Privacy, Polarization, and Passage of Divisive Laws

Benjamin Johnson
Chair of Cyber Trust
Technical University of Munich
Munich, Germany
benjamin.johnson@tum.de

Paul Laskowski
School of Information
University of California, Berkeley
Berkeley, California, USA
paul@ischool.berkeley.edu

***Abstract*—Notions of privacy are particularly salient to marginalized groups of people, especially when they find themselves disproportionately affected by the enforcement of laws. We use game theoretic modeling to explore the connections between privacy, polarization, and the divisiveness of laws. Our framework is based on a population of citizens that may be more or less polarized. A law is defined in terms of its effect on each citizen and must gain support from a majority in order to pass. We define a notion of divisiveness which allows us to measure the extent to which a law disproportionately affects different groups of citizens. Our framework allows us to explore four distinct notions of privacy, two that result from technological measures and two that emerge from legal theory. We find that privacy can prevent the passage of certain divisive laws, but the effects depend strongly on which type of privacy is in use.**

I. Introduction

As computer systems increasingly record and analyze our purchases, locations, and behaviors, they threaten traditional conceptions of privacy. Nowhere are the risks of this technological evolution felt more acutely than in marginalized communities. In China's Xinjiang Province, the Uyghur minority is being tracked by cameras armed with face-recognition technology [32]. Authorities are alerted automatically when individuals on a watch list stray beyond a designated safe zone. Machine learning algorithms have been shown to detect sexual orientation from facial images with more accuracy than humans [44]. This is not welcome news to the LGBT community in Nigeria, where a 2014 law criminalizes same sex activities, with punishments of up to 14 years in prison. In the United States, Immigration and Customs Enforcement (ICE) is pursuing a program to flag undocumented immigrants for deportation, using data from Twitter, Facebook, and other Internet sites. An analysis by the Brennan Center for Justice contends that the criteria are discriminatory and constitute a ban on Muslims [8].

While data collection poses severe threats to groups like these, a number of researchers have recently developed countervailing technologies that can enhance privacy. These range from cryptocurrencies for pseudonymous payments, to the Tor router for anonymous Internet browsing,to trusted hardware that can conceal computations from other programs running on a computer, to differential privacy for protecting individual records when analyzing databases. If technologies that reduce privacy are a potential threat to marginalized groups, an important open question is whether these privacy-enhancing technologies may protect them. Which types of technologies should we prioritize in order to benefit sub-populations? Is it better to protect privacy through technology or through laws, or are the two avenues complementary?

Legal and policy scholars have long drawn a connection between privacy and minority groups. Gangadharan argues that members of underserved communities are more vulnerable to the harmful effects of surveillance technologies. He provides the example of modern credit profiling, which allows the exploitation of low-scoring credit holders [17]. In the 2016 decision Utah v. Streiff, the Supreme Court considered the extent to which police officers that stop citizens without proper justification can initiate a search. In a dissent arguing for more limitations on police searches, Justice Sonya Sotomayor explicitly discussed the effect on minority groups, even though the defendant in the case was white.

> But it is no secret that people of color are disproportionate victims of this type of scrutiny... For generations, black and brown parents have given their children "the talk" – instructing them never to run down the street; always keep your hands where they can be seen; do not even think of talking back to a stranger – all out of fear of how an officer with a gun will react to them [41].

Complicating the picture, Marwick and boyd argue that privacy is often especially difficult for marginalized groups to access.

> Parents argue that they have the right to surveil their children "for safety reasons." Activists who challenge repressive regimes are regularly monitored by state actors. And poor people find themselves forced to provide information in return for basic services [29].

At the heart of these arguments is a recognition that privacy protects entire groups of people. One doesn't have to have drugs in one's pocket to object to arbitrary searches by police. Privacy places limitations on police power, affecting the playing field faced by all citizens.

Even when the vast majority of police officers abide by strict ethical standards, the prospect of running into a corrupt one remains threatening. Furthermore, the ability to invade the privacy of citizens has been argued to increase incentives for governments to abuse their power. [25]

The disparate treatment of minority groups has several origins. One is in the selective enforcement of laws. An example of this can be found in the relatively high rates at which minorities face traffic stops [31], [39]. Another relates to a 2010 law in Arizona that requires police officers to demand papers proving citizenship based on a "reasonable suspicion." Critics assert that the law gives police a license to discriminate against minorities [43].

Even if police are committed to even enforcement of laws however, a further problem is that the laws themselves may be written in a way that disproportionately affects the marginalized group. A classic example is the disparity between sentences for powder cocaine, typically associated with rich white communities, and crack cocaine, which is associated with disadvantaged black communities. Before the fair sentencing act of 2010, the weight of powder cocaine needed to trigger certain federal criminal penalties was 100 times greater than the weight of crack cocaine that would trigger the same penalties. This disparity is said to be a significant factor behind the large number of African Americans that have been sentenced for drug offenses. [6]

To understand how laws and police enforcement affect disadvantaged groups, we must also understand how society is polarized between different groups of people to begin with. How are laws passed that benefit one group at the expense of another group? Moreover, can privacy protection help marginalized groups overcome their disadvantageous position?

