The Regulatory Trade-Off: Data Integration, Privacy Costs and Welfare in Digital Platforms*

Alessandro Sellitto[1,2]

[1]Department of Economics, Management and Quantitative Methods, University of Milan
[2]Department of Economics and Management, University of Pavia

This version: February 2026

Abstract

This paper studies the welfare trade-offs of data integration in digital platform markets using a model with heterogeneous privacy sensitivity, personalization preferences, and endogenous platform investment. Lowering opt-out costs improves welfare by allowing privacy-sensitive users to limit data use while preserving service quality and advertising efficiency. Structural restrictions on data integration, including caps and data siloing, increase welfare only when privacy costs outweigh quality and advertising gains. Data sharing requirements have context-dependent effects through competition and investment incentives. The framework clarifies when consumer choice dominates structural intervention in data-driven digital markets.

Keywords: digital platforms, personalized advertising, privacy, regulation, data
JEL Classification: D18, D61, K24, L12, L51, L86

---

*This paper was previously circulated under the title "Breaking Up a Digital Monopoly".
I am grateful to Bruno Carballa Smichowski, Anatole Cheysson, Francesco Decarolis, Alexandre de Cornière, Néstor Duch-Brown, Chiara Farronato, Justus Haucap, Paul Heidhues, Doh-Shin Jeon, Rafael Jiménez-Durán, Reinhold Kesler, Laura Lasio, Yassine Lefouili, Luca Macedoni, Alexander MacKay, Leonardo Madio, Andrea Mantovani, Klaus Miller, Marcella Nicolini, Hans-Theo Normann, Alexander Rasch, Carlo Reggiani, Ivan Rendo, Andrew Rhodes, André Romahn, Michelangelo Rossi, Ananya Sen, Maximilian Schaefer, Shiva Shekar, Volker Stocker, Jean Tirole, Marshall Van Alstyne, Mo Xiao, Pinar Yildirim, and Bobby Zhou for helpful comments. All errors remain mine.

1 Introduction

Digital platforms increasingly rely on extensive user data to enhance service quality, personalize content, and improve advertising efficiency. While data integration, defined as the extent to which platforms collect, aggregate, and leverage user data across distinct services or products, delivers substantial benefits for users and advertisers, it also imposes privacy costs on individuals who value control over their personal information, particularly when data are combined across services. These opposing forces place privacy regulation and competition policy at the center of debates on digital market governance.

The analysis develops a formal model in which platforms select data integration levels under advertising revenues and horizontal competition, while consumers differ in their sensitivity to personalization and privacy. Consumers can opt out of personalized advertising at a cost, and this choice shapes platform demand, advertising value, and the effective use of data. The framework captures strategic interactions between platforms, heterogeneous consumer responses, and the welfare effects of regulatory tools such as limits on data integration, data siloing requirements, mandatory data sharing, and reductions in opt-out frictions.

This paper makes three main contributions. First, it provides a unified framework that jointly considers quality improvements, privacy costs, and advertising externalities, allowing data integration to affect welfare through multiple channels, including cross-service data use. Second, it distinguishes between structural policies that restrict the level or scope of data integration and choice-based policies that govern the ease with which consumers can avoid personalization. The model shows that these instruments operate through distinct mechanisms and can generate sharply different welfare outcomes. Third, it endogenizes platform investment in data integration, highlighting how incumbency advantages, competitive pressures, and externalities can lead to either under- or overinvestment relative to the social optimum.

The main results establish a policy ranking. Reducing opt-out costs and preventing strategic degradation of service quality for consumers who refuse personalization yields robust welfare gains. These measures empower privacy-sensitive users while preserving quality and advertising benefits for others. By contrast, structural interventions, such as forced reductions in data integration or strict data siloing, improve welfare only when privacy costs are large relative to quality gains, the share of privacy-sensitive consumers is high, and advertising-side benefits are limited. Outside these conditions, such interventions can reduce welfare by curtailing efficiency-enhancing uses of data.

When platforms choose data integration endogenously, investment incentives become central. A leading platform may overinvest in data to expand market share or underinvest when it fails to internalize the full social value of quality and advertising improvements. Data sharing mandates, integration caps, and siloing requirements therefore have ambiguous effects: they may stimulate competition and reduce concentration, but they can also weaken incentives to invest and innovate. The net welfare impact depends on market asymmetries, sources of incumbency advantage, and the balance between static efficiency gains and dynamic investment distortions.

Overall, the analysis provides a structured approach to evaluating regulatory interventions in data-driven digital markets. It clarifies when policies that enhance consumer choice outperform structural remedies and when more invasive interventions may be justified. By formalizing the trade-offs between privacy, quality, competition, and investment, the framework offers guidance for designing regulation that protects consumers without undermining the productive role of data integration.

The Regulatory Landscape. Regulators worldwide are intensifying efforts to limit the market power of digital platforms, reflecting concerns over privacy, competition, and the concentration of data. As platforms increasingly integrate their services, they accumulate vast amounts of user information, raising barriers to entry, reinforcing market dominance, and enabling incumbents to leverage data across multiple markets. This convergence of antitrust and digital policy signals a fundamental reassessment of how competition law should address the distinctive features of data-driven ecosystems.

In November 2024, the U.S. Department of Justice found that Google's integration of Chrome and Search violated Section 2 of the Sherman Act by unlawfully maintaining monopolies in general search and search advertising. The DOJ initially requested the divestiture of Chrome to dismantle the bundled monopoly and curb Google's ability to collect and exploit vast amounts of user data.[1] Independent engineering assessments indicate that a divested Chrome could continue serving its existing user base of four billion people, with a browser comparable to Microsoft Edge, Apple Safari, and Mozilla Firefox.[2] On March 15, 2025, the DOJ's initial proposal was approved with further remedies, including a contingent requirement to divest Android if the initial measures failed to restore meaningful competition.[3] On September 9, 2025, the Court imposed behavioral requirements instead of structural solutions, banning exclusive search deals for default placement and requiring Google to share search results data with rivals on "commercially reasonable terms."[4] Chrome represents a privileged access point to Google Search, which Alphabet uses free of charge while paying other providers, including over \$20 billion annually to Apple by 2024, to set Google Search as the default.[5] Google leverages data from one market to steer users toward other products, including YouTube, Google Ads, Android, Gemini, and Google Maps (Heidhues et al., 2024).

The Google case exemplifies a broader global trend of regulatory scrutiny of major technology platforms. In the European Union, the Digital Markets Act (DMA) establishes a comprehensive framework for regulating digital gatekeepers. Apple has faced investigations over App Store rules and default-browser restrictions, with formal non-compliance findings issued in June 2024.[6] Meta received DMA non-compliance findings in July 2024 for its "pay-or-consent" model,[7] and Microsoft is under investigation for bundling Teams with Office.[8] In

[1]United States v. Google LLC, No. 1:20-cv-03010 (D.D.C. 2024).

[2]See The Technical Feasibility of Divesting Google Chrome, Georgetown Center for Business and Public Policy (2024).

[3]United States District Court for the District of Columbia. (2025, March 7). Executive summary of plaintiffs' revised proposed final judgment (RPFJ), Case No. 1:20-cv-03010-APM. Available at justice.gov.

[4]Judge Mehta noted that "the emergence of generative AI has changed the course of this case," observing that AI companies are now "better placed to compete with Google than any search engine developer has been in decades." See The Verge, 2025.

[5]See Reuters, 2023.

[6]European Commission, Press Release, "DMA: Commission opens non-compliance investigations against Apple," June 24, 2024. Available at ec.europa.eu.

[7]European Commission, Press Release, "DMA: Commission finds Meta's 'pay or consent' model in breach of the DMA," July 1, 2024. Available at ec.europa.eu.

[8]See European Commission case AT.40627-Microsoft/Teams. Available at ec.europa.eu.

the United States, the FTC continues challenging Meta's acquisitions of Instagram and WhatsApp as unlawful monopolization,[9] while Amazon faces investigations regarding marketplace self-preferencing.[10] Authorities in Germany, Japan, South Korea, India, Canada, Australia, and Brazil are pursuing parallel investigations into platform dominance.[11]

This coordinated enforcement demonstrates that traditional antitrust tools require adaptation to the unique features of digital markets, including network effects, multi-sided platform dynamics, data advantages, and the ability to leverage dominance across complementary services. The regulatory question is no longer whether intervention is necessary, but how to design policies that promote competition and protect consumer welfare while preserving the efficiency and quality benefits of data integration.

Structure. Section 2 reviews the related literature and clarifies the paper's contributions. Section 3 introduces the baseline model with exogenous data integration and endogenous consumer opt-out decisions. Section 4 characterizes equilibrium platform behavior and welfare across alternative opt-out regimes, highlighting the trade-off between quality gains and privacy costs. Section 5 studies the welfare effects of regulatory interventions that directly restrict platform data integration. Section 6 then endogenizes data integration choices, identifying investment distortions and comparing the private equilibrium with the social optimum. Section 7 analyzes alternative policy tools, contrasting measures that constrain how platforms can combine or leverage data (such as integration caps, siloing requirements, and data sharing mandates) with interventions that operate through consumer choice by lowering opt-out frictions. Section 8 distills the theoretical findings into a policy ranking and practical guidance for regulators. Section 9 concludes. All proofs are presented in the Appendix.

2 Related Literature

The paper contributes to several strands of literature on digital platforms, data, and privacy.

A growing body of research investigates the economic, regulatory, and privacy implications of data use on digital platforms. Theoretical contributions analyze how platforms leverage user data to enhance product quality, target advertising, and extract rents under heterogeneous consumer preferences and regulatory constraints (de Cornière and Taylor (2024); Rhodes and Zhou (2024); Lefouili et al. (2024); Bhargava et al. (2025); Bisceglia et al. (2025); Hovenkamp (2024); Krämer and Shekhar (2024); van de Waerdt (2023)), whereas empirical studies document the effects of data-related policies and regulatory interventions on market outcomes and consumer welfare (Allcott et al. (2025); Decarolis et al. (2025); Pape and Rossi (2026); Kesler et al. (2025)). This paper contributes to this strand by integrating consumer heterogeneity and endogenous platform investment into a unified framework for evaluating regulatory interventions.

Acquisti et al. (2016) and Goldfarb and Que (2023) provide comprehensive reviews of the economics of privacy, summarizing evidence on consumer behavior and regulatory frameworks.

[9]FTC v. Meta Platforms, Inc., No. 1:20-cv-03590 (D.D.C.). See FTC Press Release, 2021.

[10]FTC v. Amazon.com, Inc., No. 2:23-cv-01495 (W.D. Wash. filed September 26, 2023); European Commission preliminary investigation under Article 102 TFEU initiated in November 2020.

[11]For Germany, see Bundeskartellamt at bundeskartellamt.de. International coverage available at Competition Policy International and The Verge Tech.

Theoretical work investigates how privacy preferences shape firm incentives, pricing, and advertising strategies (Casadesus-Masanell and Hervas-Drane (2015); Johnson (2013); Rhodes and Zhou (2024); Marotta et al. (2022); Choe et al. (2025)), while empirical studies document how consumers trade off privacy for platform benefits and reveal heterogeneity in responses across populations (D'Annunzio and Menichelli (2022); Benndorf and Normann (2018); Tucker (2011)). This paper extends this literature by evaluating the welfare implications of regulatory interventions that affect the trade-offs between privacy costs and data-driven benefits.

The paper is also related to the literature on data-driven product improvement, which shows that user data can enhance algorithmic recommendations, predictive accuracy, and overall product quality (Klein et al. (2025); Schaefer and Sapi (2023); Ullrich et al. (2024); Hagiu and Wright (2023)). It further connects to research on digital ecosystems, in which platforms operate across interlinked markets and leverage cross-market interactions to influence competitive dynamics (Heidhues et al. (2024); Condorelli and Padilla (2023); Rhodes et al. (2025)). The present analysis contributes to these strands by examining how regulatory constraints on data sharing and integration shape platform incentives for quality improvement and innovation.

Foundational work on two-sided platform competition (Rochet and Tirole (2003)) underpins much of the current understanding of digital markets. This paper builds on these insights by explicitly modeling the interplay between data-driven quality, privacy costs, and regulatory interventions across multiple platform markets, providing a framework to guide policy design and the evaluation of regulatory trade-offs in digital platform ecosystems.

3 Baseline Model

3.1 Setup

Two platforms $j \in \{G, F\}$ compete for users and advertisers. Each of them has a certain data integration level d_j, which represents the extent to which they collect, aggregate, and leverage user data across distinct services or products to provide higher quality and better personalization. Larger values of d_j reflect more comprehensive data collection and integration.

Assumption 1. *Platform G has a strictly higher level of data integration than platform F, i.e.* $d_G > d_F$.

As end-users access the platform's integrated services at zero monetary cost, its profit is generated entirely from charging advertisers for access to its user base. Importantly, the data collected is not only used to target advertising but also to enhance the quality of other services, such as search results or content recommendations.[12]

Consumers are uniformly distributed along a Hotelling line $x \in [0, 1]$, with platform G at $\ell_G = 0$ and platform F at $\ell_F = 1$,[13] and differ in their sensitivity to personalized advertising. A fraction θ are customization-savvy users, deriving marginal utility $\gamma \geq 0$ from personalized ads[14]

[12] Carballa-Smichowski et al. (2025) introduce the concept of economies of scope in data reuse, referring to the benefits derived from using the same dataset for multiple purposes.

[13] Horizontal distance generates a transportation cost $t > 0$, capturing differences in consumer preferences for platform interfaces, brand loyalty, and complementary services unrelated to data integration capabilities.

[14] The assumption of a constant marginal benefit from personalization is made for analytical tractability. The main qualitative results would continue to hold if customization-savvy consumers derived a concave benefit $u(\gamma)$,

with negligible privacy costs (normalized to zero),[15] while the remaining fraction $(1-\theta)$ are privacy-savvy users, incurring privacy cost $\delta > 0$ from being exposed to personalized advertising.

Privacy-savvy consumers can opt out of personalized advertising by paying a cost $c \geq 0$.[16]

Data integration level d_j affects consumer welfare via two channels: (i) enhancing service quality, through better matching, recommendations, and platform functionality, providing a benefit $\alpha > 0$ per unit of data integration to all consumers; (ii) enabling more precise personalized advertising, which benefits customization-savvy users but imposes costs on privacy-savvy users, depending on their opt-out choices.

Assumption 2 (Net Value of Data Integration)**.** *The net value of data integration ξ depends on the consumer's privacy type and, for privacy-savvy users, on the opt-out decision. Specifically, for customization-savvy users: $\xi_C = \alpha + \gamma$; for privacy-savvy users who accept personalized advertising: $\xi_P^A = \alpha - \delta$; for privacy-savvy users who opt out: $\xi_P^O(d_j) = \alpha - \frac{c}{d_j}$.*

A consumer located at x who chooses platform j obtains utility:

$$U_{i,j}(x) = v + \xi_i d_j - t|x - \ell_j|,$$

where $i \in \{C, P^A, P^O\}$ indexes the relevant privacy type and opt-out decision and $v > 0$ denotes the baseline utility from using any platform, i.e. the intrinsic value of platform access before accounting for data integration or horizontal differentiation.

The opt-out decision is endogenous and depends on the level of data integration d_j: higher data integration increases privacy costs, making opt-out more attractive for privacy-savvy consumers. This, in turn, affects the value of targeted advertising. Each platform charges advertisers a price per impression p_j, and advertisers are heterogeneous in their willingness to pay, which depends on d_j.

An advertiser of type μ derives value $\mu\beta d_j$ from an impression on platform j, where $\beta > 0$ measures the marginal value of data integration. Advertiser types are uniformly distributed on $[0, \bar{\mu}]$. An advertiser of type μ purchases impressions on platform j if and only if $\mu\beta d_j \geq p_j$, implying that all advertisers with $\mu \geq p_j/(\beta d_j)$ participate. The mass of participating advertisers is therefore:

$$A_j(p_j, d_j) = \max\left\{0, \bar{\mu} - \frac{p_j}{\beta d_j}\right\}$$

Advertiser demand is decreasing in the price p_j and increasing in data integration d_j: higher prices exclude marginal advertisers, while greater data integration increases the value of targeting and attracts more advertisers at any given price.

with $u'(\gamma) > 0$ and $u''(\gamma) < 0$, from the intensity of personalization γ. In that case, diminishing returns to personalization would affect equilibrium thresholds but would not alter the key trade-offs identified in the model. In the extreme case where $\gamma = 0$, these consumers experience no direct benefit from personalized advertising beyond the general quality improvements captured by α, effectively reducing the model to one in which all consumers value only the quality enhancement channel of data integration.

[15] Following Choe et al. (2025), this normalization simplifies the analysis without loss of generality.

[16] The opt-out cost c encompasses both explicit costs, such as subscription fees for ad-free versions or premium tiers, and implicit costs, including the time and cognitive effort required to navigate privacy settings, reduced service functionality or degraded user experience. In practice, many platforms impose substantial friction on opt-out mechanisms through complex privacy interfaces, repeated consent requests, or feature restrictions that effectively raise the cost of avoiding personalized advertising. For tractability, I will focus only on implicit opt-out costs.

Platform j delivers one advertisement to each of its N_j consumers. Advertising revenue is therefore given by:

$$R_j(p_j, d_j, N_j) = p_j \cdot A_j(p_j, d_j) \cdot N_j.$$

A larger consumer base increases revenue by providing more impressions, while higher data integration raises the value of each impression, allowing the platform to either charge higher prices or attract more advertisers. As a monopolist on its platform, firm j faces the standard trade-off between a higher price per advertiser and a lower mass of participating advertisers.

Advertisers pay for impressions on all users, including those who opt out. The value per impression, $\mu\beta d_j$, depends on aggregate data integration rather than on individual-level targeting precision.[17]

Timing. The model unfolds in two stages. Initially, platforms are endowed with fixed data integration levels, with platform G enjoying a data advantage $d_G > d_F$ due to historical investments.[18]

In the first stage, consumers observe both platforms' data integration levels and choose a platform to maximize their utility. Privacy-savvy consumers simultaneously decide whether to opt out of personalized advertising by comparing the privacy cost δd_j to the opt-out cost c. These choices determine the distribution of consumers across platforms, yielding market shares N_G and N_F.

In the second stage, platforms set per-impression advertising prices p_G and p_F to maximize revenue given their consumer bases. Each platform balances the trade-off between higher revenue per advertiser and the resulting number of participating advertisers. Advertisers then decide whether to purchase impressions according to their type-dependent valuations, and the resulting advertiser masses determine platform revenues.

