

The Data Privacy Paradox and Digital Demand

Long Chen, Yadong Huang, Shumiao Ouyang, Wei Xiong

Online Appendix

The Survey Questionnaire

- Q1. Are you concerned about privacy issues while using online services?
- Q2. What do you think about privacy protection in Alipay?
- Q3. Are you concerned about negative impacts caused by information shared to mini-programs in Alipay?
- Q4. Will you avoid visiting mini-programs in Alipay because of privacy concerns?
- Q5. What privacy issues are you concerned about when using mini-programs in Alipay? (You may select multiple choices.)
- A. Data leakage and security;
 - B. Price discrimination by merchants;
 - C. Seductive advertising and temptation consumption;
 - D. Others
- Q6. How many times will you agree if making authorization decisions for ten mini-programs?
- Q7. How often do you regret authorizing information to mini-programs in Alipay?
- Q8. Do you agree with the arguments below?
- 1) I agree to authorize data sharing with mini-programs since it is safe in Alipay.
 - 2) I agree to authorize data sharing with mini-programs since my information has already been shared in many platforms.
 - 3) I have to share my personal data in exchange for digital services even though I am concerned about my data privacy.
 - 4) I authorize data sharing with a mini-program only when the requested information is not important.
 - 5) I tend to authorize data sharing with mini-programs that are used by my friends.
- Q9. Do you know how to change privacy settings in Alipay?
- Q10. Have you ever changed your privacy settings in Alipay?
- Q11. Do you know how to opt out from mini-programs in Alipay?
- Q12. Have you ever opted out from mini-programs in Alipay?

Figure A1: Examples of the Authorization Page

This figure presents three examples of the authorization page with different information requirements. Users need to agree to share requested information before using mini-programs.

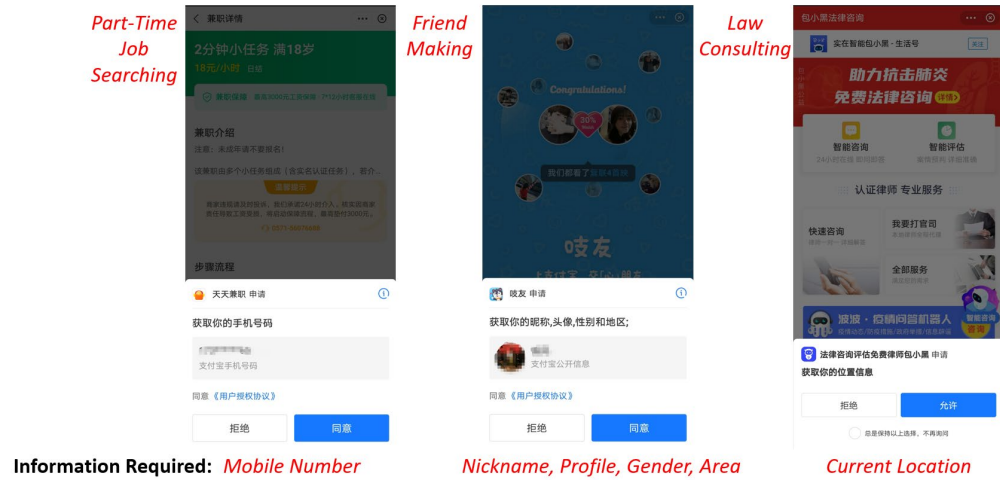


Figure A2: Completion Time Distribution (Seconds)

This figure plots the completion time distribution (seconds) in the survey conducted in July 2020 by Alipay. The vertical axis refers to the percentage of responses.

