $a_{t,j\to j'}$ uses eligible directed app-pair-market observations. The dashed vertical lines indicate the median of each distribution.

5.2 Counterfactual: Interoperability Mandate

We evaluate a counterfactual interoperability mandate that requires each app to grant non-discriminatory outgoing interoperability to all recipient apps within a given category, regardless of ownership. In practice, this policy can be implemented by requiring platforms to expose their interoperable functionalities via standardized APIs or SDKs accessible to all eligible third-party developers. Let $\mathcal{G}_{t,j,c} = \{j' \neq j : \text{cat}(j') = c\}$ denote the set of recipient apps in category c for sender app j. We use superscript IM to refer to the quantity under the new equilibrium solutions of latent interoperability under the interoperability mandate.

The mandate imposes the equality constraints: for all $j' \in \mathcal{G}_{t,j,c}$,

$$a_{t,j\to j'} = a^{IM}_{t,j,c}. \tag{5.1}$$

The directed interoperability chosen by app j can vary across recipient categories, but not across apps within the same recipient category. (5.1) is equivalent to requiring sender app to establish an equal probability of query towards all apps within the same recipient category.

Under the baseline specification, the constrained first-order condition for each sender-recipient-category group is[11]

$$\sum_{j'\in\mathcal{G}_{t,j,c}} MR^{IM}_{t,j\to j'} = |\mathcal{G}_{t,j,c}| \exp(\rho_{f(j)} a^{IM}_{t,j,c}).$$

The total pairwise synergy entering counterfactual demand is

$$\Gamma^{IM}_{t,jj'} = \gamma_{t,jj'} + a^{IM}_{t,j,\text{cat}(j')} + a^{IM}_{t,j',\text{cat}(j)} \tag{5.2}$$

Therefore, the counterfactual equilibrium solves the nonlinear system

$$\sum_{j'\in\mathcal{G}_{t,j,c}} MR_{t,j\to j'}\left(\delta_t, \Gamma^{IM}_t; \theta_D\right) = |\mathcal{G}_{t,j,c}| \exp(a^{IM}_{t,j,c})$$

for all sender-recipient-category groups (j, c), together with (5.2). Using the new equilibrium latent interoperability, we compute the probability of establishing a query link, the density of the query network, the demand for apps, and the consumer surplus, and compare them to those in the observed scenario.

[11] Note that the ownership does not change under the interoperability mandate. Consequently, the scale parameters $\kappa_{jj'}$ in cost remain unchanged and their normalization remains legitimate in the counterfactual exercise, i.e., the value of $\gamma_{tjj'}$s remain unchanged.

Table 5: Counterfactual Effects of the Interoperability Mandate

	Category-conditional mandate
$\Delta a_{t,j\to j'}$, overall	0.0226
Within-owner	-0.0267
Cross-owner	0.0551
Query-network density (pp), overall	3.8172
Within-owner	-6.5431
Cross-owner	5.0642
Consumer surplus (%)	-1.9351
Due to within-owner $a_{t,j\to j'}$	-0.3640
Due to cross-owner $a_{t,j\to j'}$	-1.6086
App owners' profit (%)	-4.4093
Due to within-owner $a_{t,j\to j'}$	-2.0093
Due to cross-owner $a_{t,j\to j'}$	-2.6586

Notes: Entries are medians of changes relative to the baseline equilibrium. The mandate requires each app to provide the same interoperability level to eligible recipient apps within a given category, regardless of ownership. Changes in query-network density are in percentage points and are summarized across markets; changes in latent interoperability are summarized across eligible directed app-pair-market observations. Changes in consumer surplus and aggregate app owners' profit are expressed in percent. Within-owner and cross-owner contributions to consumer surplus and profit use a two-group Shapley decomposition. We evaluate each outcome at baseline, changing within-owner pairwise synergies only, changing cross-owner synergies only, and changing both. Each contribution averages the group's incremental effects when introduced first and second. Changes in interoperability provision costs are allocated to the corresponding ownership group. The contributions add up in each market, but their medians need not add up to the median total effect. Consumer surplus is measured by the normalized mixed-logit expected log-sum, and profit is the aggregate usage-based revenue index net of interoperability costs. Unidentified baseline cost levels on exempt directions are normalized to zero when computing percentage profit changes.

Table 5 summarizes the equilibrium outcomes of the interoperability mandate. The mandate increases the median level of directional interoperability across app pairs, but its effects differ sharply by ownership. Median interoperability falls by 0.0267 for within-owner directions and increases by 0.0551 for cross-owner directions. The expected density of the overall query network rises by 3.82 percentage points, reflecting a 5.06-percentage-point increase in cross-owner query density despite a 6.54-percentage-point decline within ownership. Thus, the mandate expands the overall query network while reallocating interoperability across ownership boundaries. Despite this expansion, median consumer surplus decreases by approximately 1.94%, and aggregate app owners' profit falls by 4.41%. The Shapley decomposition in Table 5 indicates that changes in both within-owner and cross-owner interoperability contribute to these losses, with the larger contributions coming from cross-owner changes.

Why does the expansion of interoperability coexist with a deterioration in consumer surplus and profit? To understand this result, we first examine how the mandate changes firms' interoperability choices. In the observed unconstrained equilibrium, app owners can differentiate their interoperability toward individual recipients, taking into account their ownership relationships and the demand for different app combinations. The mandate restricts this differentiation by requiring a common interoperability level toward all recipient apps within each category, regardless of ownership or popularity. Consequently, some app pairs gain interoperability while others lose it.

