BIS Working Papers
No 1187

Nothing to hide? Gender and age differences in willingness to share data

by Olivier Armantier, Sebastian Doerr, Jon Frost, Andreas Fuster and Kelly Shue

Monetary and Economic Department

May 2024

JEL classification: C8, D8.

Keywords: data, privacy, CCPA, fintech, big tech, survey of consumer expectations.

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Nothing to hide? Gender and age differences in willingness to share data*

Olivier Armantier Sebastian Doerr Jon Frost
Andreas Fuster Kelly Shue

May 2024

Abstract

Many digital applications rely on the willingness of users to voluntarily share personal data. Yet some users are more comfortable sharing data than others. To document these differences, we draw on questions to a representative sample of U.S. households added to the New York Fed's Survey of Consumer Expectations. We find that women are less willing than men, and older individuals less willing than the young, to share their financial transaction data in exchange for better offers on financial services. Across these groups, there are significant differences in attitudes, such as willingness to take financial risks, concerns that data will become publicly available, and concerns around personal safety. Responses suggest that privacy regulation can increase the willingness to share data, but effects do not differ by gender.

JEL classification: C8, D8.

Keywords: data, privacy, CCPA, fintech, big tech, survey of consumer expectations.

*Olivier Armantier: Federal Reserve Bank of New York (olivier.armantier@ny.frb.org); Sebastian Doerr: Bank for International Settlements, Monetary and Economic Department, and Centre for Economic Policy Research (CEPR) (sebastian.doerr@bis.org); Jon Frost: Bank for International Settlements, Monetary and Economic Department, and Cambridge Centre for Alternative Finance (CCAF) (jon.frost@bis.org); Andreas Fuster: EPFL, Swiss Finance Institute and CEPR (andreas.fuster@epfl.ch); Kelly Shue: Yale School of Management (kelly.shue@yale.edu). The authors would like to thank Stijn Claessens, Leonardo Gambacorta, Rod Garratt, Marina Sanchez del Villar and Hyun Song Shin for helpful comments. The views expressed here are those of the authors only and not necessarily those of the Bank for International Settlements, the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

The digital economy is made possible by the ubiquity of data — and particularly personal data. The use of such data can reduce search costs, verification costs, and other frictions (Goldfarb and Tucker, 2019) and thus allow for better and more personalized services. For example, smartphones transmit geolocation data, supporting everything from ride-hailing apps like Uber to various health apps that record footsteps or sleep patterns. Social media applications collect highly valuable data on individuals' contacts and social connections (Graham, 2015). In financial services, the ability to "port" data through screen scraping apps and open banking has allowed for greater competition and better offers on services such as credit (He et al., 2023; Berg et al., 2022; Nam, 2022).

Yet as the volume of personal data has grown, so too have concerns about how data are used. A growing literature shows how the protection of privacy interacts, in sometimes subtle ways, with consumer welfare (Acquisti et al., 2016; Jones and Tonetti, 2020; Cong et al., 2021). And individuals have a range of specific concerns. They worry about data being harvested for unwanted advertising, or for price discrimination, i.e. to set prices close to a user's willingness to pay (Bar-Gill, 2021; Croxson et al., 2022). Alternatively, they may worry about a data breach, when their personal information is leaked or becomes publicly available online.[1] In some cases, leaking of personal information could have a harmful impact on personal reputation, and individuals may worry about the impact of the sharing of certain data on their personal safety (Armantier et al., 2021).[2] And finally, even where some individuals think they have "nothing to hide," their own actions may impinge on the privacy of others, for instance when their data helps to derive information about their contacts or those similar to them (Acemoglu et al., 2022; Bergemann et al., 2022; Liu et al., 2023). This externality means that others may have an interest in how they treat their personal data.

The balance between the efficient use of personal data and appropriate

[1] An example is the 2017 Equifax breach, in which names, birth dates, addresses, social security numbers, and other information of over 160 million U.S., British, and Canadian consumers were accessed in a cyber-attack.

[2] For instance, individuals may worry about violence and harassment by former partners, estranged family members, or strangers; about theft and kidnapping by criminals; and about threats from political authorities.

protection of user privacy is thus an important issue for consumer welfare and public policy (Acquisti et al., 2016). Yet one aspect has received comparatively little attention: what if the willingness to share data differs across demographic groups? This question has important implications. If, for instance, one group is structurally more willing to share personal data than another, it may see higher adoption of specific digital technologies. In turn, the datasets being used to develop products, personalize services, and price credit may be biased due to an over-representation of this group relative to the group that is less willing to share data. At the same time, sharing data can lead to better products and services. For example, it can improve loan market outcomes through better screening. These benefits could be particularly large for individuals from traditionally under-served groups, including minority and low-income applicants, as current credit scores do not paint an accurate picture of their future creditworthiness (Blattner and Nelson, 2021; Di Maggio et al., 2022; Doerr et al., 2023).

This paper assesses preferences toward sharing personal data based on a survey of U.S. consumers. It draws on special questions that were added to the Federal Reserve Bank of New York (FRBNY) Survey of Consumer Expectations, fielded in January 2022. The SCE is a rotating panel of roughly 1,300 nationally representative U.S. household heads. New respondents are drawn each month to match demographic targets from the American Community Survey, and they stay on the panel for up to 12 months before rotating out. Beyond questions on privacy preferences, the survey includes a wealth of detailed demographic information, including the respondent's gender, race, age, income, education, financial literacy, and willingness to take risks (Armantier et al., 2017).

Our main finding is a significant gender gap in the willingness of individuals to share either financial transactions data or geolocation and social media data. When asked about sharing with a hypothetical credit card company, women consistently express less willingness to share such data than men, and report that they would demand a higher dollar figure for doing so. These differences are robust to a battery of individual controls including race/ethnicity, income, and education.

We also find strong differences in preferences or attitudes plausibly related to willingness to share data. In particular, women are much less willing to take financial risks (in line with Borghans et al. (2009); Croson and Gneezy

(2009)). They are more likely to worry about negative consequences if data are to become public, including higher costs and risks to personal safety. Female respondents in our sample also display lower financial literacy/numeracy. These differences, which are statistically significant, explain around 40% of the gender gap in the willingness to share data.

We also examine attitudes towards privacy along the age distribution. Older respondents are significantly less willing to take risks and worry significantly more about negative consequences or that their data become publicly available. They are also significantly less willing to pay a fee to continue using online banking or social media. However, controlling for the various factors does not materially reduce the gap in the willingness to share data between older and younger respondents.

Finally, we analyze whether preferences toward data sharing are influenced by privacy regulation. To test this, we "prime" one group of participants by asking questions about the California Consumer Privacy Act (CCPA), which introduces monetary compensation for consumers who suffer a data breach. Among respondents that are shown information about the CCPA, there is substantial disagreement about whether such a framework would give them more confidence to use online services that require sharing data. In particular, female respondents are more likely to agree that the CCPA would make them more comfortable. Additionally, those respondents that are positive toward the CCPA subsequently indicate a lower required compensation for sharing data when we ask them to assume that the CCPA framework would be in place in their state. However, the overall effect of the CCPA on willingness to share data of men vs women is not statistically significant.

Overall, our findings underscore the importance of gender and age in data privacy concerns. They also suggest that privacy legislation can be helpful in reducing consumer harms. But privacy laws may not be sufficient to close the gap between men and women or to fully internalize externalities related to differential willingness to share personal data.

Related literature. Our findings have a bearing on current debates around data privacy legislation and regulation of personal data in financial services. They also contribute to three strands of research.

First, they contribute to a growing body of studies looking at people's willingness to share data. Earlier work found a "privacy paradox" (Athey et al., 2017) — a gap between people's self-reported value of their privacy and their actual behaviors in protecting it. Yet more recent evidence suggests that while the paradox can arise in some circumstances, people's attitudes and behaviours are in other cases more aligned (Acquisti et al., 2020; Solove, 2021). Meanwhile, a series of recent studies finds that consumers value their privacy and hence demand a price for sharing their data (Wathieu and Friedman, 2007; Tang, 2020; Fernandez Vidal and Medine, 2020; Bijlsma et al., 2022; Bian et al., 2023). The price demanded by users in our study is higher than in other studies, potentially because we were asking about sharing a full year of geolocation, social media, or financial transaction data, which is much more extensive than the simple details (name, address, etc.) used in other studies. Our finding that women and older respondents demand a higher price to share data is in line with Cvrcek et al. (2006), but we provide additional evidence on the differences in attitudes and other factors that may explain this difference.

