

Introductory Readings in Formal Privacy for Economists

John M. Abowd, Ian M. Schmutte, William Sexton, and Lars Vilhuber

2019-04-02

Contents

1	Overview	1
2	Background	2
3	Formal Privacy	2
4	Policy and Official Statistics	3
5	Statistical Disclosure Limitation	3
6	Economics of Privacy	4
7	Value of Privacy and Data Accuracy	4
8	References	5

1 Overview

The purpose of this document is to provide scholars with a comprehensive list of readings relevant to the economic analysis of formal privacy, and particularly its application to public statistics. Statistical agencies and tech giants are rapidly adopting formal privacy models which make the tradeoff between privacy and data quality precise. The question then becomes, how much privacy loss should they allow? J. Abowd and Schmutte (2019) argue that this choice ultimately depends on how decision makers weigh the costs of privacy loss against the benefits of higher-quality data. Making progress on these questions requires familiarity with new tools from computer science and statistics, the objectives and policy environment within which statistical agencies operate, along with the economic analysis of information.

We have organized these references into a reading course focused on 10-15 primary references in each of six different topics:

- Formal Privacy
- Policy and Official Statistics
- Statistical Disclosure Limitation
- Economics of Privacy
- Value of Privacy and Data Accuracy

In the remainder of this document, for each topic, we provide a very brief description of the papers in the reading course and why we selected them. In each case, we orient the reader to the key issues, concepts, and tools in each topic. In addition to this reading course, we have also curated a much more extensive list of references, available here.

1.1 Contributing

We encourage interested readers and researchers to use these readings for their classes and training, modifying it as needed. We encourage users to fork this document at <https://github.com/labordynamicsinstitute/>

privacy-bibliography. This document is licensed under the Creative Commons Attribution-NonCommercial 4.0 International Public License (CC BY-NC 4.0). Should that license not meet your needs, contact us. We would appreciate it if any improvements, corrections, and other contributions were sent back to us, for instance via the “Issues” feature on Github.

1.2 Support

We gratefully acknowledge the support of Alfred P. Sloan Foundation Grant G-2015-13903 and NSF Grant SES-1131848.

2 Background

We start by providing a short list of background references that frame a particular set of research topics. J. Abowd and Schmutte (2019) show how to combine formal privacy models with the classic theory of public goods to understand and guide decisions about privacy protection and data dissemination. For the neophyte, Nissim et al. (2018) provide a non-technical introduction to formal privacy models in general, and differential privacy in particular. Heffetz and Ligett (2014) also introduce differential privacy, but targeted toward economists. After reading these introductory treatments, you should review the textbook treatment of differential privacy in Dwork and Roth (2014), focusing on Chapters 1–3. We also recommend consultation of the very fine tutorial “Differential Privacy in the Wild: A tutorial on current practices and open challenges” by Michael Hay, Xi He, and Ashwin Machanavajjhala.

To provide a concrete grounding in practical issues of privacy, Jones (2017) summarizes the history of privacy breaches and privacy protection in the U.S. statistical system. It is important to ask how formal privacy models do, and do not, capture common language and legal interpretations of privacy. As such, we also recommend a review of some of the laws governing data privacy in the U.S., e.g. 13 U.S. Code (1954) and 44 U.S. Code (2002) (Confidential Information Protection and Statistical Efficiency Act, also known as CIPSEA). A quick review of the Harvard Data Privacy Lab website (Harvard Data Privacy Lab 2018) can provide a sense of how differential privacy is being implemented in various settings. Goroff (2015) also provides a very useful overview of the key issues in this reading course for a lay audience. J. M. Abowd (2017) and J. M. Abowd (2016) survey formal privacy systems being implemented at the Census Bureau.

3 Formal Privacy

The literature on formal privacy is vast. Here, we focus on those papers that will help orient the reader to the key ideas and tools of differential privacy, particularly those relevant to the economic problem of determining optimal privacy protection when publishing population statistics. For the purpose of this reading course and associated bibliography, we associate formal privacy as a literature emerging in computer science out of cryptography. In the section on “Statistical Disclosure Limitation”, we recommend additional readings from the SDL literature, which has a distinct origin in statistics.

