Tor and the City: MSA-level Correlates of Interest in Anonymous Web Browsing

Andrew M. Lindner*

Gina Pryciak

Jamie Elsner

Department of Sociology, Skidmore College (Saratoga Springs, NY)

ABSTRACT

In the 21st century, Americans have put more of their lives online, while the U.S. government has expanded its mass surveillance apparatus. Interest in anonymity-granting technologies like the The Onion Router (Tor) has grown substantially as citizens seek to protect their privacy. However, this same technology can be used to engage in illegal activity on the Dark Web. This study examines how interest in the Dark Web, public attention to the 2013 Snowden revelations, and metro area political ideology affect public interest in Tor. We link data from multiple sources including Google Trends, the American Community Survey, the Bureau of Labor Statistics, and the Cooperative Congressional Election Study for the 49 largest U.S. MSAs from 2006-2015 ($n=490$). Broadly, we find that metro areas with liberal citizen ideology and greater interest in the Dark Web were more likely to search for Tor. When controlling for the level of interest in the Dark Web, the Snowden revelations of 2013 had no significant impact on interest in Tor. These findings tend to suggest that the lure of the Dark Web and left-leaning ideological contexts offer stronger explanations for interest in anonymity-granting technology than the public attention brought to mass surveillance by the Snowden revelations.

KEYWORDS: media sociology, mass surveillance, technology, the Internet

*Corresponding author

INTRODUCTION

Early Internet utopians conceived of the Internet as a platform for new forms of expression, thought, and discussion, creating conversations across boundaries of space and status. Individuals of all types, not just those who had made it through the traditional gatekeepers, could share their ideas with a broader public or find communities of the like-minded (Marwick 2013). The seeming capacity to be anonymous online would allow people to express radical or marginalized views, engage in forms of activism, and explore interests without suffering stigma in their workplaces or physical communities. While the downsides of anonymity, like the ability to troll others without consequence, were almost immediately apparent, the illusion of an anonymous Internet for most users held for about a little more than a decade of its existence (Nagle 2017).

In 2013, the popular notion of the Internet as an anonymous space came crashing down with increasing discoveries of the extent of state mass surveillance online. The tipping point was the release of a digital dossier leaked by Edward Snowden, a former National Security Agency (NSA) contractor, showing that the U.S. government, through its intelligence agencies, had engaged in mass surveillance of its own citizens (Lyon 2015). Since that time, the public has become increasingly aware of the ways that both governments and corporations invade digital privacy and engage in various forms of mass surveillance, whether it be human or algorithmic (Keiber 2015; Lyon 2015; Rainie and Madden 2015).

Snowden's 2013 revelations posed surveillance as a public concern, leading at least some members of the public to consider using anonymity-granting technology. These tools include anti-tracking browser extensions (e.g., Ghostery), search engines committed to digital privacy (e.g.,

DuckDuckGo), and VPN servers to re-route traffic (a simple form of content masking). A more secure form of anonymity-granting technology for members of the general public is The Onion Router (Tor). The Tor browser, which allows people to connect to web servers through virtual tunnels run by its volunteers rather than connected directly to a centralized server, has become one of the most widely-used anonymity-granting technologies¹ (Chawki 2009). Within the United States, there are 400,000 active monthly users of Tor (Jardine 2018). Tor is funded by the U.S. State Department, Mozilla, and the National Science Foundation and others due to its benefits to whistleblowers and dissidents of repressive regimes globally (Jardine 2018).

There are multiple possible motivations for becoming interested in a technology like Tor and these motivations are non-exclusive. As noted, guarding one’s anonymity against surveillance is one possible motivation for using Tor. A second reason for using Tor is less about the actual technological benefit and more about taking an ideological stance against surveillance. Finally, the Tor browser is also the dominant way of accessing the Dark Web, or the portion of the Internet not indexed by search engines and devoted to illegal activities². The Dark Web consists primarily of anonymous forums for sharing illegal pornography and cryptomarkets where drugs, guns, stolen identities, and malware code are sold (Jardine 2015). A growing body of research has shown that

¹ Using this protocol, Tor itself has no access to the unique IP addresses nor Internet service providers (ISPs) of its users. Tor volunteers set up thousands of “relays” through which Tor users can send requests for the website they intend to visit. This heavily encrypted request is then relayed through at least other two computers, shielding the IP address of the original users.

² There is no consensus on the terminology to refer to various forms of Internet content. We use “Deep Web” to refer to all portions of the web not indexed by search engines and “Dark Web” to mean the portion of the Deep Web devoted to illegal activities.

the bulk of traffic on the Dark Web is directed to drug sales and especially child abuse imagery sites (Chertoff 2017; Owenson and Savage 2015).

Even as we have begun to learn more about the content of the Dark Web, the anonymous nature of Tor makes it very difficult to know exactly who is using this anonymity-granting technology and for what reasons. Existing research has sought to better understand Tor users by examining variation in Tor use by country (Jardine 2018) and interest in Tor by U.S. states (Lindner and Xiao 2018). In this study, drawing on Google Trends data, we examine interest in Tor in U.S. Metropolitan Statistical Areas (MSAs). An MSA refers to the set of counties surrounding an urbanized core of at least 50,000 people that have “a high degree of social and economic integration with the core as measured through commuting ties” (U.S. Census Bureau 2010).

To a greater extent than either countries or states, metro areas are important spaces for cultural diffusion in many realms, including the adoption of new technologies. As L'Huillier et al. (2011) point out, people often formulate interest in virtual communities through introductions made in physical communities. While physical communities can range from rural to urban areas, the sheer scale of urban metro areas allows for greater cultural innovation, a wider range of interest groups and sub-communities, as well as diffusion across groups (Malecki 2003). The persistent “digital divide” between rural and urban areas means that cities often have faster Internet speeds at lower costs and higher rates of technological accessibility than rural areas (Malecki 2003; Nedovic-Budic and Williams 2013). Metro areas, then, serve as both economic and social hubs for technological advancement (Zook 2008). Backed by capital investors, tech

companies and adjacent firms locate themselves in cities, subsidizing the technological infrastructure. Socially, these hubs attract highly technologically-skilled employees who contribute to the culture and social integration in physical spaces (Zook 2008). MSAs serve as contexts that can inspire new interest in a whole range of cultural practices, from the types of food people eat to the kinds of technological practices in which they engage.