Second, our study contributes to literature on financial technology (fintech) and financial inclusion. Several studies emphasize the potential of fintech to include underserved groups, including women (Philippon, 2019; Demirguc-Kunt et al., 2022). Yet, with a survey of 27,000 individuals in 28 countries, Chen et al. (2023) find a statistically significant fintech gender gap in use of fintech products and services. Doerr et al. (2022) show that around the world, older generations are less likely to use digital payments and fintech than younger generations, and similar findings are obtained by Aldasoro et al. (2024) for the use of generative artificial intelligence tools. Our results complement these findings and show that differences in willingness to share data may be one part of the explanation for the gender gap. Our results thereby also inform the debate on central bank digital currencies (CBDCs) and the extent to which they need to ensure privacy (Garratt and Van Oordt, 2021; Ahnert et al., 2022; Auer et al., 2022; Agur et al., 2023). Our findings suggest that without adequate privacy protection, women and older citizens may be less likely to adopt CBDCs.[3]

Third, our study informs the debate on policy approaches to data protection. For instance, Godinho de Matos and Adjerid (2022) study the impact of the European Union General Data Protection Regulation (GDPR) on consumer and firm behavior. Canayaz et al. (2022) study the impact of the CCPA

[3]Based on a randomized online survey experiment, Choi et al. (2022) show that female users would be more willing to use a CBDC if it preserves privacy.

on the market for personal data, while Doerr et al. (2023) analyze the effect of the CCPA on users' willingness to share data with banks and fintechs. These studies help to inform the optimal design of data protection laws. Yet, if there are strong differences in preferences toward data sharing within society, this may form a challenge to the definition of common rules.[4]

Relative to the existing literature, we focus specifically on gender and age, as well as the willingness to share data. We use a representative survey of consumers in the United States, and assess how the impact of privacy legislation may affect attitudes. Our study uncovers relevant patterns that further research can build on.

The rest of this paper is organized as follows. Section 2 introduces our survey data and variable definitions. Section 3 presents our empirical strategy and results. Section 4 concludes.

2 The Survey of Consumer Expectations

We investigate the attitudes towards data privacy of Americans in the Survey of Consumer Expectations (SCE). The SCE is a high-quality monthly, internet-based survey designed and conducted by the Federal Reserve Bank of New York (FRBNY) and fielded by the private firm NielsenIQ. Launched in 2013, the SCE has been used extensively to help researchers and policymakers understand how expectations are formed and how they affect consumer behaviour.

The SCE uses a 12-month rotating panel of roughly 1,300 nationally representative U.S. household heads. New respondents are drawn each month to match demographic targets from the American Community Survey (ACS), and they stay on the panel for up to 12 months before rotating out. The survey's main aim is to collect expectations for a wide range of economic outcomes (e.g. inflation, income, spending, household finance, employment, and housing). The survey includes detailed demographic information, including the respondent's gender, race, age, income, education, financial literacy, and willingness to take risks (Armantier et al., 2017). The SCE aims to be representative of a U.S. household head with respect to education, income, age, and region, in line with ACS target values.

[4]See Collis et al. (2021), Lin (2022), and Prince and Wallsten (2022) for further evidence on heterogeneity in the valuation of personal data.

To understand how consumers value their data privacy and what determines their willingness to share data, the January 2022 survey contained an additional module.[5] The module asked detailed questions on respondents' attitudes towards data privacy, for example how much they trust different counterparties to safeguard their data, or whether users think that sharing data could have negative consequences for them.

To elicit consumers' willingness to share data, we ask them the following question: "Imagine you were to sign up for a new credit card. The credit card company has approved your application and is now offering you a sign-up bonus (in the form of money credited to your card account) if you provide the company with access to your full bank transaction history from the past year. Please select for each of the following amounts whether you'd be willing to share this data." Respondents are then shown the following amounts: \$20, \$50, \$100, \$250, \$500, \$1000, \$2500, and \$5000 with the options "No, do not share data" and "Yes, share the data" for each amount. The survey also asks the same question about respondents' "geolocation and social media data" instead of their "full bank transaction history." The survey interface was designed such that respondents were alerted in case their selections violated monotonicity – e.g. somebody who is willing to share their data for \$500 should also be willing to do so at any higher amount. Therefore, we observe for every respondent a single "switching point" (except if they say no to all provided amounts).

To understand what determines users' willingness to share data, the survey then asks them whether they have concerns about sharing their personal data. To this end, respondents were asked: "Are you concerned that sharing your personal data could have negative consequences for you?"; "Are you concerned about companies using this information to charge you more money for other goods or services?"; and "Are you concerned that your personal data might become publicly available?". To answer each question, the respondent had to use a Likert scale from 1 (totally disagree) to 7 (totally agree). We further ask "What are you specifically concerned about if your personal data were to become publicly available?" with the answer options "My personal safety," "Negative effects on my reputation," "Identity theft," and "Abuse of

[5]An earlier, similar module was fielded in September 2020 and analyzed in Armantier et al. (2021), with a focus on which types of firms consumers trust with their data, and how willingness to share data was affected by the Covid-19 pandemic period.

my data for unintended purposes (in the news or media, for political agenda, targeted ads, . . .).” In addition, we ask to what extent consumers agree with the following statement: “Even if I have no immediate concerns about my reputation or safety, I do not want to share my data because ‘my data are nobody’s business’.”

A randomly selected half of respondents was shown information and asked questions about the California Consumer Privacy Act (CCPA) before proceeding with the questions on data sharing. We defer a detailed discussion of this “CCPA treatment” to Section 3.3.

Finally, we ask respondents how they value products that use digital financial technology in the areas of online banking, digital payments or social media. Specifically, we ask: “Imagine you now had to pay an annual fee in order to keep using [*online banking*] / [*digital payment technologies*] / [*social media*]. How much would you be willing to pay for the coming year?” Users are then shown the following amounts: $10, $20, $50, $100, $250, $500, $1000, and $2500 with the options “No, would not pay” and “Yes, would pay” for each amount.

Summary statistics Our final sample has information on questions related to data sharing and privacy for 1,106 respondents. Table 1 shows summary statistics for the main variables from the survey. The average age of respondents is 50. About 85% of respondents are White, 9.6% are Black, 7% are Hispanic of any race, and 4% are Asian. Regarding other characteristics, 57% of respondents have a bachelor’s degree or higher, 35% have an income above $100,000, and 58% are working full-time, with a further 11% working part-time. 71% own their primary residence. The analyses below will use weights to make the sample representative of U.S. household heads in terms of education, income, age and region.

The bottom half of the table summarizes attitudes and proxies for preferences such as risk aversion or general trust that we will use as additional controls in what follows. While these variables will be discussed in more detail later, for now we note that there is substantial variation in almost all of them.

3 Empirical strategy and results

To investigate how attitudes towards sharing data differ across genders and with age, we estimate the following regression at the respondent (i) level:

$$y_i = \beta\, female_i + \gamma\, age_i + \rho\, controls_i + \varepsilon_i. \quad (1)$$

As the dependent variable y_i we will use different measures of consumers' willingness to share data or concerns about data sharing. Depending on the outcome variable, we will estimate either ordered logit or binary logit models.[6] All regressions are weighted by the provided sample weights to ensure that our sample is representative. Standard errors are robust. Our main coefficients of interest indicate to what extent male and female respondents (β) as well as older and young respondents (γ) differ in their attitudes.

We include a rich set of individual level controls, which we will refer to as 'demographic controls': dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Under the label 'socioeconomic controls,' we further control for the respondent's educational attainment (9 categories), a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. Finally, all regressions control for whether the respondent was randomly subject to the CCPA treatment (analyzed in Section 3.3).