4 Analysis

Equilibrium and Welfare

Advertising Pricing. Platform $j \in \{G, F\}$ chooses the per-impression price p_j to maximize advertising revenue:

$$\max_{p_j} R_j(p_j, d_j, N_j) = p_j \cdot A_j(p_j, d_j) \cdot N_j = p_j \cdot \left(\bar{\mu} - \frac{p_j}{\beta d_j}\right) \cdot N_j$$

The optimal price is linear in data integration:

$$p_j^* = \frac{\bar{\mu}\beta d_j}{2} \tag{1}$$

[17] This specification reflects that: (i) advertising deals often involve bulk impression purchases where individual opt-out status is not perfectly observable ex-ante, and (ii) even ads shown to opt-out users benefit from platform-level capabilities (e.g. contextual targeting) that depend on d_j. An alternative where advertiser value is zero for opt-out users would create direct revenue incentives for platforms to discourage opt-out through high costs c or strategic quality degradation λ (Appendix B), strengthening all results regarding opt-out cost reduction (Section 8).

[18] Treating d_G and d_F as exogenous allows me to focus on the welfare effects of regulatory interventions that reduce d_G. Section 6 extends the analysis to the case where data integration levels are endogenous.

as higher data integration allows platforms to charge higher prices per impression because better targeting makes impressions more valuable to advertisers. Although higher data integration raises the value of impressions, platforms increase prices proportionally, so advertiser participation remains unchanged at $A_j^* = \bar{\mu}/2$. Platform revenue is given by:

$$R_j^* = \frac{\bar{\mu}^2 \beta d_j N_j}{4}$$

which is proportional to both data integration and consumer base size. Specifically, the term $d_j N_j$ captures a complementarity between data integration and scale: data are more valuable on platforms with a larger user base, and, conversely, a larger user base generates higher revenues when combined with more advanced data integration.

Consumer Platform Choice. Consumers observe both platforms' data integration levels and choose which to join. Since platform G has higher data integration, privacy-savvy consumers face greater privacy costs on G, creating asymmetric opt-out behavior.

To ensure that both platforms retain strictly positive market shares in equilibrium, thereby ruling out corner solutions and making all three cases well-defined and meaningful, the following condition is imposed:

Assumption 3 (Interior Market Shares)**.** *The data integration gap between platforms is sufficiently small to guarantee interior market shares, i.e.* $(d_G - d_F) < \frac{t}{\max\{\xi_C, \alpha, |\xi_P^A|\}}$.

Three regimes emerge based on the opt-out cost c relative to privacy costs δd_F and δd_G:

1. **Universal Opt-Out** ($c < \delta d_F < \delta d_G$). When the opt-out cost is sufficiently low, privacy-savvy consumers find it optimal to opt out of personalized advertising on both platforms. The privacy cost from personalized advertising exceeds the opt-out cost regardless of which platform they choose. The equilibrium market shares of platform G and F are therefore given by:

$$N_G = \frac{1}{2} + \frac{d_G - d_F}{2t}(\alpha + \theta\gamma) \tag{2}$$

   and

$$N_F = 1 - N_G = \frac{1}{2} - \frac{d_G - d_F}{2t}(\alpha + \theta\gamma) \tag{3}$$

   Here, $(\alpha + \theta\gamma)$ captures the population-weighted average marginal benefit from data integration. As a result, platform G's market share exceeds one-half whenever $d_G > d_F$, with the advantage increasing in the data integration gap, the quality effect α, and the share of customization-savvy consumers θ weighted by their personalization benefit γ.

2. **Universal Acceptance** ($c > \delta d_G > \delta d_F$). When the opt-out cost exceeds even the privacy cost on the high-integration platform, privacy-savvy consumers find it optimal to accept personalized advertising on both platforms. The opt-out cost is too high relative to the privacy harm, making it preferable to bear the privacy cost rather than incur the

opt-out cost. Platforms' market shares are:

$$N_G = \frac{1}{2} + \frac{d_G - d_F}{2t}(\alpha + \theta\gamma - (1-\theta)\delta) \tag{4}$$

$$N_F = \frac{1}{2} - \frac{d_G - d_F}{2t}(\alpha + \theta\gamma - (1-\theta)\delta) \tag{5}$$

The market share advantage is reduced relative to the Universal Opt-Out regime because privacy-savvy consumers now bear privacy costs.

3. **Asymmetric Regime** ($\delta d_F < c < \delta d_G$). When the opt-out cost falls in this intermediate range, privacy-savvy consumers opt out on G but accept personalized advertising on F. The equilibrium market shares are:

$$N_G = \frac{1}{2} + \frac{1}{2t}\left[\theta\xi_C(d_G - d_F) + (1-\theta)(\alpha(d_G - d_F) + \delta d_F - c)\right] \tag{6}$$

$$N_F = \frac{1}{2} - \frac{1}{2t}\left[\theta\xi_C(d_G - d_F) + (1-\theta)\big(\alpha(d_G - d_F) + \delta d_F - c\big)\right] \tag{7}$$

Platform G's data advantage attracts consumers through quality improvements but simultaneously imposes higher privacy costs that may drive privacy-savvy consumers to opt out or switch to platform F.

For notational convenience, the three regimes will be referred to as Case 1 (Universal Opt-Out), Case 2 (Universal Acceptance), and Case 3 (Asymmetric Regime) in all subsequent sections.

Lemma 1 (Equilibrium). *The subgame perfect Nash equilibrium consists of:*

1. *advertising prices* (p_G^*, p_F^*) *given by equation* (1) *for each platform;*
2. *consumer platform choices yielding market shares* (N_G, N_F) *as in equations* (2)–(7)*, depending on the prevailing opt-out regime;*
3. *opt-out decisions by privacy-savvy consumers based on comparing* c *to* δd_j *for each platform.*

Proof. See Appendix A.1. □

Lemma 2 (Welfare). *The welfare impact of data integration is ambiguous. Higher data integration increases quality benefits* αd_j *for all consumers, personalization benefits* γd_j *for customization-savvy consumers, and advertising-side surplus through improved targeting. At the same time, it increases privacy costs* δd_j *for privacy-savvy consumers who accept personalized advertising.*

Proof. See Appendix A.2. □

Lemma 2 establishes that the welfare effects of data integration are theoretically ambiguous, reflecting a fundamental trade-off between efficiency gains and privacy costs. This ambiguity creates scope for welfare-enhancing regulation that limits data integration under certain market conditions. The next section therefore analyzes a regulatory intervention that reduces platform G's data integration level and characterizes its welfare implications through both direct and market-mediated channels.

5 Data Integration Reduction

I consider a regulatory intervention that exogenously lowers platform G's data integration level from d_G to $d'_G < d_G$, holding d_F fixed. The resulting welfare effects are decomposed into direct effects, evaluated at given market shares, and indirect effects operating through equilibrium market share adjustments.

The direct effects on consumer surplus depend on the composition of consumers on platform G and whether they choose to opt out. Customization-savvy consumers lose $\xi_C = \alpha + \gamma$ per unit reduction in d_G, reflecting the combined loss of quality and personalization benefits. Privacy-savvy consumers who opt out lose only the quality benefit α, whereas those who accept personalized advertising experience a net effect $\xi_P^A = \alpha - \delta$, which may be positive or negative depending on the relative magnitude of quality and privacy costs.

The advertising side is always negatively affected by reductions in data integration. Platform profits and advertiser surplus both depend on the data-weighted consumer base $d_G N_G + d_F N_F$. Reducing d_G decreases this base directly, by N_G, and indirectly, as some consumers switch from platform G to F.

The market share response to changes in d_G differs across regimes. In Cases 1 and 3, $\frac{dN_G}{dd_G} = \frac{\alpha+\theta\gamma}{2t} > 0$, while in Case 2, $\frac{dN_G}{dd_G} = \frac{\alpha+\theta\gamma-(1-\theta)\delta}{2t}$, which can be positive or negative depending on whether the combined quality and personalization benefits outweigh the privacy costs.

Taken together, these direct and indirect effects suggest that the welfare consequences of reducing d_G depend crucially on the opt-out regime. Proposition 1 formalizes this, highlighting when total welfare unambiguously decreases and when it may increase.

Proposition 1 (Welfare Effects of Data Integration Reduction)**.** *Consider a reduction in platform G's data integration level, d_G. The welfare effect depends on the opt-out regime:*

1. ***Universal Opt-Out:*** *total welfare unambiguously decreases. Customization-savvy consumers lose, the advertising side is harmed by reduced targeting, and privacy-savvy consumers are shielded by opt-out. The loss arises from both direct reductions in consumer surplus and advertising efficiency, as well as indirect losses from consumer switching.*

2. ***Universal Acceptance:*** *the effect on total welfare is ambiguous. Formally, welfare increases if and only if privacy costs δ are large relative to quality benefits α, privacy-savvy consumers are numerous, advertising benefits $\bar{\mu}^2\beta$ are small, the data integration gap $d_G - d_F$ is small, and transportation costs t are large.*

3. ***Asymmetric Regime:*** *total welfare unambiguously decreases. Privacy-savvy consumers on G opt out, so welfare losses come primarily from reduced consumer surplus among customization-savvy users and diminished advertising efficiency. Both direct and indirect channels contribute to the loss.*

Proof. See Appendix A.3. □

The key insight is that reducing data integration is welfare-improving only in the universal acceptance regime when privacy costs are sufficiently large. In all other regimes, maximal data integration is optimal from a total welfare perspective. This establishes that structural remedies forcing integration reduction face a high bar for justification. While total welfare may decrease with data integration reduction, the distribution of surplus changes significantly. Customization-savvy consumers unambiguously lose from reducing d_G as they benefit from both quality and personalization without offsetting privacy costs. Privacy-savvy consumers lose in the universal opt-out and asymmetric regimes but may gain in the universal acceptance regime if privacy costs exceed quality benefits ($\delta > \alpha$). Platforms and advertisers always lose from reduced targeting precision.

The regulatory intervention creates a transfer from customization-savvy consumers and the advertising side to privacy-savvy consumers, but only when $c > \delta d_G$ and $\delta > \alpha$. Policies protecting privacy-savvy consumers may harm customization-savvy consumers and reduce advertising efficiency, with the net welfare effect depending on the relative magnitudes and the distribution of consumer types.

6 Endogenous Data Integration

The preceding analysis shows how welfare and the distribution of surplus depend on platform G's data integration level, d_G, and on the opt-out regime. In the baseline framework, data integration is treated as exogenously fixed, capturing historical investments and incumbency advantages. In practice, however, platforms choose their data integration levels strategically, and these choices depend on their initial market positions.

Understanding how incumbency shapes investment incentives is therefore essential for evaluating regulatory interventions as effective policy must account not only for static welfare effects, but also for dynamic responses arising from competitive asymmetries.

Data integration is thus modeled as a costly investment undertaken prior to consumer platform selection and advertising pricing. A key asymmetry arises from platform G's incumbency advantage, which lowers its marginal cost of data integration and reflects accumulated expertise, established data infrastructure, and economies of scale from past investments.

Two departures from the baseline framework are introduced. First, platforms face convex investment costs when expanding data infrastructure, capturing the increasing marginal difficulty of integrating additional services, maintaining system-wide data consistency, and ensuring privacy compliance at scale. Platform G, however, benefits from lower marginal costs due to its incumbent position. Second, the benefits of data integration exhibit diminishing returns: each additional unit yields smaller incremental improvements in service quality and targeting precision. Together, these features generate an asymmetric equilibrium that mirrors observed platform competition, in which dominant firms exploit data advantages to sustain market leadership, while rivals face higher costs in attempting to narrow the integration gap.

6.1 Model Setup

The extended framework is formalized below, specifying costs, quality, advertising value, and consumer utility as functions of data integration levels and opt-out decisions.

Investment Cost Function with Incumbency Advantage. The cost of achieving data integration level d_j for platform $j \in \{G, F\}$ is given by:

$$C_j(d_j) = \frac{\kappa_j d_j^{1+\eta}}{1+\eta}$$

where $\kappa_j > 0$ represents platform j's cost parameter and $\eta > 0$ captures the degree of convexity. The marginal cost is $C_j'(d_j) = \kappa_j d_j^{\eta}$, which is strictly increasing in d_j.

Assumption 4. *Platform G has an incumbency advantage, i.e.* $\kappa_G < \kappa_F$.

Quality Production with Decreasing Returns. The quality benefit from data integration exhibits decreasing marginal returns: $Q(d_j) = \alpha d_j^{\phi}$,[19] where $\phi \in (0, 1)$ captures the elasticity of quality with respect to data integration, with marginal quality improvement $Q'(d_j) = \alpha\phi d_j^{\phi-1}$.

Advertising Value Function. Advertiser willingness-to-pay for impressions exhibits decreasing returns to targeting precision: $V(d_j) = \beta d_j^{\psi}$, where $\psi \in (0, 1)$ represents the elasticity of advertising value, with marginal value $V'(d_j) = \beta\psi d_j^{\psi-1}$. The baseline model corresponds to $\psi = 1$.

Consumer Utility. Depending on the type and opt-out decision, consumer utility is given by $U_{C,j}(x) = v + \xi_C \alpha d_j^{\phi} - t|x - \ell_j|$, $U_{P,j}^{O}(x) = v + \alpha d_j^{\phi} - c - t|x - \ell_j|$, and $U_{P,j}^{A}(x) = v + \alpha d_j^{\phi} - \delta d_j - t|x - \ell_j|$. Privacy costs remain linear in d_j while quality benefits exhibit diminishing returns, creating a potential interior optimum where marginal quality gains equal marginal privacy costs.

Timing. The model adds an initial stage in which platforms simultaneously choose their data integration levels (d_G, d_F) to maximize expected profits, anticipating subsequent consumer and advertising market outcomes. The subsequent stages proceed as in the baseline model.

6.2 Asymmetric Equilibrium

Lemma 3 (Asymmetric Equilibrium with Incumbency Advantage)**.** *Consider the model with endogenous data integration and an incumbency advantage. For a given privacy regime:*

1. *Optimal advertising prices are*

$$p_j^* = \frac{\bar{\mu}\beta d_j^{\psi}}{2}, \quad j \in \{G, F\}.$$

2. *Equilibrium data integration levels* (d_G^*, d_F^*) *satisfy*

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi (d_G^*)^{\psi-1} N_G + (d_G^*)^{\psi} \frac{\partial N_G}{\partial d_G}\right] = \kappa_G (d_G^*)^{\eta}, \tag{8}$$

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi (d_F^*)^{\psi-1} N_F + (d_F^*)^{\psi} \frac{\partial N_F}{\partial d_F}\right] = \kappa_F (d_F^*)^{\eta}. \tag{9}$$

[19] When $\phi = 1$, this reduces to the baseline linear specification $Q(d_j) = \alpha d_j$.

3. If $\kappa_G < \kappa_F$ and $\eta > \psi - 1$, there exists a unique interior asymmetric equilibrium with

$$d_G^* > d_F^* > 0.$$

Proof. See Appendix A.4. □

In equilibrium, platform G invests more in data integration than platform F, reflecting its marginal cost advantage. This creates a self-reinforcing dynamic: greater data integration improves service quality, which attracts more consumers, raises advertising revenue, and in turn sustains higher investment. By contrast, platform F faces a trade-off between matching G's investment at a higher marginal cost or accommodating a smaller market share.

The equilibrium data integration gap, $d_G^* - d_F^*$, is increasing in the cost asymmetry $(\kappa_F - \kappa_G)$ and in the advertising value parameter β, and decreasing in t. The share of customization-savvy consumers, θ, further amplifies the marginal return to quality improvements, thereby widening the integration gap.

6.3 Market Failures

The asymmetric equilibrium characterized in Lemma 3 shows how incumbency advantages distort investment incentives and create barriers to entry. These distortions give rise to multiple sources of inefficiency, summarized in the following proposition.

Proposition 2 (Market Failures with Incumbency Advantage)**.** *The private equilibrium (d_G^*, d_F^*) deviates from the social optimum (d_G^{SP}, d_F^{SP}) due to four externalities:*

1. *Business stealing: private gains from attracting marginal consumers represent transfers rather than social surplus.*

2. *Advertising externality: private investment captures only a fraction of total advertising value.*

3. *Consumer externality: quality benefits to infra-marginal consumers or rival users are not internalized.*

4. *Incumbency amplification: cost advantages create excessive concentration, making the data integration gap reflect both efficiency and strategic positioning.*

Proof. See Appendix A.5. □

The incumbent's lower costs generate efficiency gains, improving quality, but may also reduce competition if the rival cannot achieve sufficient scale. Investment distortions differ across firms: the incumbent may under- or overinvest depending on advertising returns and competitive intensity, while the rival's response depends on the relative size of cost asymmetries and market share gaps. Optimal policy should preserve efficiency advantages while mitigating distortions from excessive concentration.

7 Alternative Remedies

The preceding analysis shows that private platforms may under- or overinvest in data integration due to advertising externalities, competitive pressures, and incumbency advantages.

The incumbent's cost advantage generates an asymmetric equilibrium in which the dominant platform maintains superior data integration, potentially foreclosing competition and reducing consumer welfare. These distortions motivate regulatory interventions aimed at leveling the competitive playing field while preserving efficiency gains. This section examines potential remedies and their interaction with the strategic asymmetries identified in Section 6.

7.1 Regulatory Ceilings on Data Integration

One policy response to the asymmetric equilibrium described above is to limit the incumbent's data integration through a regulatory ceiling $\bar{d}_G < d_G^*$.

By constraining G's investment, the competitive gap with the rival is reduced, which can alter both platforms' incentives. Specifically, when $d_G = \bar{d}_G$, the rival adjusts its integration optimally according to its first-order condition (9).

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi d_F^{\psi-1}N_F(\bar{d}_G, d_F) + d_F^{\psi}\frac{\partial N_F}{\partial d_F}\right] = \kappa_F d_F^{\eta}.$$

Applying the implicit function theorem yields the strategic response:

$$\frac{dd_F^*}{d\bar{d}_G} = -\frac{\frac{\partial^2 \Pi_F}{\partial d_F \partial d_G}}{\frac{\partial^2 \Pi_F}{\partial d_F^2}}.$$

This expression reflects two effects: higher d_G reduces N_F, lowering marginal advertising revenue, while a larger integration gap simultaneously strengthens F's incentive to invest. At the asymmetric equilibrium, these forces can offset each other. When $\bar{d}_G$ is near d_G^*, reducing G's advantage typically boosts F's investment, whereas when $\bar{d}_G$ is substantially below d_G^*, F may instead scale back its investment.