Taking MSAs as the unit of analysis, we examine how interest in the Dark Web, the Snowden revelations, and metro area political ideology affect interest in Tor as measured by Google search data in the 49 largest U.S. MSAs from 2006-2015. While the data are ecological in nature, the results begin to help us understand whether public interest in Tor is more commonly found after awareness-raising events like the Snowden revelations and in more politically liberal metro areas. Such findings would suggest more political motivations for initial interest in Tor. Alternatively, results showing a strong correlation between interest in the Dark Web and interest in Tor would tend to suggest that interest in anonymity-granting technologies is more closely associated with the lure of illegal pursuits online. With urbanization expanding in the 21st century, this study offers the chance to better understand whether interest in Tor is more likely to be found in communities of resistance against state surveillance or in hubs of nefarious online activity.

LITERATURE REVIEW

Since the turn of the century, Americans have put more of our lives – our photos, our communications, and our financial and medical records among other data – online (Marwick 2013;

Rainie and Madden 2015). At the same time, the U.S. government has expanded its mass surveillance apparatus, while corporations increasingly harvest our data for profit (Lyon 2015; Murakami Wood and Webster 2015, Keiber 2015). But exertions of power often prompt expressions of resistance (Marx 2003). Tor, ironically partially funded by the very state that users employ the software to block, is one such manifestation of resistance to surveillance. Simultaneously, it is a tool allowing users to access the Dark Web where they can engage in illegal activities anonymously. This study explores whether interest in the Dark Web, interest in protection from surveillance, or local political ideology offers the best predictor of overall interest in using Tor. It contributes to a small, but growing literature aimed at understanding whether anonymity-granting technologies have become tools of liberation, tools for criminal enterprise, or both. In this section, we consider the reasons why people use Tor to access the Dark Web, how the Snowden revelations raised new awareness of the need for anonymity-granting technology, and how politics play into the mass surveillance debate. But, first, we describe why MSAs are a particularly useful unit for exploring these questions.

Why Cities?

The simple answer to the question “why cities?” is because we do not have a good way of systematically observing individual behavior or interviewing individuals who use Tor. The very feature of Tor that makes it useful (i.e., that it grants anonymity) makes it difficult to identify users. Moreover, even with 400,000 active monthly users, it is unlikely that a substantial sample of Tor users be included in the sample of a nationally representative survey. While adopting an ecological approach puts us at risk for the “ecological fallacy” (i.e., that the people who have trait

x in an MSA aren't the same people who have outcome y even if the two are correlated within units), it is probably the only way to study these questions at present.

More importantly though, examining MSAs allows us to consider how conditions within metropolitan areas shape the behavior of people within them. Cities play a crucial role in society, as the percentage of the population living in cities continues to increase (Massey and Tannen 2017). As highly diverse centers of industry and employment, cities tend to be educational hubs with populations that are younger and more eager to adopt new technologies (Chinn and Fairlie 2007). Further, the newest technologies and Internet services are first accessible in urban areas, with rural areas lagging as a result of limited access, higher internet costs, or lack of exposure (Malecki 2003). Competition, technological access, capital investment, and connections with other technological hubs are all features of cities that lead to technological advancement (Zook 2008; Townsend 2001; Malecki 2009; Ferreira and Gyourko 2009). Metro areas are also centers for cultural innovation, development, and diffusion, partially due to both the large supply of high-skill workers and the competitive nature of urban life (Ferreira and Gyourko 2009; Zook 2008).

Past research has shown that members of virtual communities may be introduced to these digital spaces from acquaintances in physical communities, and that the interaction of physical and digital life shape each other, blurring the lines between the two (L'Huillier et al. 2011; Nedovic-Budic and Williams 2013). Nedovic-Budic and Williams (2003) argue that virtual life and physical life have a reciprocal or cyclical relationship, rather than a simpler linear relationship. Technology is now so ingrained in everyday life, it can sometimes be difficult to determine whether an event or situation was affected by virtual influences or not (Nedovic-Budic and

Williams 2013). In their study of Dark Web virtual communities, L’Huillier et al. (2011) show that while users form communities based on shared interests, many users were originally drawn to such communities via connections in “the real world” (i.e., physical spaces).

Thus, past findings would lead us to expect that metro areas would offer fertile contexts for drawing people into new virtual spaces. Of course, metro areas differ from each other, with some being renowned as tech hubs and others have lower-skill work forces. In general, we would expect that communities with greater population density, a more tech savvy population, and high levels of education may be more likely to enter into the realm of digital life. Compared to previous studies of states, then, this paper helps us focus in on where the action is most likely to be.

Consider a state like New York where the social context of New York City is radically different from the social context of rural areas outside of Utica. New York City is much more likely to feature connections with other cities, widespread internet access, and proximity to universities or political hubs, qualities that would logically increase the usage and knowledge of Tor. Therefore, analyzing the 49 largest metro areas allows us to examine the conditions in more internally homogeneous units. For these reasons, we believe that metropolitan areas are a particularly appropriate lens for examining a set of questions about interest in anonymity-granting technology.

The Dark Web

One motivation for using Tor is to access a variety of content only available on the Dark Web.

One prominent form of Dark Web content is the cryptomarket, where goods ranging from illegal drugs, arms, and malware code are for sale, with drugs being the most commonly sold item

(Barratt 2012; Jardine 2015). Over the past decade, various cryptomarkets have risen in popularity and subsequently fallen. But the overall trend, according to one study of 35 cryptomarkets, is toward growing sales, especially in drugs (Soska and Christin 2015). With the growth of cryptomarkets, illegal activities which once took place in physical spaces, such as drug dealing, are now moving into virtual spaces. This transformation makes it more difficult to effectively police illegal activities and reduces the risk of being caught for both the buyer and seller (Broséus et al. 2016).