3.1 Gender and willingness to share data

A plot of the share of respondents and the dollar amounts they request to share data (Figure 1) suggests that women have a lower willingness to share data than men. Panel (a) looks at bank transaction histories, panel (b) at geolocation and social media data. The y-axis reports the cumulative fraction of women and men that indicate they are willing to share their data with the credit card company when offered the amount of money on the x-axis. For any given amount, fewer women indicate they are willing to share their data. For both types of data, less than half the women indicate that they would be

[6] We have verified that our main results are also robust to the use of interval regressions.

willing to share the data when offered the highest amount, $5,000. Among men, this share is about 10 percentage points higher.

We note that in general, these amounts demanded for sharing data appear very high. Existing research has argued for a "digital privacy paradox," i.e. that people's stated aversion to sharing personal data does not match their behavior (Athey et al., 2017; Acquisti et al., 2020). In our case, this is not an issue to the extent that we are not interested per se in the level of compensation required (dollar amounts), but how they compare across respondents with different characteristics. The maintained assumption in what follows is that the extent to which aversion to sharing data is overstated in surveys vs. real-world decisions does not systematically vary across gender or with age.

Table 2 investigates the relationship between the requested amount to share data and personal characteristics in the regression setup of equation (1). Estimating an ordered logit specification, column (1) shows that women require a significantly higher amount than men to share their data. When we add the rich set of demographic controls in column (2), the estimated coefficient remains highly significant and increases in magnitude. Further adding socioeconomic controls in column (3) leads to no material change in the magnitude of the estimated coefficient, which remains strongly statistically significant.[7] These patters suggest that the relationship between the willingness to share data and gender is not explained by a rich set of (observable) respondent characteristics. Column (4) uses a dummy as the dependent variable that takes on a value of one if the amount of money required to share data is at least $2,500. Estimating a logistic regression with demographic and socioeconomic controls yields a positive coefficient of 0.72, significant at the 1% level. Based on implied average marginal effects, women are about 14.7 percentage points (pp) more likely to demand at least $2,500 than men.[8]

Columns (5) and (6) repeat the estimation exercises from columns (3) and (4), but focus on respondents' willingness to share social media/geolocation data (rather than their bank history). Also for social media/geolocation data, our empirical results show that women are significantly less willing to share their data, i.e. they demand higher amounts for doing so. The differences are

[7] Based on computed average marginal effects, women are significantly less likely to indicate that they are willing to share data for any amount up to $5,000, and are about 14 percentage points more likely to select no sharing at $5,000.

[8] Overall, 57.6% of respondents (or 60.8% weighted) answered that they require $2,500 or more.

slightly smaller: in column (6), the calculated average marginal effect implies an 8.2 pp higher likelihood for women to demand at least $2,500 to share their data.

Beyond gender, the coefficient on respondents' age is positive and strongly statistically significant in each regression. These results suggest that older respondents are generally less willing to share their bank history or social media/geolocation data than younger respondents.[9]

3.2 Examining explanations for the gap

What could account for the observed gender gap in the willingness to share data? Based on previous work, a number of explanations seem plausible. First, women are on average more risk-averse than men (Borghans et al., 2009; Croson and Gneezy, 2009), so differences in general or financial risk aversion between genders could help explain the gap. For example, women could put a greater weight on the potential financial costs or downside risks of sharing data. Second, research has found that women are less trusting in general (Alesina and La Ferrara, 2002), which might extend to financial services companies storing personal data. Third are potential differences in financial literacy and numeracy (see Lusardi and Mitchell (2011) for a survey). If there are significant differences in these variables across genders, this may influence perceived costs and benefits of data sharing. Fourth are potential differences in specific concerns around sharing data (e.g. reputational costs, risks that data become public or personal safety). And finally, men and women might value the benefits of using financial technology differently (Chen et al., 2021), with implications for their willingness to share data and associated benefits. We investigate these explanations in what follows.

To start, Table 3 uses different respondent characteristics as outcome variables to see whether they vary by gender. All regressions include demographic and socioeconomic controls. Column (1) shows that women's stated willingness to take financial risks is significantly lower, and column (2) shows similar results for the willingness to take risk in general. Column (3) indicates directionally lower trust by women, though this effect is not significant.

[9]Computed average marginal effects imply that an additional year of age increases the likelihood that a respondent demands at least $2,500 to share their data by 0.7 pp for bank history data and by 0.6 pp for social media/geolocation data.

In contrast, column (4) indicates that women in our sample have significantly lower numeracy, measured in the SCE based on a standard test with five questions.[10]

Columns (5)–(7) turn to the concerns about sharing data. The dependent variables are on a scale from 1 to 7, where 7 means strongly concerned. Women appear to be significantly more concerned that sharing their personal data could have negative consequences for them, as well as about companies using this information to charge them more money for other goods or services (columns (5) and (6)). When asked "Are you concerned that your personal data might become publicly available?", column (7) shows that women are also significantly more concerned along this dimension.

Finally, columns (8) to (10) investigate the extent to which the benefits of using digital products differ across genders. In principle, women could derive a lower utility from using e.g. online banking or payment apps.[11] To this end, we ask users how much they would be willing to pay in an annual fee to keep using online banking, digital payment technologies, or social media, as described earlier. Results show no systematic gender differences across the willingness to pay for using digital financial technology, but women express a higher willingness to pay for social media. These findings are in line with Brynjolfsson et al. (2023), who find that women are willing to pay a higher amount to keep using Facebook compared to men. [12]

Given the at times large differences in attitudes in Table 3, Table 4 analyzes whether controlling for these factors can narrow or eliminate the gender gap in the willingness to share data. Column (1) focuses on the willingness to share bank history data and adds controls for respondents' risk aversion and trust. Relative to the baseline estimate (column (3) in Table 2), the estimated gap remains almost identical. In column (2), we add controls for numeracy and whether the respondent indicates that they make the financial decisions in their household (a proxy for financial literacy). Adding these controls narrows

[10]See the last page of `https://www.newyorkfed.org/medialibrary/Interactives/sce/sce/downloads/glossary/FRBNY-SCE-ChartGlossary.pdf` for the wording of the five questions.

[11]Chen et al. (2021) show for China that there is a positive correlation between the benefits of using new financial technology and concerns about data privacy. Thus, perhaps surprisingly, those users who value fintech the most may also be more worried about potential costs.

[12]Note that questions on the willingness to pay were only asked to respondents who answered that they use these services. For those that answered that they do not use these services, we set their willingness to pay to zero.

the gap somewhat.

Accounting for the gender differences in concerns about sharing data in column (3) further narrows the gap substantially; relative to the baseline, the implied marginal effect of being female on the likelihood of not being willing to share data for any of the offered amounts is reduced by almost 40%.[13] Yet it remains economically and statically significant. Finally, controlling for the (insignificant) differences in the willingness to pay for online banking or digital payment technologies does not materially affect the gap (column 4).

In columns (5) and (6), we repeat the same exercise for the willingness to share social media and geolocation data. Similar to bank history, adding the additional controls reduces but does not eliminate the gender gap (although in the final column it is only mildly statistically significant).[14]

With respect to age, the patterns are qualitatively similar. Older respondents are significantly less willing to take risks and worry significantly more about negative consequences, or that their data become publicly available. However, they are also significantly less willing to pay a fee to continue using online banking or social media (see Table 3). However, controlling for the various factors does not materially reduce the magnitude of the age coefficient in columns (1) to (4) of Table 4 and reduces it only modestly in columns (5) and (6).

3.3 The role of privacy regulation – the CCPA

More and more jurisdictions are introducing privacy protection legislation. Could such legislation help close the gender gap in the willingness to share data? To examine this aspect, the survey asked questions about the California Consumer Privacy Act (CCPA).

The CCPA is a data privacy law covering the state of California that went into effect at the beginning of 2020. It endows Californians with several rights regarding the personal information that a firm may collect about them. In particular, Californians have the right to know what personal information is being collected, whether it is being sold and to whom, and the right to ac-

[13]The average marginal effect of being female in column (3) is 8.8 pp, vs. 14 pp in the baseline.

[14]For this outcome, we control for the stated willingness to pay for social media, rather than online banking and payment apps as before.

cess their personal information, to delete it, and to opt out of the sale of such information (Camhi and Lyon, 2018).