Total welfare under the ceiling is:

$$W(\bar{d}_G) = CS(\bar{d}_G, d_F^*(\bar{d}_G)) + \frac{3\bar{\mu}^2\beta}{8}\left[\bar{d}_G^{\psi}N_G + (d_F^*(\bar{d}_G))^{\psi}N_F\right] - C_G(\bar{d}_G) - C_F(d_F^*(\bar{d}_G))$$

A necessary condition for welfare improvement is

$$\left.\frac{dW}{d\bar{d}_G}\right|_{\bar{d}_G = d_G^*} < 0.$$

The ceiling improves welfare if reductions in the incumbent's quality and advertising distortions, together with any beneficial response from the rival, outweigh the direct loss from limiting d_G. Conversely, it can reduce welfare when the incumbent underinvests, quality and advertising gains are large, or the rival decreases investment.

Overall, the effect of a ceiling is context-dependent, shaped by the magnitude and direction of pre-existing investment distortions, the strategic interaction between platforms, the trade-off between privacy costs and quality gains, and the relative incumbency advantage $\frac{\kappa_F}{\kappa_G}$.

7.2 Data Sharing and Siloing Policies

While regulatory ceilings limit the incumbent's data integration, mandatory data sharing narrows the competitive gap by allowing the rival to access a fraction $\sigma \in (0, 1]$ of the incumbent's data. Under sharing, the rival's effective integration is:

$$\tilde{d}_F = d_F + \sigma d_G,$$

and the incumbent's market share adjusts accordingly:

$$N_G = \frac{1}{2} + \frac{\alpha\bar{\xi}\left[d_G^{\phi} - \tilde{d}_F^{\phi}\right]}{2t}.$$

The platforms' FOCs become:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi d_G^{\psi-1} N_G + d_G^{\psi}\frac{\partial N_G}{\partial d_G}\right] = \kappa_G d_G^{\eta},$$
$$\frac{\bar{\mu}^2\beta}{4}\left[\psi \tilde{d}_F^{\psi-1} N_F + \tilde{d}_F^{\psi}\frac{\partial N_F}{\partial d_F}\right] = \kappa_F d_F^{\eta}.$$

Data sharing reduces the incumbent's marginal return to investment ($\partial d_G^S/\partial\sigma < 0$), while the rival's response is ambiguous, reflecting the balance between free-riding and competitive pressure. Total welfare is:

$$W^S(\sigma) = CS(d_G^S, d_F^S, \sigma) + \frac{3\bar{\mu}^2\beta}{8}\left[(d_G^S)^{\psi} N_G + (\tilde{d}_F^S)^{\psi} N_F\right] - C_G(d_G^S) - C_F(d_F^S).$$

Proposition 3 (Welfare Effects of Data Sharing)**.** *Under mandatory data sharing with fraction* σ*:*

1. *The incumbent reduces investment:* $\frac{\partial d_G^S}{\partial \sigma} < 0$.
2. *The rival's response is ambiguous:* $\frac{\partial d_F^S}{\partial \sigma}$ *depends on free-riding versus competitive pressure.*
3. *Welfare improves if the static gains from narrowing the integration gap and increasing competition outweigh the dynamic losses from reduced investment.*

Proof. See Appendix A.6. □

As with ceilings, the welfare impact of data sharing is context-dependent.

It is most beneficial when the incumbent overinvests due to strategic positioning and market concentration is high, as sharing corrects excessive investment and levels the playing field. Conversely, if the incumbent underinvests or the rival free-rides excessively, sharing can reduce welfare.

Data Siloing. An alternative regulatory approach imposes strict data siloing, prohibiting platforms from combining user data across distinct services or product lines. Unlike integration caps, which limit the overall level of data integration, siloing restricts its *scope* by eliminating cross-service complementarities.

Formally, suppose platform j operates K services indexed by $k \in \{1, \ldots, K\}$, each generating a service-specific stock of data d_{kj}. In the absence of restrictions, the platform can pool these datasets to achieve an effective level of data integration

$$d_j = \Gamma(d_{1j}, \ldots, d_{Kj}) = \sum_{k=1}^{K} \omega_k d_{kj} + \rho \sum_{k \neq \ell} d_{kj} d_{\ell j},$$

where $\omega_k \in [0, 1]$, $\sum_k \omega_k = 1$, captures the relative importance of each service, and $\rho > 0$ measures economies of scope from cross-service data reuse.

Under a siloing policy, cross-service data flows are prohibited and each dataset can be used only within its originating service. Effective data integration is therefore limited to

$$d_j^{\text{Sil}} = \sum_{k=1}^{K} \omega_k d_{kj},$$

eliminating the complementarity term $\rho \sum_{k \neq \ell} d_{kj} d_{\ell j}$.

Siloing reduces effective data integration through two channels. First, it weakens quality spillovers across services, lowering the impact of data on service quality through the αd_j channel. Second, it reduces advertisers' willingness to pay by limiting cross-context inference, thereby lowering advertising revenues through the βd_j channel. At the same time, siloing mitigates privacy harms arising from cross-service inference for privacy-sensitive users.

Proposition 4 (Welfare Effects of Data Siloing)**.** *Imposing strict data siloing reduces effective data integration by eliminating cross-service complementarities. Total welfare increases if and only if the marginal privacy harm from cross-service inference exceeds the marginal quality and advertising gains generated by scope economies.*

Proof. See Appendix A.7. □

7.3 Opt-Out Cost Reduction

Rather than constraining data integration, a complementary policy reduces the cost c of opting out, for instance through simple, transparent mechanisms, prohibiting dark patterns, or ensuring equivalent service quality for opt-out users.

In the asymmetric equilibrium with $d_G > d_F$, lowering c disproportionately benefits privacy-savvy consumers on the high-integration incumbent. A marginal reduction in c by ϵ shifts consumers toward opting out on G, increasing the data-weighted allocation:

$$\frac{\partial(d_G N_G + d_F N_F)}{\partial c} = -(d_G - d_F)(1 - \theta)\frac{1}{2t} < 0.$$

This intervention improves welfare through three channels simultaneously.

First, privacy-savvy consumers gain utility directly from lower opt-out friction. Second, consumers reallocate toward the higher-quality platform while opting out, improving matching efficiency. Third, advertising efficiency rises as the effective data-weighted user base expands, enhancing platform revenue without reducing quality benefits.

Proposition 5 (Dominance of Opt-Out Cost Reduction). *In the asymmetric equilibrium, reducing opt-out costs:*

1. *unambiguously increases welfare.*
2. *strictly dominates structural remedies: it improves consumer choice, preserves quality benefits, and enhances advertising efficiency without creating offsetting losses.*
3. *The welfare gains are largest when platforms differ substantially in data integration, privacy-savvy consumers are numerous, advertising benefits are high, and competition across platforms is weak.*

Proof. See Appendix A.8. □

Intuitively, opt-out cost reduction directly addresses the privacy externality while preserving efficiency gains from data integration. The asymmetry amplifies benefits: privacy-savvy users migrate to the high-integration platform, exercising their privacy preference without sacrificing quality, while advertisers benefit from improved targeting on the remaining user base. Unlike structural remedies, this approach delivers Pareto-improving outcomes across a wide range of parameter values.

8 Policy Recommendations

The analysis implies a hierarchy among regulatory instruments, based on their welfare performance across different market environments. Policies that expand effective consumer control over data use tend to deliver more robust welfare gains than interventions that mechanically restrict data integration. As shown in Proposition 5, reducing frictions in opting out directly addresses privacy externalities while preserving efficiency gains from data-driven improvements in service quality and advertising. When consumers can refuse personalization without experiencing degraded service, privacy-sensitive users sort according to their preferences, while users who value customization continue to benefit from richer data environments. By operating through choice rather than prohibition, this approach avoids the quality and targeting losses associated with direct restrictions on data use.

Regulatory attention should therefore focus on the *effective* accessibility of privacy options rather than their formal availability. Complex interfaces, repeated consent requests, or the withdrawal of core functionalities raise the implicit cost of opting out and undermine the disciplining role of user choice. Ensuring functional equivalence between personalized and non-personalized service versions, except where personalization is technically indispensable, limits platforms' incentives to impose strategic degradation and strengthens the sorting mechanism emphasized in the model. Such measures improve welfare across a wide range of market configurations by relaxing distortions without eliminating productive uses of data.

Structural interventions that directly restrict data integration, including integration caps and data siloing requirements, improve welfare only under more restrictive conditions. Proposition 1 and Proposition 4 show that such measures raise welfare when privacy harms from data use, particularly those arising from cross-service inference, are large relative to gains in service quality and advertising efficiency. In these environments, high levels of integration or extensive cross-service data reuse primarily magnify privacy costs without generating commensurate efficiency benefits. Outside these cases, structural restrictions can reduce welfare by weakening incentives to invest in data-driven quality improvements and advertising technologies.

Policies aimed at reshaping competitive dynamics, such as data-sharing mandates or integration limits applied to dominant platforms, must also account for their impact on investment incentives. Proposition 2 highlights externalities that can lead incumbents to overinvest or underinvest depending on whether advertising and consumer-surplus effects are internalized. As shown in Proposition 3, measures that narrow a leading platform's data advantage may stimulate rival investment and reduce concentration, but can also induce free-riding and depress aggregate investment in data generation. The net effect of such interventions depends on whether observed integration reflects efficiency advantages or strategic behavior intended to deter entry.

Overall, the results support a regulatory approach that prioritizes instruments enhancing consumer autonomy and transparency, while reserving more invasive structural remedies, such as integration caps or data siloing, for settings in which privacy harms are substantial and less distortive tools are insufficient. This hierarchy recognizes that data integration can both create consumer value and generate privacy risks, and that effective regulation should rebalance these forces while minimizing losses in productive efficiency.

9 Conclusion

This paper develops a framework to study the welfare trade-offs associated with data integration in digital platform markets, incorporating heterogeneous privacy preferences, endogenous opt-out decisions, advertising revenues, cross-service data use, and strategic investment in data capabilities. Data integration enhances service quality and advertising efficiency, but also increases privacy exposure for some users, particularly when data are combined across services. The welfare consequences of regulation therefore hinge on how policy shapes consumer sorting and firms' incentives to invest in data-driven technologies.

Policies that lower the cost of opting out of personalized data use generally dominate structural remedies. By allowing privacy-sensitive consumers to limit data processing without sacrificing core service quality, such policies reduce privacy externalities while preserving efficiency gains. Across a wide range of market conditions, this approach can generate Pareto-improving outcomes by aligning consumer choice with platform incentives.

Structural measures that restrict data integration improve welfare primarily when marginal privacy costs exceed marginal gains in quality and advertising, or when cross-service data use generates substantial privacy harms. When integration reflects efficiency-enhancing investments, these interventions can reduce total surplus by lowering service quality and advertising productivity. High levels of data integration or cross-service data reuse should therefore not be presumed harmful but the relevant criterion is whether marginal privacy harms outweigh marginal efficiency gains.

Market outcomes may nonetheless deviate from the social optimum. Incumbent platforms may overinvest in data to deter entry or underinvest when failing to internalize the broader social value of quality and advertising improvements. Entrants' incentives are shaped by cost disadvantages and strategic interactions. Competition-oriented remedies, including data sharing, integration limits, or siloing obligations, thus require careful calibration, as they may either correct or exacerbate investment distortions depending on market conditions.

Taken together, the results support a regulatory strategy that emphasizes user empowerment and the prevention of manipulative design practices, while reserving more interventionist tools for cases in which extensive data integration or cross-service data use generates persistent welfare losses.

References

Acquisti, A., Taylor, C. and Wagman, L. (2016), 'The economics of privacy', *Journal of Economic Literature* **54**(2), 442–92.

Allcott, H., Camilo Castillo, J., Gentzkow, M., Musolff, L. and Salz, T. (2025), Sources of market power in web search: Evidence from a field experiment, Working Paper 33410, National Bureau of Economic Research.

Benndorf, V. and Normann, H.-T. (2018), 'The willingness to sell personal data', *The Scandinavian Journal of Economics* **120**(4), 1260–1278.

Bhargava, H., Dubus, A., Ronayne, D. and Shekhar, S. (2025), 'The strategic value of data sharing in interdependent markets', *Management Science* **0**(0).

Bisceglia, M., Bonatti, A. and Scott Morton, F. (2025), Regulating privacy policies on digital platforms, Cowles Foundation Discussion Papers 2474, Cowles Foundation for Research in Economics, Yale University.

Carballa-Smichowski, B., Duch-Brown, N., Höcük, S., Kumar, P., Martens, B., Mulder, J. and Prüfer, P. (2025), 'Economies of scope in data aggregation: Evidence from health data', *Information Economics and Policy* **71**, 101146.

Casadesus-Masanell, R. and Hervas-Drane, A. (2015), 'Competing with privacy', *Management Science* **61**(1), 229–246.

Choe, C., Cong, J., Matsushima, N. and Shekhar, S. (2025), Privacy Regulations, Consumer Empowerment, and Versioning, Monash Economics Working Papers 2025-03, Monash University, Department of Economics.

Condorelli, D. and Padilla, J. (2023), 'Data-Driven Envelopment with Privacy-Policy Tying', *The Economic Journal* **134**(658), 515–537.

D'Annunzio, A. and Menichelli, E. (2022), 'A market for digital privacy: consumers' willingness to trade personal data and money', *Journal of Industrial and Business Economics* **49**(3), 571–598.

de Cornière, A. and Taylor, G. (2024), 'Data-driven mergers', *Management Science* **70**(9), 6473–6482.

Decarolis, F., Li, M. and Paternollo, F. (2025), ‘Competition and defaults in online search’, *American Economic Journal: Microeconomics* **17**(3), 369–414.

Goldfarb, A. and Que, V. F. (2023), ‘The Economics of Digital Privacy’, *Annual Review of Economics* **15**(1), 267–286.

Hagiu, A. and Wright, J. (2023), ‘Data-enabled learning, network effects, and competitive advantage’, *RAND Journal of Economics* **54**(4), 638–667.

Heidhues, P., Köster, M. and Kőszegi, B. (2024), A Theory of Digital Ecosystems, ECONtribute Discussion Papers Series 329, University of Bonn and University of Cologne, Germany.

Hovenkamp, E. (2024), ‘The competitive effects of search engine defaults’, *Journal of Law and Economics, Forthcoming* .

Johnson, J. P. (2013), ‘Targeted advertising and advertising avoidance’, *RAND Journal of Economics* **44**(1), 128–144.

Kesler, R., Skiera, B., Kraft, L. and Koschella, T. (2025), Existence, antecedents and consequences of non-compliance in mobile app markets, DICE Discussion Papers 419, Heinrich Heine University Düsseldorf, Düsseldorf Institute for Competition Economics (DICE).

Klein, T., Kurmangaliyeva, M., Prüfer, J. and Prüfer, P. (2025), ‘How important are user-generated data for search result quality?’, *Journal of Law Economics* **68**(3).

Krämer, J. and Shekhar, S. (2024), ‘Regulating digital platform ecosystems through data sharing and data siloing: Consequences for innovation and welfare’, *MIS Quarterly* **49**.

Lefouili, Y., Madio, L. and Toh, Y. L. (2024), ‘Privacy regulation and quality-enhancing innovation’, *The Journal of Industrial Economics* **72**(2), 662–684.

Marotta, V., Wu, Y., Zhang, K. and Acquisti, A. (2022), ‘The welfare impact of targeted advertising technologies’, *Information Systems Research* **33**(1), 131–151.

Pape, L.-D. and Rossi, M. (2026), ‘Is competition only one click away? the digital markets act’s impact on google maps’, *Marketing Science* .

Rhodes, A. and Zhou, J. (2024), ‘Personalization and Privacy Choice’, *RAND Journal of Economics, Forthcoming* .

Rhodes, A., Zhou, J. and Zhou, J. (2025), Digital Ecosystems and Data Regulation, Working paper.

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

Rosen, J. B. (1965), ‘Existence and uniqueness of equilibrium points for concave n-person games’, *Econometrica* **33**(3), 520–534.

Schaefer, M. and Sapi, G. (2023), ‘Complementarities in learning from data: Insights from general search’, *Information Economics and Policy* **65**(C).

Tucker, C. (2011), 'Social networks, personalized advertising, and privacy controls', *Journal of Marketing Research* **51**.

Ullrich, H., Hannane, J., Peukert, C., Aguiar, L. and Duso, T. (2024), Returns to data: Evidence from web tracking, Discussion Papers of DIW Berlin 2091, DIW Berlin, German Institute for Economic Research.

van de Waerdt, P. J. (2023), 'From monocle to spectacles: competition for data and "data ecosystem building"', *European Competition Journal* **19**(2), 191–225.

A Proofs

A.1 Proof of Lemma 1

Stage 2: Advertising Pricing. Platform $j \in \{G, F\}$ chooses the per-impression price p_j to maximize advertising revenue:

$$\max_{p_j} R_j(p_j, d_j, N_j) = p_j \cdot A_j(p_j, d_j) \cdot N_j = p_j \cdot \left(\bar{\mu} - \frac{p_j}{\beta d_j}\right) \cdot N_j$$

The FOC yields the optimal price $p_j^* = \frac{\bar{\mu}\beta d_j}{2}$. Advertiser demand function is $A_j^* = \bar{\mu} - \frac{p_j^*}{\beta d_j} = \bar{\mu} - \frac{\bar{\mu}}{2} = \frac{\bar{\mu}}{2}$ and the equilibrium revenue for platform j is given by $R_j^* = p_j^* \cdot A_j^* \cdot N_j = \frac{\bar{\mu}\beta d_j}{2} \cdot \frac{\bar{\mu}}{2} \cdot N_j = \frac{\bar{\mu}^2 \beta d_j N_j}{4}$.

Stage 1: Consumer Choice. Consumers observe both platforms' data integration levels and choose which to join. The equilibrium market shares depend on the opt-out regime. As in the main text, for notational convenience, the three regimes will be referred to as Case 1 (Universal Opt-Out), Case 2 (Universal Acceptance), and Case 3 (Asymmetric Regime) in all subsequent sections.