Any given cryptomarket can be shut down in a raid. But, as a category, cryptomarkets have proven quite resilient. Silk Road, founded in 2011, was a Dark Web marketplace for all things illegal, including a wide variety of drugs as well as weapons and even “assassinations for hire” (Jardine 2015). While the original Silk Road was shut down in 2013 by law enforcement agencies including the FBI, there was nothing preventing similar sites from popping back up. Likewise, a Silk Road two, followed by three, made their way onto the Dark Web not long after their predecessors were shut down (Jardine, 2015).

In addition to drug markets, Dark Web sites also host other illegal activities, including child exploitation material. With untraceable shared files and encrypted user identifications, law enforcement agencies struggle to shut down child abuse imagery sites and identify the hosts, contributing to a growing base of such sites on the Dark Web (Bleakley 2018). The content of these sites is quite disturbing and includes violent, taboo, and illegal images and video (Moore and Rid 2016). Forums also exist on the Dark Web for users with shared illegal fetishes to create untraceable ‘support’ communities for those who would otherwise not disclose their desires on the

surface web (Moore and Rid 2016). While other content does exist on the Dark Web, like alternative forms of social media (Gehl 2016), research examining the volume of traffic on Dark Web sites has shown that worldwide, the majority of Dark Web traffic goes to drug cryptomarkets and, especially, child abuse imagery web sites (Owenson and Savage 2015).

The Dark Web is also a haven for scammers and petty criminals. The anonymous nature of the Dark Web allows users to partake in illegal acts with a lower risk of producing by monetary trails or communications. Among the most common illegal practices are money laundering using cryptocurrencies, counterfeit currency trading, and the buying and selling of illegally acquired credit card information (Moore and Rid 2016).

As disturbing as some of these features of the Dark Web are, they do constitute a set of motivations for people to use Tor. In this study, we use the proxy measure of Google search popularity as proxy measure for Dark Web interest in a metro area. While many – perhaps, most – people who search for the Dark Web on Google may not go on to use the Dark Web, Google searches have been shown to be a reliable measure of public interest in a topic (DiGrazia 2014; DiGrazia 2017; Stephens-Davidowitz 2014). Given that access to illicit services on the Dark Web constitutes a clear motivation for using Tor, we hypothesize:

H₁: Holding all else constant, MSAs with greater interest in the Dark Web (as measured by Google search popularity) will have greater interest in Tor (as measured by Google search popularity).

Anonymity and the Snowden Revelations

Tor usage is also a means of shielding one’s self from surveillance, which became a heightened concern after Edward Snowden leaked NSA information in 2013. Snowden’s cache of

documents revealed a massive state surveillance apparatus that relied on regular coordination with private telecommunication and tech companies. For example, one story emerging from the Snowden leak was that the NSA had direct access to user content like emails, search histories, and messaging transcripts stored by companies including Google, Facebook, and Apple (Greenwald and MacAskill 2013). The Snowden revelation of 2013 may have been the biggest news story ever bringing the issue of digital surveillance to the public's attention (Bauman et al. 2014; Keiber 2015; Landau 2013; Lyon 2015; Murakami Wood and Wright 2015; Preibusch 2015).

While anonymity granting browsers existed long before the Snowden revelations, Tor saw a large spike in usage between 2013 and 2014, which may be attributable to growing awareness of state surveillance (Gehl 2016; Rainie and Madden 2015). As members of the public become aware of state mass surveillance, they begin to feel a greater sense of their own “exposure,” or their vulnerability to surveillance (Ball 2009). Once people are aware of surveillance, they may act in various ways to manage their surveillance. One potential response is an “avoidance move,” where people avoid behaviors that put them at risk for surveillance (Marx 2003). Some evidence suggests that members of the public were less likely to search on Google for embarrassing, incriminating, or politically sensitive topics in the wake of the Snowden revelations (Marthews and Tucker 2017). Another response is a “masking move,” in which one attempts to shield one's identity from surveillance (Marx 2003). Using Tor is what Dupont (2008) has referred to as a “distributed masking move,” in which one collaborates with others digitally to shield their identity from surveillance (pg. 273).

If the Dark Web offers a “pull” motivation for using Tor, public awareness of state mass surveillance offers a “push” motivation to engage in the “masking move” of using Tor. To be sure, these are not mutually exclusive motivations. One may desire to mask their behavior precisely because it is illegal. However, if people have a general interest in masking their behavior regardless of whether the behavior is illegal, we ought to observe MSAs with a greater level of interest in state mass surveillance having higher levels of interest in Tor, net of their interest in the Dark Web.

Operationalizing this proposition is not without complexity. Unlike the Dark Web, there is probably no single search term that people with concerns about mass surveillance used on Google. Indeed, the word “surveillance” itself has multiple meanings and a search for it is likely to bring up links to home surveillance cameras. “Mass surveillance,” as a term, veers into the territory of academic jargon and, empirically, is an extremely unpopular search term on Google³. For this reason, we use searches related to Edward Snowden, a person whose story most prominently raised concerns about mass surveillance, as a proxy measure for interest in mass surveillance.

Consequently, we hypothesize:

H₂: Holding all else constant, MSAs with greater interest in the Snowden government surveillance revelations (as measured by Google search popularity) will have greater interest in Tor (as measured by Google search popularity).

Politics and Tor

³ Google cannot render estimates for “mass surveillance” for 30 out of 50 states on an annual basis.