In this paper, we use game theoretic modeling to explore the connections between privacy, polarization, and the passage of divisive laws. Our framework is based on a population of citizens that influence what laws are passed, or what laws are maintained. A law is defined in terms of how it impacts each individual, and our model is flexible in that it allows any set of effects. We define a notion of divisiveness which allows us to measure the extent to which a law disproportionately affects different groups of citizens.

Divisiveness is not the only factor to consider when evaluating laws. A divisive law may still be justified if it significantly improves welfare. Progressive taxation is one example in which a law targets groups differently with the frequent aim of enhancing welfare. On the other hand, some laws may not be divisive at all, but may still be welfare-decreasing or unjust for other reasons. Nevertheless, we believe that divisiveness should generally be viewed as a cause for concern, especially when a law targets a marginalized group.

Our framework allows citizens to form opinions based on how a law impacts them directly, but it optionally allows them to consider the impact on others as well. This is achieved through an influence matrix that is multiplied by the direct effect of the law. This can be used to represent a concern for friends, loyalty to a larger group, or learning from a small number of influential personalities. The influence matrix also allows us to discuss how polarized society is. At the end of our analysis, we apply this framework to model a society with one majority group and one minority group.

Our model assumes that laws that are supported by a majority of citizens are passed or maintained. Although the democratic process is generally much more complicated than a simple majority vote, we believe this is a useful and tractable way to explain the types of laws that exist in society.

Using our model of how laws are enforced, we are able to identity four distinct notions of privacy protection. Two of these are technological, including strategies that citizens can take to hide features and behaviors from authorities. The other two are legal notions, depending on a judicial branch that functions as a check on enforcement procedures. We describe the function of each of these privacy protections using our two-population model of society. We find that each type of privacy protection allows a different set of laws to be passed and enforced, resulting in different effects on divisiveness. Our work supports the idea that the protection of privacy, while far from a perfect cure, has a role to play in mitigating the divisive effects of laws in a society.

II. Related work

A. Privacy and Government

This work falls within a line of research that investigates how realizations of privacy affect the balance of power between citizens and the state. Laskowski and Johnson investigate the application of surveillance technology by a government that wishes to remain in power [25]. A major takeaway is that enhanced surveillance technology increases incentives for abuse. In a similar vein, Goh provides a model of a government that may employ surveillance to lower the risk of a terrorist attack [19]. Greater surveillance carries an increased risk that citizens will learn of its existence, which increases the risk that the government loses power. Goh finds that a rational government will employ less surveillance when citizens value their privacy more, but autocratic governments will employ more surveillance than democratic ones.

A larger literature examines the relationship between citizens and the state in general. Downs [13] provides a model of political competition based on a continuum of political preferences, extending Hoteling's study of horizontal differentiation [34]. Further studies model the process by which governments are overthrown. Ginkel and Smith [18] consider factors that determine the probability of revolution in a repressive regime. Lohmann [27] describes the potential overthrow of a government through an informational cascade model. Kuran [21] attempts to explain why revolutions often take the world by surprise

with a game theoretic model of political change. These studies do not consider the effects of privacy protections in determining political outcomes.

B. Technological Privacy Protections

Members of a marginalized group may employ a variety of technologies to enhance their privacy. These may be divided into technologies that conceal specific behaviors, and technologies that maintain the secrecy of personal characteristics. In the former category are technologies like virtual private networks and Tor, which allows anonymous internet browsing [12]. Cryptocurrencies like Bitcoin allow decentralized payments while making it difficult to discover the owners of individual accounts. Smart contracts on platforms like Ethereum allow a richer set of computations while concealing the identities of contracting parties [9]. Darknet markets provide a platform on which buyers and sellers of forbidden activities may connect with each other [40].

A second category of technologies protects personal information found in databases. Today's companies increasing store a rich variety of customer data, which may include browsing history, purchases, physical location, search terms, and so on. Such data can be used to infer a variety of personal characteristics, including minority status. Moreover, aggregating information from multiple datasets can improve the ability of a government to identify desired individuals [1].

A number of techniques have been proposed to protect information contained in databases. Early attempts focused on syntactic notions like k-anonymity [42], which requires that a database entry corresponding to any individual appears multiple times. Such notions involve a relatively weak adversary model and are vulnerable to known attacks. More recently, a series of papers developed the standard of differential privacy, which requires that the probability of any given algorithm output can only change by a bounded amount if a single entry is altered [15]. Such a guarantee may often be achieved by adding a calibrated amount of noise to an algorithm [14]. Alternately, the exponential mechanism of McSherry and Talwar may be used to provide privacy protections over discrete outcomes [30].

Trusted hardware like Intel's Software Guard Extensions are designed to allow data to be processed securely, even when the underlying computer is not trusted [2]. Cheng, et al. present a system that combines trusted hardware with a blockchain to enable differentially private computations in a distributed network [10].

An ongoing debate surrounds the use of privacy-enhancing technologies, in contexts ranging from ethics [11], [28], to law [3], [22]–[24], to security-relevant effectiveness [36]. While our model abstracts from these details, we will use it to explore the impact of technologically-based privacy enhancements.

C. Legal Privacy Protections

In the United States, support for a right to privacy can be found in the constitution. In particular, the Fourth Amendment states,

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

Slobogin argues that the structure of the amendment was a direct response to the "general warrants" that the British used to engage in broad house-to-house searches [38]. More recently, courts have been relatively permissive of such dragnet searches, barring only those that are "unreasonable."