- **Case 1.** For customization-savvy consumers located at $x \in [0, 1]$, utility from choosing platform G is $U_{C,G}(x) = v + \xi_C d_G - tx$, while utility from choosing platform F is $U_{C,F}(x) = v + \xi_C d_F - t(1-x)$. Therefore, the marginal consumer satisfies $\xi_C(d_G - d_F) = t(1-2x)$, yielding $\tilde{x}_C = \frac{1}{2} + \frac{\xi_C(d_G - d_F)}{2t}$. All customization-savvy consumers with $x \le \tilde{x}_C$ choose platform G, while those with $x > \tilde{x}_C$ choose platform F. For privacy-savvy consumers who opt out on both platforms, utility from choosing platform G is $U_{P,G}^O(x) = v + \alpha d_G - c - tx$, while utility from choosing platform F is $U_{P,F}^O(x) = v + \alpha d_F - c - t(1-x)$. The marginal privacy-savvy consumer who opts out is located at $\tilde{x}_P^O = \frac{1}{2} + \frac{\alpha(d_G - d_F)}{2t}$. The equilibrium market share of platform G is given by:

$$\begin{aligned} N_G = \theta \tilde{x}_C + (1-\theta)\tilde{x}_P^O &= \theta\left(\frac{1}{2} + \frac{\xi_C(d_G - d_F)}{2t}\right) + (1-\theta)\left(\frac{1}{2} + \frac{\alpha(d_G - d_F)}{2t}\right) \\ &= \frac{1}{2} + \frac{d_G - d_F}{2t}\left[\theta\xi_C + (1-\theta)\alpha\right] = \frac{1}{2} + \frac{d_G - d_F}{2t}(\alpha + \theta\gamma) \end{aligned}$$

  while for platform F it is:

$$N_F = 1 - N_G = \frac{1}{2} - \frac{d_G - d_F}{2t}(\alpha + \theta\gamma)$$

- **Case 2.** Customization-savvy consumers behave identically to Case 1. For privacy-savvy consumers who accept personalized advertising on both platforms, utility from choosing platform G is $U_{P,G}^A(x) = v + \xi_P^A d_G - tx$, while utility from choosing platform F is $U_{P,F}^A(x) = v + \xi_P^A d_F - t(1-x)$, where $\xi_P^A = \alpha - \delta$. The marginal privacy-savvy consumer who accepts personalized advertising is located at $\tilde{x}_P^A = \frac{1}{2} + \frac{(\alpha-\delta)(d_G - d_F)}{2t}$. The net value parameter $\xi_P^A = \alpha - \delta$ can be positive, zero or negative depending on whether quality improvements dominate privacy costs. When $\alpha > \delta$, $\xi_P^A > 0$ and $\tilde{x}_P^A > 1/2$, so privacy-savvy consumers who accept personalized advertising still prefer platform G on average because quality improvements outweigh privacy concerns; when $\alpha = \delta$, they are split equally between platforms, while with $\alpha < \delta$ they prefer platform F on average because the privacy harm from platform G's superior data integration exceeds the quality benefit. Market shares are computed following the same approach as above.

- **Case 3.** Once again, customization-savvy consumers behave identically to the previous cases, as their behavior is independent of opt-out decisions. For privacy-savvy consumers, utility from choosing platform G with opt-out is $U^O_{P,G}(x) = v + \alpha d_G - c - tx$, while utility from choosing platform F with personalized advertising is $U^A_{P,F}(x) = v + \xi^A_P d_F - t(1 - x)$. Therefore, the marginal privacy-savvy consumer is $\tilde{x}^{OA}_P = \frac{1}{2} + \frac{\alpha(d_G - d_F) + \delta d_F - c}{2t}$, and platform G's market share is $N_G = \theta \tilde{x}_C + (1-\theta)\tilde{x}^{OA}_P$. This cutoff reflects three competing effects: (i) the quality advantage of platform G, $\alpha(d_G - d_F) > 0$, which attracts consumers to G; (ii) the privacy cost on platform F, $\delta d_F > 0$, which makes F less attractive and pushes consumers toward G; and (iii) the opt-out cost on platform G, $-c < 0$, which makes G less attractive and pushes consumers toward F.
  Market shares are computed following the same approach as above.

A.2 Proof of Lemma 2

Consumer Surplus

- **Case 1**. All privacy-savvy consumers opt out on both platforms. Customization-savvy consumers located at $x \in [0, \tilde{x}_C]$ choose platform G and obtain utility:

$$U_{C,G}(x) = v + \xi_C d_G - tx$$

Integrating over all customization-savvy consumers choosing G:

$$CS_{C,G} = \int_0^{\tilde{x}_C} (v + \xi_C d_G - tx)dx = \tilde{x}_C v + \tilde{x}_C \xi_C d_G - \frac{t\tilde{x}_C^2}{2}$$

Similarly, customization-savvy consumers located at $x \in (\tilde{x}_C, 1]$ choose platform F and obtain:

$$CS_{C,F} = \int_{\tilde{x}_C}^1 (v + \xi_C d_F - t(1-x))dx = (1-\tilde{x}_C)v + (1-\tilde{x}_C)\xi_C d_F - \frac{t(1-\tilde{x}_C)^2}{2}$$

Total consumer surplus for customization-savvy consumers is:

$$CS_C = CS_{C,G} + CS_{C,F} = v + \xi_C(\tilde{x}_C d_G + (1-\tilde{x}_C)d_F) - \frac{t}{2}(\tilde{x}_C^2 + (1-\tilde{x}_C)^2)$$

For privacy-savvy consumers who opt out, those located at $x \in [0, \tilde{x}^O_P]$ choose platform G:

$$CS^O_{P,G} = \int_0^{\tilde{x}^O_P} (v + \alpha d_G - c - tx)dx = \tilde{x}^O_P v + \tilde{x}^O_P(\alpha d_G - c) - \frac{t(\tilde{x}^O_P)^2}{2}$$

Privacy-savvy consumers at $x \in (\tilde{x}^O_P, 1]$ choose platform F:

$$CS^O_{P,F} = \int_{\tilde{x}^O_P}^1 (v + \alpha d_F - c - t(1-x))dx = (1-\tilde{x}^O_P)v + (1-\tilde{x}^O_P)(\alpha d_F - c) - \frac{t(1-\tilde{x}^O_P)^2}{2}$$

Total consumer surplus for privacy-savvy consumers is:

$$CS_P^O = CS_{P,G}^O + CS_{P,F}^O = v + \alpha(\tilde{x}_P^O d_G + (1 - \tilde{x}_P^O)d_F) - c - \frac{t}{2}((\tilde{x}_P^O)^2 + (1 - \tilde{x}_P^O)^2)$$

Aggregating across consumer types with weights θ and $1 - \theta$:

$$CS^{(1)} = \theta CS_C + (1 - \theta)CS_P^O$$

Substituting the cutoff values yields:

$$\begin{aligned} CS^{(1)} = \theta & \left[v + \xi_C \left(\tilde{x}_C d_G + (1 - \tilde{x}_C)d_F \right) - \frac{t}{2}(\tilde{x}_C^2 + (1 - \tilde{x}_C)^2) \right] \\ & + (1 - \theta) \left[v + \alpha \left(\tilde{x}_P^O d_G + (1 - \tilde{x}_P^O)d_F \right) - c - \frac{t}{2}((\tilde{x}_P^O)^2 + (1 - \tilde{x}_P^O)^2) \right] \end{aligned} \tag{10}$$

The transportation cost terms $\tilde{x}^2 + (1 - \tilde{x})^2$ represent the aggregate mismatch cost from consumers not being located exactly at their chosen platform. This expression is minimized when $\tilde{x} = 1/2$ and increases as market shares become more asymmetric.

- **Case 2.** When all privacy-savvy consumers accept personalized advertising, consumer surplus for customization-savvy consumers remains unchanged. For privacy-savvy consumers who accept personalized advertising, those at $x \in [0, \tilde{x}_P^A]$ choose platform G:

$$CS_{P,G}^A = \int_0^{\tilde{x}_P^A} (v + \xi_P^A d_G - tx)dx = \tilde{x}_P^A v + \tilde{x}_P^A \xi_P^A d_G - \frac{t(\tilde{x}_P^A)^2}{2}$$

Privacy-savvy consumers at $x \in (\tilde{x}_P^A, 1]$ choose platform F:

$$CS_{P,F}^A = \int_{\tilde{x}_P^A}^1 (v + \xi_P^A d_F - t(1 - x))dx = (1 - \tilde{x}_P^A)v + (1 - \tilde{x}_P^A)\xi_P^A d_F - \frac{t(1 - \tilde{x}_P^A)^2}{2}$$

Total consumer surplus for privacy-savvy consumers is:

$$\begin{aligned} CS_P^A &= CS_{P,G}^A + CS_{P,F}^A \\ &= v + \xi_P^A(\tilde{x}_P^A d_G + (1 - \tilde{x}_P^A)d_F) - \frac{t}{2}((\tilde{x}_P^A)^2 + (1 - \tilde{x}_P^A)^2) \end{aligned}$$

Total consumer surplus in Case 2 is:

$$CS^{(2)} = \theta CS_C + (1 - \theta)CS_P^A \tag{11}$$

The key difference from the previous case is that privacy-savvy consumers now experience net value $\xi_P^A = \alpha - \delta$ rather than $\alpha - c$. When $\delta < c$, privacy-savvy consumers are better off accepting personalized advertising because the privacy cost is lower than the opt-out cost.

- **Case 3.** Consumer surplus for customization-savvy consumers is identical to that derived above. For privacy-savvy consumers, those at $x \in [0, \tilde{x}_P^{OA}]$ choose platform G and opt out:

$$CS_{P,G}^{OA} = \int_0^{\tilde{x}_P^{OA}} (v + \alpha d_G - c - tx)dx = \tilde{x}_P^{OA} v + \tilde{x}_P^{OA}(\alpha d_G - c) - \frac{t(\tilde{x}_P^{OA})^2}{2}$$

Privacy-savvy consumers at $x \in (\tilde{x}_P^{OA}, 1]$ choose platform F and accept personalized advertising:

$$CS_{P,F}^{OA} = \int_{\tilde{x}_P^{OA}}^{1} (v + \xi_P^A d_F - t(1-x))dx = (1-\tilde{x}_P^{OA})v + (1-\tilde{x}_P^{OA})\xi_P^A d_F - \frac{t(1-\tilde{x}_P^{OA})^2}{2}$$

Total consumer surplus for privacy-savvy consumers is:

$$\begin{aligned} CS_P^{OA} &= CS_{P,G}^{OA} + CS_{P,F}^{OA} \\ &= v + \tilde{x}_P^{OA}(\alpha d_G - c) + (1-\tilde{x}_P^{OA})\xi_P^A d_F - \frac{t}{2}((\tilde{x}_P^{OA})^2 + (1-\tilde{x}_P^{OA})^2) \end{aligned}$$

Total consumer surplus in Case 3 is:

$$CS^{(3)} = \theta CS_C + (1-\theta) CS_P^{OA} \tag{12}$$

This case exhibits the richest structure because privacy-savvy consumers face different effective costs on each platform. Those choosing G incur the opt-out cost c, while those choosing F bear the privacy cost δd_F. The asymmetry creates a distortion in the allocation of privacy-savvy consumers across platforms.

Platform Profits

Platform profits equal advertising revenues. The profit in equilibrium for platform j is:

$$\Pi_j^* = R_j^* = \frac{\bar{\mu}^2 \beta d_j N_j}{4}$$

yielding the aggregate platform profits:

$$\Pi^* = \Pi_G + \Pi_F = \frac{\bar{\mu}^2 \beta}{4}(d_G N_G + d_F N_F) \tag{13}$$

that depend on the data-weighted consumer base $d_G N_G + d_F N_F$. Profits increase when consumers sort toward the platform with higher data integration, as this maximizes the value extracted from the advertising side.

Advertiser Surplus

Advertisers of type $\mu \geq p_j^*/(\beta d_j)$ purchase impressions on platform j. Given the optimal price $p_j^* = \frac{\bar{\mu}\beta d_j}{2}$, the marginal advertiser has type $\mu_j^* = \bar{\mu}/2$. On platform j, an advertiser of type μ derives gross value $\mu\beta d_j N_j$ from reaching N_j consumers and pays $p_j^* N_j = \frac{\bar{\mu}\beta d_j N_j}{2}$ and her net surplus is:

$$\left(\mu\beta d_j - \frac{\bar{\mu}\beta d_j}{2}\right) N_j = \beta d_j N_j \left(\mu - \frac{\bar{\mu}}{2}\right)$$

Integrating over all participating advertisers $\mu \in [\bar{\mu}/2, \bar{\mu}]$:

$$\begin{aligned} AS_j &= \int_{\bar{\mu}/2}^{\bar{\mu}} \beta d_j N_j \left(\mu - \frac{\bar{\mu}}{2}\right) d\mu = \beta d_j N_j \left[\frac{\mu^2}{2} - \frac{\bar{\mu}\mu}{2}\right]_{\bar{\mu}/2}^{\bar{\mu}} = \beta d_j N_j \left[\left(\frac{\bar{\mu}^2}{2} - \frac{\bar{\mu}^2}{2}\right) - \left(\frac{\bar{\mu}^2}{8} - \frac{\bar{\mu}^2}{4}\right)\right] \\ &= \beta d_j N_j \left[0 + \frac{\bar{\mu}^2}{8}\right] = \frac{\bar{\mu}^2 \beta d_j N_j}{8} \end{aligned}$$

Total advertiser surplus across both platforms is:

$$AS = AS_G + AS_F = \frac{\bar{\mu}^2 \beta}{8}(d_G N_G + d_F N_F) \tag{14}$$

Advertiser surplus is exactly half of platform profits, reflecting the standard result that under linear demand with uniform distribution, consumer surplus equals half of producer surplus when producers charge the monopoly price. Here, each platform acts as a monopolist on its advertising side, facing a linear demand from advertisers.

Total Welfare

Total welfare is the sum of consumer surplus, platform profits and advertiser surplus:

$$W = CS + \Pi + AS$$

Substituting equations (13) and (14):

$$W = CS + \frac{\bar{\mu}^2 \beta}{4}(d_G N_G + d_F N_F) + \frac{\bar{\mu}^2 \beta}{8}(d_G N_G + d_F N_F) = CS + \frac{3\bar{\mu}^2 \beta}{8}(d_G N_G + d_F N_F)$$

The welfare function depends on consumer surplus, which varies by regime, and the data-weighted consumer base $d_G N_G + d_F N_F$, which determines advertising-side surplus.

- **Case 1.** Substituting equation (10) and the market shares from equations (2) and (3):

$$\begin{aligned} W^{(1)} = \theta &\left[v + \xi_C \left(\tilde{x}_C d_G + (1 - \tilde{x}_C) d_F \right) - \frac{t}{2}(\tilde{x}_C^2 + (1 - \tilde{x}_C)^2) \right] \\ &+ (1 - \theta) \left[v + \alpha \left(\tilde{x}_P^O d_G + (1 - \tilde{x}_P^O) d_F \right) - c - \frac{t}{2}((\tilde{x}_P^O)^2 + (1 - \tilde{x}_P^O)^2) \right] \\ &+ \frac{3\bar{\mu}^2 \beta}{8}(d_G N_G + d_F N_F) \end{aligned}$$

  The opt-out cost c directly reduces consumer surplus for privacy-savvy consumers, representing a deadweight loss from avoiding personalized advertising. Secondly, the transportation cost terms capture the inefficiency from imperfect matching between consumers and platforms. Finally, the data-weighted consumer base term captures the advertising-side benefits, which increase when consumers sort toward the high-integration platform.

- **Case 2.** Substituting equation (11):

$$\begin{aligned} W^{(2)} = \theta &\left[v + \xi_C \left(\tilde{x}_C d_G + (1 - \tilde{x}_C) d_F \right) - \frac{t}{2}(\tilde{x}_C^2 + (1 - \tilde{x}_C)^2) \right] \\ &+ (1 - \theta) \left[v + \xi_P^A \left(\tilde{x}_P^A d_G + (1 - \tilde{x}_P^A) d_F \right) - \frac{t}{2}((\tilde{x}_P^A)^2 + (1 - \tilde{x}_P^A)^2) \right] \\ &+ \frac{3\bar{\mu}^2 \beta}{8}(d_G N_G + d_F N_F) \end{aligned}$$

  Compared to Case 1, privacy-savvy consumers in Case 2 no longer pay the opt-out cost c, but instead incur the privacy cost δd_j. The net effect on welfare depends on whether c exceeds the average privacy cost borne by these consumers.

- **Case 3.** Substituting equation (12):

$$W^{(3)} = \theta\left[v + \xi_C\left(\tilde{x}_C d_G + (1-\tilde{x}_C)d_F\right) - \frac{t}{2}(\tilde{x}_C^2 + (1-\tilde{x}_C)^2)\right]$$
$$+ (1-\theta)\left[v + \tilde{x}_P^{OA}(\alpha d_G - c) + (1-\tilde{x}_P^{OA})\xi_P^A d_F - \frac{t}{2}((\tilde{x}_P^{OA})^2 + (1-\tilde{x}_P^{OA})^2)\right]$$
$$+ \frac{3\bar{\mu}^2\beta}{8}(d_G N_G + d_F N_F)$$

  This case exhibits the most complex welfare structure. Privacy-savvy consumers choosing platform G incur the opt-out cost c, while those choosing platform F bear the privacy cost δd_F. The asymmetry in costs creates a distortion: the marginal privacy-savvy consumer is not indifferent between the two costs, leading to an inefficient allocation.

A.3 Proof of Proposition 1

Direct Effect on Consumer Surplus

Consider first the direct effect on consumer surplus, treating market shares as fixed.

- **Case 1.** The direct effect on customization-savvy consumers choosing G is $\frac{\partial CS_{C,G}}{\partial d_G} = \tilde{x}_C\xi_C > 0$, since each of them loses $\xi_C = \alpha + \gamma$ units of utility per unit reduction in d_G. For privacy-savvy consumers choosing G, the effect is $\frac{\partial CS^O_{P,G}}{\partial d_G} = \tilde{x}^O_P\alpha > 0$, as each loses α units of utility per unit reduction in d_G. The aggregate effects are proportional to $\theta\tilde{x}_C$ and $(1-\theta)\tilde{x}^O_P$ respectively. Consumers on platform F experience no direct effect since d_F is held constant. Thus, the total direct effect on consumer surplus is:

$$\left.\frac{\partial CS^{(1)}}{\partial d_G}\right|_{\text{direct}} = \theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}^O_P\alpha > 0 \tag{15}$$

- **Case 2.** The direct effect on privacy-savvy consumers choosing G is $\frac{\partial CS^A_{P,G}}{\partial d_G} = \tilde{x}^A_P\xi^A_P = \tilde{x}^A_P(\alpha-\delta)$. The sign depends on whether $\alpha > \delta$ or $\alpha < \delta$. If quality benefits dominate privacy costs ($\alpha > \delta$), then reducing d_G directly harms privacy-savvy consumers on G. If privacy costs dominate ($\alpha < \delta$), reducing d_G directly benefits these consumers by reducing their privacy exposure. The total direct effect on consumer surplus in this regime is given by:

$$\left.\frac{\partial CS^{(2)}}{\partial d_G}\right|_{\text{direct}} = \theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}^A_P\xi^A_P$$

  The sign is ambiguous and depends on the balance between quality benefits and privacy costs.