Dwork et al. (2006) is generally regarded as the first paper to formally introduce differential privacy. Its development was due, in part, to the *database reconstruction theorem* published by Dinur and Nissim (2003), which showed that most data publication mechanisms are blatantly non-private in a well-defined sense. The database reconstruction theorem has recently been shown to have very practical consequences for the statistical disclosure protections in place at the Census Bureau. The concept of *k-anonymity* developed by Sweeney (2002) is another important antecedent that bridges the formal privacy and SDL literatures.

In assessing formal privacy as a framework for modeling data publication, it is natural to consider the optimal, or maximal amount of data accuracy that can be provided while maintaining privacy. Gupta, Roth, and Ullman (2012) establish a mechanism that is universally optimal for a broad class of data users, suggesting

that technical efficiency could be guaranteed in private data publication without the need to learn about the preferences or side information of data consumers. However, Brenner and Nissim (2014) showed that their result is not generalizable to broader types of data publication. Nissim, Orlandi, and Smorodinsky (2012) provide a clear and instructive illustration of how individual preferences for privacy can be modeled in mechanism design problems.

Several papers are more directly relevant to understanding how privacy affects the work of statistical agencies. Cummings, Echenique, and Wierman (2014) argue that privacy concerns can affect the way people report data, and show how, if properly designed, privacy protection may mitigate misreporting. While there are vast number of papers on the implementation of differentially private publication algorithms, a few are particularly relevant to how statistical agencies operate. Hardt, Ligett, and McSherry (2012) and Hardt and Rothblum (2010) provide the methods and theory for publication through online query systems. One problem with these methods is that their implementation depends on the underlying data. By contrast, C. Li et al. (2015) introduce the Matrix Mechanism, which is data-independent, but also can provide high accuracy for reasonable levels of privacy. This is one of the methods under development for use with the 2020 Decennial Census. Other formal privacy systems currently in use at Census are documented in A. Machanavajjhala et al. (2008) and Haney et al. (2017).

Finally, so-called “privacy semantics” are mappings between mathematical definitions of privacy and plain language. These are really important for practitioners because there is usually a gap between how laypeople think about privacy and how it is defined in the CS literature. By way of introduction, we recommend He, Machanavajjhala, and Ding (2014), Jorgensen, Yu, and Cormode (2015).

4 Policy and Official Statistics

In January 2017, the Committee on National Statistics released a special panel report focused on developing innovations in the U.S. statistical system focused, in part, on preserving privacy (National Academies of Sciences, Engineering, and Medicine 2017). Their report is essential reading to understand the governing principles and practical needs of the statistical system, particularly as it relates to privacy modernization. For a more applied perspective, I. M. Schmutte and Vilhuber (2017) report the proceedings of an ad hoc workshop on practical privacy held at the Census Bureau. That workshop gathered together academic privacy researchers and Census domain experts to help design formal privacy systems for key data products. In such meetings, it is necessary to make sure people are speaking the same language. Prewitt (2011) describes the specific meanings of the terms “privacy” and “confidentiality” as they have historically been used at the Census Bureau.

Manski (2015) offers a framework for thinking about total error in official statistics, which refers to the various ways measured quantities may differ from the concepts of interest, including measurement error, design error, and sampling error. From this perspective, privacy protection is yet another source of error in any statistical system. Maintaining the public trust is a key factor motivating the interest of statistical agencies in privacy protection. The less people trust the system, the less likely they respond accurately, or at all. Childs, King, and Fobia (2015) discuss recent statistics on trust in official statistics and their implications for data collection. Finally, Haney et al. (2017) and Holan et al. (2010) are good examples of the sorts of implementation details one may encounter when applying statistical privacy protections in public data.

5 Statistical Disclosure Limitation

Within most national statistical systems, the primary approach to protecting respondent privacy has been *statistical disclosure limitation* or SDL. Anderson and Seltzer (2007) describes the history of threats to confidentiality in the U.S. statistical system prior to 1965. Fellegi (1972) initiated the statistical analysis of data confidentiality. Dalenius (1977) recognized that statistical agencies would need to do more than just protect against direct disclosures, and thus warned against what he called inferential disclosure. His idea was

formalized by Duncan and Lambert (1986), and provides the ultimate rationale for formal privacy in national statistics.

J. Abowd and Schmutte (2015) review the SDL methods currently in use and discuss their application to economic data. They argue that the analysis of SDL-laden data is inherently compromised because the details of the SDL protections cannot be disclosed. If they cannot be disclosed, their consequences for inference are unknowable, and, as they show, potentially large. Garfinkel (2015) discusses techniques for de-identifying data and the many ways in which modern computing tools and a data-rich environment may render effective de-identification impossible. Finally, Harris-Kojetin et al. (2005) provides the most comprehensive review of SDL methods currently in use across the U.S. statistical system.

