

Online Appendix to
“The Impact of Privacy Protection on Online
Advertising Markets”*

Miguel Alcobendas[†] Shunto J. Kobayashi[‡] Ke Shi[§] Matthew Shum[¶]

*This appendix accompanies the article “The Impact of Privacy Protection on Online Advertising Markets.”

[†]Yahoo Research

[‡]Questrom School of Business, Boston University

[§]Shanghai Advanced Institute of Finance (SAIF), Shanghai Jiao Tong University

[¶]California Institute of Technology

A Additional reduced-form results

Table A.1: Regression results of logit model of entry decision

Dependent Variable:	Entry (1)
Cookie	1.191*** (0.053)
Opt-out	0.084* (0.046)
Computer	-0.207*** (0.028)
Gender female	0.164*** (0.010)
Gender male	0.098*** (0.015)
Age 24 and below	-0.086*** (0.021)
Age 25 to 44	-0.099*** (0.020)
Age 45 to 64	-0.113*** (0.021)
Age 65 and above	-0.140*** (0.023)
Interest segments	0.057*** (0.008)
Months monetized	0.002*** (0.000)
Total revenue (normalized)	-0.030*** (0.002)
Days in database	-0.037*** (0.004)
<i>Fixed effects</i>	
Time (hour)	Yes
City	Yes
Website	Yes
Browser	Yes
Observations	2,652,282

Notes: Estimation results of auction participation using logit model with 10% of the data. The base levels for age and gender are both Unknown. Standard errors are clustered by the hour of the day, the city, and the website and are heteroskedasticity-robust. ***, **, and * indicate statistical significance at the 1, 5, and 10% levels, respectively.

B Inner loop algorithms

B.1 MPEC formulation for solving the inverse bidding functions

For notational convenience, we omit the dependence on (μ, c, u) . We approximate the inverse bidding functions $\varphi_g(b)$ as a linear combination of the first K Chebyshev polynomials scaled to the interval $[r, \bar{b}]$:

$$\varphi_g(b) = \sum_{k=0}^K w_{k,g} T_k(b; r, \bar{b}), \quad (\text{B.1})$$

where $T_k(b; r, \bar{b})$ is the Chebyshev polynomial of degree k scaled to the interval $[r, \bar{b}]$.¹

Then, we use the mathematical programming with equilibrium constraints (MPEC) approach to discipline the Chebyshev coefficients \mathbf{c} so that the first-order conditions defining the inverse equilibrium bid functions are approximately satisfied, subject to the boundary conditions (4) and (5). In addition, we impose rationality (players must bid less than their valuation) and monotonicity (bid functions are increasing) as shape constraints on the Chebyshev approximations (Hubbard and Paarsch, 2009; Hubbard, Kirkegaard, and Paarsch, 2013). Specifically, from equation (3), we define the deviation from equilibrium

$$G_g(b; \mathbf{w}, \bar{b}) = 1 - (\varphi_g(b) - b) \sum_{h=1}^G (N_h - \mathbb{1}_{h=g}) \frac{f_h(\varphi_h(b))}{F_h(\varphi_h(b))} \varphi_h'(b), \quad (\text{B.2})$$

where \mathbf{w} are the coefficients of the linear combination of Chebyshev polynomials. Let \mathcal{B} be the set of Chebyshev nodes in $[r, \bar{b}]$. For every auction t , we solve the following

¹We set \bar{v} equal to the maximum observed bid across all auctions in the data, which is 276 times the average observed bid. We linearly extrapolate the inverse bidding functions for bids above $\bar{b}(\mu, c, u)$. At our estimates, this affects only 1.33% of observed bids.

constrained optimization problem to obtain φ_g and \bar{b} :²

$$\min_{\mathbf{w}, \bar{b}} \sum_{i=1}^N \sum_{b \in \mathcal{B}} [G_g(b; \mathbf{w}, \bar{b})]^2 \quad (\text{B.3})$$

$$\text{s.t. } \varphi_g(r) = r, \varphi_g(\bar{b}) = \bar{v}, \varphi_g(b) \geq b, \varphi'_g(b) \geq 0, \text{ for } g = 1, \dots, G \text{ and } b \in \mathcal{B}.$$

B.2 Grid construction and interpolation

To avoid solving the MPEC problem separately for every (hypothetical) auction when evaluating likelihood, we exploit the fact that under our model, an auction is fully characterized by $(\mu = x'\beta, c, r, u)$, where c denotes cookie availability, $r \in \mathcal{R}$ is the reserve price, and u is the latent auction type. We solve the MPEC only on a grid over (μ, c, r, u) and interpolate across μ .