- **Case 3.** Privacy-savvy consumers on G opt out, so:

$$\left.\frac{\partial CS^{(3)}}{\partial d_G}\right|_{\text{direct}} = \theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}^{OA}_P\alpha$$

  which is positive as privacy-savvy consumers on G are shielded from privacy costs through opt-out.

Direct Effect on Advertising-Side Surplus

Platform profits and advertiser surplus both depend on the data-weighted consumer base $d_G N_G + d_F N_F$. Holding market shares constant, the direct effect is equal $\frac{\partial(d_G N_G + d_F N_F)}{\partial d_G}\Big|_{\text{direct}} = N_G$.

A unit reduction in d_G reduces the data-weighted consumer base by N_G, the number of consumers on platform G. This reduces both platform profits and advertiser surplus:

$$\frac{\partial \Pi}{\partial d_G}\Big|_{\text{direct}} = \frac{\bar{\mu}^2 \beta N_G}{4} > 0 \tag{16}$$

$$\frac{\partial AS}{\partial d_G}\Big|_{\text{direct}} = \frac{\bar{\mu}^2 \beta N_G}{8} > 0 \tag{17}$$

The reduction in data integration unambiguously harms the advertising side by reducing the targeting precision on platform G.

Indirect Effects

- **Case 1.** Differentiating equation (2) with respect to d_G yields $\frac{dN_G}{dd_G} = \frac{\alpha+\theta\gamma}{2t} > 0$. Thus, a unit reduction in d_G shifts $\frac{\alpha+\theta\gamma}{2t}$ consumers from G to F, reflecting a quality effect $\frac{\alpha}{2t}$ and a personalization effect $\frac{\theta\gamma}{2t}$. The switching consumers are marginal, located at $\tilde{x}_C$ and $\tilde{x}_P^O$, and are indifferent between platforms. Hence, to a first-order approximation, switching entails no utility change for either customization-savvy or privacy-savvy consumers. However, the market share change affects advertising-side surplus. When consumers switch from G to F, the data-weighted consumer base changes by:

  $$\frac{\partial(d_G N_G + d_F N_F)}{\partial d_G}\Big|_{\text{indirect}} = \frac{\partial(d_G N_G + d_F N_F)}{\partial N_G} \cdot \frac{dN_G}{dd_G} = (d_G - d_F) \cdot \frac{\alpha + \theta\gamma}{2t}$$

  Additionally, the indirect effects on platform profits and advertiser surplus are given by:

  $$\frac{\partial \Pi}{\partial d_G}\Big|_{\text{indirect}} = \frac{\bar{\mu}^2 \beta}{4}(d_G - d_F) \cdot \frac{\alpha + \theta\gamma}{2t} > 0$$

  $$\frac{\partial AS}{\partial d_G}\Big|_{\text{indirect}} = \frac{\bar{\mu}^2 \beta}{8}(d_G - d_F) \cdot \frac{\alpha + \theta\gamma}{2t} > 0$$

- **Case 2.** From equation (4):

  $$\frac{dN_G}{dd_G} = \frac{\alpha + \theta\gamma - (1-\theta)\delta}{2t}$$

  The sign depends on whether $\alpha + \theta\gamma > (1-\theta)\delta$. If quality benefits and personalization benefits outweigh the privacy costs imposed on privacy-savvy consumers, then platform G's market share increases with d_G. If privacy costs dominate, platform G's market share decreases with d_G: higher data integration makes G less attractive because it imposes excessive privacy costs.

  The indirect effect on advertising-side surplus is:

  $$\frac{\partial(d_G N_G + d_F N_F)}{\partial d_G}\Big|_{\text{indirect}} = (d_G - d_F) \cdot \frac{\alpha + \theta\gamma - (1-\theta)\delta}{2t}$$

and can be positive or negative depending on whether G's market share increases or decreases with d_G.

- **Case 3.** From equation (6), the market share response is more complex because the cutoff $\tilde{x}_P^{OA}$ depends on d_G in a nonlinear way. Recalling $\tilde{x}_P^{OA} = \frac{1}{2} + \frac{\alpha(d_G - d_F) + \delta d_F - c}{2t}$ and differentiating with respect to d_G yields $\frac{\alpha}{2t}$.

  For customization-savvy consumers:

$$\frac{\partial \tilde{x}_C}{\partial d_G} = \frac{\xi_C}{2t} = \frac{\alpha + \gamma}{2t}$$

  Platform G's market share is $N_G = \theta\tilde{x}_C + (1-\theta)\tilde{x}_P^{OA}$, so:

$$\frac{dN_G}{dd_G} = \theta \cdot \frac{\alpha+\gamma}{2t} + (1-\theta) \cdot \frac{\alpha}{2t} = \frac{\alpha + \theta\gamma}{2t} > 0$$

  Interestingly, this equals the market share response in Case 1. The reason is that privacy-savvy consumers on G opt out, so they respond only to quality improvements α, while those on F face privacy costs that depend only on d_F, which is held constant.

  The indirect effect on advertising-side surplus is:

$$\left.\frac{\partial(d_G N_G + d_F N_F)}{\partial d_G}\right|_{\text{indirect}} = (d_G - d_F) \cdot \frac{\alpha + \theta\gamma}{2t} > 0$$

Total Welfare Effect

- **Case 1.** Combining equations (15), (16), (17), and the indirect effects:

$$\frac{dW^{(1)}}{dd_G} = \underbrace{\theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}_P^O\alpha}_{\text{Direct CS effect}} + \underbrace{\frac{3\bar{\mu}^2\beta N_G}{8}}_{\text{Direct advertising effect}} + \underbrace{\frac{3\bar{\mu}^2\beta}{8}(d_G - d_F) \cdot \frac{\alpha+\theta\gamma}{2t}}_{\text{Indirect advertising effect through market shares}}$$

  Substituting $N_G = \frac{1}{2} + \frac{(d_G - d_F)(\alpha+\theta\gamma)}{2t}$:

$$\begin{aligned}\frac{dW^{(1)}}{dd_G} &= \theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}_P^O\alpha + \frac{3\bar{\mu}^2\beta}{8}\left[\frac{1}{2} + \frac{(d_G - d_F)(\alpha+\theta\gamma)}{2t}\right] \\ &\quad + \frac{3\bar{\mu}^2\beta}{8}(d_G - d_F) \cdot \frac{\alpha+\theta\gamma}{2t} \\ &= \theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}_P^O\alpha + \frac{3\bar{\mu}^2\beta}{16} + \frac{3\bar{\mu}^2\beta(d_G - d_F)(\alpha+\theta\gamma)}{8t}\end{aligned}$$

  All terms are positive, implying $\frac{dW^{(1)}}{dd_G} > 0$.

- **Case 2.** Similarly, the total welfare effect is:

$$\frac{dW^{(2)}}{dd_G} = \underbrace{\theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}_P^A\xi_P^A}_{\text{Direct CS effect}} + \underbrace{\frac{3\bar{\mu}^2\beta N_G}{8}}_{\text{Direct advertising effect}} + \underbrace{\frac{3\bar{\mu}^2\beta}{8}(d_G - d_F) \cdot \frac{\alpha+\theta\gamma-(1-\theta)\delta}{2t}}_{\text{Indirect advertising effect}}$$

Substituting $N_G = \frac{1}{2} + \frac{(d_G - d_F)(\alpha + \theta\gamma - (1-\theta)\delta)}{2t}$ and $\xi_P^A = \alpha - \delta$:

$$\frac{dW^{(2)}}{dd_G} = \theta\tilde{x}_C(\alpha+\gamma) + (1-\theta)\tilde{x}_P^A(\alpha-\delta) + \frac{3\bar{\mu}^2\beta}{8}\left[\frac{1}{2} + \frac{(d_G - d_F)(\alpha+\theta\gamma-(1-\theta)\delta)}{2t}\right] + \frac{3\bar{\mu}^2\beta}{8}(d_G - d_F)\cdot\frac{\alpha+\theta\gamma-(1-\theta)\delta}{2t}$$

Simplifying:

$$\frac{dW^{(2)}}{dd_G} = \theta\tilde{x}_C(\alpha+\gamma) + (1-\theta)\tilde{x}_P^A(\alpha-\delta) + \frac{3\bar{\mu}^2\beta}{16} + \frac{3\bar{\mu}^2\beta(d_G - d_F)(\alpha+\theta\gamma-(1-\theta)\delta)}{4t}$$

- **Case 3.** The total welfare effect has the same form as in Case 1:

$$\frac{dW^{(3)}}{dd_G} = \theta\tilde{x}_C\xi_C + (1-\theta)\tilde{x}_P^{OA}\alpha + \frac{3\bar{\mu}^2\beta N_G}{8} + \frac{3\bar{\mu}^2\beta}{8}(d_G - d_F)\cdot\frac{\alpha+\theta\gamma}{2t}$$

A.4 Proof of Lemma 3

Stage 3: Advertising Pricing. Platform $j \in \{G, F\}$ solves:

$$\max_{p_j} R_j(p_j, d_j, N_j) = p_j\left(\bar{\mu} - \frac{p_j}{\beta d_j^\psi}\right)N_j$$

Taking the first-order condition:

$$\frac{\partial R_j}{\partial p_j} = \left(\bar{\mu} - \frac{2p_j}{\beta d_j^\psi}\right)N_j = 0$$

Solving yields the optimal price:

$$p_j^* = \frac{\bar{\mu}\beta d_j^\psi}{2}$$

The price is increasing in data integration with elasticity $\psi < 1$, reflecting decreasing returns to targeting precision. The mass of participating advertisers is:

$$A_j^* = \bar{\mu} - \frac{p_j^*}{\beta d_j^\psi} = \bar{\mu} - \frac{\bar{\mu}}{2} = \frac{\bar{\mu}}{2}$$

and platform revenue is:

$$R_j^* = p_j^* \cdot A_j^* \cdot N_j = \frac{\bar{\mu}\beta d_j^\psi}{2}\cdot\frac{\bar{\mu}}{2}\cdot N_j = \frac{\bar{\mu}^2\beta d_j^\psi N_j}{4}$$

Stage 2: Consumer Platform Choice. In Case 1, customization-savvy consumers located at x choosing platform G obtain utility $U_{C,G}(x) = v + \xi_C\alpha d_G^\phi - tx$, while those choosing platform F obtain $U_{C,F}(x) = v + \xi_C\alpha d_F^\phi - t(1-x)$. The marginal customization-savvy consumer is therefore located at $\tilde{x}_C = \frac{1}{2} + \frac{\xi_C\alpha(d_G^\phi - d_F^\phi)}{2t}$. For privacy-savvy consumers who opt out on both platforms, the marginal consumer's location is $\tilde{x}_P^O = \frac{1}{2} + \frac{\alpha(d_G^\phi - d_F^\phi)}{2t}$.

Platform G's market share is:

$$N_G = \theta \tilde{x}_C + (1-\theta)\tilde{x}_P^O = \theta\left(\frac{1}{2} + \frac{\xi_C \alpha(d_G^\phi - d_F^\phi)}{2t}\right) + (1-\theta)\left(\frac{1}{2} + \frac{\alpha(d_G^\phi - d_F^\phi)}{2t}\right)$$
$$= \frac{1}{2} + \frac{\alpha(d_G^\phi - d_F^\phi)}{2t}[\theta\xi_C + (1-\theta)]$$

Defining $\bar{\xi} \equiv \theta\xi_C + (1-\theta) = \theta(\alpha+\gamma) + (1-\theta)$ yields:

$$N_G = \frac{1}{2} + \frac{\alpha\bar{\xi}(d_G^\phi - d_F^\phi)}{2t} \tag{18}$$

Stage 1: Data Integration Investment. Platform j chooses d_j to maximize profit net of investment costs:

$$\Pi_j(d_j, d_{-j}) = R_j^*(d_j, d_{-j}) - C_j(d_j) = \frac{\bar{\mu}^2\beta d_j^\psi N_j(d_j, d_{-j})}{4} - \frac{\kappa_j d_j^{1+\eta}}{1+\eta}$$

Taking the first-order condition with respect to d_j:

$$\frac{\partial \Pi_j}{\partial d_j} = \frac{\bar{\mu}^2\beta}{4}\left[\psi d_j^{\psi-1} N_j + d_j^\psi \frac{\partial N_j}{\partial d_j}\right] - \kappa_j d_j^\eta = 0$$

For platform G, from equation (18):

$$\frac{\partial N_G}{\partial d_G} = \frac{\alpha\bar{\xi}\phi d_G^{\phi-1}}{2t}$$

Substituting into the first-order condition:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi d_G^{\psi-1} N_G + d_G^\psi \cdot \frac{\alpha\bar{\xi}\phi d_G^{\phi-1}}{2t}\right] = \kappa_G d_G^\eta$$

Simplifying:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi d_G^{\psi-1} N_G + \frac{\alpha\bar{\xi}\phi d_G^{\phi+\psi-1}}{2t}\right] = \kappa_G d_G^\eta \tag{19}$$

Similarly, for platform F:

$$\frac{\partial N_F}{\partial d_F} = -\frac{\partial N_G}{\partial d_F} = \frac{\alpha\bar{\xi}\phi d_F^{\phi-1}}{2t}$$

The first-order condition for platform F is:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi d_F^{\psi-1} N_F + \frac{\alpha\bar{\xi}\phi d_F^{\phi+\psi-1}}{2t}\right] = \kappa_F d_F^\eta$$

Existence and Uniqueness of Asymmetric Equilibrium. Existence and uniqueness are established by verifying that the best-response functions are well defined, continuous, and satisfy appropriate monotonicity conditions.

For platform j, define the marginal benefit function:

$$MB_j(d_j, d_{-j}) \equiv \frac{\bar{\mu}^2\beta}{4}\left[\psi d_j^{\psi-1} N_j(d_j, d_{-j}) + d_j^\psi \frac{\partial N_j}{\partial d_j}\right]$$

and the marginal cost function:

$$MC_j(d_j) \equiv \kappa_j d_j^\eta$$

Step 1: Properties at boundaries. At $d_j = 0$:

$$MB_j(0, d_{-j}) = \frac{\bar{\mu}^2 \beta \psi}{4} \cdot N_j(0, d_{-j}) > 0$$

since $N_j(0, d_{-j}) > 0$ for finite d_{-j}. Meanwhile:

$$MC_j(0) = 0$$

Therefore $MB_j(0, d_{-j}) > MC_j(0)$ for all $d_{-j} \geq 0$. As $d_j \to \infty$:

$$\lim_{d_j \to \infty} \frac{MB_j(d_j, d_{-j})}{MC_j(d_j)} = \lim_{d_j \to \infty} \frac{\frac{\bar{\mu}^2 \beta}{4} \left[\psi d_j^{\psi-1} N_j + d_j^\psi \frac{\alpha \bar{\xi} \phi d_j^{\phi-1}}{2t} \right]}{\kappa_j d_j^\eta}$$

Since N_j is bounded and $\eta > \psi - 1$ (by assumption), the highest-order term in the numerator is $d_j^{\phi+\psi-1}$ while the denominator has order d_j^η. For $\eta > \max\{\psi - 1, \phi + \psi - 1\}$, the limit is zero, ensuring that $MB_j < MC_j$ for sufficiently large d_j.

Step 2: Monotonicity of marginal benefit and marginal cost. Taking derivatives:

$$\frac{\partial MB_j}{\partial d_j} = \frac{\bar{\mu}^2 \beta}{4} \left[\psi(\psi - 1) d_j^{\psi-2} N_j + 2\psi d_j^{\psi-1} \frac{\partial N_j}{\partial d_j} + d_j^\psi \frac{\partial^2 N_j}{\partial d_j^2} \right]$$

For the marginal cost:

$$\frac{\partial MC_j}{\partial d_j} = \kappa_j \eta d_j^{\eta-1} > 0$$

The sign of $\frac{\partial MB_j}{\partial d_j}$ is ambiguous due to the competing effects of decreasing returns ($\psi < 1$ implies $\psi(\psi - 1) < 0$) and market share effects. However, for the existence of equilibrium, what matters is that eventually MC grows faster than MB as d_j increases.

Step 3: Strategic interaction and best response functions. Platform j's best response function $BR_j(d_{-j})$ is implicitly defined by:

$$MB_j(d_j, d_{-j}) = MC_j(d_j)$$

By the implicit function theorem:

$$\frac{\partial BR_j}{\partial d_{-j}} = -\frac{\frac{\partial MB_j}{\partial d_{-j}}}{\frac{\partial MB_j}{\partial d_j} - \frac{\partial MC_j}{\partial d_j}}$$

For platform G:

$$\frac{\partial MB_G}{\partial d_F} = \frac{\bar{\mu}^2 \beta}{4} \psi d_G^{\psi-1} \frac{\partial N_G}{\partial d_F}$$

Since $\frac{\partial N_G}{\partial d_F} = -\frac{\alpha \bar{\xi} \phi d_F^{\phi-1}}{2t} < 0$, $\frac{\partial MB_G}{\partial d_F} < 0$, establishing that data integration levels are strategic substitutes.

Step 4: Asymmetric equilibrium with incumbency advantage. With $\kappa_G < \kappa_F$, platform G has lower marginal costs at every level of data integration. Comparing the first-order conditions at any common level d:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi d^{\psi-1}N + \frac{\alpha\bar{\xi}\phi d^{\phi+\psi-1}}{2t}\right] = \kappa_G d^\eta < \kappa_F d^\eta$$

At the symmetric point where $d_G = d_F = d$ and $N_G = N_F = 1/2$, platform G's marginal benefit exceeds its marginal cost while platform F's marginal benefit falls short of its marginal cost. This implies that in equilibrium, platform G invests more: $d_G^* > d_F^*$.

Step 5: Existence and Uniqueness. Consider the Stage 1 data integration game where platforms simultaneously choose $d_j \in [0, \bar{d}]$ to maximize $\Pi_j(d_j, d_{-j})$.

The strategy sets are compact and convex, and $\Pi_j(d_j, d_{-j})$ is continuous in (d_j, d_{-j}) and strictly concave in d_j, given the concavity of advertising revenues in data integration and the convexity of investment costs. Moreover, the cross-partial derivatives satisfy $\frac{\partial^2 \Pi_j}{\partial d_j \partial d_{-j}} < 0$, implying that data integration choices are strategic substitutes. Under the maintained assumption that own effects dominate cross effects (i.e. diagonal dominance of the Jacobian of first-order conditions), the game is diagonally strictly concave in the sense of Rosen (1965).

