

Valuating Privacy

Bernardo A. Huberman, Eytan Adar and Leslie R. Fine
HP Labs, 1501 Page Mill Road, Palo Alto CA, 94304

Abstract

In spite of the widespread concerns expressed about the importance of privacy, individuals frequently give away or sell a myriad of personal data. How and why people decide to transition their information from the private to the public sphere is poorly understood. To address this puzzle, we conducted a reverse second-price auction to identify the monetary value of private information to individuals and how that value is set. Our results demonstrate that the more undesirable the trait with respect to the group, whether perceived or actual, impacts the price demanded to reveal private information.

Privacy is a central issue of concern in the information age. Because of the ease with which data about individuals can be obtained, aggregated and dispersed, information technology can broadcast an individual's secrets to unintended recipients who in turn can use it in ways that the individual no longer controls. While it is clear why this would be a concern with financial data and genetic information that could lead to identity abuse and discrimination, it is also true for other relatively harmless information such as a person's gender, salary, age, marital status or shopping preferences.

Several survey-based techniques have already revealed correlations between individual self-disclosures and demographic data[1-4]. For example, the Jourard Self-Disclosure Questionnaire reveals to whom a person discloses information but fails to capture the specific value of that data. More recent work[5,12] has further clarified the privacy trade-offs that individuals are willing to make in order to access specific services, while pointing out the disparities between stated privacy attitudes and actions.

At the root of the decision to transition private data into the public sphere lies an issue left unaddressed in any quantitative manner; that is, how much do people truly value their secrets, and to what extent is that valuation contextual? By contextual we mean valuation that depends on the characteristics of the group learning of the private data. Our conjecture and motivation is that people are willing to reveal information whenever they feel that they are somewhat typical or positively atypical compared to the social group.

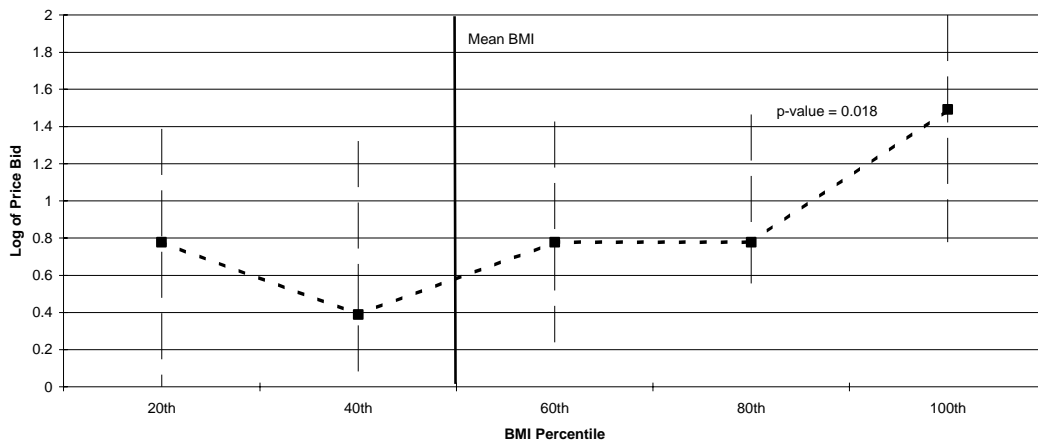
In order to test this hypothesis, we conducted experiments that revealed the true value that people place on their private data. Specifically, we tested whether desirability or undesirability of a trait is the dominant factor in dictating how a person values a piece of information. We find with great significance (in excess of 95% statistical confidence) that a linear relationship exists between the individual's trait and the price, i.e. the lesser the desirability of that trait the greater the price demanded for that information. Furthermore, we find that small deviations in a socially positive direction are associated with a lower demanded price.

By treating private information as a real good[13,14], our economic experiment was designed to determine the value of that information by offering to purchase it from subjects and reveal it to the group, in effect eliciting the individual's 'privacy calculus'[7,15,16]. Subjects were told that they would participate in a reverse second-price auction for personal data, i.e., the individual demanding the least for the information was paid the second lowest demanded price. In exchange for this compensation, and after verification, the individual had to reveal that piece of information to the other auction participants. The financially competitive nature of the auction, coupled with the fact that all participants had to anonymously submit their private data along with their demanded price, allowed us to extract the value that each individual placed on disclosing the private information.