6 Economics of Privacy

There is a large and robust literature on privacy in economics. That literature is generally focused on the value to individuals of being able to conceal private information, like a health condition, from a firm or prospective employer. The challenge to the firm is to design mechanisms, like pricing strategies, that encourage people to disclose private information. For an overview of ideas in this literature, we recommend reading Stigler (1980), Posner (1981), and Hirshleifer (1980), which appeared in an early symposium. H. Varian (1996) and Acquisti, Taylor, and Wagman (2016) both provide comprehensive surveys at different points in the development of this literature.

Very few papers tie the economics of privacy directly to official statistics. Campbell, Goldfarb, and Tucker (2015) discuss the consequences for firm profits and individual welfare of data privacy restrictions in California, which prevent firms from sharing certain types of mortgage data. Goldfarb and Tucker (2012) discuss shifts in privacy attitudes and their implications for firms. Hsu et al. (2014), address the question of optimal privacy protection from a very similar perspective to J. Abowd and Schmutte (2019). Finally, Ohm (2010) surveys the economic implications of contemporary threats to data privacy from a legal perspective.

7 Value of Privacy and Data Accuracy

One key challenge for implementing formal privacy systems lies in choosing the amount, or type, of privacy to provide. Answering this question requires some way to understand the individual and social value of privacy. Ghosh and Roth (2015) and C. Li et al. (2014) both model mechanisms for pricing private data under the assumption that individuals are only willing to disclose such information if they are paid.

Part of the social value of privacy arises from its relationship to scientific integrity. While the law of information recovery suggests that improved privacy must come at the cost of increased error in published statistics, these effects might be mitigated through two distinct channels. First, people are more truthful in surveys if they believe their data is not at risk, as Couper et al. (2008) illustrate. Second, work in computer science (Dwork et al. 2015) and statistics (Cummings et al. 2016) suggests another somewhat surprising benefit of differential privacy: protection against overfitting. A complete accounting of the costs and benefits of formal privacy systems should take these channels into account.

It is equally necessary to develop a more robust understanding of why data is valuable in the first place, the overall social cost of increasing error in public statistics. This seems to be an area in which very little comprehensive theoretical or empirical research has been done. We nevertheless recommend what seem to be good starting points.

On the theoretical side, economists studying privacy have also developed models of the value of data to firms. In these models, firms benefit from being able to tailor prices based on individual demand (C. R. Taylor 2004), or from being able to market more effectively (H. R. Varian 1998). More recently, a theoretical literature on information design has begun to consider more effective ways to manage markets for consumer information, see Bergemann, Bonatti, and Smolin (2018) and Pomatto, Strack, and Tamuz (2018). The recent literature is

related to B. D. Spencer (1985), who developed a decision-theoretic framework for modeling optimal data quality.

On the empirical side, a handful of interesting use cases suggest techniques for uncovering the value of data. For example, Card et al. (2012) and Perez-Truglia (2019) show how workers respond to pay transparency policies, which give them new information about co-worker salaries. B. Spencer and Seeskin (2015) use a calibration exercise to study the costs, measured in misallocated congressional seats, of reduced accuracy in population census data.