Therefore, the Stage 1 game admits a unique Nash equilibrium. Combined with the analysis in Steps 1–4, this equilibrium is asymmetric, with $d_G^* > d_F^* > 0$ whenever $\kappa_G < \kappa_F$.

In the unique asymmetric equilibrium, platform G achieves higher data integration $d_G^* > d_F^* > 0$ and captures a larger market share, $N_G^* > 1/2 > N_F^*$. Moreover, comparative statics with respect to $\kappa_F - \kappa_G$, β, ϕ, and t follow from the first-order conditions and have the natural signs: greater cost asymmetry, higher advertising value, and stronger quality effects widen the data integration gap, while stronger competition (higher t) mitigates it.

This completes the proof of Lemma 3. □

A.5 Proof of Proposition 2

The social planner maximizes total welfare:

$$W = CS + \Pi_G + \Pi_F + AS - C_G(d_G) - C_F(d_F)$$

In the asymmetric case with universal opt-out, consumer surplus is:

$$CS = \theta\left[v + \xi_C\alpha(N_G(d_G^*)^\phi + N_F(d_F^*)^\phi) - \frac{t}{2}(\tilde{x}_C^2 + (1-\tilde{x}_C)^2)\right]$$
$$+ (1-\theta)\left[v + \alpha(N_G(d_G^*)^\phi + N_F(d_F^*)^\phi) - c - \frac{t}{2}((\tilde{x}_P^O)^2 + (1-\tilde{x}_P^O)^2)\right]$$

The social planner's first-order condition with respect to d_G is:

$$\frac{\partial W}{\partial d_G} = \frac{\partial CS}{\partial d_G} + \frac{3\bar{\mu}^2\beta}{8}\left[\psi(d_G)^{\psi-1}N_G + (d_G)^\psi\frac{\partial N_G}{\partial d_G}\right]$$
$$+ \frac{3\bar{\mu}^2\beta}{8}(d_F)^\psi\frac{\partial N_F}{\partial d_G} - \kappa_G(d_G)^\eta = 0$$

The consumer surplus derivative consists of two components:

1. Direct quality effect on existing users:

$$\left.\frac{\partial CS}{\partial d_G}\right|_{\text{direct}} = \alpha\bar{\xi}\phi(d_G)^{\phi-1}N_G$$

2. Reallocation effect from market share changes:

$$\left.\frac{\partial CS}{\partial d_G}\right|_{\text{indirect}} = \alpha\bar{\xi}\left[(d_G)^{\phi} - (d_F)^{\phi}\right]\frac{\partial N_G}{\partial d_G} - \frac{t}{2}\frac{\partial(\tilde{x}^2 + (1-\tilde{x})^2)}{\partial d_G}$$

Since $\frac{\partial N_F}{\partial d_G} = -\frac{\partial N_G}{\partial d_G}$, the social planner's FOC becomes:

$$\alpha\bar{\xi}\phi(d_G^{SP})^{\phi-1}N_G + \alpha\bar{\xi}\left[(d_G^{SP})^{\phi} - (d_F^{SP})^{\phi}\right]\frac{\partial N_G}{\partial d_G} - \frac{t}{2}\frac{\partial\left(\tilde{x}^2 + (1-\tilde{x})^2\right)}{\partial d_G} + \frac{3\bar{\mu}^2\beta}{8}\left[\psi(d_G^{SP})^{\psi-1}N_G + (d_G^{SP})^{\psi}\frac{\partial N_G}{\partial d_G}\right] + \frac{3\bar{\mu}^2\beta}{8}(d_F^{SP})^{\psi}\frac{\partial N_F}{\partial d_G} - \kappa_G(d_G^{SP})^{\eta} = 0.$$

Comparing with platform G's private FOC from equation (19):

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi(d_G^*)^{\psi-1}N_G + \frac{\alpha\bar{\xi}\phi(d_G^*)^{\phi+\psi-1}}{2t}\right] = \kappa_G(d_G^*)^{\eta}$$

Externalities. The difference between social and private marginal benefits reflects the presence of four externalities.

1. **Advertising-side externality:** the social planner values total advertising surplus at $\frac{3\bar{\mu}^2\beta}{8}$ while platforms capture only $\frac{\bar{\mu}^2\beta}{4}$. This creates:

$$\frac{3\bar{\mu}^2\beta\psi(d_G)^{\psi-1}N_G}{8} - \frac{\bar{\mu}^2\beta\psi(d_G)^{\psi-1}N_G}{4} = \frac{\bar{\mu}^2\beta\psi(d_G)^{\psi-1}N_G}{8} > 0$$

   This externality is larger for platform G due to its larger consumer base $N_G > N_F$.

2. **Consumer-side externality:** platforms capture quality benefits only through market share effects, not from infra-marginal consumers:

$$\alpha\bar{\xi}\phi(d_G)^{\phi-1}N_G - \frac{\bar{\mu}^2\beta\alpha\bar{\xi}\phi(d_G)^{\phi+\psi-1}}{8t} > 0$$

   when $\frac{8tN_G}{(d_G)^{\psi}} > \bar{\mu}^2\beta$.

3. **Business stealing effect:** the private FOC includes market share gains that represent pure transfers:

$$\frac{\bar{\mu}^2\beta\alpha\bar{\xi}\phi(d_G)^{\phi+\psi-1}}{8t}$$

   This creates overinvestment incentives as platforms compete for market share.

4. **Incumbency amplification:** platform G's cost advantage $\kappa_G < \kappa_F$ creates excessive market concentration. The social planner would balance efficiency gains from exploiting G's cost advantage against competition benefits from maintaining a viable rival.

Direction of Distortion. Platform G underinvests ($d_G^* < d_G^{SP}$) when:

$$\alpha\bar{\xi}\phi(d_G^*)^{\phi-1}N_G + \frac{\bar{\mu}^2\beta\psi(d_G^*)^{\psi-1}N_G}{8} > \frac{\bar{\mu}^2\beta\alpha\bar{\xi}\phi(d_G^*)^{\phi+\psi-1}N_G}{4t}$$

Simplifying:

$$\frac{8t[\alpha\bar{\xi}\phi(d_G^*)^{\phi-1} + \frac{\bar{\mu}^2\beta\psi(d_G^*)^{\psi-1}}{8}]}{\bar{\mu}^2\beta\alpha\bar{\xi}\phi(d_G^*)^{\phi+\psi-1}} > N_G$$

This condition holds when $\bar{\mu}^2\beta$ and t are large, and N_G is not too large. Platform G overinvests ($d_G^* > d_G^{SP}$) when the inequality reverses, which occurs when competition is intense (t small) and business stealing effects dominate externalities. For platform F, the analysis is similar but complicated by its cost disadvantage. Platform F may overinvest to compensate for lower market share, or underinvest due to the same externalities affecting G.

This completes the proof of Proposition 2. □

A.6 Proof of Proposition 3

Under mandatory data sharing with parameter $\sigma \in (0, 1]$, platform F gains access to a fraction of platform G's data, so its effective data integration becomes:

$$\tilde{d}_F = d_F + \sigma d_G$$

Stage 2: Market Shares with Data Sharing. In Case 1 (universal opt-out), the marginal customization-savvy consumer satisfies:

$$\xi_C \alpha[(d_G)^\phi - (d_F + \sigma d_G)^\phi] = t(1 - 2\tilde{x}_C)$$

Solving:

$$\tilde{x}_C = \frac{1}{2} + \frac{\xi_C \alpha[(d_G)^\phi - (d_F + \sigma d_G)^\phi]}{2t}$$

Similarly, for privacy-savvy consumers:

$$\tilde{x}_P^O = \frac{1}{2} + \frac{\alpha[(d_G)^\phi - (d_F + \sigma d_G)^\phi]}{2t}$$

Platform G's market share becomes:

$$N_G = \frac{1}{2} + \frac{\alpha\bar{\xi}[(d_G)^\phi - (d_F + \sigma d_G)^\phi]}{2t} \tag{20}$$

Stage 1: Investment with Data Sharing. Platform G's profit function is:

$$\Pi_G^S(d_G, d_F) = \frac{\bar{\mu}^2\beta(d_G)^\psi N_G(d_G, d_F + \sigma d_G)}{4} - \frac{\kappa_G d_G^{1+\eta}}{1+\eta}$$

Taking the first-order condition:

$$\frac{\partial \Pi_G^S}{\partial d_G} = \frac{\bar{\mu}^2\beta}{4}\left[\psi(d_G)^{\psi-1}N_G + (d_G)^\psi \frac{\partial N_G}{\partial d_G}\right] - \kappa_G(d_G)^\eta = 0$$

The market share derivative now includes the dilution effect:

$$\begin{aligned}
\frac{\partial N_G}{\partial d_G} &= \frac{\alpha\bar{\xi}}{2t}\left[\phi(d_G)^{\phi-1} - \phi\sigma(d_F + \sigma d_G)^{\phi-1}\right] \\
&= \frac{\alpha\bar{\xi}\phi}{2t}\left[(d_G)^{\phi-1} - \sigma(d_F + \sigma d_G)^{\phi-1}\right]
\end{aligned}$$

The term $-\sigma(d_F + \sigma d_G)^{\phi-1}$ represents the dilution effect: when G increases d_G, platform F automatically benefits through $\tilde{d}_F = d_F + \sigma d_G$, reducing G's marginal market share gain.

Comparing with the no-sharing case where $\frac{\partial N_G}{\partial d_G}\Big|_{\sigma=0} = \frac{\alpha\bar{\xi}\phi(d_G)^{\phi-1}}{2t}$:

$$\frac{\partial N_G}{\partial d_G}\bigg|_{\sigma>0} < \frac{\partial N_G}{\partial d_G}\bigg|_{\sigma=0}$$

This establishes that data sharing unambiguously reduces platform G's marginal benefit from investment.

Effect on Platform G's Investment. Taking the total derivative of platform G's FOC with respect to σ:

$$\frac{\partial^2 \Pi_G^S}{\partial d_G^2}\frac{\partial d_G^S}{\partial \sigma} + \frac{\partial^2 \Pi_G^S}{\partial d_G \partial \sigma} = 0$$

By the implicit function theorem:

$$\frac{\partial d_G^S}{\partial \sigma} = -\frac{\frac{\partial^2 \Pi_G^S}{\partial d_G \partial \sigma}}{\frac{\partial^2 \Pi_G^S}{\partial d_G^2}}$$

The cross-partial derivative is:

$$\frac{\partial^2 \Pi_G^S}{\partial d_G \partial \sigma} = \frac{\bar{\mu}^2\beta}{4}\left[\psi(d_G)^{\psi-1}\frac{\partial N_G}{\partial \sigma} + (d_G)^{\psi}\frac{\partial^2 N_G}{\partial d_G \partial \sigma}\right]$$

From equation (20):

$$\frac{\partial N_G}{\partial \sigma} = -\frac{\alpha\bar{\xi}\phi d_G(d_F + \sigma d_G)^{\phi-1}}{2t} < 0$$

Since $\frac{\partial N_G}{\partial \sigma} < 0$ and the second-order condition requires $\frac{\partial^2 \Pi_G^S}{\partial d_G^2} < 0$:

$$\frac{\partial d_G^S}{\partial \sigma} < 0 \tag{21}$$

Platform G reduces investment in response to data sharing.

Effect on Platform F's Investment. Platform F's profit function is:

$$\Pi_F^S(d_G, d_F) = \frac{\bar{\mu}^2\beta(d_F + \sigma d_G)^{\psi} N_F(d_G, d_F + \sigma d_G)}{4} - \frac{\kappa_F d_F^{1+\eta}}{1+\eta}$$

The first-order condition is:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi(d_F + \sigma d_G)^{\psi-1}N_F + (d_F + \sigma d_G)^{\psi}\frac{\partial N_F}{\partial d_F}\right] = \kappa_F(d_F)^{\eta}$$

Taking the total derivative with respect to σ:

$$\frac{\partial d_F^S}{\partial \sigma} = -\frac{\frac{\partial^2 \Pi_F^S}{\partial d_F \partial \sigma}}{\frac{\partial^2 \Pi_F^S}{\partial d_F^2}} + \frac{\frac{\partial^2 \Pi_F^S}{\partial d_F \partial d_G}}{\frac{\partial^2 \Pi_F^S}{\partial d_F^2}} \cdot \frac{\partial d_G^S}{\partial \sigma}$$

The direct effect through σ is:

$$\frac{\partial^2 \Pi_F^S}{\partial d_F \partial \sigma} = \frac{\bar{\mu}^2 \beta}{4} \Bigg[\psi(\psi-1)(d_F + \sigma d_G)^{\psi-2} d_G N_F + \psi(d_F + \sigma d_G)^{\psi-1} \frac{\partial N_F}{\partial \sigma} + \psi(d_F + \sigma d_G)^{\psi-1} d_G \frac{\partial N_F}{\partial d_F} + (d_F + \sigma d_G)^{\psi} \frac{\partial^2 N_F}{\partial d_F \partial \sigma} \Bigg]$$

The sign of this expression is ambiguous due to competing effects:

1. **Free-riding effect:** the terms involving d_G in the effective integration $\tilde{d}_F = d_F + \sigma d_G$ increase revenue without additional cost, reducing the marginal return to proprietary investment d_F. This effect tends to make $\frac{\partial^2 \Pi_F^S}{\partial d_F \partial \sigma} < 0$;
2. **Competitive pressure effect:** As platform F's effective integration increases through sharing, its market share N_F grows, raising the base over which advertising revenue is earned. This increases the return to further investment, tending to make $\frac{\partial^2 \Pi_F^S}{\partial d_F \partial \sigma} > 0$.

The net effect depends on β, the cost disadvantage $\kappa_F - \kappa_G$, the initial integration gap $d_G^* - d_F^*$, and the elasticities ψ and ϕ.

Welfare Analysis. Total welfare under data sharing is:

$$W^S(\sigma) = CS(d_G^S(\sigma), d_F^S(\sigma), \sigma) + \frac{3\bar{\mu}^2\beta}{8}[(d_G^S)^{\psi} N_G + (d_F^S + \sigma d_G^S)^{\psi} N_F] - C_G(d_G^S(\sigma)) - C_F(d_F^S(\sigma))$$

Taking the derivative with respect to σ and evaluating at $\sigma = 0$:

$$\begin{aligned}
\left.\frac{dW^S}{d\sigma}\right|_{\sigma=0} = & \underbrace{\frac{\partial CS}{\partial \sigma}}_{\text{Direct CS effect}} + \underbrace{\frac{\partial CS}{\partial d_G}\frac{\partial d_G^S}{\partial \sigma}}_{\text{CS effect via } d_G} + \underbrace{\frac{\partial CS}{\partial d_F}\frac{\partial d_F^S}{\partial \sigma}}_{\text{CS effect via } d_F} \\
& + \underbrace{\frac{3\bar{\mu}^2\beta}{8}\left[\psi(d_G^*)^{\psi-1} N_G \frac{\partial d_G^S}{\partial \sigma} + (d_G^*)^{\psi}\frac{\partial N_G}{\partial d_G}\frac{\partial d_G^S}{\partial \sigma}\right]}_{\text{Advertising effect via } d_G} \\
& + \underbrace{\frac{3\bar{\mu}^2\beta}{8}\left[\psi(d_F^*)^{\psi-1} N_F \left(\frac{\partial d_F^S}{\partial \sigma} + d_G^*\right) + (d_F^*)^{\psi}\frac{\partial N_F}{\partial d_F}\frac{\partial d_F^S}{\partial \sigma}\right]}_{\text{Advertising effect via } d_F \text{ and } \sigma} \\
& \underbrace{- \kappa_G (d_G^*)^{\eta}\frac{\partial d_G^S}{\partial \sigma}}_{\text{Cost effect for } G} \underbrace{- \kappa_F (d_F^*)^{\eta}\frac{\partial d_F^S}{\partial \sigma}}_{\text{Cost effect for } F}
\end{aligned}$$

The direct consumer surplus effect is:

$$\frac{\partial CS}{\partial \sigma} = (1-\theta) N_F \alpha \phi (d_G^*)^{\phi} \cdot (d_G^*)^{\phi-1} > 0$$

This captures the quality improvement for privacy-savvy consumers on platform F who benefit from access to platform G's data. The direct advertising effect is:

$$\frac{3\bar{\mu}^2\beta\psi(d_F^*)^{\psi-1} N_F d_G^*}{8} > 0$$

This represents improved targeting precision on platform F.