The rights included in the CCPA directly address some of the concerns that individuals list when it comes to sharing their data, like identity theft or abuse of data. A consumer concerned with these issues can request under CCPA that her data not be sold or that her data be deleted after she finishes transacting with a firm. Therefore, the CCPA likely increased certainty around the use of personal data: by assuring consumers that they can safeguard their privacy if they choose to do so, it could increase consumers' willingness to share their data (Doerr et al., 2023).

To understand how the CCPA has affected individuals' willingness to share data, we provided a random half of the survey respondents with the following information:

> *The California Consumer Privacy Act (CCPA) ensures privacy rights for consumers in California. The law is widely considered to provide the strongest consumer data protection in the U.S.. The law provides consumers with the right to know the personal information that a business collects about them, and the right to delete such personal information. The law also provides consumers with the right to opt out of the sale of personal information to third parties. In addition, if there is a data breach and personal information is stolen (e.g. a consumer's name or driver's license number), then the consumer can sue the business for damages up to $750.*

We then asked them: "*To what extent do you agree with the following statement? Please indicate your level of agreement on a scale from 1 (do not agree at all) to 7 (completely agree): If the CCPA was in place in my state, then it would give me greater confidence to use online services that require sharing of my personal data.*"

Importantly, only half of the respondents saw the CCPA prompt, and did so directly before they were asked about the required amounts to share their data with the credit card company. In particular, those who were shown the CCPA question were also shown an altered version of the willingness-to-share question: "*Imagine that the legal framework of the CCPA was in place in your state and imagine you were to sign up for a new credit card. The credit card company has approved your application and is now offering you a sign-up bonus (in the form of money credited to your card account) if you provide the company with access to your* *full bank transaction history from the past year."*

This randomization allows us to investigate the extent to which agreeing with the statement about the CCPA correlates with respondents' willingness to share data. Figure 2 shows that among the 554 respondents that were shown the CCPA prompt, about 25% responded 3 or lower, i.e. do not agree with the statement. In contrast, 55% selected a value of 5 or higher, with the remaining 20% selecting the intermediate value of 4. These patterns suggest that, on average, privacy regulation in the spirit of the CCPA gives individuals greater confidence to use online services that require the sharing of personal data. Importantly, the histogram also shows that female respondents tend to agree with the statement more strongly, and the difference in distributions is statistically significant ($p = 0.04$, Wilcoxon–Mann–Whitney test). This suggests that privacy-protecting rules might disproportionately affect women's confidence to use online service that feature data sharing.

In Table 5, we study whether showing the CCPA prompt to a respondent affects their willingness to share data. Column (1) shows that there is no average effect on respondents' willingness to share their bank history data. This regression specification corresponds to the one from column (2) of Table 4.[15]

However, column (2) shows that if respondents agreed with the statement that the CCPA would give them greater confidence, then they require significantly lower amounts to share their data. As noted above, female respondents are more likely to agree with this statement; however, column (3) indicates that the differential effect of the CCPA treatment on female respondents is not statistically significant.[16] The final three columns of the table show that the qualitative patterns are the same for the question on social media and geolocation data sharing.

Taken together, these results suggest that, as long as respondents believe that the CCPA protects their data, the policy has a positive effect on individuals' willingness to share data. However, there is no differential effect between

[15]Note that in Table 4 we were also controlling for the CCPA treatment, but without displaying the coefficient. In this section, we opt to use the specification without the controls for concerns about potential risks from sharing data because those questions were asked after the CCPA prompt.

[16]Interaction terms are not straightforward to interpret in nonlinear models in general (Ai and Norton, 2003), and this is particularly true for ordered logit models. However, our conclusions are unchanged if we transform the model into a binary logit as earlier and evaluate marginal effects in the different ways suggested by Dow et al. (2019)—the CCPA $\times$ Female interaction effect is never close to statistically significant.

men and women.

To provide insights into which subgroups of the population state that privacy regulation would give them more confidence to use online services requiring them to share their data, we regress CCPA agreement measures on various respondent-level characteristics in Table 6. Column (1) estimates an ordered logit regression with the level of agreement with the CCPA statement (on a scale from 1–7) as the outcome variable. It shows that women agree significantly more (in line with the histogram discussed earlier). Married respondents and – to a lesser extent in terms of significance – minority respondents agree less. Interestingly, neither respondent age nor any of the "behavioral" characteristics like risk aversion, trust, or numeracy are significantly associated with the outcome. Using a dummy for agreement (at least response 4 out of 7) in column (2) provides a qualitatively similar picture.

3.4 Additional tests

Socioeconomic characteristics. We now investigate whether the gender gap in willingness to share data varies with socioeconomic characteristics, namely income, education, or financial literacy. We estimate Equation (1), but interact the female dummy with dummies for respondents with incomes above $100,000, a bachelor degree or higher, or high numeracy (a score of 5 out of 5, achieved by 41% of respondents). Since interaction effects are difficult to interpret in nonlinear models like logits or ordered logits, Table 7 shows results from linear probability models, i.e. ordinary least squares (OLS) regressions, using a dummy for whether a respondent indicated they would not share their data for less than $2,500 as dependent variable.[17] Column (1) shows that the gender gap in the amount required to share data does not significantly change with respondents' income. Column (2) reports a similar picture for education, and column (3) for numeracy. In all three specifications, the coefficient on the interaction term of the female dummy with the measure of socioeconomic status is insignificant. When performing a principal component analysis (PCA) and extracting the first principal component of education, income, and numeracy, and interacting the PCA measure with the gender dummy in column (4), we again obtain an insignificant interaction term. Higher socioeconomic status tends to increase willingness to share, but except for numeracy, this effect

[17] We obtain qualitatively similar results in binary logit or ordered logit models.

is not statistically significant.

Concerns. The survey asked respondents the following question: "What are you specifically concerned about if your personal data were to become publicly available?" The answer options were "My personal safety," "negative effects on my reputation," "identity theft," and "abuse of my data for unintended purposes (in the news or media, for political agenda, targeted ads, ...)." Most respondents are concerned about ID theft (92%), followed by abuse for unintended purposes (64%), personal safety (50%) and reputation (25%). Table 8 investigates to what extent these concerns about sharing data differ across genders or by age. When asked about what they are specifically concerned about (yes or no questions, columns (1)–(4)), women worry significantly more about their personal safety (while older respondents worry less about this aspect). There are no statistically significant differences for reputational concerns, identity theft, or data abuse.[18] Finally, in column (5), we study determinants of respondents' agreement with the statement "my data are nobody's business" (which could be seen as a non-instrumental preference for privacy). There are no significant gender differences, but older respondents are more likely to agree more strongly with this statement.

These results suggest that personal safety concerns may be the most distinguishing factor driving differential privacy concerns of men and women in our sample. For older respondents, the aversion to sharing data appears to be more of a matter of principle.

4 Conclusion

Willingness to share personal data is a prerequisite to access a growing range of services across the digital economy. Yet we show that willingness to share such data differs by gender: in our survey of U.S. households, women consistently report being more concerned than men about sharing their data on financial transactions or social media activity and geolocation data. This may relate to gender differences in risk aversion or financial literacy, which have been documented in existing work and which we also find in our sample. Yet

[18]The number of observations varies across columns because some categorical controls perfectly determine the outcome. Also, only respondents who indicated that they were at least somewhat concerned if their personal data were to become public were asked these questions, but this applied to all but 19 respondents.

it could also relate to specific concerns that data will become publicly available (e.g. in a data breach) or – crucially – to concerns around personal safety. Our empirical exercise lends support to these explanations. We further show that older individuals are also less willing to share their data, perhaps as a matter of principle.

An implication of these differences in the willingness to share data is that, over time, the data sets being used for digital services may have fewer observations for women or older individuals. This could result in biased samples and outcomes that are not in the interest of the underrepresented groups, e.g. in lending decisions, financial advice, health applications, and many more. This requires further care on the part of developers to explicitly test models, including those built on big data, for demographic biases, and to seek out remedies.

Yet our study also holds grounds for hope. Data privacy protections such as the CCPA, which give individuals more control over their data and introduce recourse in the case of data breaches, may increase trust and willingness to share data. Further research will be needed to assess the effectiveness of such rules over time, and any differential impact by gender or other demographic characteristics.

References

Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar (2022). Too much data: Prices and inefficiencies in data markets. *American Economic Review: Microeconomics*.