The warrant clause of the Fourth Amendment provides the strongest protection against individual searches, requiring a standard of probable cause. There are however exceptions to this rule, and police often do not need a warrant to conduct a search. Even so, a search ordinarily requires an individual suspicion of wrongdoing in order to be found reasonable [16].

The so-called "exclusionary rule" provides protection against the collection of evidence in violation of the Fourth Amendment. It prevents such evidence from being used in court, regardless of whether a defendant actually committed a crime [35]. This provides a disincentive for police to engage in searches without proper justification.

During oral argument in Utah v. Strieff, which considered limitations to the exclusionary rule, Justice Sonya Sotomayor described the behavior that it restrains:

> What stops us from becoming a police state and just having the police stand on the corner down here and stop every person, ask them for identification, put it through, and if a warrant comes up, searching them? [41]

D. Polarization

There is at least some evidence that most voters in the United States could be considered moderate in their policy positions [26]. Nevertheless, the United States Congress has passed a number of divisive laws, many of which have been challenged and overturned by the US Supreme Court. Divisive laws have also been passed in European countries, such as those against face covering, pejoratively dubbed "burka bans."

The United States congress in particular has become increasingly polarized over the last 40 years (see Figure 1). Researchers have posited a number of reasons for this phenomenon [4] [33], ranging from a polarized electorate, to southern realignment, to gerrymandering, to the evolution of modern primary elections, to economic inequality, to money in politics, to the media environment, or to congress-based factors such as congressional rule changes, majority

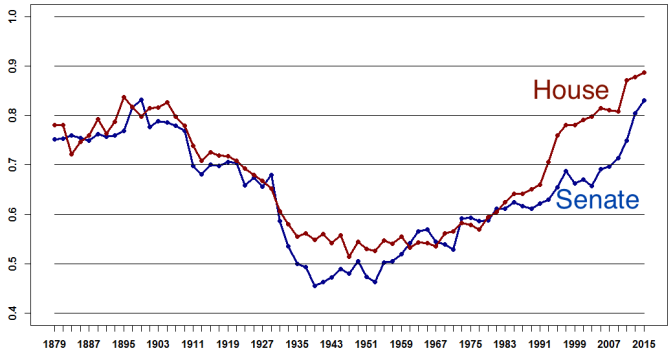

Figure 1. **Distance between Parties in the US House and Senate, 1879-2015**

party agenda control, party pressures, teamsmanship, or the breakdown of bipartisan norms. All of these issues are discussed in [33]. More culturally-specific theories involving authoritarianism are also prevalent [20].

A more economically-driven explanation derives from the notion of information cascades. An information cascade occurs when people receive a noisy informational signal and observe the behavior of friends and colleagues to inform decision-making. Although agents are individually rational, they may find it optimal to rely on the information they derive from previous agents, ignoring their private signals [7].

The notion that people exhibit herding behavior in predictable circumstances has been around for decades [37]. For example, researchers at Iowa State University conducted 259 interviews with farmers who had largely refused offers to adopt drought-resistant seed corn during the Great Depression and Dust Bowl. They found that the slow rate of adoption was due to "how farmers valued the opinion of their friends and neighbors instead of the word of a salesman" [5].

We include a notion of influenced behavior in our model as a way to describe the polarization of society. Our model does not mandate that citizens consider how a law affects other citizens, but merely allows it.

III. Model

A. Definitions

As we use the terms divisiveness, polarization, and privacy throughout the paper, we take care to define them here.

- *Divisiveness* of a law refers to the extent to which the different citizens value the law differently – measured, for example, by the standard deviation in valuations.
- *Polarization* of a society refers to the grouping of individuals into sets that have little influence over each other, or that prioritize the welfare of insiders over that of outsiders.
- *Privacy* refers to an individual's ability to possess property (physical or intellectual) that is free from inspection or search.

B. Overview

Our general model presents an incentive structure that relates the polarization of a society to the passage of divisive laws. At a high level, our model captures the following dynamics.

- Each individual begins with an initial evaluation of the law. This represents the direct benefit of the law to that individual.
- Each individual may (optionally) be influenced by others. The polarization of society is captured in an influence matrix.
- An individual's support for a law is found by combining their initial evaluations with those of their influencers.
- Laws that gather positive support from a majority of citizens are passed or maintained.

C. Citizen Influence

For each pair of citizens $i, j \in \{1, \dots N\}$, define

$$a_{ij} \in [0, 1]$$

to be the influence of person j on person i. We may think of a_{ij} also as the affinity person i has toward person j.

Each individual will arrive at a level of support for the law based on a weighted average of the valuations of their influencers, where the weighting is determined by the influence parameters a_{ij}. To be consistent with the notion of weighted average, we thus require

$$\sum_j a_{ij} = 1.$$

D. Citizen Valuations of Laws

Given a law L and a citizen $i \in \{1, \dots N\}$, let

$$V_L(i) \in \mathbb{R}$$

represent the *direct impact* of the law on person i. We assume that the direct impact of L is unbounded because laws may save a life or take away life from individuals impacted by them.

While the direct impact represents an initial evaluation of a law, we intuitively expect a person's view to evolve after some time, based on the views and experiences of their influences. To capture this phenomenon, we separately define the *support* of citizen i for law L to be

$$U_L(i) = \sum_j a_{ij} V_L(j). \tag{1}$$

If an individual's direct valuation for a law is perfectly aligned with the direct valuations of their influencers, then this person's support for a law will simply equal $V_L(i)$. In other cases, a law may have little or no direct impact on the individual, but their influences may lead them to have strongly positive or strongly negative support for the law.

