

MIT Open Access Articles

Shifts in Privacy Concerns

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Goldfarb, Avi, and Catherine Tucker. "Shifts in Privacy Concerns." *American Economic Review* 102.3 (2012): 349–353.

As Published: <http://dx.doi.org/10.1257/aer.102.3.349>

Publisher: American Economic Association

Persistent URL: <http://hdl.handle.net/1721.1/75861>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Shifts in Privacy Concerns[†]

By AVI GOLDFARB AND CATHERINE TUCKER*

The information and communications technology revolution has led to a seismic shift in the storage and availability of personal consumer data for commercial purposes. This has led privacy advocates to call increasingly for regulation to protect consumers' privacy. In order to shape how regulation should look, however, it is important to understand how consumer privacy concerns have evolved over time alongside this increase in the intensity of data usage.

To address this need, we measure how consumers' privacy concerns have changed using three million observations collected by a market research company from 2001–2008, covering whether consumers chose to protect their privacy by not revealing their income in an online survey. There are two key patterns in the data: (1) Refusals to reveal information have risen over time, and (2) Older people are much less likely to reveal information than are younger people. Our data further suggest that though younger respondents have become somewhat more private over time, the gap between younger and older people is widening.

There are several possible explanations for these trends. These include an increase in experience with information technology (and awareness of potential privacy concerns), a change in the composition of the online population, a change in the composition of survey respondents, a change in income, and a change in underlying preferences for privacy. We find evidence that a change in underlying

preferences for privacy explains a substantial part of the trend, though this does not rule out the possibility that the other explanations can explain the remainder. Specifically, we document that there has been an increase in the number of contexts in which consumers perceive privacy concerns to be relevant. Consumers were always privacy-protective in contexts where they were answering personal questions about financial and health products, but they are increasingly becoming privacy-protective in contexts that have not been highlighted as standard areas of privacy concern (such as answering questions about consumer packaged goods or movies). This finding can also explain the observed differences in privacy-protective behavior across age groups: the difference in information revelation between personal and nonpersonal contexts is largest for older people. Therefore, as more topics are seen as potentially personal, the difference between the total amount revealed by older and younger people grows.

Consistent with recent discussions in law and philosophy, these results suggest a change in consumers' perceptions over what contexts warrant privacy. Zittrain (2009), Solove (2008), and Nissenbaum (2010) all discuss how changes in technology change our perception of the public and private sphere. Taylor (2004) develops an economic model of information sharing across firms that provides one explanation of why the context in which data is used might matter: if data is used to price-discriminate, and changes in technology mean that firms can sell the data to other firms even outside their sector, then many rational consumers will choose to keep information private in order to prevent future price discrimination. Our results contribute to the broader policy discussion, where the economics literature has tended to focus on the costs (Goldfarb and Tucker 2011c, forthcoming), leaving the discussion of benefits to survey research conducted by legal scholars such as (Hoofnagle et al. 2010).

*Goldfarb: Associate Professor of Marketing, Rotman School of Management, University of Toronto, 105 St. George St., Toronto, Ontario (e-mail: agoldfarb@rotman.utoronto.ca); Tucker: Associate Professor of Marketing, MIT Sloan School of Management, 100 Main St., Cambridge, MA (e-mail: cetucker@mit.edu). We thank Markus Mobius and seminar participants at Queen's University and the 2012 AEA meetings for helpful comments. This research was supported by an NBER Economics of Digitization and Copyright Initiative Research Grant and by SSHRC grant 486971.

[†] To view additional materials, visit the article page at <http://dx.doi.org/10.1257/aer.102.3.349>.