Figure B.1 shows that the solved Chebyshev coefficients vary smoothly in μ , with no numerical irregularities across grid points. This smoothness ensures that the cubic spline interpolation used in Step 2 closely matches the MPEC solutions, making the interpolation accurate and allowing us to avoid repeatedly solving the MPEC during likelihood evaluation.

Step 1: grid over (μ, c, r)

For each cookie status $c \in \{0, 1\}$ and each reserve price $r \in \mathcal{R}$:

1. Compute $\mu_t = x'_t\beta$ for all auctions with $(c_t, r_t) = (c, r)$ and determine the empirical support $[\underline{\mu}_{cr}, \bar{\mu}_{cr}]$.
2. Construct an evenly spaced grid $\{\mu_{cr}^{(1)}, \dots, \mu_{cr}^{(M)}\}$ over this support.³
3. For each grid point $\mu_{cr}^{(m)}$ and each latent type $u = 1, \dots, U$, solve the MPEC problem (B.3) to obtain

$$\{w_{k,g}(\mu_{cr}^{(m)}, c, r, u)\}_{k=0:K, g=1:G} \quad \text{and} \quad \bar{b}(\mu_{cr}^{(m)}, c, r, u).$$

²In the implementation, we use the first $K = 5$ order Chebyshev polynomials and 500 Chebyshev nodes for \mathcal{B} to numerically approximate the inverse bid functions. These specifications are sufficiently flexible for approximations in our setting.

³We set the grid size as 60 points for our estimation.

Step 2: cubic spline interpolation

For each (c, r, u) , we apply cubic spline interpolation in μ to the equilibrium quantities obtained in Step 1—specifically, the Chebyshev coefficients $\{w_{k,g}\}$ and the upper bid support \bar{b} . For every type g and coefficient k , we fit a spline to the grid values

$$w_{k,g}(\mu_{cr}^{(m)}, c, r, u), \quad m = 1, \dots, M,$$

and similarly for $\bar{b}(\mu_{cr}^{(m)}, c, r, u)$. This yields smooth functions

$$w_{k,g}(\mu, c, r, u), \quad k = 0, \dots, K, \quad g = 1, \dots, G,$$

and

$$\bar{b}(\mu, c, r, u),$$

defined for all μ in the empirical support for (c, r) .

Step 3: constructing $\varphi_{gtu}(b)$ for likelihood evaluation

For any auction with characteristics (μ_t, c_t, r_t, u) , we evaluate the interpolated Chebyshev coefficients at μ_t :