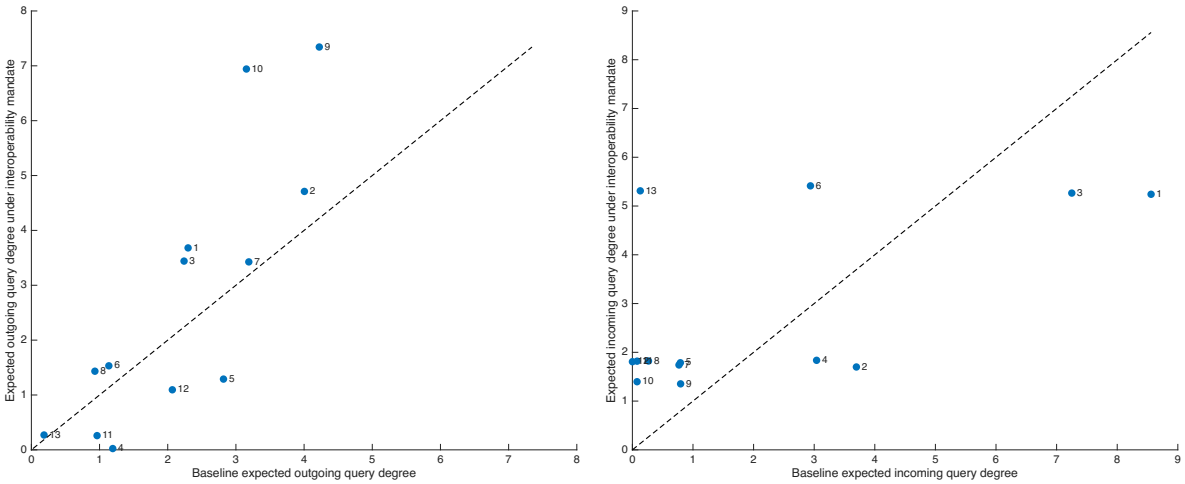
Figure 9 illustrates the resulting changes in expected query degrees at the app level. Panel (a) shows that leading apps, such as WeChat and TikTok (Apps 1 and 2), increase their expected outgoing query degrees under the mandate. These apps can no longer differentiate their outgoing interoperability among recipients in the same category, including affiliated apps and competing apps. The increase in outgoing queries is therefore consistent with the broader shift toward cross-owner interoperability documented in Table 5. At the same time, the mandate also restricts smaller apps' ability to differentiate their interoperability toward recipients of different popularity. As shown in Panel (b), WeChat and TikTok receive fewer expected incoming queries under the mandate. Taken together, these changes suggest that the policy alters not only the overall density of the query network but also the allocation of interoperability across apps.

This reallocation helps explain the negative effects on consumer surplus and profit. In particular, the mandate may increase interoperability among many relatively less important app combinations while reducing interoperability among combinations that consumers use more frequently. The overall query network can become denser even when the latter changes are economically more consequential. To investigate this possibility, we divide app-pair-market observations into deciles according to their baseline joint-usage shares $s^m_{tjj'}$. Figure 10 compares the distributions of changes in interoperability and expected query probabilities between app pairs in the highest decile and those in deciles 1–9. Relative to the latter group, the highest-decile pairs exhibit more pronounced reductions in the sum of the two directional latent interoperability levels. Their changes in expected directional query probabilities are also shifted toward negative values. Thus, the mandate weakens interoperability for some of the app combinations most frequently used by consumers, even as it expands the overall query network.

We next examine the implications of this uneven reallocation for consumer surplus and firms' profits. An important property of our demand model is that the

Figure 9: Changes in Expected Query Degrees under the Interoperability Mandate

(a) Expected Outgoing Query Degree (b) Expected Incoming Query Degree

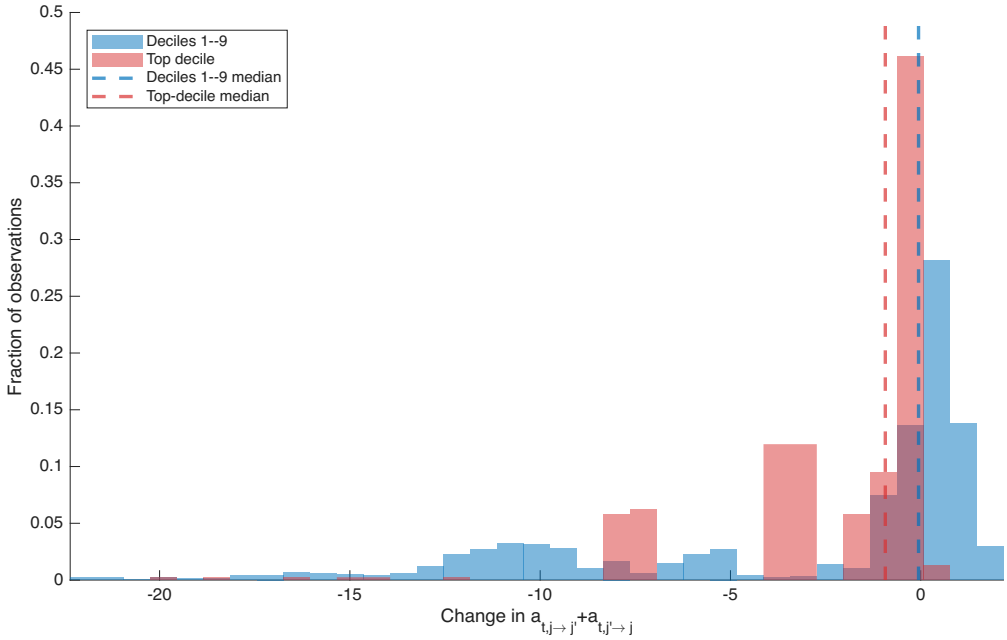


Notes: The figure compares the expected number of outgoing queries (Panel (a)) and incoming queries (Panel (b)) for each app under the baseline equilibrium and the interoperability mandate. Expected outgoing (incoming) query degree is calculated by summing the predicted query probabilities over all eligible receiving (sending) apps within each market and then averaging across markets. Each point represents an app, with its baseline value on the horizontal axis and its counterfactual value on the vertical axis. The dashed 45-degree line indicates no change. Points above (below) the line indicate an increase (decrease) in expected query degree under the mandate.

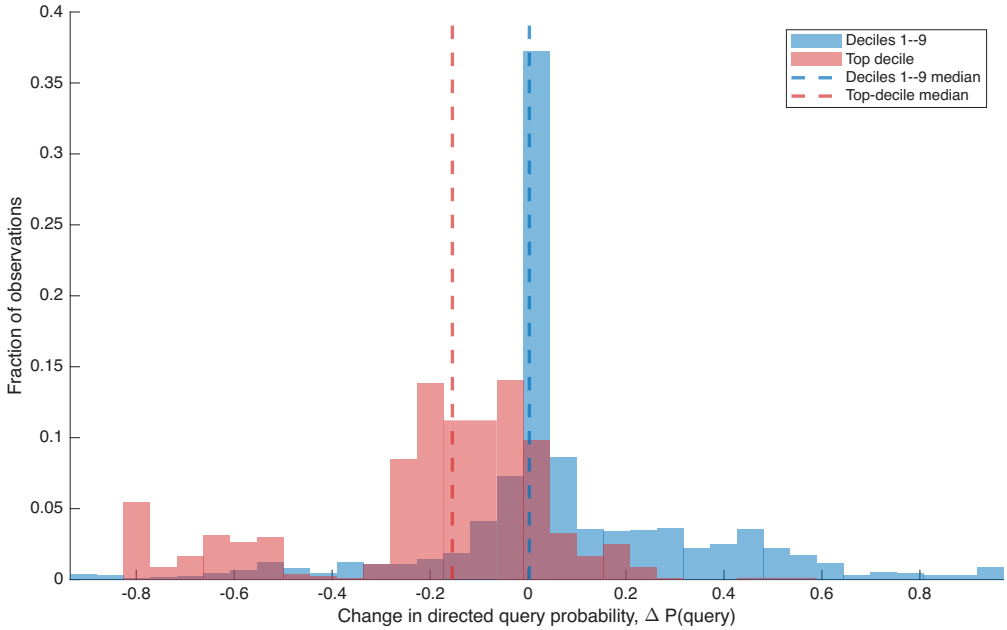
marginal effect of pairwise demand synergy $\Gamma_{tjj'}$ on consumer surplus is the corresponding joint-usage share $s^m_{tjj'}$. Therefore, holding other demand synergies fixed, an increase in the interoperability of a given app pair cannot reduce consumer surplus. The negative aggregate effect must instead reflect sufficiently important reductions in demand synergies elsewhere. Moreover, changes in the synergies of frequently used app pairs can have particularly large effects on consumer surplus.