For example, most Americans may experience little direct benefit from a ban on muslims, but they may be influenced by those that are directly affected to oppose this policy.

IV. Analysis

A. A Law Valuation Functional Form

In this section, we specify a functional form for V_L in order to discuss the effects of various privacy protections. While this reduces the generality of our model, it allows us to distinguish between the effects of different protection mechanisms.

We imagine that three conditions must apply for citizen i to be punished by a law:

1) As the text of the law is written, it specifies that citizen i is engaged in unlawful (criminal) behavior. We assume this occurs with probability $C_L(i)$.
2) Citizen i is searched by the authorities. We assume this occurs with probability $S_L(i)$.
3) A search on citizen i is successful in finding evidence. Conditional on the previous conditions, we assume this occurs with probability $F_L(i)$. For simplicity, we will assume this function is a constant and write it as F_L.

The probability that citizen i is punished in the context of law L is therefore $S_L(i)C_L(i)F_L$. Let P_L be the punishment for violating the law. Then each citizen's individual utility based only on the law's expected punishment on them may be given by

$$S_L(i)C_L(i)F_L \cdot P_L.$$

We may also compute the total number of citizens punished by the law as

$$\sum_j S_L(j)C_L(j)F_L.$$

As far as the benefits of a law go, we assume that each individual i suffers a cost $l_L(i)$ each time a particular crime is committed. This may be interpreted to include direct effects (e.g. the individual is targeted by theft or violence) as well as indirect effects (e.g. the prevalence of a crime makes the individual feel less welcome in their community). For the current analysis, we assume that the number of crimes is reduced by exactly the number of people punished. In other words, we do not allow for deterrent effects. If such effects exist, they may multiply the effects of each search, and they may cause non-linearities in the valuation functions. We defer these interesting issues to a future analysis.

Given these assumptions, we can write the initial valuation of law L to person i as the difference between the benefit gained from less crime and the individual's expected punishment in case she breaks the law. That is,

$$V_L(i) = l_L(i) \sum_j S_L(j)C_L(j)F_L - S_L(i)C_L(i)F_L \cdot P_L \tag{2}$$

B. Four Types of Privacy Protections

Our valuation model allows us to compare four types of privacy protections, the first two of which are technological in nature, and the second two of which are legal.

1) Attribute Privacy Enhancement: Attribute privacy is centered on the notion that a person can conceal personal characteristics, which authorities may use to identify them as someone likely to commit a crime. *Attribute privacy enhancement serves to prevent or reduce targeted searches.* An example in which the attributes of individuals are fully or nearly anonymized is the case in which users of TOR hide their online activities. We may model this type of privacy protection by stipulating that $S_L(i)$ is a constant.

$$S_L(i) = S_L \tag{3}$$

For a given text of a law, which implicitly specifies for each citizen i the probability with which they are a criminal with respect to the law, restrictions on S_L may prevent a majority from emerging to support the law. In practice, when individual-targeted searches are prevented, we may expect authorities to restrict searches to a minority of the population, thereby ensuring that the law's valuation $V_L(i)$ is positive for a majority of individuals.

In this case, the shape of V_L is entirely determined by the distribution of criminal behavior. Thus the number of individuals that support the law depends only on the likelihood of criminality $C_L(i)$ and how polarized society is.

2) Search Privacy Enhancement: Search privacy in our context represents the idea that a citizen may use technology to prevent the discovery of evidence in the event that she is searched. We represent this type of privacy protection as a decrease in the parameter, F_L,

$$F_L < 1 \tag{4}$$

representing the chance that evidence is found when a citizen breaking the law is searched. In our valuation model, F_L appears in both the positive and negative components of V_L so that it does not change the number of citizens supporting the law.

Nevertheless, this type of privacy protection will affect welfare, scaling it towards zero. Search privacy enhancement may therefore be welfare-benefitting in the case of divisive laws for which welfare is negative.

3) Dragnet (High Volume) Search Prohibition: Our first legal notion of privacy protection corresponds to the idea that searches should not be widespread in a society. Commentary discussed in the introduction involving the recent supreme court case Utah v. Strieff exemplifies this notion; and it is encoded in our valuation model by requiring that the fraction of citizens who are searched is less than some bound,

$$\frac{\sum_j S_L(j)}{n} < m. \tag{5}$$

Laws that don't meet this requirement may be declared unenforceable. Dragnet search prohibition prevents the enforcement of laws against a large fraction of a population.

4) Low Evidence Search Prohibition: Another legal notion of privacy protection supposes that authorities must have individualized reasonable suspicion to conduct a search. This is similar to the notion of dragnet search prohibition, defined above, but the focus is not on the total proportion of searches with respect to a population, but rather on how well-targeted the searches are. We may encode this notion of privacy protection by requiring that a certain fraction of searches result in the finding of evidence,

$$\frac{\sum_j S_L(j)C_L(j)F_L}{\sum_j S_L(j)} > \rho \tag{6}$$

C. A Two-Party Influence Framework

To further explore issues of privacy, we will place additional structure on citizen affinity. In this section, we analyze the special case that there are two subsets of citizens, labeled G_0 and G_1. Citizens in G_0 are assumed to form the majority with population n_0. Citizens in G_1 form the minority with population $n_1 < n_0$. We will also refer to members of G_0 and G_1 as type 0 and type 1, respectively. Fixing L for the moment, we assume that all members of a group G_k ($k \in \{0, 1\}$) share the same probability of being a criminal, which we denote by C_k, as well as the same probability of being searched, which we denote by S_k. We also assume all citizens share the same direct cost for each crime committed, labeled l.