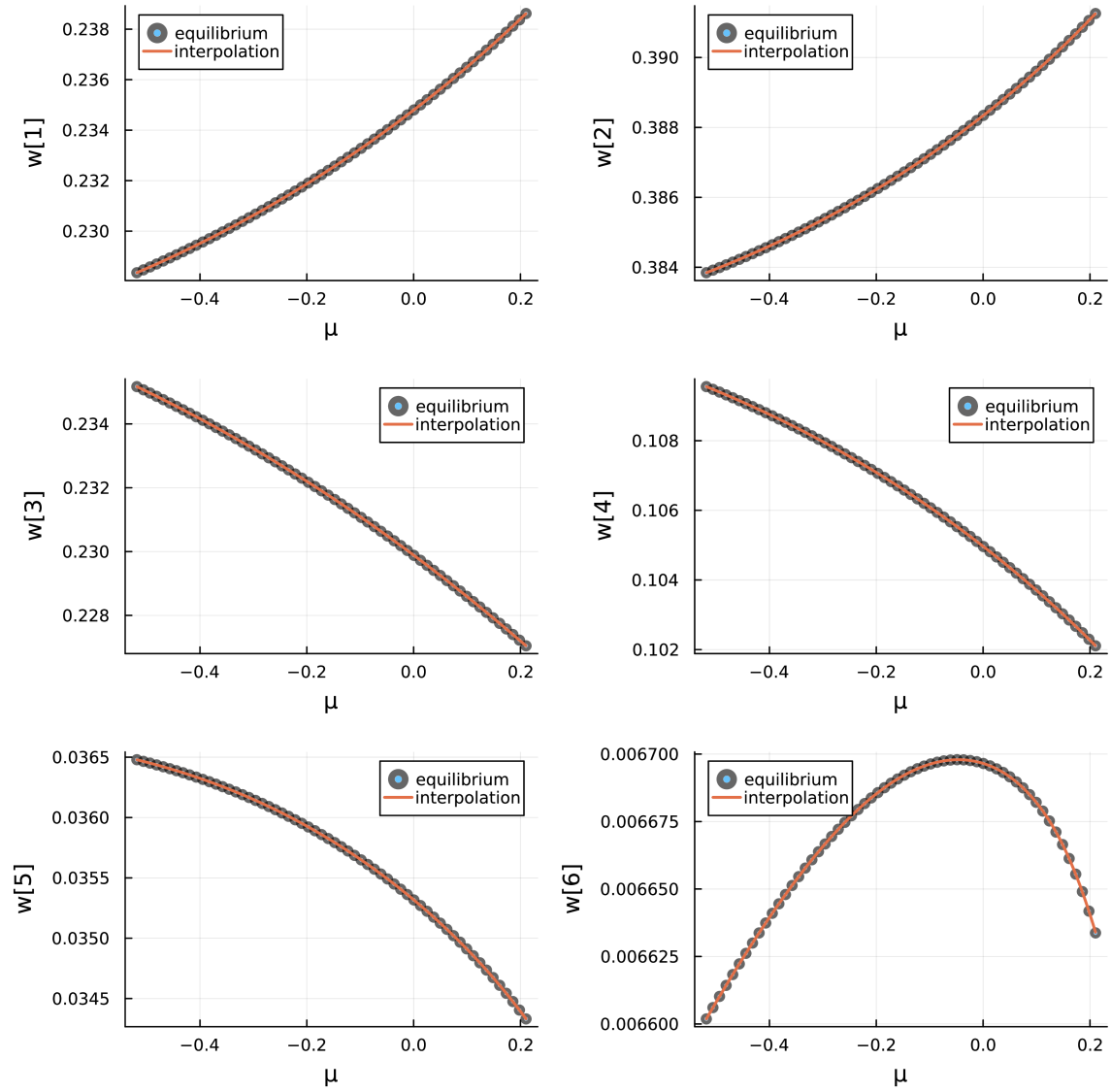
$$w_{k,g}(\mu_t, c_t, r_t, u), \quad k = 0, \dots, K, \quad g = 1, \dots, G.$$

Substituting these values into the Chebyshev expansion yields the inverse bidding functions

$$\varphi_{gtu}(b) = \sum_{k=0}^K w_{k,g}(\mu_t, c_t, r_t, u) T_k(b; r, \bar{b}(\mu_t, c_t, r_t, u))$$

and their derivatives $\varphi'_{gtu}(b)$, which are used in the likelihood evaluation.