We considered weight and age as an example of a type of privacy that most people value, which one can verify instantly and does not have financial or identity-theft repercussions. A post experimental questionnaire, which presented hypothetical bidding scenarios on financial data, also asked all participants questions about their attitudes towards privacy, self-perception of weight, beliefs about the other players in the room, who they knew in the session and how well.

Of the two auctions, the one for weight (127 participants) displays the strongest effects. Figure 1 depicts the relationship between weight (normalized as the Body Mass Index, or BMI), binned by percentile, and (log) price requested to make that information public. A Kruskal-Wallis ANOVA1 test (used throughout)

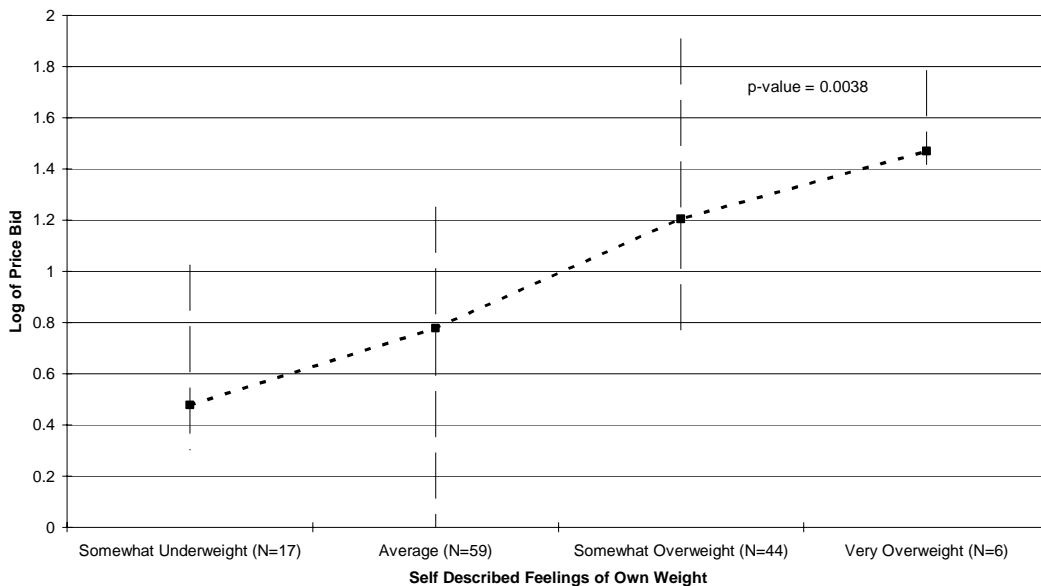
Figure 1
Kruskal-Wallis Anova1 Analysis of
Log of Price Bid, in Bins based on BMI Percentiles



reveals statistical significance ($p = 0.019$) with a distinct visual trend in the average price as a function of BMI. Those individuals weighing slightly below average, an “ideal” weight by cultural standards did not require a lot of money for publicizing this. On the other hand, those who weighed more and who may fear embarrassment or stigmatization[17-19] demanded more. Interestingly, while a characteristic such as weight can be visually inferred by anyone, it is still considered private. Such behavior is potentially linked to both our internal (potentially false) beliefs on how the group perceives us as well as self-perception[20,21].

To test the impact of self-perception factors, each subject completed a survey indicating if they believed themselves to be “very under, somewhat under, average, somewhat over, or very over” in relation to the average weight of the other subjects. Binned by these categories, the results are even more striking than actual weight. Once again, those who perceived themselves to be very underweight indicated that they would reveal weight information for a small amount of money. As perception of weight relative to average increases so does the price demanded. As can be seen in Figure 2, the slightly higher price demanded by the lowest weight group in Figure 1 disappears when binning by perception. This suggests that while certain subjects had a low weight in

Figure 2
Kruskal-Wallis Anova1 Analysis of
Log of Price Bid, in Bins based on Feelings on Weight



reality, they did not perceive themselves as such and priced their information accordingly.