Within each group, we further assume that all citizens share the same level of support for the law. For example, one way for this to happen would be if the influence function a_{ij} itself were constant within each of the four cases $i, j \in G_0$; $i \in G_0, j \in G_1$; $i \in G_1, j \in G_0$; and $i, j \in G_1$. In this case, regardless of any individual's initial valuation of a law, the resulting final support for the law would be uniform within each group.

As another example, we could assume that each of the two subgroups of citizens included a single influential thought leader, labeled $t_0 \in G_0$ and $t_1 \in G_1$. Suppose that every individual only had positive affinity for these two thought leaders, and the vector of affinities within each group was uniform. In terms of final support for a law, this scenario is indistinguishable from the example above.

Regardless of how support is formed, the average affinity an individual in group G_k has for an individual in group $G_{k'}$ may be computed as

$$A_{k,k'} = \frac{1}{n_k n_{k'}} \sum_{i \in G_k, j \in G_{k'}} a_{ij} \qquad k, k' \in \{0, 1\}. \tag{7}$$

Our last assumption encodes some degree of polarization in this society. We assume that each group has a higher inter-group affinity than a cross-group affinity. That is,

$$A_{00} > \frac{1}{N} = \frac{1}{n_0 + n_1} > A_{01}, \text{ and} \tag{8}$$

$$A_{11} > \frac{1}{N} = \frac{1}{n_0 + n_1} > A_{10}. \tag{9}$$

We may compute $p_0 = C_0 S_0 F$ as the fraction of the majority that are punished, and $p_1 = C_1 S_1 F$ as the fraction of the minority that are punished. We use Figure 2 to plot these quantities for different possible laws. On this graph, the x-axis represents p_0 and the y-axis represents p_1. In our model, the text of the law specifies the maximum fraction of each type of citizen that can be punished, C_0 and C_1, which occurs when $F = 1$ and everyone is searched. This region is represented by the outermost rectangle in the Figure, labeled A.

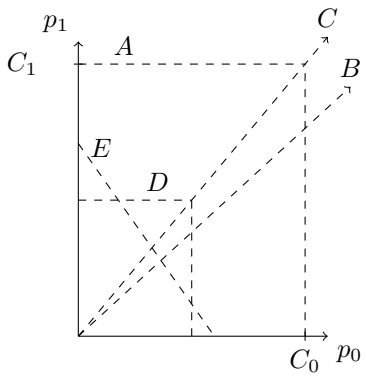

Figure 2. A plot representing the space of possible laws in terms of how many majority members, and how many minority members are punished.

Applying our more restrictive framework to Equation 2, the valuation of an individual in group G_k may be computed as

$$V_k = l \sum_j S(j)C(j)F - S(i)C(i)F \cdot P \tag{10}$$

$$= l\,(n_0 p_0 + n_1 p_1) - p_k P \tag{11}$$

Since more than half of the population is in G_0, a law may pass whenever the support of each individual in group G_0 is non-negative. Following Equation 1, we may compute the support U_0 from individuals in group G_0 as follows.

$$\begin{aligned} U_0 &= n_0 A_{00} V_0 + n_1 A_{01} V_1 \\ &= n_0 A_{00}[l(p_0 n_0 + p_1 n_1) - p_0 P] \\ &\quad + n_1 A_{01}[l(p_0 n_0 + p_1 n_1) - p_1 P] \\ &= l(p_0 n_0 + p_1 n_1) - n_0 A_{00} p_0 P - n_1 A_{01} p_1 P \\ &= [l - P A_{00}] n_0 p_0 + [l - P A_{01}] n_1 p_1 \end{aligned} \tag{12}$$

Our assumption that $A_{00} > A_{01}$ implies that there are 3 possibilities for the signs of the coefficients on p_0 and p_1.

1) $\partial U_0 / \partial p_0 \leq 0$ and $\partial U_0 / \partial p_1 \leq 0$. This will happen whenever $l \leq A_{01} P$, meaning that the benefit from

the law is small relative to the size of the punishment multiplied by the affinity of the majority for the minority.

2) $\partial U_0/\partial p_0 \leq 0$ and $\partial U_0/\partial p_1 > 0$. This will happen whenever $A_{01}P < l \leq A_{00}P$, or for intermediary amounts of benefit. In this case, the majority will tend to support laws that punish the minority more and the majority less.
3) $\partial U_0/\partial p_0 > 0$ and $\partial U_0/\partial p_1 > 0$. This will happen whenever $A_{00}P < l$, or for high amounts of benefit. In this case, the majority will tend to support laws that punish both groups as much as possible.

The middle possibility, in which majority support increases with p_1 but decreases with p_0 is of particular interest, since this is the case in which the majority would favor the most divisive treatment of the two groups.