8 References

- 13 U.S. Code. 1954. “USC: Title 13 - Census Act.” https://www.law.cornell.edu/uscode/pdf/lii_usc_TI_13.pdf.
- 44 U.S. Code. 2002. “Confidential Information Protection and Statistical Efficiency Act.” http://www.law.cornell.edu/topn/confidential_information_protection_and_statistical_efficiency_act_of_2002.
- Abowd, John M. 2016. “Why Statistical Agencies Need to Take Privacy-Loss Budgets Seriously, and What It Means When They Do.” *The 13th Biennial Federal Committee on Statistical Methodology (FCSM) Policy Conference*, December. <https://digitalcommons.ilr.cornell.edu/ldi/32/>.
- . 2017. “How Will Statistical Agencies Operate When All Data Are Private?” *Journal of Privacy and Confidentiality* 7 (3). doi:10.29012/jpc.v7i3.404.
- Abowd, John, and Ian Schmutte. 2019. “An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices.” *American Economic Review* 109 (1): 171–202. doi:10.1257/aer.20170627.
- Abowd, John, and Ian M. Schmutte. 2015. “Economic Analysis and Statistical Disclosure Limitation.” *Brookings Papers on Economic Activity*, 221–67. doi:10.1353/eca.2016.0004.
- Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. “The Economics of Privacy.” *Journal of Economic Literature* 54 (2): 442–92. doi:10.1257/jel.54.2.442.
- Anderson, Margo, and William Seltzer. 2007. “Challenges to the Confidentiality of US Federal Statistics, 1910-1965.” *Journal of Official Statistics* 23 (1): 1. <https://www.scb.se/contentassets/ff271eeeca694f47ae99b942de61df83/challenges-to-the-confidentiality-of-u.s.-federal-statistics-1910-1965.pdf>.
- Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin. 2018. “The Design and Price of Information.” *American Economic Review* 108 (1): 1–48. doi:10.1257/aer.20161079.
- Brenner, Hai, and Kobbi Nissim. 2014. “Impossibility of Differentially Private Universally Optimal Mechanisms.” *SIAM Journal on Computing* 43 (5): 1513–40. doi:10.1137/110846671.
- Campbell, James, Avi Goldfarb, and Catherine Tucker. 2015. “Privacy Regulation and Market Structure.” *Journal of Economics & Management Strategy* 24 (1): 47–73. doi:10.1111/jems.12079.
- Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. “Inequality at Work: The Effect of Peer Salaries on Job Satisfaction.” *American Economic Review* 102 (6): 2981–3003. doi:10.1257/aer.102.6.2981.
- Childs, Jennifer Hunter, Ryan King, and Aleia Fobia. 2015. “Confidence in U.S. Federal Statistical Agencies.” *Survey Practice* 8 (5). doi:10.29115/sp-2015-0024.
- Couper, Mick P, Eleanor Singer, Frederick G Conrad, and Robert M Groves. 2008. “Risk of Disclosure, Perceptions of Risk, and Concerns About Privacy and Confidentiality as Factors in Survey Participation.” *Journal of Official Statistics* 24 (2): 255. <http://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/>

risk-of-disclosure-perceptions-of-risk-and-concerns-about-privacy-and-confidentiality-as-factors-in-survey-participation.pdf.

Cummings, Rachel, Federico Echenique, and Adam Wierman. 2014. “The Empirical Implications of Privacy-Aware Choice.” *CoRR* abs/1401.0336. <http://arxiv.org/abs/1401.0336>.

Cummings, Rachel, Katrina Ligett, Kobbi Nissim, Aaron Roth, and Zhiwei Steven Wu. 2016. “Adaptive Learning with Robust Generalization Guarantees.” *CoRR* abs/1602.07726. <http://arxiv.org/abs/1602.07726>.

Dalenius, Tore. 1977. “Towards a Methodology for Statistical Disclosure Control.” *Statistik Tidskrift* 15: 429–44. doi:10.1145/320613.320616.

Dinur, Irit, and Kobbi Nissim. 2003. “Revealing Information While Preserving Privacy.” In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, 202–10. PODS ’03. New York, NY, USA: ACM. doi:10.1145/773153.773173.

Duncan, George, and Diane Lambert. 1986. “Disclosure-Limited Data Dissemination.” *Journal of the American Statistical Association* 81 (393): 10–18. doi:10.1080/01621459.1986.10478229.

Dwork, Cynthia, and Aaron Roth. 2014. “The Algorithmic Foundations of Differential Privacy.” *Foundations and Trends in Theoretical Computer Science* 9 (3-4): 211–407. doi:10.1561/04000000042.

Dwork, Cynthia, Vitaly Feldman, Moritz Hardt, Toni Pitassi, Omer Reingold, and Aaron Roth. 2015. “Generalization in Adaptive Data Analysis and Holdout Reuse.” In *Advances in Neural Information Processing Systems 28*, edited by C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, 2341–9. <http://papers.nips.cc/paper/5993-generalization-in-adaptive-data-analysis-and-holdout-reuse.pdf>.

Dwork, Cynthia, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. “Calibrating Noise to Sensitivity in Private Data Analysis.” In *Proceedings of the Third Conference on Theory of Cryptography*, 265–84. TCC’06. Berlin, Heidelberg: Springer-Verlag. doi:10.1007/11681878_14.