Acquisti, A., L. Brandimarte, and G. Loewenstein (2020). Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology 30*(4), 736–758.

Acquisti, A., C. Taylor, and L. Wagman (2016). The economics of privacy. *Journal of Economic Literature 54*(2), 442–492.

Agur, M. I., M. A. Ari, and M. G. Dell'Ariccia (2023). Bank competition and household privacy in a digital payment monopoly. *IMF Working paper*.

Ahnert, T., P. Hoffmann, and C. Monnet (2022). The digital economy, privacy, and cbdc. *ECB Working Paper*.

Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123–129.

Aldasoro, I., O. Armantier, S. Doerr, L. Gambacorta, and T. Oliviero (2024). Survey evidence on gen AI and households: job prospects amid trust concerns. *BIS Bulletin* (86).

Alesina, A. and E. La Ferrara (2002). Who trusts others? *Journal of Public Economics 85*(2), 207–234.

Armantier, O., S. Doerr, J. Frost, A. Fuster, and K. Shue (2021). Whom do consumers trust with their data? US survey evidence. *BIS Bulletin* (42).

Armantier, O., G. Topa, W. Van der Klaauw, and B. Zafar (2017). An overview of the Survey of Consumer Expectations. *FRBNY Economic Policy Review* (23-2), 51–72.

Athey, S., C. Catalini, and C. Tucker (2017). The digital privacy paradox: Small money, small costs, small talk. *NBER Working Paper*.

Auer, R., J. Frost, L. Gambacorta, C. Monnet, T. Rice, and H. S. Shin (2022). Central bank digital currencies: motives, economic implications, and the research frontier. *Annual review of economics 14*, 697–721.

Bar-Gill, O. (2021). Price discrimination with consumer misperception. *Applied Economics Letters 28*(10), 829–834.

Berg, T., A. Fuster, and M. Puri (2022). FinTech Lending. *Annual Review of Financial Economics 14*, 187–207.

Bergemann, D., A. Bonatti, and T. Gan (2022). The economics of social data. *The RAND Journal of Economics 53*(2), 263–296.

Bian, B., X. Ma, and H. Tang (2023). The supply and demand for data privacy: Evidence from mobile apps. *Working Paper*.

Bijlsma, M., C. Cruijsen, and N. Jonker (2022, 05). Consumer willingness to share payments data: Trust for sale? *Journal of Financial Services Research*, 1–40.

Blattner, L. and S. Nelson (2021). How costly is noise? Data and disparities in consumer credit. Working paper.

Borghans, L., J. J. Heckman, B. H. H. Golsteyn, and H. Meijers (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association 7*(2-3), 649–658.

Brynjolfsson, E., A. Collis, A. Liaqat, D. Kutzman, H. Garro, D. Deisenroth, N. Wernerfelt, and J. J. Lee (2023). The digital welfare of nations: New measures of welfare gains and inequality. *NBER Working Paper*.

Camhi, R. and S. Lyon (2018). What is the california consumer privacy act? *Risk Management 65*(9), 12–14.

Canayaz, M., I. Kantorovitch, and R. Mihet (2022). Consumer privacy and the value of consumer data. Working paper, HEC Lausanne.

Chen, L., Y. Huang, S. Ouyang, and W. Xiong (2021). The data privacy paradox and digital demand. Working Paper 28854, National Bureau of Economic Research.

Chen, S., S. Doerr, J. Frost, L. Gambacorta, and H. S. Shin (2023). The fintech gender gap. *Journal of Financial Intermediation 54*(101026).

Choi, S., Y. S. Kim, B. Kim, and O. Kwon (2022). Central bank digital currency and privacy: A randomized survey experiment. *BIS Working Paper*.

Collis, A., A. Moehring, A. Sen, and A. Acquisti (2021). Information frictions and heterogeneity in valuations of personal data. *Working Paper*.

Cong, L. W., D. Xie, and L. Zhang (2021). Knowledge accumulation, privacy, and growth in a data economy. *Management Science 67*(10), 6480–6492.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–474.

Croxson, K., J. Frost, L. Gambacorta, and T. Valletti (2022, 08). Platform-based business models and financial inclusion: policy trade-offs and approaches. *Journal of Competition Law & Economics*.

Cvrcek, D., M. Kumpost, V. Matyas, and G. Danezis (2006). A study on the value of location privacy. In *Proceedings of the 5th ACM workshop on Privacy in electronic society*, pp. 109–118.

Demirguc-Kunt, A., L. Klapper, D. Singer, and S. Ansar (2022, June). The Global Findex Database 2021 : Financial inclusion, digital payments, and resilience in the age of COVID-19. Technical report, The World Bank.

Di Maggio, M., D. Ratnadiwakara, and D. Carmichael (2022). Invisible primes: Fintech lending with alternative data. Working Paper 29840, National Bureau of Economic Research.

Doerr, S., J. Frost, L. Gambacorta, and H. Qiu (2022). Population ageing and the digital divide. *SUERF Policy Briefs 270*.

Doerr, S., L. Gambacorta, L. Guiso, and M. Sanchez del Villar (2023). Privacy regulation and fintech lending. *BIS Working Paper 1103*.

Dow, W. H., E. C. Norton, and J. T. Donahoe (2019). Stata tip 134: Multiplicative and marginal interaction effects in nonlinear models. *The Stata Journal 19*(4), 1015–1020.

Fernandez Vidal, M. and D. Medine (2020, June). Study shows Kenyan borrowers value data privacy, even during pandemic. Blog, Consultative Group to Assist the Poor (CGAP).

Garratt, R. J. and M. R. Van Oordt (2021). Privacy as a public good: a case for electronic cash. *Journal of Political Economy 129*(7), 2157–2180.

Godinho de Matos, M. and I. Adjerid (2022). Consumer consent and firm targeting after GDPR: The case of a large telecom provider. *Management Science 68*(5), 3330–3378.

Goldfarb, A. and C. Tucker (2019). Digital economics. *Journal of Economic Literature 57*(1), 3–43.

Graham, B. S. (2015). Methods of identification in social networks. *Annual Review of Economics 7*(1), 465–485.

He, Z., J. Huang, and J. Zhou (2023). Open banking: Credit market competition when borrowers own the data. *Journal of Financial Economics 147*(2), 449–474.

Jones, C. I. and C. Tonetti (2020). Nonrivalry and the economics of data. *The American Economic Review 110*(9), 2819–2858.

Lin, T. (2022). Valuing intrinsic and instrumental preferences for privacy. *Marketing Science 41*(4), 663–681.

Liu, Z., M. Sockin, and W. Xiong (2023). Data privacy and algorithmic inequality. *NBER Working Paper*.

Lusardi, A. and O. S. Mitchell (2011). Financial literacy around the world: an overview. *Journal of Pension Economics & Finance 10*(4), 497–508.

Nam, R. J. (2022). Open Banking and customer data sharing: Implications for fintech borrowers. Working paper.

Philippon, T. (2019, September). On fintech and financial inclusion. Working Paper 26330, National Bureau of Economic Research.

Prince, J. T. and S. Wallsten (2022). How much is privacy worth around the world and across platforms? *Journal of Economics & Management Strategy 31*(4), 841–861.

Solove, D. J. (2021). The myth of the privacy paradox. *Geo. Wash. L. Rev. 89*, 1.

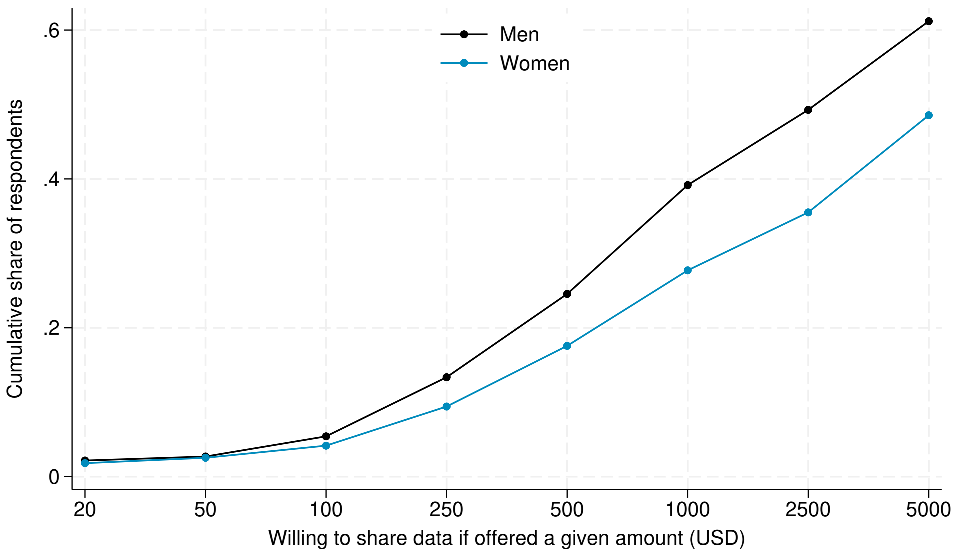
Tang, H. (2020). The value of privacy: Evidence from online borrowers. Working paper.