Figure 10: Changes in Latent Interoperability and Query Probabilities

(a) Change in Latent Interoperability

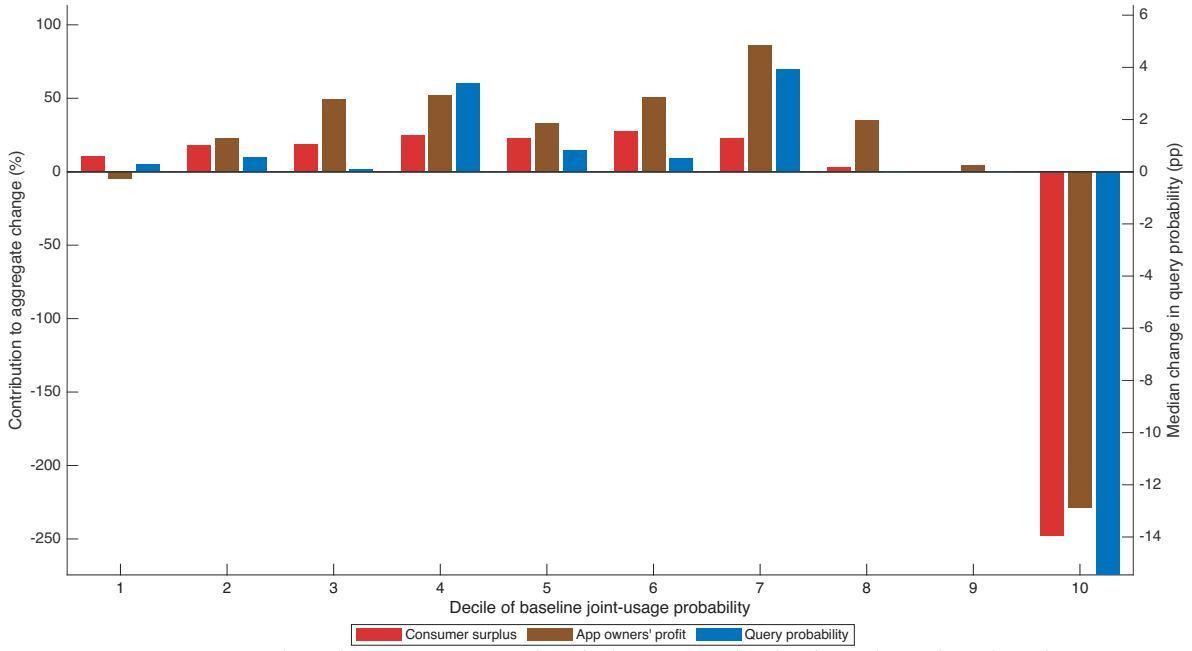


(b) Change in Query Probability



Notes: App-pair-market observations are divided into deciles according to their baseline joint-usage probabilities $s^{m,0}_{tjj'}$. The figure compares the highest decile (red) with deciles 1–9 (blue). Panel (a) reports changes in the sum of the two directional latent interoperability levels, $\Delta(a_{t,j\to j'} + a_{t,j'\to j})$, with one observation per eligible unordered app pair and market. Panel (b) reports changes in predicted query probabilities for each eligible directed link; both directions of an app pair are included separately when eligible and inherit the same baseline joint-usage decile. Dashed vertical lines indicate group medians.

Figure 11: Decomposition of Welfare Changes



Notes: App-pair-market observations are divided into ten deciles based on their baseline joint-usage probabilities $s_{tjj'}^{m,0}$. The red and brown bars report, respectively, the contributions of each decile to changes in consumer surplus and aggregate app owners' profit (left axis). Contributions to consumer surplus and the revenue component of profit are obtained from an exact path-integral decomposition of mixed-logit demand; profit contributions also subtract changes in directional interoperability costs (Appendix D.2). Within each decile, pair-level contributions are summed and averaged across markets, then expressed as percentages of the absolute average change in the corresponding aggregate outcome. The blue bars report the median change in predicted directional query probability, in percentage points, for eligible links in each decile (right axis). Negative values indicate losses or declines.

This reasoning provides the economic motivation for the exercise in Figure 11, where we decompose the changes in consumer surplus and aggregate app owners' profit into contributions from app pairs in each decile of baseline joint usage (left axis). We also report the median change in expected query probability in each decile (right axis). The decomposition uses the exact path-integral approach described in Appendix C.2, which accounts for the adjustment of demand along the path from the baseline to the counterfactual equilibrium. The results show that changes affecting pairs in the first nine deciles generally contribute positively to consumer surplus and profit. However, these gains are more than offset by the substantial negative contributions from the highest decile. The expected query probability also declines particularly strongly among these frequently used app pairs.

The decomposition therefore explains how an increase in overall interoperability can coexist with negative aggregate economic outcomes. Although the mandate increases interoperability for many app combinations, it also weakens interoperability

among combinations that are especially important in baseline demand. The resulting losses from these combinations dominate the benefits generated elsewhere. Importantly, this mechanism depends on the distribution of interoperability changes and their demand implications, rather than simply on the total number of query links created by the mandate.

Overall, these counterfactual results highlight the complexity of designing interoperability mandates in digital markets. A policy that requires more uniform interoperability can expand the query network and reduce discrimination in access across ownership boundaries, but it may also weaken interoperability among app combinations that generate substantial value for consumers and firms. Welfare consequences therefore depend not only on the extent of interoperability required by the policy, but also on which apps are subject to the mandate and how other apps respond strategically. This raises the question of whether a more targeted mandate could achieve similar improvements in interoperability with smaller adverse effects on consumer surplus and profit. To complement the present exercise, we are also examining a mandate that applies only to the leading app in each category, such as WeChat and TikTok, while allowing smaller apps to choose their interoperability independently. Comparing these policies will help clarify the role of policy coverage and firms' strategic responses in determining the effects of interoperability regulation.