Finally, we turn to political motivations for interest in Tor. If Dark Web users want to access something and surveillance-wary users want to protect their privacy, it is also possible to adopt Tor as a type of political statement. While the protection of individual privacy rights has historically been an important issue to small government conservatives, since the post-9/11 transformation of the intelligence community into a mass surveillance apparatus under the administration of George W. Bush, anti-surveillance attitudes have become more consistently an issue of the left (Ackerman and Siddiqui 2015; Bauman et al. 2014; Landau 2013). Though voters of both political parties tend to oppose the mass collection of private data (Ackerman and Siddiqui 2015), much of the activism and mobilization on anti-surveillance issues has come from the political left (Postigo 2012). Of course, urban areas are politically left-leaning relative to the country as a whole and that is particularly true of largest MSAs. By focusing on MSAs in this study, we exclude rural areas and smaller metro areas, which tend to lean more conservative.

For people living in more liberal-minded MSAs, there are a couple of potential conditions that may increase the probability of having one's interest in anonymity-granting technology piqued. First, given that more of the mobilization around anti-surveillance issues is organizationally and culturally situated on the left, liberal MSAs are more likely to host demonstrations, pamphleting, petition-gathering, and other social movement activities that could raise individuals' awareness of surveillance and the personal interventions one can take to manage it. Two of the most prominent organizations opposing state mass surveillance, The Electronic Frontier Foundation (EFF), and the American Civil Liberties Union (ACLU), are based in two fairly liberal-leaning cities, San Francisco and New York respectively. The largest anti-surveillance

mobilizations in recent years have occurred in New York and Washington D.C. (Bauman et al. 2014). Living in a political context where people are mobilizing around surveillance may well increase one’s feeling of “exposure” and create an interest in anonymity-granting technology.

Alternatively, interest in anonymity-granting technology like Tor may be a cultural manifestation of left-leaning politics in much the same way as pro-gun attitudes are an expression of Christian nationalist identities (Whitehead et al. 2018). That is to say, that espousing an interest in one’s privacy, reading up on Tor, and maybe even downloading it may be less about providing an actual shield for one’s privacy than about affiliating one’s self with an ideological cause. Likewise, for people who identify as liberal, there is a possibility that searching for “Tor” on Google is less about “masking moves” and more about signaling a connection to a political identity.

Whether as a consequence of left mobilization or as the result of culture practice among self-identifying liberals, we hypothesize:

H₃: Holding all else constant, MSAs with more liberal mean citizen political ideology will have greater interest in Tor (as measured by Google search popularity).

Controls

In addition to the three hypothesized measures, we control for three clusters of control measures that may well be correlated with both the independent variables and the dependent variables. First, we control for demographic measures of the metro areas including population, median income, racial composition, age distribution, level of education, and the level of inequality in each MSA. Second, we control for industries in the MSA that might use Tor *en masse*. In

particular, Tor’s own web site describes itself as used by journalists, political actors, and people in the tech industries. We control for the proportion of the local workforce employed in these industries. Though we do not offer formal hypotheses regarding these measures, it is easy to imagine that metro areas with younger, more affluent, more educated populations might be more interested in Tor. Similarly, cities with many journalists, political workers, or tech workers would have more people whose jobs provide them with incentive to use anonymity-granting technology.

METHODS

Following Lindner and Xiao’s (2018) research in U.S. states, we employ an ecological approach at the MSA-level to examine patterns of public interest in Tor. The sample for this study consists of the 49 largest U.S. metropolitan statistical areas (MSAs) with annual observations from 2006 through 2015. We use a combination of Google search data from Google Trends, demographic measures from the American Community Survey (ACS), employment measures from the Bureau of Labor Statistics (BLS), and political data aggregated from the Cooperative Congressional Election Study (CCES). Complete descriptive statistics for all measures can be found in Table 1.

The dependent measure is a metro-year measure of the Google search volume for the topic “Tor.” Data are publically-available through Google Trends, a web site which provides normalized, relative interest data on a per-query data. For each new query, the time or place with the peak popularity for that search term is given a value of 100. All other values are reported as percentages of that interest. For example, a metro area with a value of 15 has fifteen percent of

the search volume for Tor as the peak. Because the time data was reported in months and aggregated into years by the researchers, no metro-year has a value of 100 (i.e., peak search volume for Tor) for the whole year. In our data, Raleigh-Durham, NC in 2014 had the highest relative search volume for Tor (52.1), and Rochester, NY in 2006 had the lowest (5.76). Due to a skewed distribution, the measure of Tor is logged.

Two of the independent variables are also derived from Google Trends. First, to measure interest in the Dark Web, we took the metro-year average of the Google search popularity for two terms: “Dark Web” and “Deep Web.” Likewise, to assess metro area interest in Edward Snowden, we used the Google Trends topic, “Edward Snowden, American computer professional.” Both of these measures were logged.

This study makes no assumption about what happens after people search for these terms. Do they become Tor users? Do they access the Dark Web? Or do they digitally shield themselves from states and corporations? We do not know. Rather, given a substantial body of literature showing that Google searches correlate with public interest in various topics (DiGrazia 2014; DiGrazia 2017; Stephens-Davidowitz 2014), we understand higher relative search popularity in a metro-year to mean greater public interest in that topic.

We also explore how political ideology in the metro-year is associated with interest in Tor. To do so, we make use of a MSA ideology scale constructed using data from the CCES and, for the purpose of case-weighting, the ACS. The Common Content portion of the CCES is an annual nationally representative survey, administered online by YouGov. In election years, the sample is collected in pre- and post-election waves and the sample has ranged from 32,800 to 64,600

respondents. Given the CCES’s large overall sample size, in most years, there are at least fifty-three U.S. MSAs with a minimum of 100 respondents. Major metro areas like New York, Chicago, and Los Angeles typically have over a thousand CCES respondents in a given year. These samples are sufficient to render MSA-level estimates of a survey measure with a reasonable degree of confidence.⁴ For this study, we use the CCES’s 5-point ideology scale, which is measured as follows: “In general, how would you describe your own political viewpoint?”

- 1 Very liberal
- 2 Liberal
- 3 Moderate
- 4 Conservative
- 5 Very conservative
- 6 Not sure
- 8 Skipped

Cases with missing data (6 and 8) on the ideology were dropped from the sample.