Wathieu, L. and A. A. Friedman (2007). An empirical approach to understanding privacy valuation. HBS Marketing Research Paper 07-075.

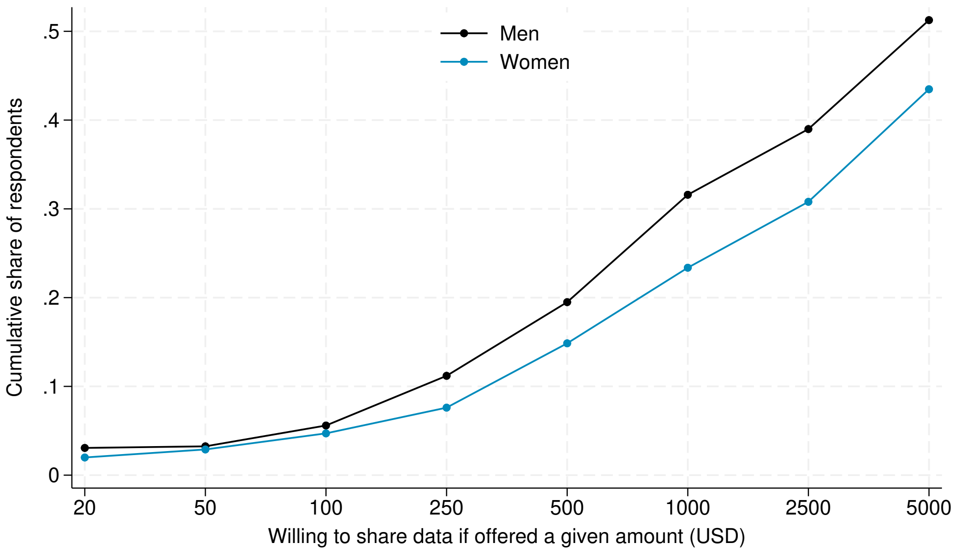
A Figures and tables

Figure 1: **Women are less willing than men to share their data**

(a) Bank transaction history

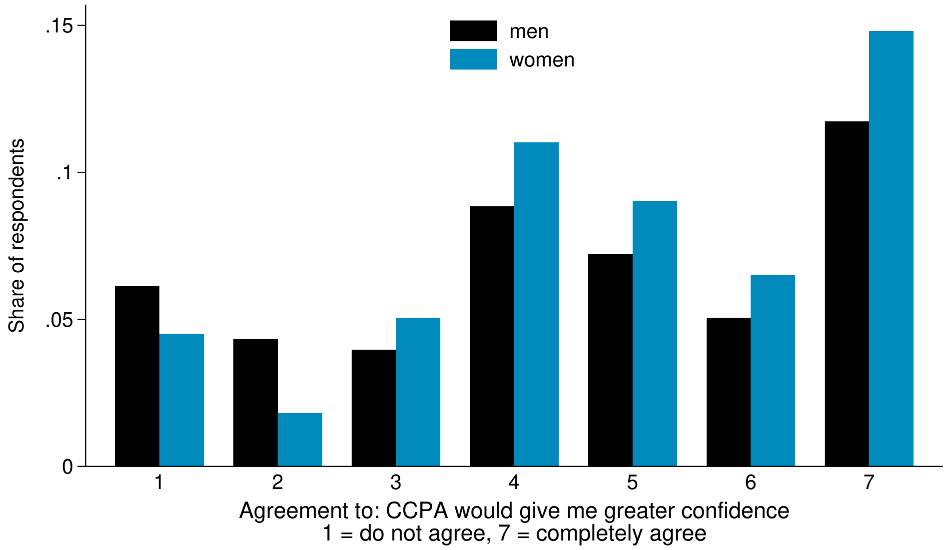


(b) Geolocation and social media data



Note: This figure shows the share of male and female respondents that indicated that they would be willing to share their bank transaction history (panel a) or geolocation and social media data (panel b) with a credit card company if offered the USD amount shown on the x-axis.

Figure 2: **Agreement with the CCPA statement**



Note: This figure shows the share of male and female respondents for each level of agreement with the statement "To what extent do you agree with the following statement? Please indicate your level of agreement on a scale from 1 (do not agree at all) to 7 (completely agree): If the CCPA was in place in my state, then it would give me greater confidence to use online services that require sharing of my personal data."

Table 1: **Summary statistics – covariates**

Variable	Obs	Mean	Std. Dev.	Min	Max	P25	P50	P75
Age (years)	1106	49.967	15.382	18	94	37	49.5	62
White (0/1)	1106	.847	.36	0	1	1	1	1
Hispanic (0/1)	1106	.07	.255	0	1	0	0	0
Black (0/1)	1106	.095	.293	0	1	0	0	0
Asian (0/1)	1106	.041	.198	0	1	0	0	0
Education: bachelor or more (0/1)	1106	.571	.495	0	1	0	1	1
Income above 100k (0/1)	1106	.346	.476	0	1	0	0	1
Working full-time (0/1)	1106	.58	.494	0	1	0	1	1
Working part-time (0/1)	1106	.112	.316	0	1	0	0	0
Owner of primary residence (0/1)	1106	.709	.454	0	1	0	1	1
Married (0/1)	1106	.609	.488	0	1	0	1	1
Lives alone (0/1)	1106	.262	.44	0	1	0	0	1
Numeracy score (0-5)	1106	3.938	1.16	0	5	3	4	5
Willingness to take financial risks (1-7)	1104	3.611	1.549	1	7	2	4	5
Willingness to take daily risks (1-7)	1105	3.746	1.448	1	7	3	4	5
Makes financial decisions in household (0/1)	1106	.576	.494	0	1	0	1	1
General trust in people (1-7)	1106	3.14	1.528	1	7	2	3	4
Has been subject to data breach (0/1)	1106	.612	.487	0	1	0	1	1
Concern: negative personal conseq. (1-7)	1106	5.38	1.678	1	7	4	6	7
Concern: higher costs (1-7)	1106	5.105	1.724	1	7	4	5	7
Concern: publicly available (1-7)	1105	5.695	1.497	1	7	5	6	7
WTP for online banking 20USD+ (0/1)	1106	.36	.48	0	1	0	0	1
WTP for payment apps 20USD+ (0/1)	1106	.173	.378	0	1	0	0	0
WTP for social media 20USD+ (0/1)	1106	.15	.357	0	1	0	0	0

Note: This table shows summary statistics (observations, mean, standard deviation, minimum, maxmimum, as well 25th, 50th and 75th percentile) of the main variables. Sample weights are not applied.