6 Conclusion

This paper develops a structural framework for measuring digital interoperability, understanding firms' incentives to provide it, and evaluating its welfare consequences. We construct a directional measure based on query declarations in app code, which record whether one app requests visibility of another. Using data from the Chinese mobile-app market, we document that query links are associated with greater joint usage, are disproportionately directed toward larger apps, and are more prevalent between apps under common ownership.

Our structural framework combines consumer demand for app bundles, firms' choices of directional interoperability, and a threshold technology linking latent interoperability to observed queries. Bringing usage and query data together allows us to distinguish strategic interoperability from residual pairwise demand synergies and to estimate the preferences and costs underlying firms' choices. The framework thus connects an observable feature of digital products to the economic

incentives governing interoperability and provides a basis for evaluating how regulation changes both firms' decisions and consumer outcomes.

Our counterfactual illustrates a central challenge for interoperability policy. A mandate requiring equal outgoing interoperability across recipients within a category expands cross-owner connectivity, yet reduces consumer surplus and aggregate firm profit in our application. Firms reallocate interoperability away from some of the app combinations consumers use most frequently, and losses on those connections outweigh gains elsewhere. The welfare consequences of a mandate therefore depend on which connections improve or deteriorate as firms respond, as well as on the overall extent of connectivity. Although our analysis holds app availability, quality, and ownership fixed, it demonstrates the importance of evaluating interoperability mandates through their effects on consumer demand and firms' strategic choices. More broadly, the paper provides a framework for moving from the measurement of digital connections to an assessment of their economic value and the consequences of regulating them.

Our study also opens avenues for examining other dimensions of digital interoperability. Query declarations capture whether one app requests visibility of another, while additional information in app code could help characterize the scope and nature of interactions between them. For example, Android intents and intent filters describe requested actions and the functions available to handle them. Together with information on whether these functions are accessible to other apps, these features could support complementary measures of interoperability, distinguishing connections that enable different functions, such as sharing content or handing off tasks. We leave the extension of our framework to these measures to future research.

References

Adner, Ron, Chen, Jianqing, & Zhu, Feng. 2020. Frenemies in Platform Markets: Heterogeneous Profit Foci as Drivers of Compatibility Decisions. *Management Science*, **66**(6), 2432–2451.

Aridor, Guy. 2025. Measuring Substitution Patterns in the Attention Economy: An Experimental Approach. *The RAND Journal of Economics*, **56**(3), 302–324.

Armstrong, Mark. 1998. Network Interconnection in Telecommunications. *The Economic Journal*, **108**(448), 545–564.

Augereau, Angelique, Greenstein, Shane, & Rysman, Marc. 2006. Coordination

versus Differentiation in a Standards War: 56K Modems. *The RAND Journal of Economics*, **37**(4), 887–909.

Bourreau, Marc, & Kraemer, Jan. 2022. Interoperability in Digital Markets. *SSRN Electronic Journal*.

Bresnahan, Timothy F, Davis, Jason P, & Yin, Pai-Ling. 2014. Economic value creation in mobile applications. *Pages 233–286 of: The changing frontier: Rethinking science and innovation policy*. University of Chicago Press.

Brown, Ian. 2020 (July). *Interoperability as a tool for competition regulation*.

Brunnermeier, Markus K., Limodio, Nicola, & Spadavecchia, Lorenzo. 2023 (Sept.). *Mobile Money, Interoperability, and Financial Inclusion*.

Chen, Jiawei, Doraszelski, Ulrich, & Harrington Jr, Joseph E. 2009a. Avoiding market dominance: Product compatibility in markets with network effects. *The RAND Journal of Economics*, **40**(3), 455–485.

Chen, Jiawei, Doraszelski, Ulrich, & Harrington, Jr., Joseph E. 2009b. Avoiding market dominance: product compatibility in markets with network effects. *The RAND Journal of Economics*, **40**(3), 455–485.

Chen, Zhijun, & Rey, Patrick. 2026. *A Theory of Conglomerate Mergers*.

Crawford, Gregory S., Crémer, Jacques, Dinielli, David, Fletcher, Amelia, Heidhues, P., Schnitzer, Monika, Scott Morton, F. M., & Seim, Katja. 2023. Equitable Interoperability: The "Super Tool" of Digital Platform Governance. *Yale Journal on Regulation*.

Crémer, Jacques, Rey, Patrick, & Tirole, Jean. 2000. Connectivity in the Commercial Internet. *The Journal of Industrial Economics*, **48**(4), 433–472. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-6451.00132.

Dix, Rebekah, Moran, Kelsey, & Nguyen, Thi Mai Anh. 2025. Costs of Technological Frictions: Evidence from EHR (Non-)Interoperability.

Doganoglu, Toker, & Wright, Julian. 2006. Multihoming and compatibility. *International Journal of Industrial Organization*, **24**(1), 45–67.

Ekmekci, Mehmet, White, Alexander, & Wu, Lingxuan. 2025. Platform Competition and Interoperability: The Net Fee Model. *Management Science*, **71**(10), 8842–8864.

Ershov, Daniel, Laliberté, Jean-William, Marcoux, Mathieu, & Orr, Scott. 2025. Estimating complementarity with large choice sets: An application to mergers. *The RAND Journal of Economics*, **56**(4), 689–707.

European Commission. Joint Research Centre. 2021. *The EU digital markets act: a report from a panel of economic experts.* LU: Publications Office.

Farrell, Joseph, & Saloner, Garth. 1992. Converters, compatibility, and the control of interfaces. *The Journal of Industrial Economics*, 9–35.

Frahan, Lancelot de, Goodman, Joseph, Holz, Justin, List, John, McKay, Evan, McMenamin, Niall, Mogstad, Magne, Sadoff, Sally, & Sider, Hal. 2025. *Consumer Demand and Market Competition with Time-Intensive Goods*.

Gale, David, & Nikaido, Hukukane. 1965. The Jacobian matrix and global univalence of mappings. *Mathematische Annalen*, **159**(2), 81–93.

Genakos, Christos, Kühn, Kai-Uwe, & Reenen, John Van. 2018. Leveraging Monopoly Power by Degrading Interoperability: Theory and Evidence from Computer Markets. *Economica*, **85**(340), 873–902. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecca.12257.

Gentzkow, Matthew. 2007. Valuing new goods in a model with complementarity: Online newspapers. *American Economic Review*, **97**(3), 713–744.

Gross, Daniel P. 2020. Collusive Investments in Technological Compatibility: Lessons from U.S. Railroads in the Late 19th Century. *Management Science*, **66**(12), 5683–5700.

Han, Sang, park, sungho, & Oh, Wonseok. 2016. Mobile App Analytics: A Multiple Discrete-Continuous Choice Framework. *Management Information Systems Quarterly*, **40**(4), 983–1008. 00031.

Heidhues, Paul, Köster, Mats, & Kőszegi, Botond. 2024. *A Theory of Digital Ecosystems*.