If condition 1 holds above, no law within rectangle A can have majority support, except for the origin. This means that nobody will be punished and citizens are unaffected. If condition 3 holds above, every law within rectangle A has majority support, so a law may pass for any values of S_0, S_1, and F. For the more interesting condition 2, the previous equation divides rectangle A into two regions. This is depicted by line B in Figure 2. Every law that falls on this line or above can have majority support.

Now the support of an individual in the minority group G_1 may be expressed as

$$U_1 = [l - PA_{11}]n_1p_1 + [l - PA_{10}]n_0p_0 \tag{13}$$

We now compute the divisiveness of a law L as the standard deviation,

$$\begin{aligned} s_L &= \sqrt{var(U_L)} = \sqrt{E\left[\left(U_L - \overline{U_L}\right)^2\right]} \\ &= \sqrt{\frac{n_0\left(U_0 - \overline{U_L}\right)^2 + n_1\left(U_1 - \overline{U_L}\right)^2}{n_0 + n_1}}, \end{aligned}$$

where $\overline{U_L} = \frac{n_0U_0 + n_1U_1}{n_0+n_1}$ is the average support for the law. After applying some algebra, and using the fact that $U_0 > U_1$ in the region where laws have majority support, we obtain

$$\begin{aligned} s_L &= \frac{\sqrt{n_0n_1}}{n_0 + n_1}\left|U_0 - U_1\right| \\ &= \frac{\sqrt{n_0n_1}}{n_0 + n_1}\left((A_{10} - A_{00})Pn_0p_0 + (A_{11} - A_{01})Pn_1p_1\right). \end{aligned}$$

Now since we know $A_{10} - A_{00} < 0$ and $A_{11} - A_{01} > 0$, we also have

$$\partial s_L/\partial p_0 < 0 \text{ and } \partial s_L/\partial p_1 > 0. \tag{14}$$

In other words, in the region of laws with majority support, divisiveness goes up with the number of minority members punished and down with the number of minority members punished. This means that the law most favored by the majority, which is in the upper left corner of rectangle A, is also the most divisive.

D. Effects of Privacy Protections

Having outlined some basic properties of our model, we turn our attention to the effects of different types of privacy protections.

1) Effects of Attribute Privacy Enhancement: We begin with attribute privacy enhancement, which we understand to mean restrictions on the search probabilities S_k. We model the extreme case of anonymity, represented as $S_0 = S_1$. The set of possible laws is represented by line C in Figure 2. Note that line C passes through the point (C_0, C_1) at the upper right of rectangle A. A more moderate version of attribute privacy enhancement may place bounds on the ratio of S_1 to S_0. For example, we may specify,

$$0 < \underline{r} < S_1/S_0 < \bar{r} < \infty \tag{15}$$

Since divisiveness is maximized at the point $(0, C_1)$, attribute privacy protection necessarily reduces the maximum amount of divisiveness since it disallows laws at this point. As the figure is depicted, the slope of line C is greater than the slope of line B,

This means that any law that falls inside line C and rectangle A will also have majority support. The majority would offer the most support to laws that fall further to the upper right of the line. It is also possible, for the slope of line C to be less than the slope of line B,

In this case, no law on line C will have majority support except for a law at the origin. Nevertheless, divisiveness decreases since having no law or a law at the origin brings divisiveness to zero.

2) Effects of Search Privacy Enhancement: We interpret search privacy enhancement to mean that not all individuals who are searched while committing a crime are caught,

$$F < 1 \tag{16}$$

This condition limits the maximum number of people, of either type, who may be punished. This is depicted by rectangle D in Figure 2. Like attribute privacy enhancement, search privacy enhancement reduces the maximum amount of divisiveness that can be caused by a law. Specifically, the maximum divisiveness is reduced by a factor of F. Unlike attribute privacy enhancement, search privacy enhancement never results in a scenario in which a previously enforced law can no longer have majority support. As long as S_0 and S_1 are unchanged, reducing F still results in a law with majority support. Divisiveness and welfare are simply scaled downwards.

3) Effects of Dragnet Search Prohibition: To model a policy of dragnet search prohibition, we impose the condition,

$$\frac{n_0S_0 + n_1S_1}{n_0 + n_1} < m \tag{17}$$

The maximum value of S_0 can be attained when $S_1 = 0$, allowing $S_0 = m(n_0 + n_1)/n_0$. Similarly, when $S_0 = 0$, S_1 can attain a value of $S_1 = m(n_0 + n_1)/n_1$. Note that the maximum value of S_0 is smaller than the maximum value of S_1. This feature is plotted as line E in Figure 2. Laws below this line are acceptable under the dragnet search standard.

As the figure demonstrates, dragnet search prohibition may reduce the maximum possible divisiveness of a law. The majority will still favor a law that is at the top corner of the acceptable region, which represents the maximum amount of divisiveness for a given amount of searches. If the bound m is not set low enough, the maximum amount of divisiveness may not be reduced at all.

4) Effects of Low Evidence Search Prohibition: Low evidence search prohibition requires a certain proportion of searches to turn up evidence that an individual is breaking the law. We represent this by writing,

$$\frac{\sum_j S(j)C(j)F}{\sum_j S(j)} = \frac{n_0 S_0 C_0 F + n_1 S_1 C_1 F}{n_0 S_0 + n_1 S_1} > \rho.$$

Rearranging, we obtain

$$(n_1 C_1 F - \rho n_1) S_1 > -(n_0 C_0 F - \rho n_0) S_0.$$

When ρ is close enough to 0, the left hand side is non-negative and the right hand side is non-positive, so the constraint does not bind. The set of possible laws is therefore not reduced. On the other hand, when ρ is close enough to 1, the left hand side is non-positive and the right hand side is non-negative, so the constraint binds unless $S_0 = S_1 = 0$. No law can pass the privacy requirement except for the trivial one that doesn't punish anyone.