Using the FOCs from equations (8) and (9) and the envelope theorem, the terms involving $\frac{\partial d_G^S}{\partial \sigma}$

and $\frac{\partial d_F^S}{\partial \sigma}$ can be simplified. From platform G's FOC:

$$\frac{\bar{\mu}^2\beta}{4}\left[\psi(d_G^*)^{\psi-1}N_G + (d_G^*)^{\psi}\frac{\partial N_G}{\partial d_G}\right] = \kappa_G(d_G^*)^{\eta}$$

Therefore:

$$\begin{aligned}
&\frac{3\bar{\mu}^2\beta}{8}\left[\psi(d_G^*)^{\psi-1}N_G + (d_G^*)^{\psi}\frac{\partial N_G}{\partial d_G}\right] - \kappa_G(d_G^*)^{\eta} \\
&= \frac{3}{2}\cdot\frac{\bar{\mu}^2\beta}{4}\left[\psi(d_G^*)^{\psi-1}N_G + (d_G^*)^{\psi}\frac{\partial N_G}{\partial d_G}\right] - \kappa_G(d_G^*)^{\eta} \\
&= \frac{3}{2}\kappa_G(d_G^*)^{\eta} - \kappa_G(d_G^*)^{\eta} \\
&= \frac{1}{2}\kappa_G(d_G^*)^{\eta} > 0
\end{aligned}$$

This represents the advertising externality that platform G fails to internalize. The social planner values advertising surplus at $\frac{3\bar{\mu}^2\beta}{8}$ while platform G captures only $\frac{\bar{\mu}^2\beta}{4}$. Similarly, for platform F:

$$\frac{3\bar{\mu}^2\beta}{8}\left[\psi(d_F^*)^{\psi-1}N_F + (d_F^*)^{\psi}\frac{\partial N_F}{\partial d_F}\right] - \kappa_F(d_F^*)^{\eta} = \frac{1}{2}\kappa_F(d_F^*)^{\eta} > 0$$

Substituting these simplifications:

$$\begin{aligned}
\left.\frac{dW^S}{d\sigma}\right|_{\sigma=0} &= (1-\theta)N_F\alpha\phi(d_G^*)^{2\phi-1} + \frac{3\bar{\mu}^2\beta\psi(d_F^*)^{\psi-1}N_F d_G^*}{8} \\
&\quad + \frac{1}{2}\kappa_G(d_G^*)^{\eta}\frac{\partial d_G^S}{\partial \sigma} + \frac{1}{2}\kappa_F(d_F^*)^{\eta}\frac{\partial d_F^S}{\partial \sigma} \\
&\quad + \frac{\partial CS}{\partial d_G}\frac{\partial d_G^S}{\partial \sigma} + \frac{\partial CS}{\partial d_F}\frac{\partial d_F^S}{\partial \sigma}
\end{aligned}$$

Conditions for Welfare Improvement. Data sharing improves welfare if and only if:

$$(1-\theta)N_F\alpha\phi(d_G^*)^{2\phi-1} + \frac{3\bar{\mu}^2\beta\psi(d_F^*)^{\psi-1}N_F d_G^*}{8} > -\Delta W^{investment}$$

where:

$$\Delta W^{investment} = \left[\frac{\partial CS}{\partial d_G} + \frac{1}{2}\kappa_G(d_G^*)^{\eta}\right]\frac{\partial d_G^S}{\partial \sigma} + \left[\frac{\partial CS}{\partial d_F} + \frac{1}{2}\kappa_F(d_F^*)^{\eta}\right]\frac{\partial d_F^S}{\partial \sigma}$$

Since $\frac{\partial d_G^S}{\partial \sigma} < 0$ from equation (21), the first term is negative (welfare loss from reduced G investment) when:

$$\frac{\partial CS}{\partial d_G} + \frac{1}{2}\kappa_G(d_G^*)^{\eta} > 0$$

This condition holds when platform G underinvests in the no-sharing equilibrium, i.e., when the marginal social benefit of G's investment exceeds its marginal cost. The second term depends on the sign of $\frac{\partial d_F^S}{\partial \sigma}$. Specifically, if $\frac{\partial d_F^S}{\partial \sigma} > 0$ (competitive pressure dominates free-riding), the welfare effect is positive when platform F underinvests; if $\frac{\partial d_F^S}{\partial \sigma} < 0$, the welfare effect is negative when platform F underinvests.

Therefore, data sharing improves welfare when:

1. **Static gains dominate:** $(1-\theta)N_F$ is large (many privacy-savvy consumers on F) and d_G^* is large (substantial data to share), making the direct CS and advertising effects large;

2. **Platform G overinvests:** the term $\frac{\partial CS}{\partial d_G} + \frac{1}{2}\kappa_G(d_G^*)^\eta < 0$, so reducing G's investment generates a welfare gain;

3. **Platform F increases investment:** $\frac{\partial d_F^S}{\partial \sigma} > 0$ and platform F underinvests ($\frac{\partial CS}{\partial d_F} + \frac{1}{2}\kappa_F(d_F^*)^\eta > 0$), so sharing induces welfare-improving investment by F.

Conversely, data sharing reduces welfare when platform G underinvests in the no-sharing equilibrium, platform F reduces investment in response to sharing (free-riding effect), and static gains are small relative to dynamic losses from reduced investment.

This completes the proof of Proposition 3. □

A.7 Proof of Proposition 4

Platform j operates K services indexed by $k \in \{1, \dots, K\}$, each generating service-specific data d_{kj}. With unrestricted pooling, effective data integration is:

$$d_j = \Gamma(d_{1j}, \dots, d_{Kj}) = \sum_{k=1}^{K} \omega_k d_{kj} + \rho \sum_{k \neq \ell} d_{kj} d_{\ell j}$$

where $\omega_k \in [0, 1]$ with $\sum_{k=1}^{K} \omega_k = 1$ captures the relative importance of service k, and $\rho > 0$ measures economies of scope in cross-service data reuse. Under siloing, cross-service data flows are prohibited, restricting effective integration to:

$$d_j^{\text{Sil}} = \sum_{k=1}^{K} \omega_k d_{kj}$$

The welfare impact of siloing is evaluated by comparing outcomes under unrestricted pooling ($\rho > 0$) versus siloing ($\rho = 0$).

Investment Incentives. Platform j's profit function with service-specific data investments is:

$$\Pi_j(d_{1j}, \dots, d_{Kj}) = \frac{\bar{\mu}^2 \beta}{4} \left[\Gamma(d_{1j}, \dots, d_{Kj})\right]^\psi N_j - \sum_{k=1}^{K} \frac{\kappa_{kj} d_{kj}^{1+\eta}}{1+\eta}$$

where $\kappa_{kj} > 0$ is the marginal cost parameter for service k and platform j. The FOC for investment in service k is:

$$\frac{\partial \Pi_j}{\partial d_{kj}} = \frac{\bar{\mu}^2 \beta \psi}{4} \left[\Gamma(d_{1j}, \dots, d_{Kj})\right]^{\psi - 1} \frac{\partial \Gamma}{\partial d_{kj}} N_j - \kappa_{kj} d_{kj}^\eta = 0 \tag{22}$$

Computing the partial derivative of Γ with respect to d_{kj}:

$$\frac{\partial \Gamma}{\partial d_{kj}} = \omega_k + \rho \sum_{\ell \neq k} d_{\ell j}$$

Under unrestricted pooling ($\rho > 0$), the marginal benefit from investing in service k includes both the direct effect ω_k and the complementarity effect $\rho \sum_{\ell \neq k} d_{\ell j}$. Under siloing ($\rho = 0$), only the direct effect remains. Taking the cross-partial derivative:

$$\frac{\partial^2 \Gamma}{\partial d_{kj} \partial \rho} = \sum_{\ell \neq k} d_{\ell j} \geq 0 \tag{23}$$

with strict inequality when $d_{\ell j} > 0$ for some $\ell \neq k$. This establishes that the marginal return to investing in service k is weakly increasing in ρ. By the implicit function theorem applied to equation (22):

$$\frac{\partial d_{kj}^*}{\partial \rho} = -\frac{\frac{\partial^2 \Pi_j}{\partial d_{kj} \partial \rho}}{\frac{\partial^2 \Pi_j}{\partial d_{kj}^2}}$$

The numerator is:

$$\frac{\partial^2 \Pi_j}{\partial d_{kj} \partial \rho} = \frac{\bar{\mu}^2 \beta \psi}{4} N_j \left[(\psi - 1) \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\psi - 2} \frac{\partial \Gamma}{\partial d_{kj}} \frac{\partial \Gamma}{\partial \rho} + \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\psi - 1} \frac{\partial^2 \Gamma}{\partial d_{kj} \partial \rho} \right]$$

Since $\frac{\partial \Gamma}{\partial \rho} = \sum_{k \neq \ell} d_{kj} d_{\ell j} > 0$ and $\frac{\partial^2 \Gamma}{\partial d_{kj} \partial \rho} > 0$ (from equation (23)), both terms in the numerator are positive when $\psi > 0$. The denominator satisfies $\frac{\partial^2 \Pi_j}{\partial d_{kj}^2} < 0$ by the second-order condition for profit maximization. Therefore:

$$\frac{\partial d_{kj}^*}{\partial \rho} > 0$$

Cross-service complementarities increase investment incentives in each service. Siloing ($\rho = 0$) reduces investment relative to unrestricted pooling.

Consumer Surplus. Consumer surplus aggregates utility across all consumer types and platforms. In general form:

$$CS = \sum_{i \in \mathcal{I}} \sum_{j \in \{G,F\}} \theta_i N_{ij} \left[v + \xi_i \alpha \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\phi} - OC_i - TC_{ij} \right]$$

where $\mathcal{I}$ indexes consumer types, θ_i is the population share of type i, N_{ij} is the mass of type-i consumers on platform j, ξ_i is the net value of data integration for type i (accounting for personalization and privacy), OC_i is the opt-out cost (if applicable), and TC_{ij} represents transportation costs. Taking the derivative with respect to ρ:

$$\begin{aligned} \frac{\partial CS}{\partial \rho} &= \sum_{i \in \mathcal{I}} \sum_{j \in \{G,F\}} \theta_i N_{ij} \xi_i \alpha \phi \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\phi - 1} \frac{\partial \Gamma}{\partial \rho} \\ &= \sum_{j \in \{G,F\}} \left(\sum_{i \in \mathcal{I}} \theta_i N_{ij} \xi_i \right) \alpha \phi \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\phi - 1} \sum_{k \neq \ell} d_{kj} d_{\ell j} \end{aligned} \tag{24}$$

Define the population-weighted average net value on platform j as:

$$\bar{\xi}_j \equiv \frac{\sum_{i \in \mathcal{I}} \theta_i N_{ij} \xi_i}{\sum_{i \in \mathcal{I}} \theta_i N_{ij}} = \frac{\sum_{i \in \mathcal{I}} \theta_i N_{ij} \xi_i}{N_j} \tag{25}$$

Then equation (24) becomes:

$$\frac{\partial CS}{\partial \rho} = \sum_{j \in \{G,F\}} N_j \bar{\xi}_j \alpha \phi \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\phi - 1} \sum_{k \neq \ell} d_{kj} d_{\ell j}$$

The sign of $\frac{\partial CS}{\partial \rho}$ depends on whether the average net value $\bar{\xi}_j$ is positive or negative on each platform. When quality benefits dominate privacy costs on average ($\bar{\xi}_j > 0$), consumer surplus increases with cross-service complementarities. When privacy costs dominate ($\bar{\xi}_j < 0$), consumer surplus may decrease.

Privacy Costs. Privacy costs are borne by consumers who accept personalized advertising (those who do not opt out). Let $\mathcal{I}_{\text{accept},j}$ denote the set of consumer types accepting personalized advertising on platform j. Aggregate privacy costs are:

$$PC = \sum_{j \in \{G,F\}} \sum_{i \in \mathcal{I}_{\text{accept},j}} \theta_i N_{ij} \delta \cdot \Gamma(d_{1j}, \ldots, d_{Kj})$$

Taking the derivative:

$$\frac{\partial PC}{\partial \rho} = \sum_{j \in \{G,F\}} \sum_{i \in \mathcal{I}_{\text{accept},j}} \theta_i N_{ij} \delta \cdot \frac{\partial \Gamma}{\partial \rho} = \sum_{j \in \{G,F\}} \left(\sum_{i \in \mathcal{I}_{\text{accept},j}} \theta_i N_{ij} \right) \delta \sum_{k \neq \ell} d_{kj} d_{\ell j}$$

Define the population share accepting personalized advertising on platform j as:

$$\bar{\theta}_{\text{accept},j} \equiv \frac{\sum_{i \in \mathcal{I}_{\text{accept},j}} \theta_i N_{ij}}{N_j}$$

Then:

$$\frac{\partial PC}{\partial \rho} = \sum_{j \in \{G,F\}} N_j \bar{\theta}_{\text{accept},j} \delta \sum_{k \neq \ell} d_{kj} d_{\ell j} > 0$$

Cross-service complementarities unambiguously increase privacy costs for consumers who accept personalized advertising.

Advertising-Side Surplus. Platform profits and advertiser surplus jointly depend on the data-weighted consumer base. Total advertising-side surplus is:

$$AS + \Pi = \frac{3\bar{\mu}^2 \beta}{8} \sum_{j \in \{G,F\}} \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\psi} N_j$$

Taking the derivative:

$$\begin{aligned} \frac{\partial (AS + \Pi)}{\partial \rho} &= \frac{3\bar{\mu}^2 \beta \psi}{8} \sum_{j \in \{G,F\}} \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\psi - 1} N_j \frac{\partial \Gamma}{\partial \rho} \\ &= \frac{3\bar{\mu}^2 \beta \psi}{8} \sum_{j \in \{G,F\}} N_j \left[\Gamma(d_{1j}, \ldots, d_{Kj}) \right]^{\psi - 1} \sum_{k \neq \ell} d_{kj} d_{\ell j} > 0 \end{aligned}$$

Cross-service complementarities unambiguously increase advertising-side surplus by enhancing targeting precision.

Investment Costs. Total investment costs are:

$$TC = \sum_{j \in \{G,F\}} \sum_{k=1}^{K} \frac{\kappa_{kj} d_{kj}^{1+\eta}}{1 + \eta}$$

Taking the derivative:

$$\frac{\partial TC}{\partial \rho} = \sum_{j \in \{G,F\}} \sum_{k=1}^{K} \kappa_{kj} d_{kj}^{\eta} \frac{\partial d_{kj}}{\partial \rho}$$

Using the first-order condition (22):

$$\kappa_{kj} d_{kj}^{\eta} = \frac{\bar{\mu}^2 \beta \psi}{4} \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\psi - 1} \frac{\partial \Gamma}{\partial d_{kj}} N_j$$

By the envelope theorem, the direct effect of ρ on costs through induced changes in d_{kj} vanishes to first order at the optimal investment levels. The welfare calculation therefore focuses on the direct effects through quality, privacy, and advertising channels.

Total Welfare. Total welfare is:

$$W = CS - PC + AS + \Pi - TC$$

Taking the derivative with respect to ρ and applying the envelope theorem:

$$\begin{aligned}
\frac{dW}{d\rho} &= \frac{\partial CS}{\partial \rho} - \frac{\partial PC}{\partial \rho} + \frac{\partial (AS + \Pi)}{\partial \rho} \\
&= \sum_{j \in \{G,F\}} N_j \bar{\xi}_j \alpha \phi \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\phi - 1} \sum_{k \neq \ell} d_{kj} d_{\ell j} \\
&\quad - \sum_{j \in \{G,F\}} N_j \bar{\theta}_{\text{accept},j} \delta \sum_{k \neq \ell} d_{kj} d_{\ell j} \\
&\quad + \frac{3\bar{\mu}^2 \beta \psi}{8} \sum_{j \in \{G,F\}} N_j \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\psi - 1} \sum_{k \neq \ell} d_{kj} d_{\ell j}
\end{aligned}$$

Factoring out $\sum_{k \neq \ell} d_{kj} d_{\ell j}$ for each platform:

$$\begin{aligned}
\frac{dW}{d\rho} = \sum_{j \in \{G,F\}} N_j \sum_{k \neq \ell} d_{kj} d_{\ell j} \Bigg[& \bar{\xi}_j \alpha \phi \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\phi - 1} \\
& - \bar{\theta}_{\text{accept},j} \delta + \frac{3\bar{\mu}^2 \beta \psi}{8} \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\psi - 1} \Bigg]
\end{aligned}$$

Siloing improves welfare if and only if $\frac{dW}{d\rho} < 0$, which requires:

$$\bar{\theta}_{\text{accept},j} \delta > \bar{\xi}_j \alpha \phi \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\phi - 1} + \frac{3\bar{\mu}^2 \beta \psi}{8} \left[\Gamma(d_{1j}, \ldots, d_{Kj})\right]^{\psi - 1} \tag{26}$$

for at least one platform j with sufficient weight in the welfare function.

Condition (26) trades off three marginal effects.

First, the term $\bar{\theta}_{\text{accept},j} \delta$ captures the privacy cost borne by consumers who accept personalized advertising when cross-service data use expands. This effect is stronger when a larger share of users is exposed to personalized ads and when privacy losses δ are substantial. Second, the term $\bar{\xi}_j \alpha \phi \left[\Gamma(\cdot)\right]^{\phi - 1}$ measures the marginal improvement in service quality generated by pooling data across services. This channel is more important when average net valuations $\bar{\xi}_j$ and quality sensitivity α are high, and when diminishing returns to data-driven quality improvements are weak. Third, $\frac{3\bar{\mu}^2 \beta \psi}{8} \left[\Gamma(\cdot)\right]^{\psi - 1}$ captures the marginal gain in advertising efficiency arising from richer user profiles. This effect is amplified by high advertiser willingness to pay and weak diminishing returns in targeting.

Siloing is therefore more likely to increase welfare when privacy costs are large relative to the quality and advertising benefits of data integration, when a sizable fraction of consumers accepts personalized advertising, and when cross-service data use primarily facilitates privacy-invasive profiling rather than efficiency-enhancing improvements in service quality or ad targeting. Moreover, strong diminishing returns to data integration (i.e., low ϕ and ψ) limit the marginal value of complementarities, further tilting the welfare comparison in favor of siloing. Conversely, when cross-service data pooling generates substantial quality improvements or advertising efficiency gains that dominate privacy harms, imposing data siloing reduces aggregate welfare.

Regime-Specific Implications. The welfare effect varies across opt-out regimes:

- **Case 1:** when all privacy-savvy consumers opt out on both platforms, $\bar{\theta}_{\text{accept},j} = \theta$ (only customization-savvy consumers accept advertising). Privacy costs are minimal, so $\frac{dW}{d\rho} > 0$ and siloing unambiguously reduces welfare.
- **Case 2:** when opt-out costs are prohibitively high, $\bar{\theta}_{\text{accept},j} = 1 - \theta + \theta = 1$ on both platforms (all consumers exposed to personalized advertising, though customization-savvy consumers experience no privacy costs). Here, $\bar{\theta}_{\text{accept},j}\delta = (1-\theta)\delta$, and siloing may improve welfare if $(1-\theta)\delta$ is sufficiently large.
- **Case 3:** when privacy-savvy consumers opt out on platform G but accept on platform F, $\bar{\theta}_{\text{accept},G} = \theta$ and $\bar{\theta}_{\text{accept},F} \in (\theta, 1)$. The welfare effect depends on the platform-specific balance between privacy harms and efficiency gains.

This completes the proof of Proposition 4. □

A.8 Proof of Proposition 5

Consider a marginal reduction in opt-out cost from c to $c - \varepsilon$ in Case 3.

Effect on Consumer Platform Choice. The marginal privacy-savvy consumer satisfies:

$$v + (\alpha - \lambda)d_G - c - t\tilde{x}_P^{OA} = v + \xi_P^A d_F - t(1 - \tilde{x}_P^{OA})$$

Taking the total derivative with respect to c:

$$-1 - t\frac{\partial \tilde{x}_P^{OA}}{\partial c} = -t\frac{\partial(1 - \tilde{x}_P^{OA})}{\partial c} = t\frac{\partial \tilde{x}_P^{OA}}{\partial c}$$

Solving:

$$\frac{\partial \tilde{x}_P^{OA}}{\partial c} = -\frac{1}{2t} < 0$$

Reducing the opt-out cost causes the marginal privacy-savvy consumer to shift toward platform G (where opt-out is now more attractive). Platform G's market share is $N_G = \theta\tilde{x}_C + (1-\theta)\tilde{x}_P^{OA}$. Since customization-savvy consumers are unaffected by c:

$$\frac{\partial N_G}{\partial c} = (1-\theta)\frac{\partial \tilde{x}_P^{OA}}{\partial c} = -\frac{1-\theta}{2t} < 0$$

Reducing c shifts $(1-\theta)\frac{\varepsilon}{2t}$ privacy-savvy consumers from platform F to platform G.