To aggregate these individual-level survey responses into MSA-level estimates, we merged each case with ACS 5-year data on the demographic composition of each respondent’s MSA. Then, we constructed case-weights for the CCES respondents by race, age, and education to better reflect the MSA’s actual demographic composition.⁵ In most cases, the CCES sample was a higher proportion white, more highly educated, and older than the MSA. Then, using the CCES participants’ responses to the ideology measure, we estimated a weighted mean for each of the 49

⁴ While the CCES does not record respondents’ MSA, it does include their home county and the associated Federal Information Processing Standards (FIPS) code. By merging an MSA variable with CCES data, it is possible to identify respondents’ MSAs.

⁵ Race was weighted by the MSA proportion of white, black, and other residents. Education was weighted by the proportion of no high school, less than college degree, and college degree holders in the MSA. Age was weighted by the MSA proportion in three groups: under 35 years old, 35-64, and 65 and older.

MSAs in each year of the CCES. Finally, for the non-election years, we interpolated ideology scores by taking the mean of the prior and following year.

The resulting measure represents the mean ideology of people in a given metro area in a given year, not the actual distribution of all people in those MSAs. Thus, both an MSA where all people were moderates (3s) and a highly polarized MSA with an equal number of “Very Liberal” (1) and “Very Conservative” people would have a mean of 3. Full Stata code for this measure and a validation of the measure using Presidential Election data are available at Open Science Framework (OSF): <https://osf.io/n53ad/>. For the purposes of this study, we recoded the measure to be in line in with the direct of our hypothesis (i.e., higher values are more liberal MSAs). The MSA ideology scale ranged from 1.00 (Birmingham, AL in 2008) to 2.11 (San Francisco-Oakland, CA in 2008).

As controls, demographic measures in this study include metro-year population (in ten thousands), household median income (in thousand dollars), percent white, percent over 35 years-old, percent with college degree or more, and the Gini coefficient, a measure of inequality. All measures were drawn from the ACS 5-year estimates (2006-2015). The percentage measures and the Gini coefficient are represented on a 0-100 scale.

We also control for whether local industries are associated with increased Tor interest. In particular, journalists, tech workers, and political actors all have reasons to use Tor. Using data from the BLS, we measure the respective percentages of journalism, tech jobs, and legislative staff in a metro area.

Analytical Approach

Due to the hierarchical nature of these data, we estimated models with random effects on MSA.⁶ To address the nonlinear time trend of Tor interest, we use natural cubic splines with three knots. These splines fit cubic polynomials, which are joined at knots selected by the researchers (Beck, Katz, and Tucker 1998). The advantage is to model time trends using fewer degrees of freedom than in a less parsimonious model with dummy variables for every single year. In this case, we selected knots constructed in a three year sequence from 2009 to 2011 to reflect rises and dips in what is otherwise an upward linear trend toward greater interest in Tor.⁷ In a subsequent series of models assessing the effect of Snowden interest, we restrict the sample to the years 2013-2015 (the time period in which Edward Snowden became well-known). In these models, we employ dummy variables for 2014 and 2015.

RESULTS

The first hypothesis (H_1) was that MSAs with greater interest in the Dark Web would have more interest in Tor (as measured by Google search popularity). Table 2 displays the results of the random effects model for the logged Tor measure (2006-2015). Controlling for other conditions of the MSA, a one percent increase in the relative popularity of searches for the Dark Web in a metro area is associated with a 0.21% increase in the popularity of searches for Tor

⁶ We conducted a Hausman test to compare the choice of a random effects rather than a fixed effects model. The non-significant result ($p > .05$) suggests a random effects model is preferable because MSA errors are not uniquely correlated with the regressors.

⁷ The splines were created in Stata 15.1 using the *mk spline* package (<https://www.stata.com/manuals13/rmkspline.pdf>). In alternative analyses not shown here, we tried different numbers of knots and locations, but they did not change the substantive results. These analyses are available from the corresponding author upon request.

($p < .001$). This result supports the H_1 , suggesting that the allure of the Dark Web may be an important source of interest in Tor.

The second hypothesis (H_2) was that metro areas with greater interest in the Snowden government surveillance revelations would have greater interest in Tor. Figure 1 shows the scatter plots for the relationship between the logged Snowden measure and the logged Tor measure in each of the years after the Snowden revelations. In each year, there is a bivariate correlation between the Snowden measure and Tor of approximately 0.5 (0.511 in 2013, 0.488 in 2014, and 0.519 in 2015). In a series of random effects models presented in Table 3, we model this relationship with controls. In model 1, which does not include a control for searches for the Dark Web, we observe a positive statistically significant relationship between logged search popularity for Snowden and logged search popularity for Tor ($p < .05$). Controlling for other conditions of the MSA, a one percent increase in the relative popularity of searches for Snowden in a metro area is associated with a 0.15% increase in the popularity of searches for Tor. However, in model 2 in Table 3, when controlling for the popularity of searches for the Dark Web, the Snowden measure has no significant effect on Tor search popularity. In other words, while there is a bivariate correlation between interest in Snowden and interest in Tor, it appears to be a spurious association. MSAs with greater interest in the Dark Web took particular interest in the Snowden story and are also more interested in Tor. For this reason, we reject H_2 .

The third hypothesis (H_3) held that MSAs with more liberal mean citizen political ideology will have greater interest in Tor. In Table 2 (2006-2015), we observe a positive statistically significant relationship between how liberal a metro area is and the relative popularity of Tor

searches ($p<.05$). For each one unit increase in political liberalism of the MSA, there is a 28.7% increase in the relative popularity of searches for Tor⁸. Similarly, in the complete model for just the years post-Snowden (model 2 in Table 3), for each one unit increase in political liberalism of the MSA, there is an 18.8% increase in the relative popularity of searches for Tor ($p<.05$). While these magnitudes seem relatively large, bear in mind that the difference from the most liberal MSA to the most conservative MSA was only 1.11. Thus, a one unit increase represents a large shift in the ideology of the metro area. Nonetheless, these results support H₃, indicating that there is greater public interest in Tor in more liberal metro areas.