While weaker than the trends noted in the weight auctions, the results of the age auction showed similar tendencies (age data was gathered for seven of the ten experiments, representing 88 participants). An analysis of the log price bid in bins based on subject age (range of 23-62, average 40) showed a slight increase with age ($p = .17$). However, it is notable that for the two extreme bins we do find a significant log price difference (.665 or \$3.62 versus 1.28 or \$18.05, $p = .0297$) implying that the very young subjects are more willing to reveal their age than the older ones and that the large (middle) population segments have similar privacy demands. In contrast to the weight auctions, the smaller demand differences by different age group segments may also indicate that age information is less sensitive than weight. This interpretation is supported by the difference in average demand price for the two auctions (\$57.56 for age versus \$74.06 for weight).

In contrast to age information, which appears to be less privacy sensitive than weight, we also studied price demands for salary, spousal salary, credit rating, and savings. As part of the survey participants were asked to imagine they were participating in auctions for this data and to indicate how much they would demand. In these simulated auctions, the percentage of individuals demanding more than \$100 was 48%, 36%, 24%, 38% for salary (77 participants), spousal salary (52), credit rating (78), and savings (77) respectively. All are relatively high in comparison to weight and age auctions where only 5.5% and 3.5% respectively demanded more than \$100. This additional cost may be in part related to social taboos that prevent open discussion of information such as salary in order to prevent conflict, and indicates that setting correct auction limits is critical. Further, because many subjects knew each other in a professional context they may be forced to evaluate the potential future impacts of revealing financial information. It is also likely that just as the BMI normalization was a better metric than raw weight, factors such as occupation, years on the job, etc. may help to explain noisy trends in the pricing data.

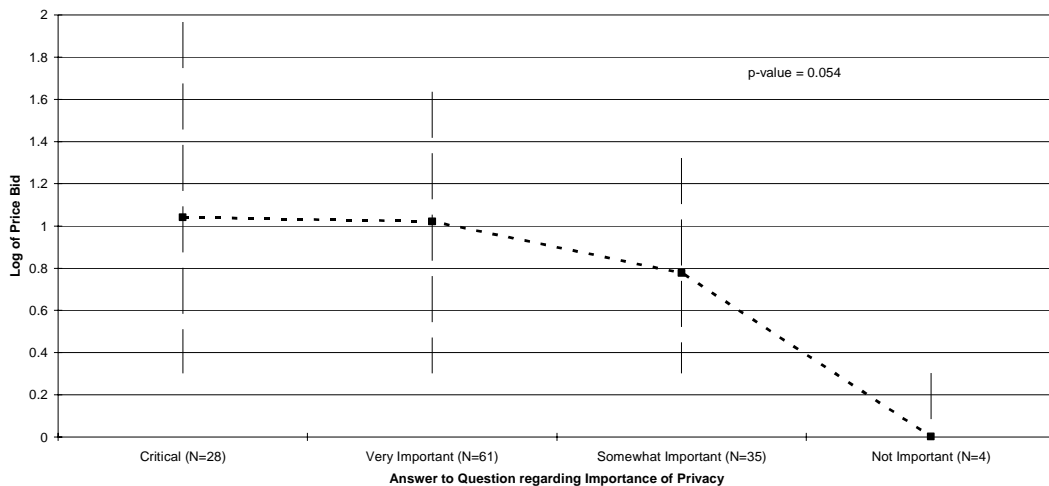
It is also worth considering how general privacy attitudes impact the price demanded to reveal private information. The post-auction survey asked the

general question, “How important to you is your personal privacy information...” with options for critical, very important, somewhat important, and unimportant. Figure 3 depicts the weight prices binned by these categories. While not insignificant ($p = .056$), general privacy attitudes are clearly not as strong as other factors.