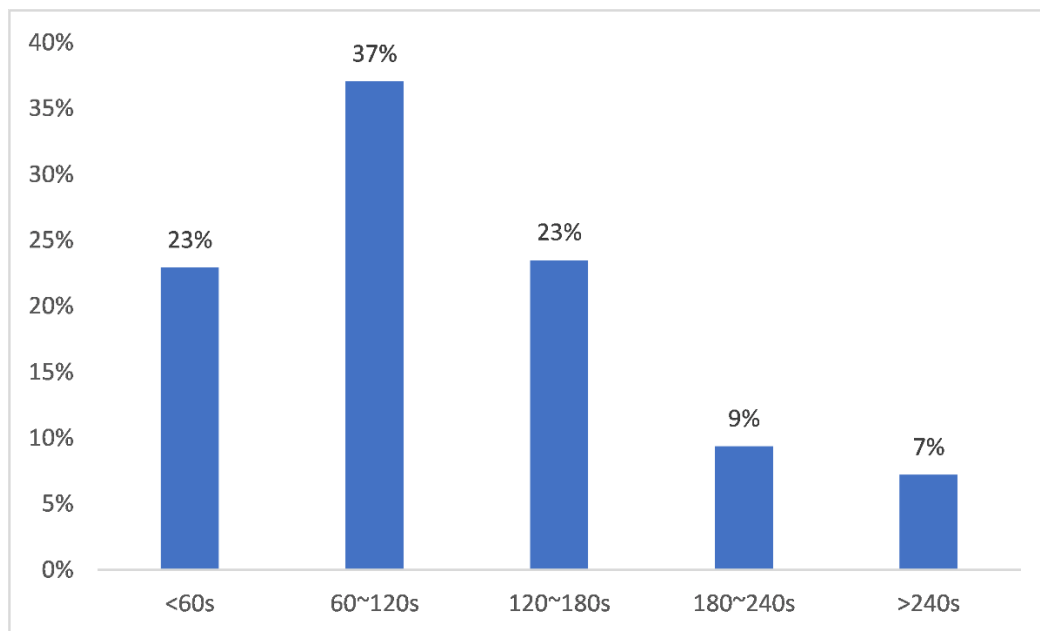


Figure A3: Age Distribution (Years)

This figure plots the age distribution of respondents from the survey conducted in July 2020 by Alipay. The vertical axis refers to the percentage of responses.

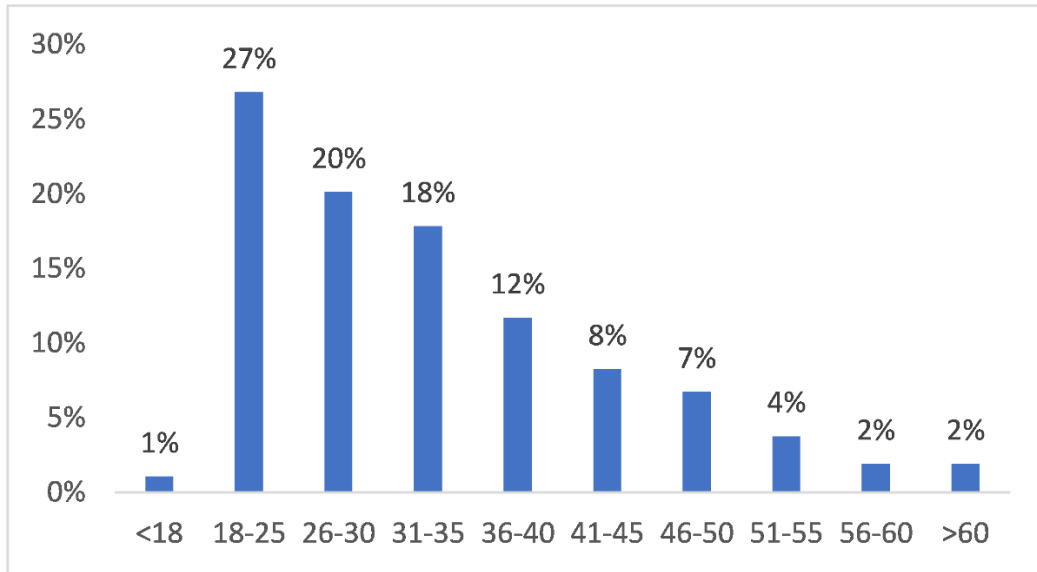


Figure A4: Distribution Across Provinces

This figure plots the linear correlation between distribution of respondents in the survey and distribution of the population across provinces. The vertical axis refers to the percentage of responses; the horizontal axis refers to the percentage of resident population (2019).