Fellegi, I. P. 1972. “On the Question of Statistical Confidentiality.” *Journal of the American Statistical Association* 67 (337). American Statistical Association: 7–18. doi:10.2307/2284695.

Garfinkel, Simson. 2015. “De-Identification of Personal Information.” Internal Report 8053. Internal Report. National Institute of Standards; Technology. doi:10.6028/nist.ir.8053.

Ghosh, Arpita, and Aaron Roth. 2015. “Selling Privacy at Auction.” *Games and Economic Behavior* 91: 334–46. doi:10.1016/j.geb.2013.06.013.

Goldfarb, Avi, and Catherine Tucker. 2012. “Shifts in Privacy Concerns.” *American Economic Review* 102 (3): 349–53. doi:10.1257/aer.102.3.349.

Goroff, Daniel L. 2015. “Balancing privacy versus accuracy in research protocols.” *Science* 347 (6221): 479–80. doi:10.1126/science.aaa3483.

Gupta, Anupam, Aaron Roth, and Jonathan Ullman. 2012. “Iterative Constructions and Private Data Release.” In *Proceedings of the 9th International Conference on Theory of Cryptography*, 339–56. TCC’12. Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-28914-9_19.

Haney, Samuel, Ashwin Machanavajjhala, John M. Abowd, Matthew Graham, Mark Kutzbach, and Lars Vilhuber. 2017. “Utility Cost of Formal Privacy for Releasing National Employer-Employee Statistics.” In *Proceedings of the 2017 ACM SIGMOD International Conference on Management of Data*. SIGMOD ’17. Association for Computing Machinery. doi:10.1145/3035918.3035940.

Hardt, Moritz, and Guy N. Rothblum. 2010. “A Multiplicative Weights Mechanism for Privacy-Preserving Data Analysis.” *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. Los Alamitos, CA, USA: IEEE Computer Society, 61–70. doi:10.1109/FOCS.2010.85.

Hardt, Moritz, Katrina Ligett, and Frank McSherry. 2012. “A Simple and Practical Algorithm for Differentially Private Data Release.” In *Advances in Neural Information Processing Systems 25*, 2339–47. <http://papers>.

- nips.cc/paper/4548-a-simple-and-practical-algorithm-for-differentially-private-data-release.pdf.
- Harris-Kojetin, Brian A., Wendy L. Alvey, Lynda Carlson, Steven B. Cohen, Steve H. Cohen, Lawrence H. Cox, Robert E. Fay, et al. 2005. “Statistical Policy Working Paper 22: Report on Statistical Disclosure Limitation Methodology.” Research Report. U.S. Federal Committee on Statistical Methodology.
- Harvard Data Privacy Lab. 2018. “Harvard Data Privacy Lab Homepage.” <https://dataprivacylab.org/>. doi:N/A.
- He, Xi, Ashwin Machanavajjhala, and Bolin Ding. 2014. “Blowfish Privacy: Tuning Privacy-Utility Trade-Offs Using Policies.” In *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, 1447–58. Association for Computing Machinery. doi:10.1145/2588555.2588581.
- Heffetz, Ori, and Katrina Ligett. 2014. “Privacy and Data-Based Research.” *Journal of Economic Perspectives* 28 (2): 75–98. doi:10.1257/jep.28.2.75.
- Hirshleifer, Jack. 1980. “Privacy: Its Origin, Function, and Future.” *The Journal of Legal Studies*, 649–64. doi:10.1086/467659.
- Holan, Scott H., Daniell Toth, Marco A. R. Ferreira, and Alan F. Karr. 2010. “Bayesian Multiscale Multiple Imputation with Implications for Data Confidentiality.” *Journal of the American Statistical Association* 105 (490): 564–77. doi:10.1198/jasa.2009.ap08629.
- Hsu, Justin, Marco Gaboardi, Andreas Haeberlen, Sanjeev Khanna, Arjun Narayan, Benjamin C. Pierce, and Aaron Roth. 2014. “Differential Privacy: An Economic Method for Choosing Epsilon.” *2014 IEEE 27th Computer Security Foundations Symposium*, July, 398–410. doi:10.1109/CSF.2014.35.
- Jones, Christa. 2017. “Nonconfidential Memorandum on Census Bureau Privacy Breaches.” Memorandum to file. doi:10.5281/zenodo.1208758.
- Jorgensen, Z., T. Yu, and G. Cormode. 2015. “Conservative or Liberal? Personalized Differential Privacy.” In *Proceedings of the 2015 IEEE 31st International Conference on Data Engineering*, 1023–34. doi:10.1109/ICDE.2015.7113353.
- Li, Chao, Daniel Yang Li, Gerome Miklau, and D A N Suciu. 2014. “A Theory of Pricing Private Data.” *ACM Transactions on Database Systems* 39 (4): 34:1–34:27. doi:10.1145/2448496.2448502.
- Li, Chao, Gerome Miklau, Michael Hay, Andrew McGregor, and Vibhor Rastogi. 2015. “The Matrix Mechanism: Optimizing Linear Counting Queries Under Differential Privacy.” *The VLDB Journal* 24 (6): 757–81. doi:10.1007/s00778-015-0398-x.
- Machanavajjhala, A., D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber. 2008. “Privacy: Theory Meets Practice on the Map.” In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, 277–86. doi:10.1109/ICDE.2008.4497436.
- Manski, Charles F. 2015. “Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years After Morgenstern.” *Journal of Economic Literature* 53 (3): 631–53. doi:10.1257/jel.53.3.631.
- National Academies of Sciences, Engineering, and Medicine. 2017. *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*. Committee on National Statistics. Washington, DC: National Academies Press. doi:10.17226/24652.
- Nissim, Kobbi, Claudio Orlandi, and Rann Smorodinsky. 2012. “Privacy-Aware Mechanism Design.” In *Proceedings of the 13th ACM Conference on Electronic Commerce*, 774–89. EC ’12. New York, NY, USA: ACM. doi:10.1145/2229012.2229073.
- Nissim, Kobbi, Thomas Steinke, Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, David R. O’Brien, and Salil Vadhan. 2018. “Differential Privacy: A Primer for a Non-Technical Audience.” *Privacy Law Scholars Conference 2017*. doi:N/A.
- Ohm, Paul. 2010. “Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization.”