Iaria, Alessandro, & Wang, Ao. 2020. Identification and estimation of demand for bundles.

Iaria, Alessandro, & Wang, Ao. 2026. An empirical model of quantity discounts with large choice sets. *Journal of the European Economic Association*, jvag033.

Jacobides, Michael G, Cennamo, Carmelo, & Gawer, Annabelle. 2024. Externalities and complementarities in platforms and ecosystems: From structural solutions to endogenous failures. *Research Policy*, **53**(1), 104906.

Kades, Michael, & Morton, Fiona Scott. 2020. Interoperability as a competition remedy for digital networks. *Washington Center for Equitable Growth Working Paper Series*.

Katz, Michael L, & Shapiro, Carl. 1985. Network externalities, competition, and compatibility. *The American Economic Review*, **75**(3), 424–440.

Knittel, Christopher R., & Stango, Victor. 2009. HOW DOES INCOMPATIBILITY AFFECT PRICES?: EVIDENCE FROM ATM'S. *The Journal of Industrial Economics*, **57**(3), 557–582.

Knittel, Christopher R., & Stango, Victor. 2011. Strategic incompatibility in ATM markets. *Journal of Banking & Finance*, **35**(10), 2627–2636.

Lee, Robin S. 2013. Vertical integration and exclusivity in platform and two-sided markets. *American Economic Review*, **103**(7), 2960–3000.

Leung, Hareton, & Fan, Zhang. 2002. Software cost estimation. *Pages 307–324 of: Handbook of Software Engineering and Knowledge Engineering: Volume II: Emerging Technologies*. World Scientific.

Li, Jing. 2023. Compatibility and investment in the us electric vehicle market. *Unpublished manuscript, MIT*.

Manenti, Fabio M, & Somma, Ernesto. 2008. One-way compatibility, two-way compatibility and entry in network industries. *International Journal of the Economics of Business*, **15**(3), 301–322.

OECD. 2021. *Data portability, interoperability and digital platform competition*. Tech. rept. OECD Competition Committee Discussion Paper.

Park, Sangin. 2004. Quantitative analysis of network externalities in competing technologies: The VCR case. *Review of Economics and Statistics*, **86**(4), 937–945.

Peitz, Martin, & Sato, Susumu. 2025. Asymmetric Platform Oligopoly. *The RAND Journal of Economics*, **n/a**(n/a). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.70008.

Wang, Ao. 2021. Demand with Complementarity: A Market-Share Approach. *Available at SSRN 3785148*.

Whinston, Michael D. 1990. Tying, Foreclosure, and Exclusion. *The American Economic Review*, 837–859.

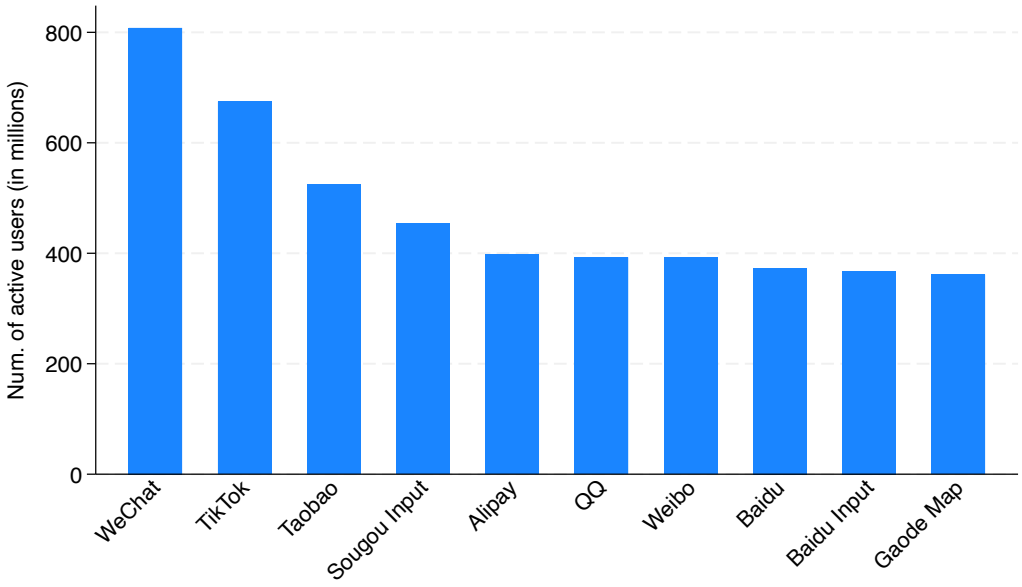
Yuan, Han. 2026. Competing for time: A study of mobile applications. *International Journal of Industrial Organization*, **105**(Apr.), 103247.

Appendices for Online Publication Only

Ao Wang, Jian Xie, Han Yuan

A Additional Tables and Figures

Figure A.1: Active Android users of the top 10 apps at the end of 2023



Notes: The figure plots the number of active Android users (in millions) for the top 10 apps at the end of 2023. The apps are WeChat, TikTok, Taobao, Sougou Input, Alipay, QQ, Weibo, Baidu, Baidu Input, and Gaode Map.

B Proof of Proposition 4.1

Let $\eta_t = (\delta_t^\top, \Gamma_t^\top)^\top$. Define a matrix W of dimension $\left(J + \binom{N}{2}\right) \times |\mathbf{J} \cup \mathbf{B}|$ whose columns correspond to alternatives $\mathbf{b} \in \mathbf{J} \cup \mathbf{B}$. For bundle $\mathbf{b}$, its app-level rows and app-pair-level rows are, respectively,

$$W_{j,\mathbf{b}} = \mathbf{1}\{j \in \mathbf{b}\}, \qquad W_{(j,j'),\mathbf{b}} = \mathbf{1}\{j, j' \in \mathbf{b}\}, \quad j < j'.$$

The individual indirect-utility vector, denoted by v_i, can then be written as

$$v_i = W^\top \eta_t + \left(\sum_{j \in \mathbf{b}} \beta_i x_{tj} \right)_{\mathbf{b} \in \mathbf{J} \cup \mathbf{B}}.$$

The availability of singleton and two-app bundles implies that W has full row rank. To see this, one can re-order its columns so that the alternatives j and (j, j') for any $j, j' \in \mathbf{J}$ and $j < j'$ are the first columns. Then, the corresponding submatrix in W is square and has the following form:

$$\begin{pmatrix} \mathbf{I}_{J \times J} & (\mathbf{1}\{j = r \text{ or } j = l\})_{j \in \mathbf{J}, (r,l), r<l} \\ 0 & \mathbf{I}_{\binom{J}{2} \times \binom{J}{2}} \end{pmatrix}.$$

Because this submatrix has full row rank, W is then of full row rank.