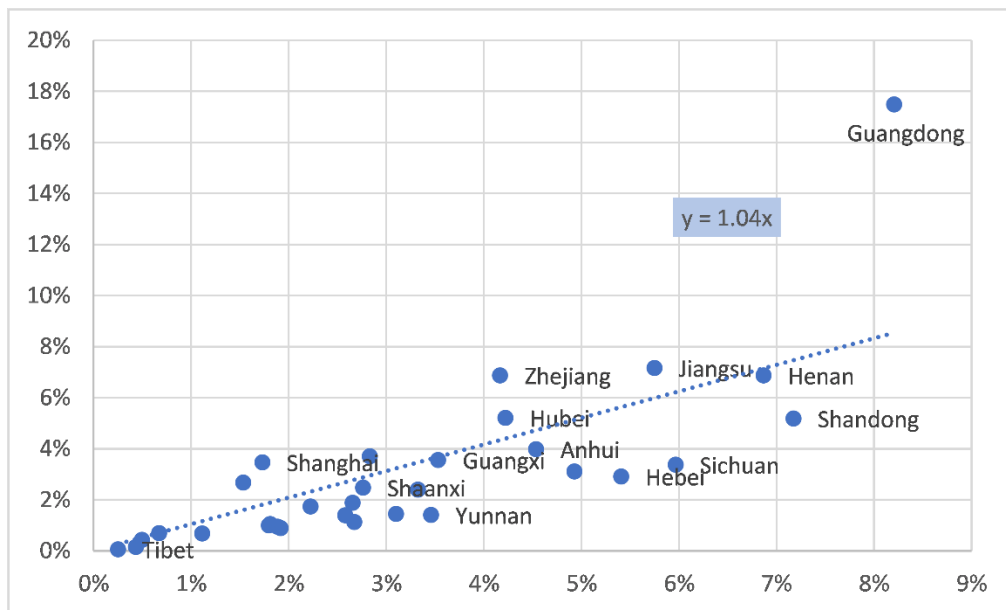


Figure A5: Digital Age Distribution (Years)

This figure plots the digital age distribution of respondents from the survey conducted in July 2020 by Alipay. The vertical axis refers to the percentage of responses; digital age is defined as the length of time since a user registered with Alipay.

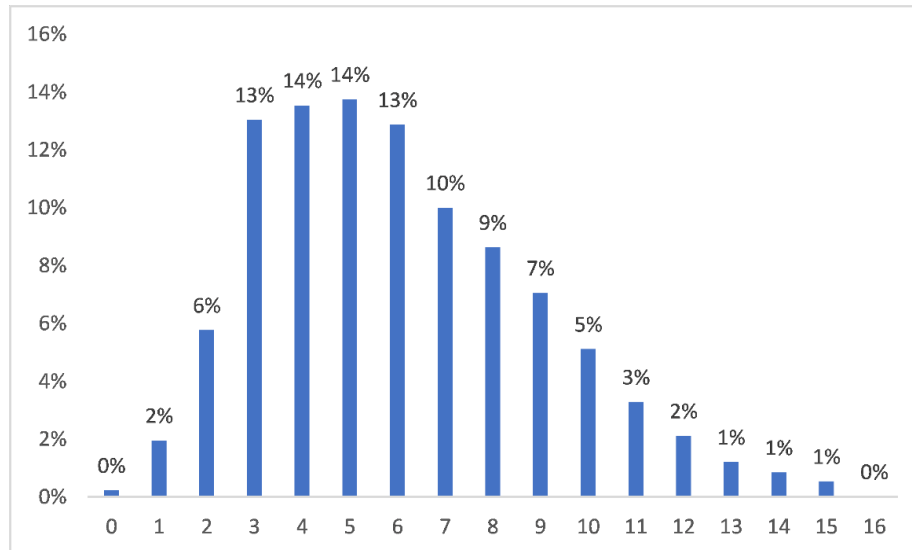


Figure A6: Cancellations During the 2017 Footprint Report Incident

The figure plots the fraction of Alipay users that canceled at least one mini-program authorizations in each day around the 2017 Annual User Footprint Report Incident, which happened on January 3, 2018.