I. Data

In economics, privacy has commonly been modeled as the transfer of information between consumer and firm and explored why consumers may find that undesirable (Posner 1981; Acquisti and Varian 2005; Hui and Png 2006). A key challenge in identifying how privacy concerns, when measured even in this narrow sense of information revelation, have changed, is that the ways in which people can reveal information have grown over time. Therefore, it is difficult to identify an underlying propensity to reveal information while controlling for the increased opportunities to reveal, changing potential benefits of revealing information, and changing risks and costs to revealing. In order to address this challenge, we look at an individual instance of information revelation and examine it over time. Specifically, we look at the propensity to reveal income in a web-based survey by a marketing research company that received over three million individual survey responses from 2001 to 2008. The aim of this database is to provide comparative guidance to advertisers about which types of ad design are effective and to benchmark current campaigns against past campaigns. This database is among the main sources used by the online media industry to benchmark online display ad effectiveness. We have used the data to measure advertising performance directly in other studies, including Goldfarb and Tucker (2011a, b, c).

Despite its primary marketing focus, the database is particularly useful for understanding changes in underlying privacy preferences over time because the benefits and costs for revealing this information relate to inherent preferences for privacy rather than to any explicit gain or risk associated with the information. Furthermore, the precise question asked has remained constant over time.

The survey began with various questions about the likelihood of purchase for multiple different products and brands. In this paper, we focus on what these different product categories meant for the context of the survey. Was the survey asking about purchase preferences for health insurance? Or was the survey asking about purchase preferences for detergent?

After answering questions on the product category, the respondent is asked to provide demographic information, specifically age, gender,

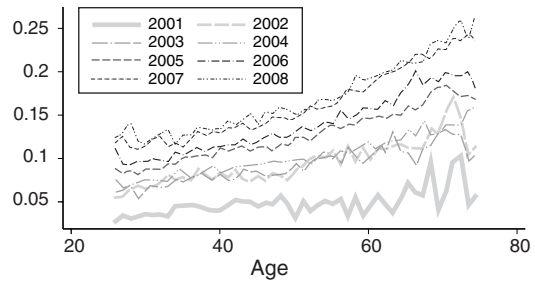


FIGURE 1. FRACTION REFUSING TO REVEAL INCOME BY AGE AND YEAR

Internet use, zip code, and income. When asked to give their income level, respondents were given the option of “Prefer not to say.” Refusals to provide income are several times higher than other refusals: 15 percent on average for income and less than 0.5 percent for all other questions. We use this decision to keep income information private as our core dependent variable.

II. Data Analysis

Figure 1 displays an initial graphic representation of the relationship between age, privacy-seeking, and year. The y-axis is the fraction of respondents who choose to keep their income private, and the x-axis is age. Each line represents the respondents in a particular year. The figure shows three patterns that will be confirmed in the econometric analysis. First, all lines are upward sloping: older people are more privacy-protective than younger people. Second, the lines steadily rise over time: privacy is increasing over time across all age groups. Third, though perhaps somewhat less clearly in the picture, the difference across age groups is increasing over time.

In order to better document and understand these patterns, we turn to regression analysis. Our econometric estimation approach examines the relationship between individual-, survey-, and location-level factors and whether the respondent reveals their income. The dependent variable is refusal to reveal income, and therefore measures whether respondents choose to keep their income private. We use a linear regression model that can be seen as a simple cost-benefit model of keeping information private. The regression controls for month of the year, county fixed effects, website category fixed

TABLE 1—PRIVACY CONCERNS, DEMOGRAPHICS, AND TIME TRENDS

	(1)	(2)
Years since 2001	0.013*** (0.00022)	0.0017*** (0.00051)
Age	0.0023*** (0.000029)	0.00068*** (0.000061)
Age × years		0.00026*** (0.000011)
Female	0.046*** (0.00073)	0.045*** (0.00073)
Average payroll (00,000)	-0.011 (0.014)	0.0088 (0.013)
Month fixed effects	Yes	Yes
Site category fixed effects	Yes	Yes
Product type fixed effects	Yes	Yes
County fixed effects	Yes	Yes
Observations	3,176,706	3,176,706
Log-likelihood	-985,945.6	-985,538.6
R ²	0.02	0.02

***Significant at the 1 percent level.
 **Significant at the 5 percent level.
 *Significant at the 10 percent level.

effects, survey type fixed effects, gender, and annual payroll at the county level.