The more interesting region is when ρ lies between $C_0 F$ and $C_1 F$. In the case that $C_1 > C_0$, meaning that the text of the law targets behavior that is more common for individuals of type 1, we can write the constraint as

$$\frac{S_1}{S_0} > \frac{\rho n_0 - n_0 C_0 F}{n_1 C_1 F - \rho n_1}, \tag{18}$$

where the fraction is arranged so that the numerator and denominator are positive. This constraint is depicted by line F in Figure 2. Every law above this line is permitted under the search specificity rule. Unfortunately, this rule does nothing to reduce the maximum possible divisiveness of a law. In fact, it may have the opposite effect, allowing only laws for which members of the minority are much more likely to be searched than members of the majority.

In part, this conclusion may be the result of limitations of our model. We have assumed that all members within a group have equal odds of breaking the law and equal odds of being searched. A more detailed model might allow individuals to vary within each group. Authorities may be able to select a smaller fraction of each group, choosing individuals with a higher probability of breaking the law, thereby increasing the proportion of searches that turn up evidence. In such a model, search specificity privacy may be expected to limit the maximum divisiveness of a law. Nevertheless, we believe that our central insight, that search specificity privacy limits searches on the majority more than on the minority, will continue to hold.

V. Discussion and Conclusion

Our study is an attempt to understand privacy, not at the level of individual incentives, but at the level of communities and the relationships they have to each other. Though our modeling framework is exploratory, it reveals some of the complexity inherent in this setting. We were able to relate privacy to polarization and the divisiveness of laws, but found that outcomes depend critically on what notion of privacy protections are in play. Privacy protections enforced through technology can have dramatic effects on what laws are enforced, but all laws may be rendered ineffective whether divisive or not. A privacy-enhancing technology works equally well, after all, whether it is concealing sexual behavior between consenting adults, or a plot to assassinate a leader in government. Legal notions of privacy protections hold the promise of more precise judgments. Courts can identify disadvantaged groups and extend protections to them, without extending those same protections to say, serial killers. Yet our model predicts that here too, privacy protections are not perfectly tailored to prevent the enforcement of divisive laws. Further research is needed to assess how privacy laws may work in conjunction with anti-discrimination laws and policies, to better protect marginalized groups.

In the future, we plan to extend our model to capture more features of the legal system and its relation to privacy. One important addition will be a description of how authorities gather information on potential law breakers. A detailed model might describe citizens with a collection of attributes, each of which may carry information about a citizen's propensity to violate a specific law. Some attributes may be hidden with technological privacy-enhancing measures, some may be the topic of legal protection, and others may remain public and available to police as they direct their investigations.

We would further like to model the issues that arise when police have the right to search a large number of people, but only enough resources to choose a small number. The ability to arbitrarily select which citizens to search carries a significant amount of power, even when the number of searches remains small. A more complete model would separate legality from the actual performance of a search in order to highlight these issues.

As technology advances, many established notions of privacy protection face considerable pressure to evolve. Location monitoring, deep packet inspection, linking of consumer databases, and face recognition are just a few of the threats to our ability to control our personal information. We hope that studies like ours will help spur discussion about the role privacy protection plays in

maintaining balances of power, and how future definitions of privacy may best be structured to protect vulnerable groups and individuals.

Acknowledgments: We are grateful to the reviewers for their comments. Our research was supported substantially by a grant from the Center for Long-Term Cybersecurity at UC Berkeley. The research activities of Benjamin Johnson are supported by the German Institute for Trust and Safety on the Internet (DIVSI).