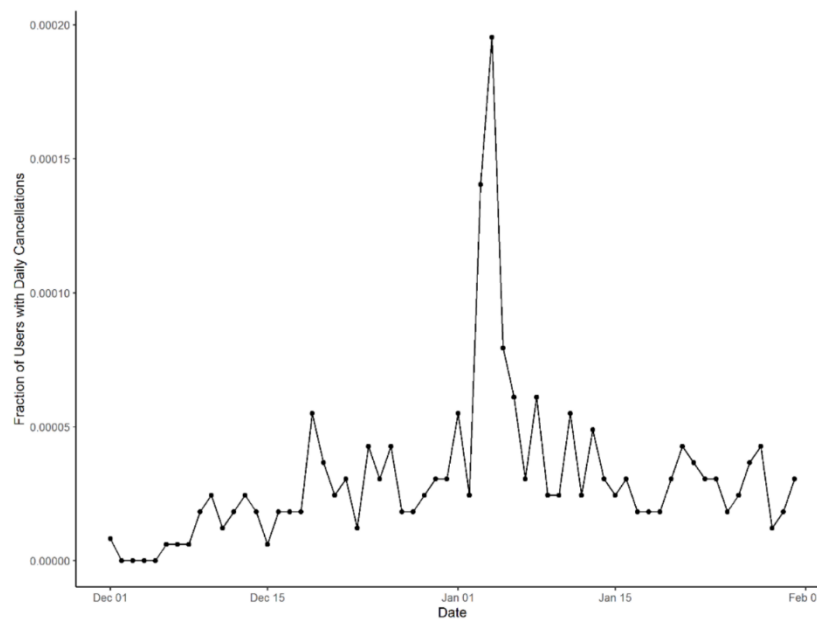


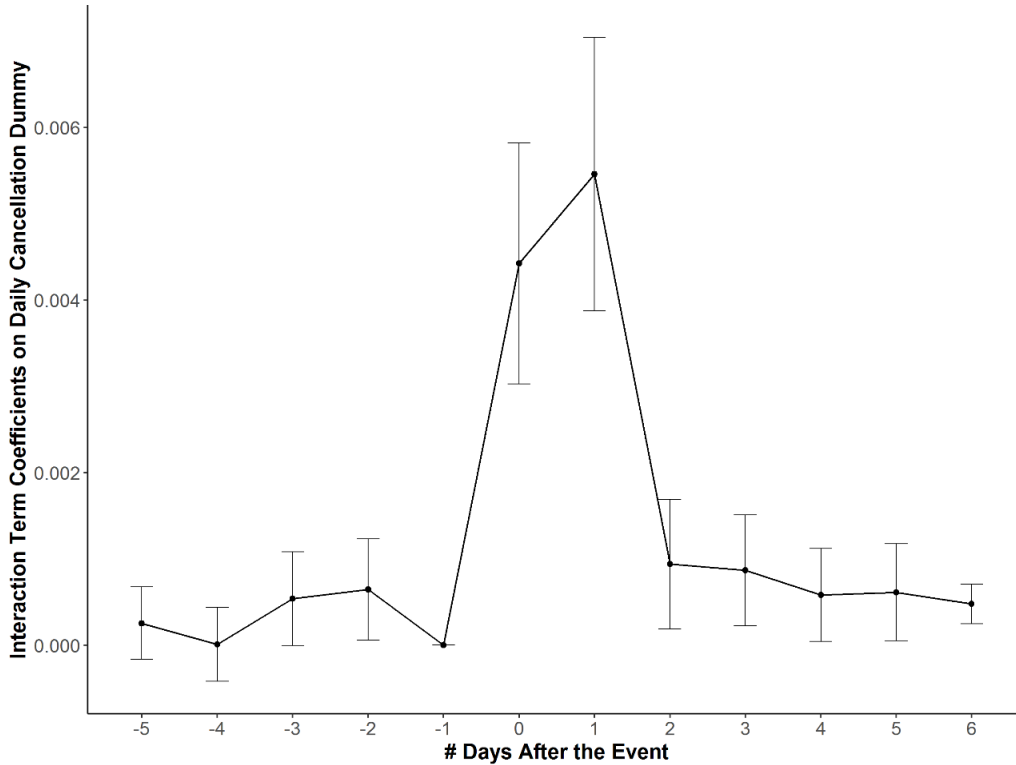
Figure A7: Activeness and Response to the 2017 Footprint Report Incident