Table 1 documents the basic correlations between the respondent characteristics and refusals to provide income information. Consistent with Figure 1, column 1 shows that the propensity to keep income private rises over time and that privacy increases with age. The interaction term in column 2 shows that the older people are becoming relatively more private over time. The estimated effects are economically large and statistically relevant. Column 1 suggests that refusals increase by 1.3 percentage points per year relative to an average of 13 percent. The age effects are also large: refusals rise 4.6 percentage points when age increases from 25 to 45.

Next, we explore a possible driver of the increase in refusals over time, and why this increase is sharper for older people. In the companion paper (Goldfarb and Tucker 2012), we rule out explanations based on internet technology use. Specifically, we document that the people who spend the most time online are the most likely to reveal their income. This does not change over time. Therefore, the result is not driven by growing familiarity with the online world.

Instead, we argue that changing perceptions of what is private help to explain the trend.

TABLE 2—BROADENING PRIVACY CONCERNS MAY EXPLAIN THIS

	(1)	(2)
Nonpersonal topic	-0.024*** (0.0023)	-0.012* (0.0062)
Nonpersonal topic × years	0.0032*** (0.00033)	0.0013 (0.0011)
Nonpersonal topic × age		-0.00031** (0.00014)
Nonpersonal topic × age × years		0.000051** (0.000024)
Age × years		0.00022*** (0.000021)
Years since 2001	0.010*** (0.00036)	0.00045 (0.00094)
Age	0.0023*** (0.000029)	0.00091*** (0.00012)
Female	0.045*** (0.00073)	0.045*** (0.00073)
Average payroll (00,000)	-0.0067 (0.013)	0.014 (0.012)
Month fixed effects	Yes	Yes
Site category fixed effects	Yes	Yes
Product type fixed effects	Yes	Yes
County fixed effects	Yes	Yes
Observations	3,176,706	3,176,706
Log-likelihood	-985,877.5	-985,456.8
R ²	0.02	0.02

***Significant at the 1 percent level.
 **Significant at the 5 percent level.
 *Significant at the 10 percent level.

Applying Acquisti, John, and Loewenstein (2011), who showed that the product questions that appear early in a survey can generate privacy-protective responses, we use the topic of questions early in the survey to identify situations that do and do not obviously warrant privacy. As people become increasingly aware of the various uses of online data, it is possible that they become increasingly private in situations that do not obviously warrant privacy. Under this hypothesis, the increase in privacy would be sharpest in nonpersonal contexts such as surveys about entertainment, games, and consumer packaged goods. In contrast, people have always known that financial- and health-related contexts require privacy and we would expect less change over time.

The first row of column 1 of Table 2 confirms that people are indeed more likely to keep income information private when the survey is about a personal topic. The second row shows

that this is decreasing over time: as we move from 2001 to 2008, the difference between personal and nonpersonal topics narrows. The coefficient values suggest that the difference in whether people reveal income in personal and nonpersonal topics disappears completely after eight years. Put another way, 24 percent ($0.0032/(0.0032 + 0.0100)$) of the increase in privacy can be explained by decreasing willingness to reveal income information in connection with “nonpersonal” topics. Column 2 shows the full set of interactions with age and date. The three-way interaction between age, date, and nonpersonal topic suggests that the increase in privacy for nonpersonal topics is sharpest for older people. Interestingly, the two-way interaction between age and nonpersonal topic is negative, suggesting that, early in the sample, the difference between older and younger people is larger for personal topics than nonpersonal topics.

This generates a likely partial explanation for the observed rise in privacy over time: people are increasingly treating all online information sharing the same. This is consistent with Nissenbaum (2010): that technological advances in the ability to track, aggregate, and disseminate information have led to a change in the contexts that can be seen as private.