- UCLA Law Review* 57: 1701. <https://www.uclalawreview.org/broken-promises-of-privacy-responding-to-the-surprising-failure-c>
- Perez-Truglia, Ricardo. 2019. “The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment.” *SSRN*, February. doi:10.2139/ssrn.2657808.
- Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2018. “The Cost of Information.” 1812.04211. arXiv. <https://arxiv.org/abs/1812.04211>.
- Posner, Richard A. 1981. “The Economics of Privacy.” *The American Economic Review* 71 (2). JSTOR: 405–9. <https://www.jstor.org/stable/1815754>.
- Prewitt, Kenneth. 2011. “Why It Matters to Distinguish Between Privacy & Confidentiality.” *Journal of Privacy and Confidentiality* 3 (2): 41–47. doi:10.29012/jpc.v3i2.600.
- Schmutte, Ian M., and Lars Vilhuber, eds. 2017. *Proceedings from the 2016 NSF-Sloan Workshop on Practical Privacy*. Labor Dynamics Institute. Cornell University. <https://digitalcommons.ilr.cornell.edu/ldi/33/>.
- Spencer, Bruce D. 1985. “Optimal Data Quality.” *Journal of the American Statistical Association* 80 (391): 564–73. doi:10.1080/01621459.1985.10478155.
- Spencer, Bruce David, and Zachary H. Seeskin. 2015. “Effects of Census Accuracy on Apportionment of Congress and Allocations of Federal Funds.” *JSM Proceedings, Government Statistics Section*, 3061–75. <https://www.ipr.northwestern.edu/publications/papers/2015/ipr-wp-15-05.html>.
- Stigler, George J. 1980. “An Introduction to Privacy in Economics and Politics.” *Journal of Legal Studies* 9 (4): 623–44. doi:10.2307/724174.
- Sweeney, L. 2002. “Achieving K-Anonymity Privacy Protection Using Generalization and Suppression.” *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems* 10 (5): 571–88. doi:10.1142/s021848850200165x.
- Taylor, Curtis R. 2004. “Consumer Privacy and the Market for Customer Information.” *The RAND Journal of Economics* 35 (4): 631–50. doi:10.2307/1593765.
- Varian, Hal. 1996. “Economic aspects of personal privacy.” *Topics in Regulatory Economics and Policy*, no. 3: 1–12. doi:10.1007/978-1-4419-0038-8.
- Varian, Hal R. 1998. “Markets for Information Goods.” <http://people.ischool.berkeley.edu/~hal/Papers/japan/index.html>.