The figures plot the $\beta_{H-L,\tau}$ coefficients estimated in the following regression with the bands indicating 95% confidence intervals:

$$\begin{aligned} \text{Daily Cancellation Dummy}_{i,t} &= \alpha_0 + \sum_{\substack{\tau=-5, \\ \tau \neq -1}}^5 \beta_{H-L,\tau} \cdot \text{Heavy User}_i \cdot \mathbb{I}(t = \tau) + \beta_{H-L,6} \cdot \text{Heavy User}_i \cdot \mathbb{I}(t \geq 6) \\ &+ \sum_{\substack{\tau=-5, \\ \tau \neq -1}}^5 \beta_{\tau} \cdot \mathbb{I}(t = \tau) + \beta_6 \cdot \mathbb{I}(t \geq 6) + \delta_i + \varepsilon_{i,t}, \end{aligned}$$

where *Daily Cancellation Dummy*_{*i,t*} is a dummy variable indicating whether user *i* has canceled at least one mini-program during day *t*; *t* corresponds to the number of days after January 3, 2018, (the day of the incident); *Heavy User*_{*i*} is a dummy indicating whether the user *i* has used more mini-programs than 75% of the population as of November 30, 2017; δ_i is the individual fixed effects; and $\varepsilon_{i,t}$ is the error term that varies across individuals and over time. Panel A covers a representative random sample of 100,000 Alipay users without any filtering, and Panel B covers only the Alipay users who canceled at least one mini-program before November 30, 2017, in the representative random sample. The data are at individual and daily levels. The sample period ranges from December 29, 2017, through January 31, 2018.

Panel A: Representative Random Sample: Full Sample



Panel B: Representative Random Sample: Users Having Canceled Before November 30, 2017