III. Conclusions

This paper examines how privacy concerns have changed over time and across age groups and argues that the increases in consumer privacy concerns may be explained by a widening in scope of the contexts in which privacy is relevant. As with any study, our results have several limitations. While we document an increase in privacy over time and with age, a key question is the extent to which the decision not to answer a potentially intrusive question on a survey is a good proxy for privacy-protective behavior or for a consumer having privacy concerns in general. While Cho and Larose (1999), Day (1975), and others argue that privacy violations can arise from sampling, contact, and response methods in online surveys, at best we study one particular measure of privacy. Furthermore, our explanation for a substantial fraction of the trend over time relies on the particular context of the online survey topic. Any attempt to extrapolate to other contexts is necessarily speculative. Finally, we

leave three-quarters of the trend unexplained, and cannot reject the possibility that some of this is due to changes in who is online and who is responding to the survey.

Despite these limitations, we provide novel behavior-based evidence of the existence of, and the reasons for, increasing privacy concerns among consumers. Our results suggest that attention to a particular type of privacy is increasing. As more and more contexts appear to be private, people are revealing less information in the same situations. Furthermore, because older people are particularly private in personal contexts, the difference between older and younger people has widened.

REFERENCES

- Acquisti, Alessandro, Leslie John, and George Loewenstein.** Forthcoming. “The Impact of Relative Standards on the Propensity to Disclose.” *Journal of Marketing Research*.
- Acquisti, Alessandro, and Hal R. Varian.** 2005. “Conditioning Prices on Purchase History.” *Marketing Science* 24 (3): 367–81.
- Cho, H., and R. Larose.** 1999. “Privacy Issues in Internet Surveys.” *Social Science Computer Review* 17 (4): 421–34.
- Day, George S.** 1975. “The Threats to Marketing Research.” *Journal of Marketing Research* 12 (4): 462–67.
- Goldfarb, Avi, and Catherine Tucker.** 2011a. “Advertising Bans and the Substitutability of Online and Offline Advertising.” *Journal of Marketing Research* 48 (2): 207–27.
- Goldfarb, Avi, and Catherine Tucker.** 2011b. “Online Display Advertising: Targeting and Obtrusiveness.” *Marketing Science* 30 (3): 389–404.
- Goldfarb, Avi, and Catherine Tucker.** 2011c. “Privacy Regulation and Online Advertising.” *Management Science* 57 (1): 57–71.
- Goldfarb, Avi, and Catherine Tucker.** 2012. “Technology, Age, and Shifting Privacy Concerns.” Unpublished.
- Goldfarb, Avi, and Catherine Tucker.** Forthcoming. “Privacy and Innovation.” In *Innovation Policy and the Economy*. Vol. 12, edited by Josh Lerner and Scott Stern. Chicago: University of Chicago Press.
- Hoofnagle, Chris Jay, Jennifer King, Su Li, and Joseph Turow.** 2010. “How Different are Young Adults from Older Adults When It Comes to

- Information Privacy Attitudes and Policies?" Unpublished.
- Hui, Kai-Lung, and I. P. L. Png.** 2006. "The Economics of Privacy." In *Economics and Information Systems*, edited by Terrence Hendershott, 471–97. Amsterdam: Elsevier.
- Nissenbaum, Helen.** 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Palo Alto, CA: Stanford University Press.
- Posner, Richard A.** 1981. "The Economics of Privacy." *American Economic Review* 71 (2): 405–09.
- Solove, Daniel J.** 2008. *Understanding Privacy*. Cambridge, MA: Harvard University Press.
- Taylor, Curtis R.** 2004. "Consumer Privacy and the Market for Customer Information." *RAND Journal of Economics* 35 (4): 631–50.
- Zittrain, Jonathan.** 2009. *The Future of the Internet—And How to Stop It*. New Haven, CT: Yale University Press.