In terms of the controls, most of the demographic and labor market measures were non-significant ($p>.05$). In the complete model, only the percent white in the MSA had a statistically significant relationship with Tor search popularity. For each one percent increase in the share of whites in a metro area, there was a 0.3% increase in the relative popularity of Tor searches ($p<.05$). This finding was not observed in the model for only the years 2013-2015. Unsurprisingly, there were also statistically significant time dynamics in the relative popularity of Tor. The relative popularity of Tor largely increased throughout the period from 2006 through 2015, but there were dips in that overall trend in 2010 and 2012.

DISCUSSION

⁸ To calculate the percent change with a logged outcome variable and a non-transformed independent variable, we exponentiate the unstandardized coefficient and then calculate the percent change. Here, $\exp(0.252)=1.286$; $(1.286-1)*100=28.7\%$.

The findings of this study show that MSAs with liberal leanings and greater interest in the Dark Web are more likely to be interested in Tor. While there is a statistically significant bivariate correlation between interest in the Snowden revelations and interest in Tor, it is a spurious correlation, explained away by the higher level of interest in the Dark Web in such cities. Given that the research design is ecological and observes MSAs rather than individuals, we must be cautious in interpretation. However, these findings tend to suggest that the lure of the Dark Web and left-leaning ideology offer stronger explanations for interest in anonymity-granting technology than the public attention brought to mass surveillance by the Snowden revelations.

There are real reasons to find these results troubling. Cinematic portrayals of the Dark Web sometimes portray it as a cool, if somewhat naughty cyber-utopia for techno-punks. Though there is some evidence of purely ideological, even noble motivations for using the Dark Web (Gehl 2016), the truth of the Dark Web is largely a criminal digital space. Sites that promote unregulated, illegal drug and firearms sales and child abuse imagery forums garnered the bulk of the traffic on the Dark Web (Owenson and Savage 2015). While Tor was and continues to be operated as a project in digital freedom, it also enables criminal activity. In this study, that MSA-level interest in the Dark Web is a stronger predictor of Tor interest than the more philosophical motivations associated with the Snowden story. This finding is in keeping with past research that has demonstrated that Tor has more harms than benefits.

These results runs contrary to Lindner and Xiao’s (2018) state-level study, which found interest in Snowden rather than the Dark Web to be a stronger predictor of Tor search popularity. The difference in these results may be partially about the populations they examine (i.e., U.S.

states vs. MSAs). One potential interpretation is that the Dark Web is the greater driver of interest in Tor in metro areas. By contrast, strains of anti-government sentiment in suburban and rural areas that are included in the state-wide populations may make for more “purely”-motivated masking moves in non-urban counties.

It is, perhaps, telling that the Snowden interest was higher in MSAs with greater interest in the Dark Web. A longtime refrain of apologists for mass surveillance has been “what’s the problem if you have nothing to hide?” (Ball 2009). But it is possible that those with something to hide had a greater interest in the Snowden story than those with more mundane browser histories. These motivations are not mutually exclusive and deviance behavior increases the incentive to engage in masking moves (Marx 2003).

As in the Lindner and Xiao (2018) study of U.S. states, we found almost no evidence of demographic characteristics of places accounting for differences in Tor interest. Much research has documented that “digital divides” between various demographic and social groups continue (Hargittai and Hsieh 2013; Hargittai and Jennrich 2016), but the consistent findings of both Lindner and Xiao (2018) and the current study are that most differences would be found within-communities not between-communities. For advocates who seek to expand the use of anonymity-granting technologies, our findings would suggest there is as much work to do in affluent, educated cities like Boston as in less well-to-do cities like Tampa. The one exception to this pattern is that MSAs with higher percentage white population tend to have greater interest in Tor (as measured by search popularity), controlling for other factors. Communities of color have often been disproportionately targeted for state surveillance (Beutin 2017), but these findings offer some

indication that anonymity-granting technologies may not be reaching the people who might most need such digital affordances.

What is clear is that more liberal cities tend to have greater Tor search popularity. As we have suggested, interest in Tor could be a type of ideological stance against surveillance. But it could also be because there are more left activists who use anonymity-granting technology in their work. Or that media sources in left-leaning metro areas are more likely to mention anonymity-granting technologies. Another possibility is that Tor is a kind of fashionable cultural thing in liberal spaces in a way it is not in conservative communities.

Limitations

As with any study using Google search data, this research employs a proxy measure. We do not know how a Google search for “Tor” fits into a larger pattern of behavior. It may be the first step that leads to regular use of anonymity-granting technology. Or it may reflect a certain cluelessness about the technology. Nonetheless, past research has shown that Google search popularity is a reliable indicator of real world interest, so we suspect the same is true for these terms. But we cannot know with certainty. Unfortunately, anonymous technologies are an inherently difficult thing to study.

Another limitation of the current study is the use of Edward Snowden as our measure of concern with mass surveillance. The Snowden revelations are just one story and they happened in relatively recent history, giving us very few years of observations. Future research ought to

investigate how public reactions to news stories about corporate violations of privacy (e.g., Facebook) and high profile hacks of corporate databases affect interest in Tor.

CONCLUSION

This study offers an initial examination of interest in the anonymity-granting technology, Tor, in the largest 49 U.S. metro areas using Google search data. In an age of both digital lives and mass surveillance, it is important to consider whether technologies like Tor are being used as a tool of resistance or a space for criminal activity. Our findings indicate that MSAs with liberal citizen ideology and greater interest in the Dark Web were more likely to search for Tor. When controlling for the level of interest in the Dark Web, the Snowden revelations of 2013 had no significant impact on interest in Tor. Likewise, there are few demographic differences between metro areas with greater and lesser interest in Tor. Taken together, these findings tend to support previous findings showing that the illicit activities of the Dark Web are Tor's best pitch, but that interest in the technology also leans to the political left.