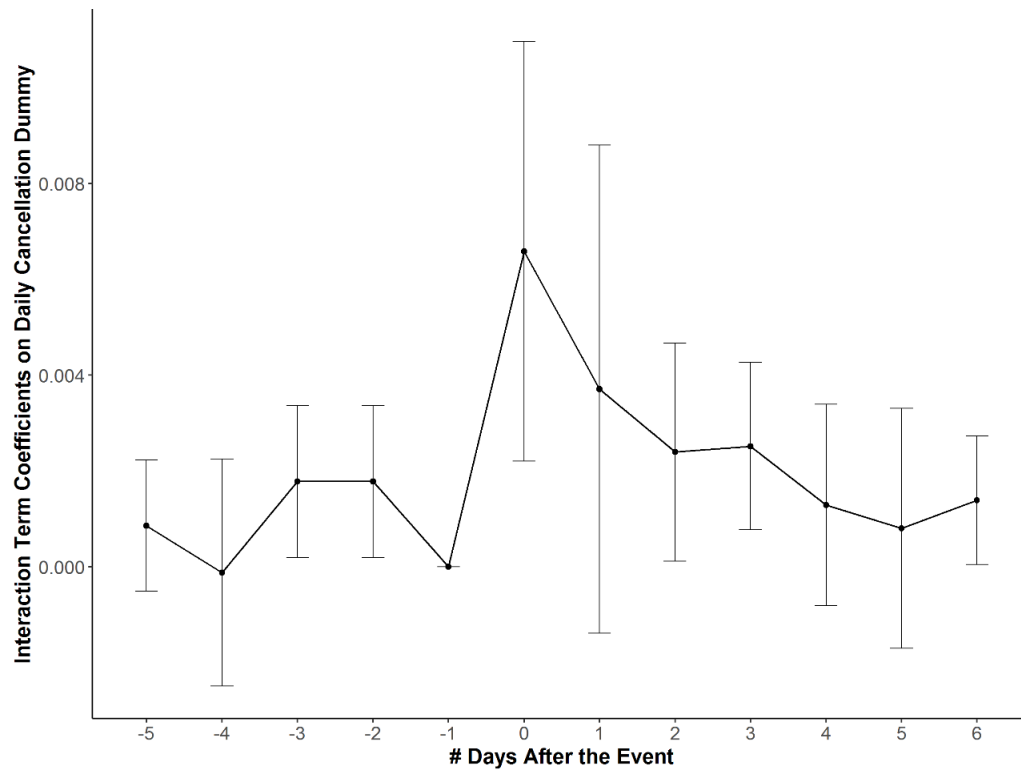


Figure A8: Time Trend in Demand for Digital Services

This figure shows the monthly time series of the average number of visited pages per active mini-program of Alipay users in three groups based on their responses to the survey question “*Are you concerned about negative impacts caused by information shared to mini-programs in Alipay?*” The sample covers the pre-survey period and focuses on the mini-programs that each user uses at least once during the corresponding months.

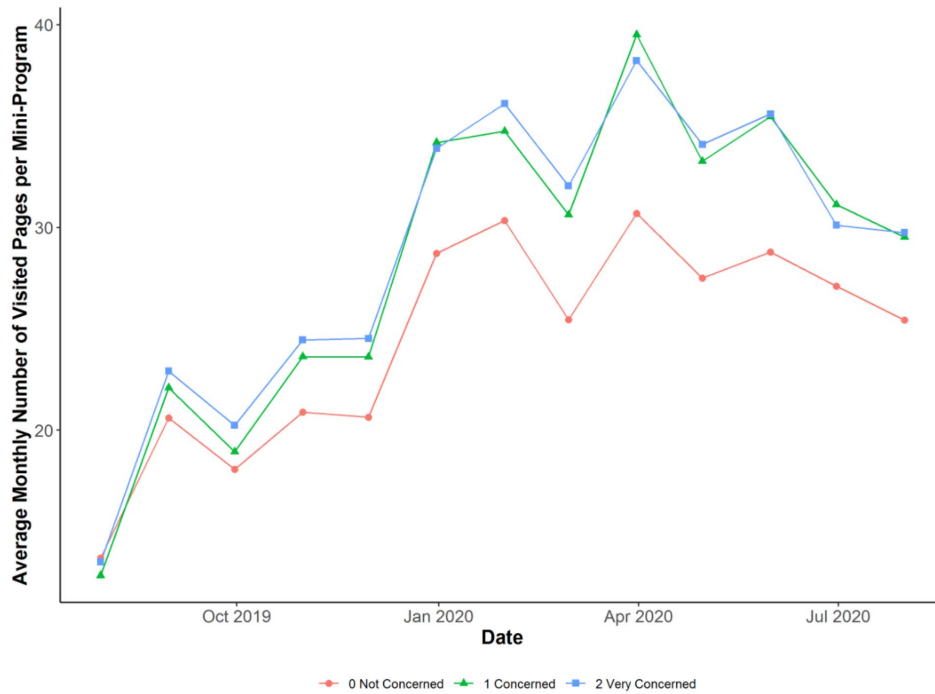


Table A1: Determinants of Data-Sharing Authorizations in Survey

Panel A summarizes the responses of the respondents to five statements. The respondents are split into two groups, one for those whose answers to the survey question “*Are you concerned about negative impacts caused by information shared to mini-programs in Alipay?*” were “concerned” or “very concerned,” and the other group for those whose answers to this survey question were “not concerned.” Panel B shows the regression results. The dependent variable takes a value of 1 if a respondent agrees with a statement. We denote ***, **, and * as the 1%, 5%, and 10% confidence levels, respectively. We report standard errors in parentheses.