Table 2: **Compensation required to share bank history and social media data**

VARIABLES	(1) ord log BH amount	(2) ord log BH amount	(3) ord log BH amount	(4) logit BH > 2.5k	(5) ord log SM amount	(6) logit SM > 2.5k
Female (0/1)	0.617***	0.679***	0.652***	0.716***	0.390***	0.400**
	(0.141)	(0.142)	(0.144)	(0.161)	(0.142)	(0.162)
Age (years)	0.042***	0.040***	0.037***	0.035***	0.028***	0.028***
	(0.005)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Observations	1,106	1,106	1,106	1,106	1,106	1,106
Demographic Controls	-	✓	✓	✓	✓	✓
Socioeconomic Controls	-	-	✓	✓	✓	✓
Pseudo R2	0.0451	0.0491	0.0582	0.105	0.0385	0.0596

Note: This table reports results for Equation (1). Columns (1)–(3) and (5) report results from ordered logit regressions, columns (4) and (6) from logistic regressions. Columns (1)–(3) use the dollar amount respondents require to share their bank history (BH) as the dependent variable. Column (4) uses a dummy as the dependent variable that takes on a value of one if the amount of money required to share bank history data is at least $2,500. Columns (5) uses the dollar amount respondents require to share their social media data as dependent variable. Column (6) uses as the dependent variable a dummy that takes on a value of one if the amount of money required to share social media/geolocation data is at least $2,500. *Female* is a dummy with a value of one if the respondent is female. *Age* is respondent's age in years. Demographic controls include dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Socioeconomic controls include the respondent's educational attainment, a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. All regressions control for whether the respondent was randomly subject to the CCPA treatment. All regressions are weighted and use robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: **Individual characteristics and the correlation with gender**

VARIABLES	(1) ord log fin risk	(2) ord log gen risk	(3) ord log trust	(4) ord log numeracy	(5) ord log neg cons	(6) ord log costs	(7) ord log publ avail	(8) ord log onl bank amt	(9) ord log pay app amt	(10) ord log soc med amt
Female (0/1)	-0.407***	-0.281**	-0.129	-0.663***	0.287**	0.231*	0.351**	0.006	0.100	0.492***
	(0.136)	(0.137)	(0.134)	(0.146)	(0.138)	(0.132)	(0.140)	(0.138)	(0.158)	(0.176)
Age (years)	-0.014**	-0.022***	-0.005	0.005	0.010*	0.007	0.017***	0.002	-0.021***	-0.021***
	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.007)
Observations	1,104	1,105	1,106	1,106	1,106	1,106	1,106	1,106	1,106	1,106
Demographic Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Socioeconomic Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Pseudo R2	0.0325	0.0152	0.0317	0.106	0.0250	0.0212	0.0315	0.0256	0.0360	0.0351

Note: This table reports results for Equation (1), estimated with ordered logit regressions. Columns (1) uses respondents' willingness to take financial risks as the dependent variable. Columns (2) uses respondents' willingness to take risks in general as the dependent variable. Columns (3) uses respondents' level of general trust as the dependent variable. Columns (4) uses respondents' numeracy level as the dependent variable. Columns (5) uses respondents' level of concern about negative consequences from sharing data as the dependent variable. Columns (6) uses respondents' level of concern about higher monetary costs from sharing data as the dependent variable. Columns (7) uses respondents' level of concern about data becoming publicly available as the dependent variable. Columns (8) uses the dollar amount respondents require to not use online banking as the dependent variable. Columns (9) uses the dollar amount respondents require to not use digital payments technologies as the dependent variable. Columns (10) uses the dollar amount respondents require to not use social media as the dependent variable. *Female* is a dummy with a value of one if the respondent is female. *Age* is respondent's age in years. Demographic controls include dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Socioeconomic controls include the respondent's educational attainment, a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. All regressions control for whether the respondent was randomly subject to the CCPA treatment. All regressions are weighted and use robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: **Required compensation to share data, controlling for further factors**

VARIABLES	(1) ord log BH amount	(2) ord log BH amount	(3) ord log BH amount	(4) ord log BH amount	(5) ord log SM amount	(6) ord log SM amount
Female (0/1)	0.640***	0.560***	0.452***	0.465***	0.318**	0.283*
	(0.144)	(0.149)	(0.157)	(0.159)	(0.146)	(0.160)
Age (years)	0.035***	0.036***	0.035***	0.035***	0.027***	0.024***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
Observations	1,104	1,104	1,104	1,104	1,104	1,104
Demographic Controls	✓	✓	✓	✓	✓	✓
Socioeconomic Controls	✓	✓	✓	✓	✓	✓
Risk av. & trust	✓	✓	✓	✓	✓	✓
Fin. literacy	-	✓	✓	✓	✓	✓
Concerns	-	-	✓	✓	-	✓
Use benefit	-	-	-	✓	-	✓
Pseudo R2	0.0637	0.0716	0.0964	0.105	0.0478	0.0846

Note: This table reports results for Equation (1), estimated with ordered logit regressions. Columns (1)–(4) use the dollar amount respondents require to share their bank history (BH) as dependent variable. Columns (5)–(6) use the dollar amount respondents require to share their social media/geolocation data as dependent variable. *Female* is a dummy with a value of one if the respondent is female. *Age* is respondent's age in years. Demographic controls include dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Socioeconomic controls include the respondent's educational attainment, a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. All regressions control for whether the respondent was randomly subject to the CCPA treatment. All regressions are weighted and use robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: **Required compensation to share data given privacy legislation**

	(1)	(2)	(3)	(4)	(5)	(6)
	ord log	ord log	ord log	ord log	ord log	ord log
VARIABLES	BH amount	BH amount	BH amount	SM amount	SM amount	SM amount
Female (0/1)	0.560***	0.601***	0.598***	0.318**	0.334**	0.310
	(0.149)	(0.148)	(0.207)	(0.146)	(0.145)	(0.195)
CCPA treatment	0.022	0.407**	0.057	0.218	0.444**	0.211
	(0.148)	(0.195)	(0.194)	(0.147)	(0.193)	(0.200)
CCPA treatment and agrees		-0.689***			-0.410*	
		(0.201)			(0.210)	
Female × CCPA			-0.074			0.015
			(0.279)			(0.290)
Observations	1,104	1,104	1,104	1,104	1,104	1,104
Demographic Controls	✓	✓	✓	✓	✓	✓
Socioeconomic Controls	✓	✓	✓	✓	✓	✓
Risk av. & trust	✓	✓	✓	✓	✓	✓
Fin. literacy	✓	✓	✓	✓	✓	✓
Pseudo R2	0.0717	0.0765	0.0717	0.0479	0.0497	0.0479

Note: This table reports results for Equation (1), estimated with ordered logit regressions. Columns (1)–(3) use the dollar amount respondents require to share their bank history as dependent variable. Columns (4)–(6) use the dollar amount respondents require to share their social media/geolocation data as dependent variable. *Female* is a dummy with a value of one if the respondent is female. *CCPA treatment* is a dummy with a value of one if the respondent was shown the CCPA statement. *CCPA treatment and agrees* is a dummy with a value of one if the respondent was shown the CCPA statement and agrees with it (4 or higher). Demographic controls include dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Socioeconomic controls include the respondent's educational attainment, a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. All regressions control for whether the respondent was randomly subject to the CCPA treatment. All regressions are weighted and use robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: **CCPA correlations**

VARIABLES	(1) ord log CCPA agreement (1-7)	(2) logit CCPA agrees (>4)
Female (0/1)	0.430***	0.520**
	(0.163)	(0.221)
Age (years)	-0.003	-0.000
	(0.006)	(0.009)
Minority racial or ethnic group (0/1)	-0.374*	-0.327
	(0.207)	(0.273)
Owner of primary residence (0/1)	-0.113	0.036
	(0.189)	(0.258)
Working full-time (0/1)	-0.082	-0.087
	(0.203)	(0.264)
Working part-time (0/1)	-0.196	-0.339
	(0.270)	(0.368)
Married (0/1)	-0.560**	-0.767**
	(0.248)	(0.371)
Income above 100k (0/1)	0.216	0.243
	(0.202)	(0.258)
Education: bachelor or more (0/1)	0.052	0.114
	(0.180)	(0.210)
Lives alone (0/1)	-0.092	-0.248
	(0.263)	(0.373)
Willingness to take financial risks (1-7)	-0.025	-0.041
	(0.060)	(0.079)
Willingness to take daily risks (1-7)	0.076	0.061
	(0.060)	(0.082)
General trust in people (1-7)	0.016	-0.006
	(0.052)	(0.071)
Numeracy score (0-5)	0.088	0.137
	(0.072)	(0.101)
Makes financial decisions in household (0/1)	0.053	-0.106
	(0.191)	(0.266)
Has been subject to data breach (0/1)	0.088	0.263
	(0.157)	(0.213)
Lives in California (0/1)	0.014	-0.089
	(0.293)	(0.405)
Observations	552	552
Pseudo R2	0.0127	0.0323