References

[1] Acquisti, A., Gross, R.: Predicting social security numbers from public data. Proceedings of the National academy of sciences 106(27), 10975–10980 (2009)
[2] Anati, I., Gueron, S., Johnson, S., Scarlata, V.: Innovative technology for cpu based attestation and sealing. In: Proceedings of the 2nd international workshop on hardware and architectural support for security and privacy. vol. 13. ACM New York, NY, USA (2013)
[3] Bankston, K., Soltani, A.: Tiny constables and the cost of surveillance: Making cents out of united states v. jones. Yale Law Journal Online 123 (2014)
[4] Barber, M., McCarty, N.: Causes and consequences of polarization. Solutions to Political Polarization in America 15 (2015)
[5] Beal, G.M., Bohlen, J.M., et al.: The diffusion process. Agricultural Experiment Station, Iowa State College (1957)
[6] Beaver, A.L.: Getting a fix on cocaine sentencing policy: reforming the sentencing scheme of the anti-drug abuse act of 1986. Fordham L. Rev. 78, 2531 (2009)
[7] Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of political Economy pp. 992–1026 (1992)
[8] Brennan Center for Justice: Ice extreme vetting initiative: A resource page, https://www.brennancenter.org/analysis/ice-extreme-vetting-initiative-resource-page. Last retrieved May 15, 2018.
[9] Buterin, V., et al.: A next-generation smart contract and decentralized application platform. white paper (2014)
[10] Cheng, R., Zhang, F., Kos, J., He, W., Hynes, N., Johnson, N., Juels, A., Miller, A., Song, D.: Ekiden: A platform for confidentiality-preserving, trustworthy, and performant smart contract execution. arXiv preprint arXiv:1804.05141 (2018)
[11] Diffie, W., Landau, S.: Privacy on the Line: The Politics of Wiretapping and Encryption. The MIT Press, updated and expanded edition edn. (Feb 2010), http://www.worldcat.org/isbn/0262514001
[12] Dingledine, R., Syverson, P.F.: Privacy enhancing technologies. Springer (2003)
[13] Downs, A.: An economic theory of political action in a democracy. The journal of political economy pp. 135–150 (1957)
[14] Dwork, C., McSherry, F., Nissim, K., Smith, A.: Calibrating noise to sensitivity in private data analysis. In: Theory of Cryptography Conference. pp. 265–284. Springer (2006)
[15] Dwork, C., Roth, A., et al.: The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science 9(3–4), 211–407 (2014)
[16] Esmaili, S.: Searching for a needle in a haystack: The constitutionality of police dna dragnets. Chi.-Kent L. Rev. 82, 495 (2007)
[17] Gangadharan, S.P.: Digital inclusion and data profiling. First Monday 17(5) (2012)
[18] Ginkel, J., Smith, A.: So you say you want a revolution a game theoretic explanation of revolution in repressive regimes. Journal of Conflict Resolution 43(3), 291–316 (1999)
[19] Goh, B.: Prosperity and security: A political economy model of internet surveillance (2015)
[20] Hetherington, M.J., Weiler, J.D.: Authoritarianism and polarization in American politics. Cambridge University Press (2009)
[21] Kuran, T.: Sparks and prairie fires: A theory of unanticipated political revolution. Public choice 61(1), 41–74 (1989)
[22] Landau, S.: National Security on the Line. Social Science Research Network Working Paper Series (Apr 2009), http://ssrn.com/abstract=1166155
[23] Landau, S.: Making sense from snowden: What's significant in the nsa surveillance revelations. IEEE Security and Privacy 11(4), 54–63 (2013)
[24] Landau, S.: Making sense of snowden, part ii : What's significant in the nsa revelations. IEEE Security and Privacy 12(1), 62–64 (2014)
[25] Laskowski, P., Johnson, B., Maillart, T., Chuang, J.: Government surveillance and incentives to abuse power. In: Proceedings (online) of the 13th Workshop on the Economics of Information Security (WEIS) (2014)
[26] Layman, G.C., Carsey, T.M., Horowitz, J.M.: Party polarization in american politics: Characteristics, causes, and consequences. Annu. Rev. Polit. Sci. 9, 83–110 (2006)
[27] Lohmann, S.: The dynamics of informational cascades: The monday demonstrations in leipzig, east germany, 1989–91. World politics 47(01), 42–101 (1994)
[28] Lyon, D.: Surveillance as social sorting: privacy, risk, and digital discrimination. Psychology Press (2002)
[29] Marwick, A.E., danah boyd: Privacy at the margins. International Journal of Communication 12, 9 (2018)
[30] McSherry, F., Talwar, K.: Mechanism design via differential privacy. In: Foundations of Computer Science, 2007. FOCS'07. 48th Annual IEEE Symposium on. pp. 94–103. IEEE (2007)
[31] Novak, K.J.: Disparity and racial profiling in traffic enforcement. Police Quarterly 7(1), 65–96 (2004)
[32] Phillips, T.: China testing facial-recognition surveillance system in xinjiang – report | world news | the guardian, https://www.theguardian.com/world/2018/jan/18/china-testing-facial-recognition-surveillance-system-in-xinjiang-report. Last retrieved May 15, 2018.
[33] Poole, K.T., Rosenthal, H.: The polarization of american politics. The Journal of Politics 46(04), 1061–1079 (1984)
[34] Press, O.: Hotelling (1929),? stability in competition? Economic Journal 39
[35] Re, R.M.: The due process exclusionary rule. Harvard Law Review pp. 1885–1966 (2014)
[36] Schneier, B.: It for oppression. Security & Privacy, IEEE 11(2), 96–96 (2013)
[37] Shiller, R.J.: Conversation, information, and herd behavior. The American Economic Review 85(2), 181–185 (1995)
[38] Slobogin, C.: Government dragnets. Law and Contemporary Problems 73(3), 107–143 (2010)
[39] Smith, M.R., Petrocelli, M.: Racial profiling? a multivariate analysis of police traffic stop data. Police Quarterly 4(1), 4–27 (2001)
[40] Soska, K., Christin, N.: Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. In: USENIX Security Symposium. pp. 33–48 (2015)
[41] Sotomayor, S.: Utah v strieff, 579 u.s., 136 s. ct. 2056 (2016) (sotomayor dissenting)
[42] Sweeney, L.: k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10(05), 557–570 (2002)
[43] Union, A.C.L.: Frequently asked questions about the arizona racial profiling law, https://www.aclu.org/other/frequently-asked-questions-about-arizona-racial-profiling-law. Last retrieved May 15, 2018.
[44] Wang, Y., Kosinski, M.: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of personality and social psychology 114(2), 246 (2018)