Figure B.1: Chebyshev coefficients and their interpolation

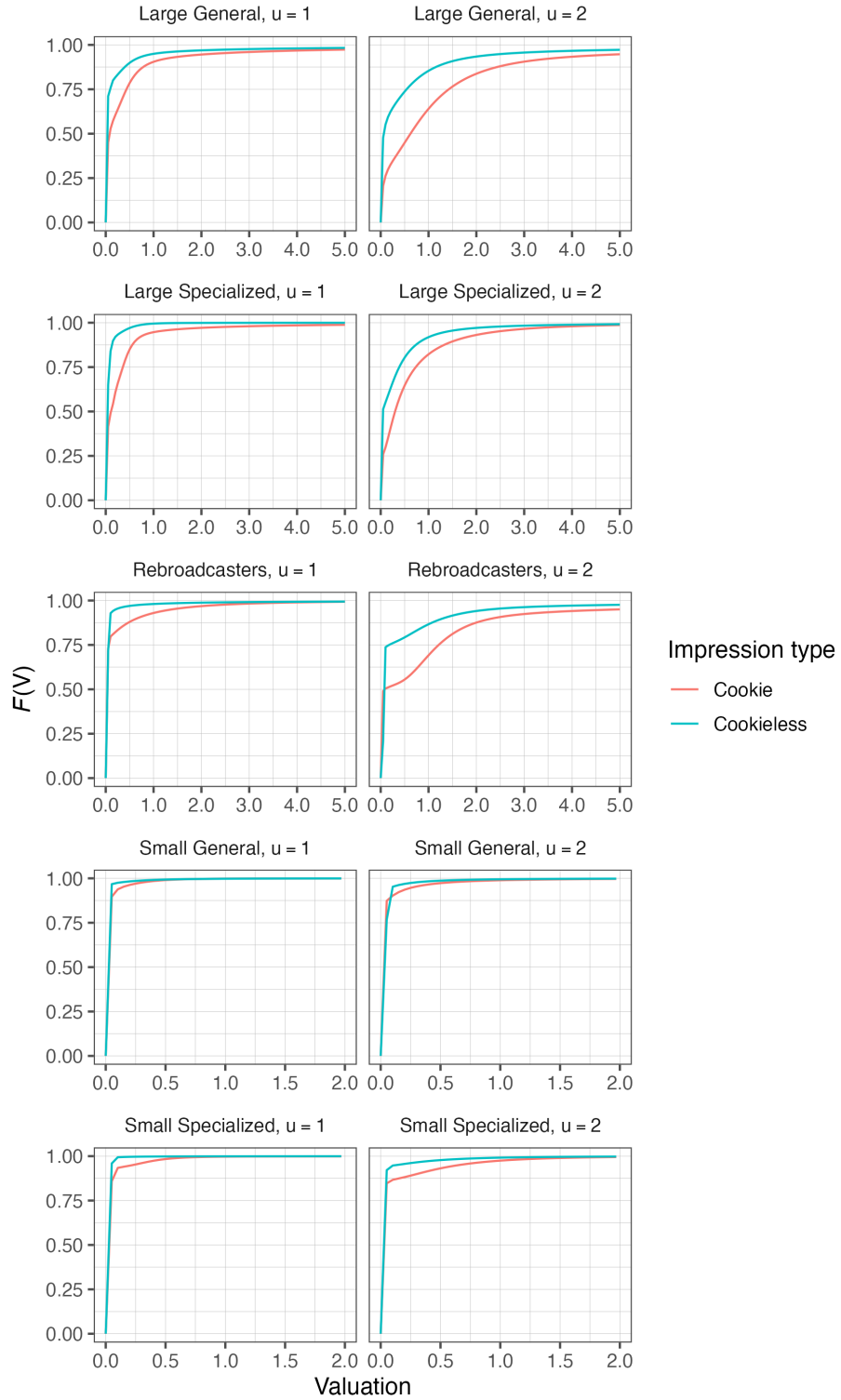


Notes: This set of plots shows the Chebyshev coefficients and their interpolation for large general-purpose DSP type ($g = 1$) in auctions for cookie impressions $c = 1$ and latent auction type $u = 1$ with the lowest reserve price $r = 0.064$.

C Residual valuation in valuation distributions

Figure C.1 shows the empirical distributions of residual valuations, defined as log valuations net of the estimated single-index term. Distributions are reported separately for cookie and cookieless impressions and for each latent auction type $u \in \{1, 2\}$.

Figure C.1: Residual valuations



Notes: This figure shows the empirical distributions of residual valuations, defined as log valuations net of the estimated single-index term. Distributions are reported separately for cookie and cookieless impressions and for each latent auction type $u \in \{1, 2\}$.