Let $\sigma_i(v_i; \theta_e) = (\sigma_{i\mathbf{b}}(v_i; \theta_e))_{\mathbf{b} \in \mathbf{J} \cup \mathbf{B}}$. By the definitions of marginal and pairwise joint-usage shares,

$$\sigma_t^m(\delta_t, \Gamma_t; \theta_D) = W \int \sigma_i(v_i; \theta_e) \, dF(\beta_i; \theta_F).$$

Consequently,

$$\frac{\partial \sigma_t^m(\delta_t, \Gamma_t; \theta_D)}{\partial \eta_t^\top} = W \left[\int \frac{\partial \sigma_i(v_i; \theta_e)}{\partial v_i^\top} \, dF(\beta_i; \theta_F) \right] W^\top.$$

Under the additive random-utility structure, individual choice probabilities are the gradient of a convex surplus function in the utility indices. Their Jacobian is therefore symmetric positive semidefinite. Condition 2 of Proposition 4.1 strengthens this property to positive definiteness for a set of consumer types with positive probability. The matrix inside brackets is consequently positive definite. Since W has full row rank, the Jacobian of σ_t^m with respect to η_t is also symmetric positive definite. Since this Jacobian is symmetric positive definite, it is also positive quasi-definite. Together with the convex domain in Condition 1 of Proposition 4.1, Theorem 6 of Gale & Nikaido (1965) implies that $\sigma_t^m(\cdot; (x_{tj})_{j \in \mathbf{J}}, \theta_D)$ is globally one-to-one in (δ_t, Γ_t).

C Empirical application: Details

C.1 Bundle Size Calibration: A Diagnostic Statistic

We use the observed marginal app shares and pairwise joint-usage shares to construct a simple diagnostic for the bundle size required by the data. Recall that s^m_{tj} denotes the marginal usage share of app j in market t, and $s^m_{tjj'}$ denotes the joint-usage share of apps j and j'.

Suppose that consumers choose bundles containing at most L apps. Conditional on app j being used, a consumer can use at most $L-1$ additional apps. It follows that, for every app j with $s^m_{tj} > 0$,

$$\sum_{j' \neq j} s^m_{tjj'} \leq (L-1) s^m_{tj}. \tag{C.1}$$

Equivalently,

$$\frac{\sum_{j' \neq j} s^m_{tjj'}}{s^m_{tj}} \leq L - 1. \tag{C.2}$$

The condition has a direct interpretation. The numerator in (C.2) measures the total pairwise co-usage involving app j. A consumer who uses app j and chooses a bundle of size at most L can contribute to at most $L-1$ pairwise joint-usage events involving j. Hence, aggregate pairwise co-usage involving app j cannot exceed $(L-1)$ times its marginal usage.

For each market t, this yields the necessary lower bound

$$L \geq 1 + \max_{j : s^m_{tj} > 0} \frac{\sum_{j' \neq j} s^m_{tjj'}}{s^m_{tj}}. \tag{C.3}$$

We define the market-specific diagnostic

$$L^{\text{nec}}_{t,\min} = \left\lceil 1 + \max_{j : s^m_{tj} > 0} \frac{\sum_{j' \neq j} s^m_{tjj'}}{s^m_{tj}} \right\rceil. \tag{C.4}$$

A common maximum bundle size across markets must therefore satisfy

$$L \geq \max_t L^{\text{nec}}_{t,\min}. \tag{C.5}$$

The diagnostic in (C.4) provides only a necessary condition. Satisfying (C.1) does not by itself guarantee that the observed marginal and pairwise shares can be ra-

tionalized by a nonnegative distribution over bundles of size at most L. Thus, $L^{\text{nec}}_{t,\min}$ should be interpreted as a convenient lower bound rather than as a sufficient rationalizability condition. In the empirical application, our computation shows $\max_t L^{\text{nec}}_{t,\min} = 6$, implying that the necessary lower bound on the maximum bundle size is 6. We therefore calibrate $L = 6$. With this choice, the demand inverse converges and matches the observed marginal app shares and pairwise joint-usage shares in all markets. The resulting set of nonempty bundles contains $\sum_{\ell=1}^{6} \binom{13}{\ell} = 4{,}095$ alternatives.

C.2 Decomposition of consumer welfare and firms' profits

We divide all app-pair-market observations into ten deciles according to their baseline joint-usage probabilities $s^m_{t,jj'}$, and further distinguish between within-owner and cross-owner app pairs. The same deciles are used for both ownership groups and both outcomes.

To construct the contributions to consumer welfare, consider the linear path between the baseline and counterfactual demand synergies, $\Gamma_t(\lambda) = \Gamma^0_t + \lambda(\Gamma^{IM}_t - \Gamma_t)$, where $\lambda \in [0,1]$. The exact contribution of app pair (j, j') to the consumer surplus change in market t is

$$CS_{t,jj'} = \Delta\Gamma_{t,jj'} \int_0^1 s^m_{t,jj'}\big(\Gamma_t(\lambda)\big)\, d\lambda, \tag{C.6}$$

where $\Delta\Gamma_{t,jj'} = \Gamma^{IM}_{t,jj'} - \Gamma_{t,jj'}$. These contributions sum to the total welfare change in each market.

Similarly, let $R_t(\Gamma_t) = \sum_j s^m_{tj}(\Gamma_t)$ denote aggregate app owners' revenue. The contribution of each app pair to the change in aggregate profit is

$$\Pi_{t,jj'} = \Delta\Gamma_{t,jj'} \int_0^1 \frac{\partial R_t(\Gamma_t(\lambda))}{\partial \Gamma_{t,jj'}}\, d\lambda - \Delta C_{t,jj'}, \tag{C.7}$$

where $\Delta C_{t,jj'}$ is the sum of the changes in interoperability provision costs in the two directions $j \to j'$ and $j' \to j$.

For each decile d and ownership group o, the corresponding bar is computed by summing the pair-level contributions over all observations in that group and

dividing by the number of markets:

$$W_{d,o} = \frac{1}{T} \sum_{(t,j,j') \in \mathcal{D}_{d,o}} W_{t,jj'}, \qquad W \in \{CS, \Pi\}, \tag{C.8}$$

where $\mathcal{D}_{d,o}$ denotes the set of app-pair-market observations belonging to decile d and ownership group o. Thus, each bar measures the group's contribution to the average change in consumer welfare or aggregate profit across markets.

