

Autonomy and the Collection of Personal Data: Measuring the Privacy Impact Google's Privacy Policy Change

James C. Cooper

Associate Professor of Law
Director, Program on Economics & Privacy
Antonin Scalia Law School, George Mason University

October 3, 2016

Abstract:

Privacy is complex and multidimensional. As such, invasions of privacy often involve subjective harm, which is inherently difficult to measure. One important dimension of privacy is autonomy: a space in which one can be free to make certain choices away from the gaze and judgment of society. Not only does autonomy give rise to personal growth, but it also can be instrumental to society. Although there is a great deal of support in the privacy law scholarship for the notion that a reduction in autonomy from unwanted observation is harmful, there is no empirical evidence to support this claim. This paper is an attempt to fill this gap in the literature by estimating the impact of Google's 2012 privacy policy change on the volume of sensitive searches—those that involve terms most would like to keep private. The theory is simple. After March 1, 2012, Google combined user information across platforms, meaning that it would have a more comprehensive view of its users. Some may want to avoid this intrusion and forego using Google to search for sensitive topics, such as those dealing with sexuality, sensitive health conditions, or controversial political views. In this manner, a reduction in sensitive search is an indirect measure of reduced autonomy.

This study relies on Google Trends data to measure weekly search volume at the state level. Using a difference-in-difference approach, non-sensitive search volume is employed as a benchmark against which to measure changes in sensitive search volume that resulted from Google's 2012 policy change. Regression analysis suggests that there was a short-term (1 month) reduction of sensitive search relative to non-sensitive search volume of about 4 percent, but there is no statistically measurable difference looking at six-month or two-month windows. Surprisingly, there also doesn't appear to be any difference between high- and low-privacy demand states—measured by the prevalence of state-level privacy legislation. These results are robust to different samples of sensitive search terms, although I cannot rule out the possibility that seasonality is playing some part in the measured short-term decrease in sensitive search.

The results suggest that consumer choice in privacy works: those who were uncomfortable with the Google's new policy of combining data reduced their use of Google for sensitive searches. However, that the reduction was small and transient indicates that any reduction in autonomy was small and perhaps swamped by customization resulting from the cross-platform data sharing. More generally, the empirical results are also in line with a host of research suggesting that consumers are not terribly concerned with the type of data sharing involved in the day-to-day functioning of the online ecosystem that relies on advertising.