D Counterfactual 3: platform data sharing

The third counterfactual simulates a scenario where, after third-party cookies are removed from Chrome, DSPs still receive the demographic information that Yahoo observes for its own users (gender and age categories). All other impression-level information that DSPs could condition on in the baseline, such as detailed browsing histories, cross-site signals, or richer targeting attributes, is no longer available and is treated as unobserved by the bidders. Cookie availability is therefore effectively zero for all impressions. This counterfactual isolates the incremental value of platform-level demographic sharing: it represents the most informative scenario Yahoo could implement using its own first-party footprint while still falling far short of the granular behavioral data that third-party cookies previously provided.

Table [D.1a](#) reports the effects of platform-level demographic sharing. Under the asymmetric-ban scenario, average bids fall from \$0.91 to about \$0.79 (a 13% decline), and participation drops from 7.3 bidders to roughly 4.9. Publisher revenue declines from \$2.17 to \$1.53 (a reduction of around 30%), while bidder surplus falls from \$2.01 to about \$1.68 (−16%). Overall, the simulated outcomes remain qualitatively similar to those under Cookiepocalypse (Table [7a](#)) and fall well short of the status quo, indicating that platform-provided demographic data offers limited, if any, ability to mitigate the efficiency and welfare losses generated by the removal of third-party cookies.

The redistribution of welfare under platform data sharing closely resembles that observed under Cookiepocalypse. As shown in Table [D.1b](#), the Big Tech DSP wins about 12% of auctions and earns roughly \$30,200 in surplus under the asymmetric ban, which is only slightly below the \$30,400 it attains under Cookiepocalypse and still substantially more than any other DSP group. Other DSP groups exhibit only limited improvement. Large specialized DSPs, for example, earn about \$3,890, a modest improvement over the \$3,720 they receive under Cookiepocalypse, but still well below the status quo. Overall, demographic sharing provides only a very limited, if any, remedy for the redistributive and competitive distortions generated by the loss of third-party cookies.

Table D.1: Counterfactual 3: Platform Data Sharing by Yahoo

(a) Simulated outcome

	Status quo	Symmetric ban	Asymmetric ban
Bid	0.908 (1.303)	0.775 (1.125)	0.786 (1.357)
No. bidders	7.269 (4.212)	4.811 (3.571)	4.943 (3.562)
Publisher revenue	2.169 (2.706)	1.476 (2.050)	1.529 (2.665)
Bidder surplus	2.011 (4.298)	1.662 (4.133)	1.684 (4.143)

(b) Welfare distribution

	Status quo	Symmetric ban	Asymmetric ban
Panel A: Winning frequency			
Big Tech DSP	-	-	0.118
Large general-purpose	0.071	0.076	0.073
Small general-purpose	0.003	0.004	0.004
Rebroadcaster	0.041	0.036	0.035
Large specialized	0.032	0.029	0.027
Small specialized	0.002	0.002	0.002
Panel B: Surplus			
Big Tech DSP	-	-	30,200
Large general-purpose	26,200	22,900	22,300
Small general-purpose	235	250	242
Rebroadcaster	14,100	12,400	12,200
Large specialized	7,280	3,770	3,890
Small specialized	204	154	150
Panel C: Full-information surplus			
Big Tech DSP	-	-	30,200
Large general-purpose	-	29,400	28,500
Small general-purpose	-	324	312
Rebroadcaster	-	15,800	15,400
Large specialized	-	7,370	7,220
Small specialized	-	262	248

Notes: Simulated results are based on 100,000 auctions randomly drawn from the data. The Big Tech DSP is drawn from the large general-purpose DSP group. For Chrome impressions, auction characteristics are masked for all bidders in the symmetric ban scenario and are available exclusively to the Big Tech DSP in the asymmetric ban scenario. For each scenario, valuations are updated according to counterfactual characteristics, and outcomes are recomputed using the equilibrium characterization.

References

- HUBBARD, T. P., R. KIRKEGAARD, AND H. J. PAARSCH (2013): “Using Economic Theory to Guide Numerical Analysis: Solving for Equilibria in Models of Asymmetric First-Price Auctions,” *Computational Economics*, 42(2), 241–266.
- HUBBARD, T. P., AND H. J. PAARSCH (2009): “Investigating bid preferences at low-price, sealed-bid auctions with endogenous participation,” *International Journal of Industrial Organization*, 27(1), 1–14.