To make the two panels comparable, we normalize each contribution by the absolute value of the corresponding average aggregate change:

$$\widetilde{W}_{d,o} = 100 \frac{W_{d,o}}{|\frac{1}{T} \sum_t (W_t^{IM} - W_t^0)|}, \qquad W \in \{CS, \Pi\}. \tag{C.9}$$

C.3 Additional Results

Post-MLE IV regression. We report the estimates of the post-MLE regressions in Table C.1 for the mixed-logit demand specification. The OLS estimate of σ_τ is larger than the IV estimate. This difference may reflect the endogenous response of interoperability to unobserved functional complementarity. Intuitively, greater functional complementarity increases the attractiveness of joint app usage, which may raise the marginal revenue from providing additional interoperability. Firms consequently have stronger incentives to invest in interoperability. In this case, unobserved functional complementarity is positively correlated with the equilibrium interoperability index, leading to an upward bias in the OLS estimate of σ_τ.

Table C.1: Post-MLE Scale Estimation: OLS versus IV

	OLS	IV
Scale parameter σ_τ	0.3616	0.2259
	(0.0531)	(0.1191)
Pair fixed effects	Yes	Yes
Market fixed effects	Yes	Yes
R^2	0.7843	0.7840
First-stage F-statistic	–	201.14
First-stage partial R^2	–	0.1254

Notes: The number of observations is 2915. Standard errors in parentheses are two-way clustered by app pair and market. Both regressions use the same sample. IV uses two excluded instruments.

D A theoretical justification for convex cost

In this appendix, we provide a theoretical motivation for Assumption 1. We show how convex interoperability costs can generate an asymmetry consistent with the data: dominant apps query non-dominant apps less frequently than non-dominant apps query dominant ones.

To fix ideas, consider two app categories, denoted by 1 and 2. In the first category, a user chooses either a dominant app A or an outside option 0. In the second category, the user chooses between two apps, B and C. The user therefore chooses one bundle from $\{AB, AC, 0B, 0C\}$.

Let $a_{X\to Y}$ denote the interoperability level from app X to app Y. Following (3.4), the deterministic component of utility from bundle XY is

$$\delta_{XY} = \delta_X + \delta_Y + \Gamma_{XY},$$

where δ_X captures the standalone quality of app X. For $X \in \{B, C\}$, the synergy between A and X is

$$\Gamma_{AX} = \gamma_{AX} + a_{A\to X} + a_{X\to A}.$$

Interoperability with the outside option is held fixed and absorbed into γ_{0B} and γ_{0C}. The indirect utilities of the four bundles are

$$\begin{aligned}
u_{AB} &= \delta_A + \delta_B + \gamma_{AB} + a_{A\to B} + a_{B\to A} + \varepsilon_{AB},\\
u_{AC} &= \delta_A + \delta_C + \gamma_{AC} + a_{A\to C} + a_{C\to A} + \varepsilon_{AC},\\
u_{0B} &= \delta_0 + \delta_B + \gamma_{0B} + \varepsilon_{0B},\\
u_{0C} &= \delta_0 + \delta_C + \gamma_{0C} + \varepsilon_{0C},
\end{aligned}$$

where $(\varepsilon_{AB}, \varepsilon_{AC}, \varepsilon_{0B}, \varepsilon_{0C})$ are bundle-specific idiosyncratic shocks.

Write

$$a_{A\leftrightarrow X} = a_{A\to X} + a_{X\to A}, \qquad X \in \{B, C\}.$$

The user chooses the bundle delivering the highest utility. Let

$$q_{XY}(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta, \gamma) = \Pr(u_{XY} = \max\{u_{AB}, u_{AC}, u_{0B}, u_{0C}\})$$

denote the choice probability of bundle XY, where $\delta = (\delta_A, \delta_B, \delta_C, \delta_0)$ and $\gamma = (\gamma_{AB}, \gamma_{AC}, \gamma_{0B}, \gamma_{0C})$. We suppress the dependence on γ below. The four choice probabilities sum to one. We assume that bundle demands are differentiable and

strictly substitutable: an increase in the utility index of one bundle reduces the choice probabilities of the other bundles.

We now turn to the supply side. Let $R_X(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta)$ denote the revenue generated by app $X \in \{A, B, C\}$ that is differentiable in $(a_{A\leftrightarrow B}, a_{A\leftrightarrow C})$. Consider independent ownership. Each app chooses its outgoing interoperability levels to maximize its revenue net of costs, taking the other apps' choices as given. Following the separably additive specification in the main text, let

$$C_A(a_{A\to B}, a_{A\to C}) = c_A(a_{A\to B}) + c_A(a_{A\to C}), \qquad C_X(a_{X\to A}) = c_X(a_{X\to A}), \quad X \in \{B, C\},$$

where c_X is a differentiable, increasing cost function. App A solves

$$(a^*_{A\to B}, a^*_{A\to C}) \in \arg\max_{(a_{A\to B}, a_{A\to C})} \left\{ R_A(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) - C_A(a_{A\to B}, a_{A\to C}) \right\}. \tag{D.1}$$

Similarly, each app $X \in \{B, C\}$ solves

$$a^*_{X\to A} \in \arg\max_{a_{X\to A}} \left\{ R_X(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) - C_X(a_{X\to A}) \right\}. \tag{D.2}$$

Proposition D.1. *Suppose that $c_A = c_B = c_C = c$, where c is differentiable, increasing, and strictly convex. At any interior Nash equilibrium, for each $X \in \{B, C\}$,*

$$\frac{\partial R_A}{\partial a_{A\leftrightarrow X}} < \frac{\partial R_X}{\partial a_{A\leftrightarrow X}} \quad \Longleftrightarrow \quad a^*_{A\to X} < a^*_{X\to A},$$

where the derivatives are evaluated at the equilibrium interoperability levels.

Proposition D.1 links differences in marginal revenue to directional interoperability. The condition $\frac{\partial R_A}{\partial a_{A\leftrightarrow X}} < \frac{\partial R_X}{\partial a_{A\leftrightarrow X}}$ states that an additional unit of bilateral interoperability generates less revenue for the dominant app A than for app X. With identical, strictly convex link-specific costs, A therefore chooses less interoperability toward X than X chooses toward A. Under a common, strictly increasing query-threshold distribution in the two directions, this ordering also implies a lower query probability from A to X. The result thus provides a mechanism that rationalizes the query asymmetry documented in Section 2.1 and a theoretical motivation for Assumption 1. Note that the curvature of costs is essential to this implication. If the common link-specific cost function were instead weakly concave, its marginal cost would be nonincreasing. The same strict marginal-revenue inequality would imply therefore imply $a^*_{A\to X} \geq a^*_{X\to A}$, which is the opposite of the observed query asymmetry.