REFERENCES

- Ackerman, Spencer, and Sabrina Siddiqui. 2015. NSA Surveillance Opposed by American Voters from all Parties, Poll Finds. *The Guardian*, 18 May. (<https://www.theguardian.com/us-news/2015/may/18/us-voters-broadly-opposed-nsa-surveillance>).
- Ball, Kirstie. 2009. Exposure: Exploring the subject of surveillance. *Information, Communication & Society* 12(5):639-657.
- Barber, Benjamin R. 2013. *If Mayors Ruled the World: Dysfunctional Nations, Rising Cities*. New Haven: Yale University Press.
- Barratt, Monica J. 2012. Silk Road: eBay for Drugs. *Addiction* 107(3): 683-683.
- Bauman, Zygmunt, Didier Bigo, Paulo Esteves, Elspeth Guild and Vivienne Jabri. 2014. After Snowden. Rethinking the Impact of Surveillance. *International Political Sociology* 8(2):121-144.
- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *Journal of Political Science* 42:1260-288
- Beutin, Lyndsey P. 2017. Racialization as a way of seeing: the limits of counter-surveillance and Police reform. *Surveillance & Society* 15(1):5-20.
- Bleakley, Paul. 2018. Watching the Watchers: Taskforce Argos and the Evidentiary Issues Involved with Infiltrating Dark Web Child Exploitation Networks. *The Police Journal: Theory, Practice and Principles*:32258.
- Broséus, Julian, Damien Rhumorbarbe, Caroline Mireault, Vincent Ouellette, Frank Crispino and David Décary-Héту. 2016. Studying Illicit Drug Trafficking on Darknet Markets: Structure and Organisation from a Canadian Perspective. *Forensic Science International* 264:7-14.
- Chawki, Mohamed. 2009. Anonymity in Cyberspace: Finding the Balance between Privacy and Security. *International Journal of Technology Transfer and Commercialisation* 9(3):183-199.
- Chertoff, Michael. 2017. A Public Policy Perspective of the Dark Web. *Journal of Cyber Policy* 2(1):26-38.

- Chinn, Menzie D. and Robert W. Fairlie. 2007. The determinants of the global digital divide: a cross-country analysis of computer and internet penetration. *Oxford Economic Papers* 59(1):16-44.
- DiGrazia, Joseph. 2014. The Tea Party Movement: Right-Wing Mobilization in the Age of Obama. July. <http://www.riss.kr/pdu/ddodLink.do?id=T13702907>
- . 2017. Using Internet Search Data to Produce State-Level Measures: The Case of Tea Party Mobilization. *Sociological Methods & Research* 46(4):898-925.
- Dupont, Benoit. 2008. Hacking the panopticon: Distributed online surveillance and resistance. In *Surveillance and governance: crime control and beyond*, Volume 10 (pp. 257-278), Mathieu Deflem and Jeffrey T. Ulmer (ed.). Bingley, UK: Emerald Group Publishing Limited.
- Ferreira, Fernando, and Joseph Gyourko. 2009. Do political parties matter? Evidence from US cities. *The Quarterly journal of economics* 124(1): 399-422.
- Gehl, Robert W. 2016. Power/Freedom on the Dark Web: A Digital Ethnography of the Dark Web Social Network. *New Media & Society* 18(7):1219-1235.
- Hargittai, Eszter, and Yuli P. Hsieh. 2013. Digital Inequality. In *The Oxford Handbook of Internet Studies*, 129–50. Oxford, UK: Oxford University Press.
- Hargittai, Eszter, and Kaitlin Jennrich. 2016. The Online Participation Divide. In *The Communication Crisis in America, And How to Fix It*, edited by Mark Lloyd and Lewis A. Friedland, 199–213. New York: Palgrave Macmillan.
- Jardine, Eric. 2015. The Dark Web Dilemma: Tor, Anonymity and Online Policing. *Policy File* (21):1-13.
- . 2018. Tor, what is it Good for? Political Repression and the use of Online Anonymity-Granting Technologies. *New Media & Society* 20(2):435-452.
- Keiber, Jason. 2015. Surveillance Hegemony. *Surveillance & Society* 13(2):168-181.
- Landau, Susan. 2013. Making Sense from Snowden: What's Significant in the NSA Surveillance Revelations. *IEEE Security & Privacy* 11(4):54-63.
- L'Huillier, Gaston, Hector Alvarez, Sebastián Ríos and Felipe Aguilera. 2011. Topic-Based Social Network Analysis for Virtual Communities of Interests in the Dark Web. *ACM SIGKDD Explorations Newsletter* 12(2):66-73.

- Lindner, Andrew M. and Tongtian Xiao. 2018. "Subverting Surveillance or Accessing the Dark Web? Interest in the Tor Anonymity Network in U.S. States, 2006-2015." Working paper, *SocArXiv*. <https://doi.org/10.31235/osf.io/79mek>.
- Lyon, David. 2015. The Snowden Stakes: Challenges for Understanding Surveillance Today. *Surveillance & Society* 13(2):139-152.
- Malecki, Edward J. 2003. Digital development in rural areas: potentials and pitfalls. *Journal of rural studies* 19(2): 201-214.
- Malecki, Edward J. 2009. The Economic Geography of the Internet's Infrastructure. *Economic Geography* 78(4): 399-424.
- Marwick, Alice E. 2013. *Status Update*. New Haven: Yale University Press
- Massey, Douglas S., and Jonathan Tannen. 2017. Suburbanization and Segregation in the United States: 1970-2010. *Ethnic and Racial Studies* 86(4):1-18.
- Marthews, Alex, and Catherine Tucker. 2017. Government Surveillance and Internet Search Behavior. *Social Science Research Network*.
- Marx, Gary T. 2003. A tack in the shoe: Neutralizing and resisting the new surveillance. *Journal of Social Issues*, 59(2):369-390.
- Moore, Daniel, and Thomas Rid. 2016. Crypolitik and the Darknet. *Survival: Global Politics and Strategy* 58(1):7-38.
- Murakami Wood, David, and Steve Wright. 2015. Before and After Snowden. *Surveillance & Society* 13(2):132-138.
- Nagle, Angela. 2017. *Kill all Normies*. New York: John Hunt Publishing.
- Nedovic-Budic, Zorica, and Kate Williams. 2013. Becoming a uCity: The case of Chicago. *Journal of Urban Management* 2(2):3-26.
- Owenson, Gareth, and Nick Savage. 2015. The Tor Dark Net. *Policy File*.
- Postigo, Hector. 2012. Cultural production and the digital rights movement: Framing the right to participate in culture. *Information, communication & society*, 15(8):1165-1185.
- Preibusch, Sören. 2015. Privacy Behaviors After Snowden. *Communications of the ACM*, Apr 23, pp. 48.