Panel A: Summary of Responses to Survey Statements

	Count	Share	Count	Share	Total
	Agree		Disagree		
<i>Q1: I agree to authorize data sharing with mini-programs because it is safe in Alipay.</i>					
Concerned or very concerned	3,918	42%	5,331	58%	9,249
Not concerned	1,308	80%	318	20%	1,626
<i>Q2: I agree to authorize data sharing with mini-programs because my information has already been shared in many platforms.</i>					
Concerned or very concerned	1,083	12%	8,166	88%	9,249
Not concerned	493	30%	1,133	70%	1,626
<i>Q3: I have to share my information in exchange for digital services even though I have concerns about my data privacy.</i>					
Concerned or very concerned	6,030	65%	3,219	35%	9,249
Not concerned	913	56%	713	44%	1,626
<i>Q4: I only authorize data sharing with mini-programs when the requested data are not important.</i>					
Concerned or very concerned	1,852	20%	7,397	80%	9,249
Not concerned	485	30%	1,141	70%	1,626
<i>Q5: I tend to authorize data sharing with mini-programs that are used by my friends.</i>					
Concerned or very concerned	4,042	44%	5,207	56%	9,249
Not concerned	942	58%	684	42%	1,626

Panel B: Regression Analysis

Agree with	Q1	Q2	Q3	Q4	Q5
	(1)	(2)	(3)	(4)	(5)
Concerned or Very Concerned _i	-0.320*** (0.011)	-0.203*** (0.013)	0.083*** (0.014)	-0.096*** (0.014)	-0.158*** (0.014)
Digital Experience _i	-0.001*** (0.0002)	-0.001*** (0.0001)	0.0003** (0.0001)	-0.001*** (0.0001)	-0.00001 (0.0002)
Age _i	0.002*** (0.001)	0.001** (0.0004)	0.0005 (0.0005)	0.004*** (0.0005)	-0.001 (0.001)
City FE	Y	Y	Y	Y	Y
Gender FE	Y	Y	Y	Y	Y
Observations	8,658	9,637	9,780	9,356	9,110
Adjusted R ²	0.070	0.052	0.013	0.019	0.014

Table A2: Summary Statistics of the Random Sample

This table reports summary statistics of a representative random sample of 100,000 Alipay users. Panel A reports user information in three parts. The first part reports the general information. *Privacy Setting Changed*, a proxy measure for privacy concerns, is a dummy variable equal to 1 if a user changed their privacy setting at least once between May 2017, and April 2020, and 0 otherwise. *Digital Experience* is the number of months since the user first registered on Alipay, and *Age* is the user's physical age in July 2020. The second part covers data sharing with mini programs, including the number of authorized, entered, and canceled mini-programs over the pre-survey period of July 2019 to July 2020; the *Cancellation Rate* of used mini-programs between July 2019 and July 2020; and the *Has Canceled* status over the period of January 2013 to July 2020. The third part reports summary statistics of monthly use variables of Alipay users in each mini-program during the pre-survey period from July 2019 to July 2020, including number of active days, number of uses, number of launches, and number of visited pages. Use variables are winsorized at the 1% and 99% levels.

	N	Mean	Std	Min	p25	Median	p75	Max
General information								
Privacy Setting Changed _i	98,679	0.09	0.28	0.00	0.00	0.00	0.00	1.00
Digital Experience _i (month)	99,600	60.69	36.81	0.00	32.00	55.00	82.00	190.00
Age _i (year)	97,876	36.61	12.89	1.00	27.00	34.00	46.00	120.00
Data sharing with mini programs								
# Authorized Mini-Programs _i	100,000	2.40	3.52	0.00	0.00	1.00	3.00	136.00
# Entered Mini-Programs _i	100,000	3.02	4.59	0.00	0.00	2.00	4.00	248.00
Has Canceled _i	99,995	0.12	0.32	0.00	0.00	0.00	0.00	1.00
# Cancellations _i	98,674	0.30	1.45	0.00	0.00	0.00	0.00	61.00
Cancellation Rate _i	98,674	0.01	0.07	0.00	0.00	0.00	0.00	1.00
Monthly mini-program use								
# Active Days _{it}	3,036,555	0.27	1.59	0.00	0.00	0.00	0.00	27.00
# Uses _{it}	3,036,555	0.34	2.21	0.00	0.00	0.00	0.00	40.00
# Launches _{it}	3,036,555	1.10	6.90	0.00	0.00	0.00	0.00	123.00
# Visited Pages _{it}	3,036,555	3.06	19.96	0.00	0.00	0.00	0.00	342.00