Proof of Proposition D.1. At an interior equilibrium, app A's first-order conditions

are

$$\frac{\partial R_A}{\partial a_{A\leftrightarrow B}} = c'(a^*_{A\to B}),\ \frac{\partial R_A}{\partial a_{A\leftrightarrow C}} = c'(a^*_{A\to C}). \tag{D.3}$$

For each $X \in \{B, C\}$, the corresponding first-order condition is

$$\frac{\partial R_X}{\partial a_{A\leftrightarrow X}} = c'(a^*_{X\to A}). \tag{D.4}$$

All derivatives are evaluated at the equilibrium choices. Because c is differentiable and strictly convex, c' is strictly increasing. Therefore,

$$\frac{\partial R_A}{\partial a_{A\leftrightarrow X}} < \frac{\partial R_X}{\partial a_{A\leftrightarrow X}} \quad\Longleftrightarrow\quad c'(a^*_{A\to X}) < c'(a^*_{X\to A}) \quad\Longleftrightarrow\quad a^*_{A\to X} < a^*_{X\to A}.$$

This proves the result. □

E Convex cost: A piecewise-linear example

This appendix illustrates how the threshold-based query technology in (3.13) can generate the convex costs in Assumption 1. Conditional on an implementation threshold, costs are piecewise linear; averaging over heterogeneous thresholds yields strictly convex expected costs.

Let τ denote the implementation threshold and consider the separably additive cost specification in (4.4), with

$$c_j(a; \theta_H, \theta_L) = \mathbb{E}_\tau \left[\theta_L a + (\theta_H - \theta_L)(a - \tau)_+\right],$$

where $(x)_+ = \max\{x, 0\}$ and $\theta_H > \theta_L > 0$. Assume that the expectation is finite. Conditional on τ, implementation below the threshold has marginal cost θ_L, while implementation above the threshold has the higher marginal cost θ_H. This specification captures the idea that supporting richer interactions, associated with a query declaration, requires more costly development and maintenance at the margin.

If τ has a continuous distribution function F_τ, expected marginal cost is

$$\partial_a c_j(a; \theta_H, \theta_L) = \theta_L + (\theta_H - \theta_L) F_\tau(a).$$

This expression is positive and nondecreasing in a: higher interoperability makes it more likely that implementation requires the higher marginal cost. If, in addition, τ has a continuous, strictly positive density over the relevant range of interoperability,

then

$$\partial_a^2 c_j(a;\theta_H,\theta_L) = (\theta_H - \theta_L) f_\tau(a) > 0.$$

Thus, although implementation costs are piecewise linear conditional on a threshold, expected costs are increasing and strictly convex.

F Revenue specifications: Examples

We continue with the four-alternative, two-category setting described in Appendix D, where users choose among the bundles $\{AB, AC, 0B, 0C\}$. Using the same notation, we illustrate two revenue specifications corresponding to advertising-based and subscription-based business models.

Example 1: Advertising-based revenue. Suppose app A monetizes user attention through advertising. Let $m_A > 0$ denote advertising revenue per user, which is independent of $(a_{A\leftrightarrow B}, a_{A\leftrightarrow C})$. A natural revenue specification is

$$R_A(a_{A\leftrightarrow B}, a_{A\leftrightarrow C};\delta) = m_A\left[q_{AB}(a_{A\leftrightarrow B}, a_{A\leftrightarrow C};\delta) + q_{AC}(a_{A\leftrightarrow B}, a_{A\leftrightarrow C};\delta)\right].$$

Thus, app A's revenue is proportional to the total demand for bundles containing A.

For each $X \in \{B, C\}$, the substitution property of bundle demands implies

$$\frac{\partial q_{0B}}{\partial a_{A\leftrightarrow X}} < 0, \qquad \frac{\partial q_{0C}}{\partial a_{A\leftrightarrow X}} < 0.$$

Since the four bundle probabilities sum to one,

$$\frac{\partial R_A}{\partial a_{A\leftrightarrow X}} = m_A\left(\frac{\partial q_{AB}}{\partial a_{A\leftrightarrow X}} + \frac{\partial q_{AC}}{\partial a_{A\leftrightarrow X}}\right) = -m_A\frac{\partial(q_{0B}+q_{0C})}{\partial a_{A\leftrightarrow X}} > 0.$$

Greater interoperability with either app therefore increases app A's advertising revenue.

Example 2: Subscription-based revenue. Suppose instead that app A follows a subscription-based business model. Let $p_A > 0$ denote its subscription price, with the marginal utility of income normalized to one. The bundle utilities are

$$u_{AB} = \tilde{\delta}_{AB} - p_A + a_{A\leftrightarrow B} + \varepsilon_{AB},$$

$$
\begin{aligned}
u_{AC} &= \tilde{\delta}_{AC} - p_A + a_{A\leftrightarrow C} + \varepsilon_{AC}, \\
u_{0B} &= \delta_{0B} + \varepsilon_{0B}, \\
u_{0C} &= \delta_{0C} + \varepsilon_{0C},
\end{aligned}
$$

where $\tilde{\delta}_{AB} = \delta_A + \delta_B + \gamma_{AB}$ and $\tilde{\delta}_{AC} = \delta_A + \delta_C + \gamma_{AC}$ collect the baseline utility components excluding A's price and bilateral interoperability. Similarly, $\delta_{0X} = \delta_0 + \delta_X + \gamma_{0X}$ for $X \in \{B, C\}$.

If other apps also charge subscription prices, we hold those prices fixed in this example and subtract them from the corresponding bundle indices. For instance, app B's price enters both $\tilde{\delta}_{AB}$ and δ_{0B}. We continue to use δ to denote the collection of utility primitives, suppressing dependence on these fixed prices.

Given p_A, app A's revenue is

$$
\tilde{R}_A(p_A, a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) = p_A \left[q_{AB}(p_A, a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) + q_{AC}(p_A, a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) \right].
$$

Suppose app A chooses its subscription price optimally:

$$
p_A^*(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) \in \arg\max_{p_A > 0} \tilde{R}_A(p_A, a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta).
$$

Define its maximized revenue as

$$
R_A(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta) = \tilde{R}_A(p_A^*(a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta), a_{A\leftrightarrow B}, a_{A\leftrightarrow C}; \delta).
$$

Assuming an interior optimal price and differentiability of maximized revenue, the envelope theorem gives, for each $X \in \{B, C\}$,[12]

$$
\frac{\partial R_A}{\partial a_{A\leftrightarrow X}} = p_A^* \left(\frac{\partial q_{AB}}{\partial a_{A\leftrightarrow X}} + \frac{\partial q_{AC}}{\partial a_{A\leftrightarrow X}} \right) = -p_A^* \frac{\partial (q_{0B} + q_{0C})}{\partial a_{A\leftrightarrow X}} > 0,
$$

where the demand derivatives are evaluated at $p_A = p_A^*$ and hold prices fixed. Thus, greater interoperability with either app also increases app A's maximized subscription revenue.

[12]The envelope theorem eliminates the effect through app A's own optimal price. If other apps' subscription prices also adjust to interoperability, the derivative of equilibrium revenue includes additional terms through those prices. The sign established here therefore applies when other apps' prices are held fixed.