Our survey also attempted to extract the number of auction participants known to each subject. In the weight auction, those individuals who were in the top 50th percentile in terms of demanded price on the average knew 36% of others present, whereas the bottom 50th percentile knew 23% ($p = .05$), suggesting that individuals are less reluctant to reveal information to an anonymous audience (“phenomenon of the stranger” [21]). Unfortunately, most subjects were familiar with approximately the same number of people and so no effect is seen when binning based on the percent of people known versus price. Thus the effect, given our population, can at best be considered weak.

We found very slight behavioral variations between genders. For example, in the weight auction, men on average demanded a log price of .847 (\$6.03) whereas women demanded 1.13 (\$12.49, $p = .15$). Of the seven individuals demanding “infinity,” six were women. Examining the trends in price demanded as a function of perceived weight, both curves display a marked upward trend, however the male trend appears better defined ($p = .0037$ for the male trend, $p = .2$ for female). For women, this result may be due in part to the distribution

Figure 3
Kruskal-Wallis Anova1 Analysis of
Log of Price Bid, in Bins based Privacy Attitudes



of responses (e.g. only one subject considered herself underweight). Female subjects believing themselves to be “average” displayed a broad variation in price, and comparing only the “somewhat over” to the “somewhat under” is significant at $p = .099$.

In conclusion, we conducted a reverse second-price auction to identify the monetary value of private information to individuals and how that value is set. Our results demonstrate that the more undesirable the trait with respect to the group, whether perceived or actual, impacts the price demanded to reveal private information.

These results also help explain the apparent paradox that individuals frequently give away or sell a myriad of personal data in spite of their stated concerns about privacy. Recent debates on privacy issues ranging from financial information[23] to genetic and medical data[24,25] to surveillance[26] require a careful consideration of how individuals choose to reveal their private information. Our results, which highlight the strongly contextual nature of this decision, also suggest possible ways that could be used in order to increase the level of comfort that people experience when revealing private data.

Methods

Data Collection

In total 127 individuals (59% male), recruited through local colleges and company mailing lists, participated in 10 separate sessions. Five of the sessions were mixed gender, three were female only, and two were male only. In all sessions we conducted the weight auction and for seven of the sessions the age auction (88 participants, 57% male). The subjects were paid a nominal fee (\$25) for their attendance plus auction earnings. In all auctions prices were limited to a range of \$0 - \$100 as well as “infinity” to indicate that \$100 would not be enough for the individual to reveal information to others.

All subjects were given a randomly assigned identifier and no records were kept linking individuals to this number. The experiment was fully explained to the subjects and a consent form was signed. Subjects were free to leave or not participate. Each auction form and the survey contained this ID. For the weight auction, gender, height, price, and weight were collected. For age

auctions only age, gender and price were collected. The bid with the lowest price was declared the winner (or a randomly selected bid if there was a tie). Weight was validated through a scale and age through a driver's license. To enforce truthful revelation, subjects were required to be within five pounds of the weight listed on their bid forms.

Data Analysis

All "infinity" bids were recoded to a randomly selected number between \$100 and \$2000. The log of the price was used to prevent large variations. BMI was calculated as weight (in Kg) / height² (in cm). The Kruskal-Wallis ANOVA¹ test was used throughout as well as Tukey's Honestly Significantly Different (HSD) test for pairwise comparisons of binned data.