- Rainie, Lee, and Mary Madden. 2015. Americans' Privacy Strategies Post-Snowden.
<http://www.pewinternet.org/2015/03/16/Americans-Privacy-Strategies-Post-Snowden/>
- Soska, Kyle and Nicolas Christin. 2015. Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem. Paper presented at the *24th USENIX Security Symposium* Washington, D.C., August 12-14:33-48.
- Stephens-Davidowitz, Seth. 2014. The Cost of Racial Animus on a Black Candidate: Evidence using Google Search Data. *Journal of Public Economics* 118:26-40.
- Townsend, Anthony M. 2001. The Internet and the Rise of the New Network Cities, 1969–1999. *Environment and Planning B: Planning and Design* 28(1): 39–58.
- U.S. Census Bureau. 2010. Geographic Terms and Concepts.
<https://www.census.gov/geo/reference/terms.html>
- Van Hout, Marie C., and Tim Bingham. 2013. Responsible Vendors, Intelligent Consumers: Silk Road, the Online Revolution in Drug Trading. *International Journal of Drug Policy* 25(2):183-189.
- Whitehead, Andrew L., Landon Schnabel, and Samuel L. Perry. 2018. Gun control in the crosshairs: Christian nationalism and opposition to stricter gun laws. *Socius* 4: 2378023118790189.
- Zook, Matthew, 2008. *The geography of the internet industry: Venture capital, dot-coms, and local knowledge*. New York: John Wiley & Sons.

Table 1. Descriptive Statistics

	<i>n</i>	Mean	SD	Median	Min	Max
<i>Google Search Popularity Measures</i>						
Tor (Log)	490	2.931	0.593	2.973	1.751	3.953
Dark Web (Log)	490	0.621	0.715	0.565	-0.924	2.147
Snowden (Log)	147	1.773	0.430	1.627	1.164	2.857
<i>Demographic Measures</i>						
Population (in ten thousands)	490	333.169	326.837	213.927	99.566	2018.230
Median Income (in thousands)	490	59.129	9.044	57.829	44.843	94.510
Gini Coefficient (0-100)	490	46.073	1.863	46.000	41.100	51.630
% White	490	72.103	9.404	73.185	47.202	89.389
% Over 35 Years Old	490	47.544	3.006	47.474	40.411	57.123
% College Degree+	490	32.012	5.475	31.084	20.174	49.349
Metro Ideology	490	1.540	0.177	1.552	1.000	2.111
<i>Employment Measures</i>						
% Tech Jobs	490	0.747	0.072	0.758	0.451	0.976
% Journalism Jobs	490	0.025	0.010	0.025	0.000	0.080
% Legislative Jobs	490	0.007	0.007	0.005	0.000	0.044

Table 2. Metro Random Effects Regression of Independent Variables on Google Search Volume for “Tor” (Log), 2006-2015

	Tor (Log)
Population (in ten thousands)	-0.000 (0.000)
Median Income (in thousands)	0.002 (0.002)
Gini Coefficient (0-100)	-0.000 (0.008)
% White (0-100)	0.003* (0.001)
% Over 35 Years Old (0-100)	0.003 (0.004)
% College Degree+ (0-100)	0.001 (0.004)
% Tech Jobs (0-100)	0.074 (0.182)
% Journalism Jobs (0-100)	0.311 (1.215)
% Legislative Jobs (0-100)	0.177 (1.465)
Google Searches for Dark Web (Log)	0.215*** (0.034)
Metro Ideology	0.252*** (0.069)
Cubic Spline 1	0.273*** (0.009)
Cubic Spline 2	-0.141*** (0.015)
Constant	-546.745*** (18.442)

n=490; Standard errors in parentheses; * p<.05, ** p<.01, *** p<.001

Table 3. Metro Random Effects Regression of Independent Variables on Google Search Volume for “Tor” (Log), 2013-2015

	(1) Tor (Log)	(2) Tor (Log)
Population (in ten thousands)	-0.000 (0.000)	-0.000 (0.000)
Median Income (in thousands)	-0.001 (0.004)	-0.001 (0.003)
Gini Coefficient (0-100)	-0.016 (0.010)	-0.011 (0.008)
% White (0-100)	0.001 (0.002)	0.000 (0.001)
% Over 35 Years Old (0-100)	-0.007 (0.006)	-0.005 (0.005)
% College Degree+ (0-100)	0.002 (0.005)	0.004 (0.004)
% Tech Jobs (0-100)	-0.033 (0.199)	-0.123 (0.164)
% Journalism Jobs (0-100)	-0.627 (1.371)	-0.769 (1.185)
% Legislative Jobs (0-100)	1.832 (2.071)	2.261 (1.694)
Google Searches for Snowden (Log)	0.155* (0.071)	0.071 (0.064)
Google Searches for Dark Web (Log)	-- --	0.419*** (0.062)
Metro Ideology	0.228* (0.110)	0.173* (0.091)
Year 2013	0.000 (.)	0.000 (.)
Year 2014	0.301*** (0.061)	0.191*** (0.057)
Year 2015	0.321*** (0.063)	-0.059 (0.081)
Constant	3.762*** (0.701)	3.250*** (0.584)

n=147; Standard errors in parentheses; * p<.05, ** p<.01, *** p<.001

Figure 1. Searches for “Tor” (Log) and “Edward Snowden” (Log) By Year, 2013-2015