Note: This table reports conditional correlations between the level of agreement with the CCPA statement and respondent characteristics (columns 1), or whether respondents strongly agree with the CCPA statement and respondent characteristics (column 2). Columns (1) estimates an ordered logit regression, while column (2) estimates a logistic regressions. All regressions are weighted and use robust standard errors. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: **Gender differences and socioeconomic characteristics**

VARIABLES	(1) income BH > 2.5k	(2) education BH > 2.5k	(3) numeracy BH > 2.5k	(4) PCA BH > 2.5k
Female (0/1)	0.145***	0.138***	0.063	0.153***
	(0.040)	(0.046)	(0.060)	(0.033)
Socio indicator	0.027	-0.042	-0.144**	-0.027
	(0.052)	(0.046)	(0.060)	(0.031)
Female × socio indicator	-0.027	-0.001	0.115	0.029
	(0.068)	(0.060)	(0.072)	(0.023)
Observations	1,104	1,104	1,104	1,104
R-squared	0.149	0.149	0.152	0.151
Demographic Controls	✓	✓	✓	✓
Socioeconomic Controls	-	-	-	-

Note: This table reports variations of Equation (1) estimated with OLS regressions. Columns (1)–(4) use a dummy as the dependent variable that takes on a value of one if the amount of money required to share bank history (BH) data is at least $2,500. *Female* is a dummy with a value of one if the respondent is female. In column (1) *socio indicator* is a dummy that takes on a value of one for respondents with incomes above $100,000. In column (1) it is a dummy that takes on a value of one for respondents with a bachelor's degree or higher. In column (3) it is a dummy that takes on a value of one for respondents with high numeracy (a score of 5 out of 5). In column (4) it is first principal component of education, income, and numeracy. Demographic controls include dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Socioeconomic controls include the respondent's educational attainment, a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. All regressions control for whether the respondent was randomly subject to the CCPA treatment. All regressions are weighted and use robust standard errors. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: **Concerns**

VARIABLES	(1) logit pers safe	(2) logit reput	(3) logit ID theft	(4) logit abuse	(5) ord logit nobody's bus
Female (0/1)	0.313**	-0.153	-0.211	-0.024	0.106
	(0.153)	(0.167)	(0.293)	(0.164)	(0.149)
Age (years)	-0.022***	-0.008	0.019*	-0.009	0.018***
	(0.006)	(0.008)	(0.011)	(0.006)	(0.006)
Observations	1,086	1,077	1,071	1,086	1,106
Demographic Controls	✓	✓	✓	✓	✓
Socioeconomic Controls	✓	✓	✓	✓	✓
Pseudo R2	0.0461	0.0240	0.0901	0.0449	0.0598

Note: This table reports results for Equation (1). Columns (1)–(4) reports results from logistic regressions, while column (5) reports results from an ordered logit regression. Columns (1) uses respondents' concern about their personal safety when data become publicly available as the dependent variable. Columns (2) uses respondents' concern about negative effects on their reputation when data become publicly available as the dependent variable. Columns (3) uses respondents' concern about identity theft when data become publicly available as the dependent variable. Columns (4) uses respondents' concern about abuse my data for unintended purpose when data become publicly available as the dependent variable. Columns (5) uses respondents' agreement with the statement "my data are nobody's business" as the dependent variable. *Female* is a dummy with a value of one if the respondent is female. *Age* is respondents age in years. Demographic controls include dummies for whether the respondent owns their primary residence, whether they are married, whether they belong to a racial or ethnic minority, whether they are working (with separate dummies for full-time and part-time work), and whether they are living alone. Socioeconomic controls include the respondent's educational attainment, a quadratic function of the household's income category, and whether they have been subject to a data breach in the past. All regressions control for whether the respondent was randomly subject to the CCPA treatment. All regressions are weighted and use robust standard errors. *** p<0.01, ** p<0.05, * p<0.1.

B Online appendix

Table 9: **Summary statistics – covariates (weighted)**

Variable	Obs	Mean	Std. Dev.	Min	Max	P25	P50	P75
Age (years)	1106	51.807	15.984	18	94	38	53	64
White (0/1)	1106	.842	.365	0	1	1	1	1
Hispanic (0/1)	1106	.086	.281	0	1	0	0	0
Black (0/1)	1106	.101	.302	0	1	0	0	0
Asian (0/1)	1106	.028	.166	0	1	0	0	0
Education: bachelor or more (0/1)	1106	.347	.476	0	1	0	0	1
Income above 100k (0/1)	1106	.284	.451	0	1	0	0	1
Working full-time (0/1)	1106	.506	.5	0	1	0	1	1
Working part-time (0/1)	1106	.114	.318	0	1	0	0	0
Owner of primary residence (0/1)	1106	.659	.474	0	1	0	1	1
Married (0/1)	1106	.576	.494	0	1	0	1	1
Lives alone (0/1)	1106	.28	.449	0	1	0	0	1
Has been subject to data breach (0/1)	1106	.549	.498	0	1	0	1	1
Willingness to take financial risks (1-7)	1104	3.481	1.613	1	7	2	3	5
Willingness to take daily risks (1-7)	1105	3.72	1.541	1	7	3	4	5
General trust in people (1-7)	1106	3.005	1.569	1	7	2	3	4
Numeracy score (0-5)	1106	3.706	1.246	0	5	3	4	5
Makes financial decisions in household (0/1)	1106	.575	.495	0	1	0	1	1
Concern: negative personal conseq. (1-7)	1106	5.417	1.712	1	7	5	6	7
Concern: higher costs (1-7)	1106	5.199	1.726	1	7	4	6	7
Concern: publicly available (1-7)	1105	5.734	1.533	1	7	5	6	7
WTP for online banking 20USD+ (0/1)	1106	.31	.463	0	1	0	0	1
WTP for payment apps 20USD+ (0/1)	1106	.153	.36	0	1	0	0	0
WTP for social media 20USD+ (0/1)	1106	.129	.335	0	1	0	0	0

Note: This table shows summary statistics (observations, mean, standard deviation, minimum, maxmimum, as well 25th, 50th and 75th percentile) of the main variables. Observations are weighted to correspond to target values from the American Community Survey. WTP = willingness to pay.

Previous volumes in this series

1186 May 2024	Unconditional convergence in the Mexican manufacturing sector (1988-2018)	Alex Rivadeneira
1185 May 2024	Allocative efficiency and the productivity slowdown	Lin Shao and Rongsheng Tang
1184 May 2024	Determinants of currency choice in cross-border bank loans	Lorenz Emter, Peter McQuade, Swapan-Kumar Pradhan and Martin Schmitz
1183 May 2024	Why DeFi lending? Evidence from Aave V2	Giulio Cornelli, Leonardo Gambacorta, Rodney Garratt and Alessio Reghezza
1182 April 2024	Reserve requirements as a financial stability instrument	Carlos Cantú, Rocio Gondo and Berenice Martinez
1181 April 2024	Synthetic controls with machine learning: application on the effect of labour deregulation on worker productivity in Brazil	Douglas K G Araujo
1180 April 2024	Incentive-Compatible Unemployment Reinsurance for the Euro Area	Alexander Karaivanov, Benoit Mojon, Luiz Pereira da Silva, Albert Pierres Tejada and Robert M Townsend
1179 April 2024	The impact of artificial intelligence on output and inflation	Iñaki Aldasoro, Sebastian Doerr, Leonardo Gambacorta and Daniel Rees
1178 April 2024	Finternet: the financial system for the future	Agustín Carstens and Nandan Nilekani
1177 March 2024	Pre-publication revisions of bank financial statements: a novel way to monitor banks?	Andre Guettler, Mahvish Naeem, Lars Norden and Bernardus F Nazar Van Doornik
1176 March 2024	The effect of Covid pension withdrawals and the Universal Guaranteed Pension on the income of future retirees in Chile	Carlos Madeira
1175 March 2024	Unmitigated disasters? Risk- sharing and macroeconomic recovery in a large international panel	Goetz von Peter, Sebastian von Dahlen, and Sweta Saxena
1174 March 2024	The impact of information and communication technologies on banks, credit and savings: an examination of Brazil	Flavia Alves

All volumes are available on our website www.bis.org.