1. Cozby, Paul, "Self-Disclosure: A Literature Review," *Psychological Bulletin*, 79(2):73-91.
2. Dindia, Kathryn, and Mike Allen, "Sex-Differences in Self-Disclosure: A Meta-Analysis," *Psychological Bulletin*, 112(1):106-124.
3. Jourard, Sidney M., *Self-Disclosure: An Experimental Analysis of the Transparent Self*, Wiley-Interscience, New York, 1971.
4. Petronio, Sandra (ed.), *Balancing the Secrets of Private Disclosures*, Lawrence Erlbaum Associates, Mahwah, New Jersey, 2000.
5. Acquisiti, Alessandro, and Jens Grossklags, "Losses, Gains, and Hyperbolic Discounting: An Experimental Approach to Information Security Attitudes and Behavior," Second Workshop on Economics and Information Security, University of Maryland, College Park, MD, May 29-30, 2003.
6. Hann, Il-Horn, Kai-Lung Hui, Tom S. Lee, and I.P.L Png, "Overcoming Online Information Privacy Concerns: A Comparison of Privacy Policies, Convenience, and Promotions," working paper, September 2003.
7. Houston, Franklin S., and Jule B. Gassenheimer, "Marketing and Exchange," *Journal of Marketing*, 51:3-18.
8. Jupiter Media Metrix, "70% of U.S. Consumers Worry About Online Privacy," Press Release, June 3, 2002.
9. Kleinberg, Jon, Christos H. Papadimitriou, and Prabhakar Raghvan, "On the Value of Private Information," *Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2001)*, Johan van Benthem, ed. Certosa di Pontignano, University of Siena, Italy, July 8-10, 2001. Morgan Kaufmann 2001.
10. Milne, George R., and Mary Ellen Gordon, "Direct Mail Privacy-Efficiency Trade-offs Within an Implied Social Contract Framework," *Journal of Public Policy & Marketing*, 12(2):206-215.

11. Spiekermann, Sarah, Jens Grossklags, and Bettina Berendt, "E-privacy in 2nd Generation E-Commerce: Privacy Preferences versus actual Behavior," *Proceedings of the 3rd ACM Conference on Electronic Commerce (EC-2001)*, Tampa, Florida, USA, October 14-17, 2001, ACM Press, pp. 38-47.
12. Varian, Hal R., "Economic Aspects of Personal Privacy", in *Privacy and Self-Regulation in the Information Age*, NTIA Report, 1996.
13. Altman, Irwin, *The Environment and Social Behavior: Privacy, Personal Space, Territory, Crowding*, Brooks/Cole Publishing, Monterey California, 1975.
14. Westin, Alan F. *Privacy and Freedom*, Atheneum, New York, 1970.
15. Altman, Irwin, and Dalmas A. Taylor, *Social Penetration: The Development of Interpersonal Relationships*, Holt, Rinehart and Winston, Inc., New York, 1973.
16. Laufer, Robert S., and Maxine Wolfe, "Privacy as a Concept and a Social Issue: A Multidimensional Development Theory," *Journal of Social Issues*, 33(3):22-42.
17. Goffman, Erving, "Stigma: Notes on the Management of Spoiled Identity," Prentice-Hall, Englewood Cliffs, NJ, 1963.
18. Goodwin, Cathy, "A Conceptualization of Motives to Seek Privacy for Nondeviant Consumption," *Journal of Consumer Psychology*, 1(3):261-284 (1992).
19. Wolfgang, Aaron, and Joan Wolfgang, "Exploration of Attitudes via Physical Interpersonal Distance Towards the Obese, Drug Users, Homosexuals, Police, and Other Marginal Figures," *Journal of Clinical Psychology*, 26:511-512.
20. Goffman, Erving, *The Presentation of Self in Everyday Life*, Doubleday, New York, 1959.
21. Simmel, Georg, *The Sociology of Georg Simmel*, ed. Kurt H. Wolf, The Free Press, New York, 1967.
22. Harper, Jim, and Solveig Singleton, "With a Grain of Salt: What Consumer Privacy Surveys Don't Tell Us," Working Paper, Competitive Enterprise Institute, June 1, 2001.
23. Singleton, Solveig, "California's Financial Privacy Debate: Privacy Policy for an Open Economy," Open Letter, Competitive Enterprise Institute, May 9, 2002.
24. Collins, Francis S., and James D. Watson, "Genetic Discrimination: Time to Act," *Science*, 302:745 (2003).
25. Cyranoski, David, "Dispute over data privacy halts cancer study," *Nature*, 424:359 (July 24, 2003)
26. Schneier, Bruce, "Protecting Privacy and Liberty," *Nature* 413:773 (October 25, 2001).

Acknowledgements

The authors would like to thank Sara Dubowsky, Mette Huberman, and Lada Adamic for their useful comments and suggestions.