Welfare Effects. Total welfare in Case 3 is:

$$\begin{aligned}W^{(3)} &= CS^{(3)} + \frac{3\bar{\mu}^2\beta}{8}(d_G N_G + d_F N_F) \\ &= \theta CS_C + (1-\theta)CS_P^{OA} + \frac{3\bar{\mu}^2\beta}{8}(d_G N_G + d_F N_F)\end{aligned}$$

Taking the derivative with respect to c:

$$\frac{dW^{(3)}}{dc} = (1-\theta)\frac{\partial CS_P^{OA}}{\partial c} + \frac{3\bar{\mu}^2\beta}{8}\frac{\partial(d_G N_G + d_F N_F)}{\partial c}$$

1. **Direct consumer surplus effect.** For privacy-savvy consumers who opt out (those on platform G), the direct utility effect is:

$$\frac{\partial U_{P,G}^{O}}{\partial c} = -1 < 0$$

Each opt-out user loses one unit of utility per unit increase in c. The mass of privacy-savvy consumers on G is $(1-\theta)\tilde{x}_P^{OA}$, so:

$$\left.\frac{\partial CS_P^{OA}}{\partial c}\right|_{\text{direct}} = -(1-\theta)\tilde{x}_P^{OA} < 0 \tag{27}$$

2. **Reallocation effect.** The consumers who switch from F to G in response to a reduction in c are marginal consumers located at $\tilde{x}_P^{OA}$. Before switching, their utility on F was:

$$U_{P,F}^{A}(\tilde{x}_P^{OA}) = v + \xi_P^A d_F - t(1-\tilde{x}_P^{OA})$$

After switching to G and opting out, their utility is:

$$U_{P,G}^{O}(\tilde{x}_P^{OA}) = v + \alpha d_G - c - t\tilde{x}_P^{OA}$$

At the margin, these consumers are indifferent, so the first-order effect on consumer surplus from switching is zero. However, there is a second-order effect from the fact that consumers are sorting toward the platform with higher quality (since $d_G > d_F$ implies $\alpha d_G > \alpha d_F$). The net utility gain for a switching consumer is:

$$\begin{aligned}\Delta U &= [\alpha d_G - c - t\tilde{x}_P^{OA}] - [\alpha d_F - \delta d_F - t(1-\tilde{x}_P^{OA})] \\ &= \alpha(d_G - d_F) - c + \delta d_F - t(2\tilde{x}_P^{OA} - 1)\end{aligned}$$

At the equilibrium cutoff, this equals zero. But as c decreases, more consumers find it optimal to switch, and the infra-marginal consumers (those with $x < \tilde{x}_P^{OA}$) experience a utility gain from the reduced opt-out cost. The reallocation effect on consumer surplus is therefore:

$$\left.\frac{\partial CS_P^{OA}}{\partial c}\right|_{\text{reallocation}} = 0$$

to first order.

3. **Advertising efficiency effect.** The data-weighted consumer base is $d_G N_G + d_F N_F$. Taking the

derivative:

$$\begin{aligned}\frac{\partial(d_G N_G + d_F N_F)}{\partial c} &= d_G \frac{\partial N_G}{\partial c} + d_F \frac{\partial N_F}{\partial c} \\ &= d_G \cdot \left(-\frac{1-\theta}{2t}\right) + d_F \cdot \frac{1-\theta}{2t} \\ &= -\frac{(1-\theta)(d_G - d_F)}{2t} < 0\end{aligned}$$

since $d_G > d_F$ by assumption. Reducing c (increasing $-\partial/\partial c$) shifts consumers from the low-integration platform F to the high-integration platform G, increasing the data-weighted consumer base. The advertising-side welfare effect is:

$$\frac{3\bar{\mu}^2\beta}{8}\frac{\partial(d_G N_G + d_F N_F)}{\partial c} = -\frac{3\bar{\mu}^2\beta(1-\theta)(d_G - d_F)}{16t} < 0 \tag{28}$$

Total Welfare Effect. Combining equations (27) and (28):

$$\begin{aligned}\frac{dW^{(3)}}{dc} &= -(1-\theta)\tilde{x}_P^{OA} - \frac{3\bar{\mu}^2\beta(1-\theta)(d_G - d_F)}{16t} \\ &= -(1-\theta)\left[\tilde{x}_P^{OA} + \frac{3\bar{\mu}^2\beta(d_G - d_F)}{16t}\right] < 0\end{aligned}$$

Both terms are negative, establishing that:

$$\frac{dW^{(3)}}{dc} < 0$$

Reducing the opt-out cost (decreasing c) unambiguously increases total welfare.

Pareto Dominance over Data Integration Reduction. Compare this with a policy that reduces platform G's data integration from d_G to $d'_G < d_G$. From Proposition 1, reducing d_G in Case 3 creates:

- welfare losses for customization-savvy consumers: $-\theta\tilde{x}_C\xi_C(d_G - d'_G) < 0$
- welfare losses for the advertising side: $-\frac{3\bar{\mu}^2\beta}{8}[\text{terms}] < 0$
- ambiguous effects for privacy-savvy consumers depending on α versus δ

while reducing opt-out costs:

- creates welfare gains for privacy-savvy consumers who opt out: $(1-\theta)\tilde{x}_P^{OA}\varepsilon > 0$
- creates welfare gains from improved advertising efficiency: $\frac{3\bar{\mu}^2\beta(1-\theta)(d_G-d_F)\varepsilon}{16t} > 0$
- leaves customization-savvy consumers unaffected: $\Delta CS_C = 0$
- leaves platform profits unaffected to first order through the envelope theorem

No consumer group experiences reduced utility, and two groups strictly benefit and therefore opt-out cost reduction Pareto-dominates data integration reduction.

Comparative Statics. The welfare gain from reducing opt-out costs is larger when:

1. Data integration gap $d_G - d_F$ is large. The advertising efficiency effect in equation (28) is proportional to $d_G - d_F$. When platforms are highly asymmetric, consumer reallocation toward the high-integration platform generates larger advertising-side benefits.

2. Proportion $(1-\theta)$ of privacy-savvy consumers is large. The direct consumer surplus effect is proportional to the mass of privacy-savvy consumers who benefit from reduced opt-out costs.

3. Advertising benefits $\bar{\mu}^2\beta$ are large. The value of improved targeting precision increases with advertiser willingness-to-pay.

4. t is small. When horizontal differentiation is weak, consumer reallocation is more responsive to changes in opt-out costs, amplifying both the direct CS effect and the advertising efficiency effect.

This completes the proof of Proposition 5. □

B Strategic quality degradation from opt-out

The baseline model assumes that quality improvements from data integration αd_j accrue equally to all consumers regardless of their opt-out decisions. This abstracts from an important real-world mechanism: platforms may reduce service quality for users who opt out of personalized advertising, either for legitimate technical reasons or as a strategic response to increase the effective cost of opting out.

Quality with opt-out degradation. I modify the quality function to distinguish between users who accept personalized advertising and those who opt out. For users who accept personalized advertising on platform j, quality remains:

$$Q_j^A = \alpha d_j$$

while for users who opt out of personalized advertising, quality becomes:

$$Q_j^O = (\alpha - \lambda) d_j$$

where $\lambda \geq 0$ represents the quality degradation parameter which captures both the technical degradation λ^T and the strategic degradation λ^S. On one side, quality reductions arising from the technical impossibility of providing certain personalized features without individual-level data processing (eg. personalized recommendations, customized search rankings, and tailored content feeds that fundamentally require user-specific behavioral data). On the other side, the platform can deliberately reduce the service quality in order to increase the effective cost of opting out, thereby coercing users into accepting personalized advertising (eg. by removing features that are technically feasible without personalization, degrading user interfaces for opt-out users, or limiting access to functionality).

The total degradation is $\lambda = \lambda^T + \lambda^S$. From a welfare perspective, technical degradation represents an unavoidable technological constraint, while strategic degradation represents a distortion that reduces consumer surplus without generating offsetting benefits.

Consumers. For customization-savvy consumers, utility is unchanged as they never opt out: $U_{C,j}(x) = v + \xi_C d_j - t|x - \ell_j|$. For privacy-savvy consumers who accept personalized advertising $U^A_{P,j}(x) = v + \alpha d_j - \delta d_j - t|x - \ell_j| = v + \xi^A_P d_j - t|x - \ell_j|$, while for privacy-savvy consumers who opt out, utility now reflects quality degradation $U^O_{P,j}(x) = v + (\alpha - \lambda)d_j - c - t|x - \ell_j|$. Privacy-savvy consumers choose to opt out if and only if the utility from opting out exceeds the utility from accepting personalized advertising $U^O_{P,j}(x) > U^A_{P,j}(x)$, i.e.:

$$
\begin{aligned}
v + (\alpha - \lambda)d_j - c - t|x - \ell_j| &> v + \alpha d_j - \delta d_j - t|x - \ell_j| \\
(\alpha - \lambda)d_j - c &> (\alpha - \delta)d_j \\
-\lambda d_j - c &> -\delta d_j \\
\delta d_j - \lambda d_j &> c \\
(\delta - \lambda)d_j &> c
\end{aligned}
$$

and therefore, privacy-savvy consumers opt out if and only if:

$$c < (\delta - \lambda)d_j$$

Lemma 4 (Effect of quality degradation on opt-out decision)**.** *Quality degradation increases the effective cost of opting out. Specifically, when $\lambda = 0$ (no degradation), the opt-out condition is $c < \delta d_j$, when $\lambda > 0$ (with degradation), the opt-out condition is $c < (\delta - \lambda)d_j$, which is strictly more restrictive, while when $\lambda \geq \delta$ (severe degradation), opting out is never optimal regardless of c, as the quality loss exceeds the privacy protection.*

Equilibrium regimes. The three equilibrium regimes are now determined by comparing c to the modified privacy cost net of quality degradation, $(\delta - \lambda)d_j$.

- Case 1. This regime obtains when $c < (\delta - \lambda)d_F < (\delta - \lambda)d_G$, so that privacy-savvy consumers find it optimal to opt out on both platforms despite the quality degradation. The marginal privacy-savvy consumer satisfies:

  $$v + (\alpha - \lambda)d_G - c - t\tilde{x}^O_P = v + (\alpha - \lambda)d_F - c - t(1 - \tilde{x}^O_P)$$

  Solving:

  $$\tilde{x}^O_P = \frac{1}{2} + \frac{(\alpha - \lambda)(d_G - d_F)}{2t} \tag{29}$$

  platform G's market share becomes:

  $$N_G = \theta \tilde{x}_C + (1 - \theta)\tilde{x}^O_P = \frac{1}{2} + \frac{d_G - d_F}{2t}[\theta \xi_C + (1 - \theta)(\alpha - \lambda)]$$

- Case 2. This regime obtains when $c > (\delta - \lambda)d_G > (\delta - \lambda)d_F$, so that privacy-savvy consumers find it optimal to accept personalized advertising on both platforms. The analysis proceeds identically to the baseline model, as quality degradation is irrelevant for consumers who accept personalized advertising. Platform G's market share is:

$$N_G = \frac{1}{2} + \frac{d_G - d_F}{2t}(\alpha + \theta\gamma - (1-\theta)\delta)$$

- Case 3. This regime obtains when $(\delta - \lambda)d_F < c < (\delta - \lambda)d_G$, so that privacy-savvy consumers opt out on platform G but accept personalized advertising on platform F. The marginal privacy-savvy consumer satisfies:

$$v + (\alpha - \lambda)d_G - c - t\tilde{x}_P^{OA} = v + \xi_P^A d_F - t(1 - \tilde{x}_P^{OA})$$

Solving:

$$\tilde{x}_P^{OA} = \frac{1}{2} + \frac{(\alpha - \lambda)d_G - c - \xi_P^A d_F}{2t} = \frac{1}{2} + \frac{(\alpha - \lambda)(d_G - d_F) + \delta d_F - \lambda d_F - c}{2t}$$

and simplifying:

$$\tilde{x}_P^{OA} = \frac{1}{2} + \frac{(\alpha - \lambda)(d_G - d_F) + (\delta - \lambda)d_F - c}{2t}$$

Welfare (Example). Consider Case 1. For privacy-savvy consumers who opt out on both platforms, consumer surplus is:

$$CS_P^O = (1-\theta)\left[v + (\alpha - \lambda)(\tilde{x}_P^O d_G + (1 - \tilde{x}_P^O)d_F) - c - \frac{t}{2}((\tilde{x}_P^O)^2 + (1 - \tilde{x}_P^O)^2)\right] \tag{30}$$

Total consumer surplus is:

$$CS^{(1)} = \theta CS_C + (1-\theta)CS_P^O$$

Welfare effect of quality degradation. Consider the welfare impact of increasing λ from zero to some positive value. The direct effect on consumer surplus for privacy-savvy consumers who opt out is:

$$\frac{\partial CS_P^O}{\partial \lambda} = -(1-\theta)[\tilde{x}_P^O d_G + (1 - \tilde{x}_P^O)d_F] < 0$$

This represents a pure welfare loss: quality degradation reduces consumer surplus for privacy-savvy consumers without generating any offsetting benefits. Importantly, this loss is proportional to the data-weighted presence of privacy-savvy consumers on both platforms.

Additionally, quality degradation affects the marginal privacy-savvy consumer's platform choice. From equation (29):

$$\frac{\partial \tilde{x}_P^O}{\partial \lambda} = -\frac{d_G - d_F}{2t} < 0$$

Quality degradation causes privacy-savvy consumers to shift away from platform G (which has higher data integration and thus higher absolute quality loss from degradation) toward platform F. This reallocation effect has second-order welfare implications through changes in the data-weighted consumer base.

Welfare implications of strategic vs. technical degradation. When $\lambda = \lambda^T$ reflects genuine technical constraints, quality degradation is an unavoidable consequence of opt-out. In this case, the social planner faces a constrained optimization problem where privacy-savvy consumers must trade off privacy protection against quality loss. The optimal opt-out decision for a privacy-savvy consumer balances benefit from opting out (avoiding privacy cost δd_j) and costs of opting out (paying opt-out cost c and suffering quality loss $\lambda^T d_j$). The modified net benefit from opting out is:

$$NB^O = \delta d_j - c - \lambda^T d_j = (\delta - \lambda^T) d_j - c$$

Technical degradation reduces the net benefit of opting out, making acceptance of personalized advertising relatively more attractive. This is efficient from a social welfare perspective when the quality loss reflects complementarities between personalization and service provision.

When $\lambda = \lambda^S$ reflects deliberate platform choices to degrade quality for opt-out users, the welfare implications are unambiguously negative. Strategic degradation creates a deadweight loss by:

- reducing consumer surplus: privacy-savvy consumers who opt out experience lower quality without any technological justification, representing a pure welfare loss of $(1-\theta)\lambda^S[\tilde{x}_P^O d_G + (1-\tilde{x}_P^O)d_F]$ per period;
- coercing suboptimal choices: consumers with privacy valuations $v_i \in (\delta - \lambda^S, \delta)$ are forced to accept personalized advertising despite preferring to opt out in the absence of strategic degradation, creating a distortion in consumer choice;
- increasing effective opt-out costs: strategic degradation functions as an implicit increase in the opt-out cost from c to $c + \lambda^S d_j$, reducing consumer surplus without generating offsetting benefits.

Proposition 6 (Welfare loss from strategic degradation). *Strategic quality degradation $\lambda^S > 0$ unambiguously reduces total welfare by:*

$$\Delta W = -(1-\theta)\lambda^S[\tilde{x}_P^O d_G + (1-\tilde{x}_P^O)d_F] - \textit{Distortion from coerced choices} < 0$$

where the first term represents the direct consumer surplus loss for opt-out users and the second term captures the welfare loss from consumers making suboptimal opt-out decisions.

The distinction between technical and strategic degradation has important implications for optimal regulatory policy. Regulatory authorities should prohibit strategic quality degradation by requiring platforms to provide functionally equivalent services to opt-out users whenever technically feasible. Specifically, platforms must ensure that $\lambda^S = 0$, providing the same core functionality to users who opt out of personalized advertising as to users who accept it and should bear the burden of demonstrating that the degradation is technologically unavoidable rather than strategically chosen, i.e. $\lambda^T > 0$.

Even when quality degradation is technically necessary, it should be proportionate to the data processing that is forgone through opt-out. Features not directly dependent on personalized data should remain accessible.

Proposition 7 (Dominance of anti-degradation policies). *A regulatory policy prohibiting strategic quality degradation (enforcing $\lambda^S = 0$) Pareto-dominates data integration reduction in all equilibrium regimes as privacy-savvy consumers who opt out are strictly better off (higher quality), other consumer groups are unaffected, and platform profits and advertiser surplus are* *unchanged. Therefore, enforcing service equivalence should be prioritized over structural remedies.*

Proof. Consider the welfare function with strategic degradation $\lambda^S > 0$ versus the welfare function with $\lambda^S = 0$ (ceteris paribus). From equation (30), consumer surplus for privacy-savvy consumers is strictly increasing in quality, which is decreasing in λ. Setting $\lambda^S = 0$ increases consumer surplus by:

$$\Delta CS_P^O = (1-\theta)\lambda^S[\tilde{x}_P^O d_G + (1-\tilde{x}_P^O)d_F] > 0$$

Customization-savvy consumers never opt out, so their utility is unaffected: $\Delta CS_C = 0$. Platform profits depend only on advertising revenue, which depends on market shares and data integration levels. Since prohibiting strategic degradation does not directly affect these variables (it only changes the quality received by opt-out users), platform profits are unchanged: $\Delta\Pi = 0$. Advertiser surplus similarly depends only on the data-weighted consumer base and is unaffected: $\Delta AS = 0$. Therefore, total welfare increases by $\Delta W = \Delta CS_P^O > 0$, with no group experiencing a welfare loss. This establishes Pareto dominance.

In contrast, data integration reduction from d_G to $d'_G < d_G$ creates welfare losses for customization-savvy consumers, platforms, and advertisers. Even in Case 2 where data integration reduction may increase welfare under restrictive conditions, it involves redistributing surplus rather than Pareto-improving. □