Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

______________________________ _______________

Griffin Edwards Date

Adverse Consequences of Tort and Statutory Law

By

Griffin Edwards
Doctor of Philosophy

Economics

Paul H. Rubin
Advisor

Joanna Shepherd Bailey
Committee Member

David Frisvold
Committee Member

Sara Markowitz
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Adverse Consequences of Tort and Statutory Law

By

Griffin Edwards
M.A. Economics, Emory 2010
B.S., Economics, Brigham Young University – Idaho, 2007

Advisor: Paul H. Rubin, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2011

Abstract

Adverse Consequences of Tort and Statutory Law
By Griffin Edwards

The effect of a law, whether it be through legislatures or courts, is often difficult to identify given unintended consequences that arise. One example is the seminal ruling of *Tarasoff v. Regents* that enacted a duty that required mental health providers to warn potential victims of any real threat to life made by a patient. Using a fixed effects model and exploiting the variation in the timing and style of duty to warn laws across states, I find that mandatory duty to warn laws cause an increase in homicides of 5%. These results are robust to model specifications, falsification tests, and help to clarify the true effect of state duty to warn laws.

Another is the ruling in *Chevron v. Natural Resources Defense Council*. Previous research has found both theoretically and empirically that *Chevron* favors agencies and their interpretation of statutes, but the magnitude of *Chevron*'s impact remains unclear due to possible selection issues biasing the post-*Chevron* world. I account for the possibility that incentives change both to the challenger of an agency and the agency itself post-*Chevron* by estimating a break in the trend of agency deference on the date *Chevron* was decided. This allows me to exploit the exogenous cases that were pending when *Chevron* was decided while still employing the full sample of rulings. Both parametric and nonparametric specifications of the trend in agency deference suggest that *Chevron* increased agency deference by about 20 percentage points meaning that agency will win a challenge around 80% of the time.

The third law on which I focus deals with an organized criminal firm's ability to extract monopoly rents from victim firms. Using a U.S. state panel and data on federal racketeering cases charged, I find that all else equal, a 0.1 percentage point increase in the amount of non-English speakers in a state will increase the expected number of racketeering cases per state per year by 0.8. This is weakly supported by the fact that states with fewer small businesses, and thus a higher probability of earning monopoly rents, experience less racketeering activity.

Adverse Consequences of Tort and Statutory Law

By

Griffin Edwards
M.A. Economics, Emory 2010
B.S., Economics, Brigham Young University – Idaho, 2007

Advisor: Paul H. Rubin, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Economics
2011

Table of Contents

Tables and Figures

Doing Their Duty: An empirical analysis of the unintended effect of *Tarasoff v Regents* on homicidal activity

In its landmark ruling, *Tarasoff v Regents*[1] (hereafter *Tarasoff)* set the standard of duty required of a mental health professional. According to *Tarasoff*, when a patient expresses a credible threat to life, the mental health professional incurs a legal duty to warn the potential victim. Contrary to the typical notion that a legal duty cannot be owed to a third party, *Tarasoff* stands as not only an exception to the rule of duty, but also as staple in tort law. Virtually every Torts class discusses *Tarasoff* and its implications. And though it is frequently discussed, its effect remains untested. Since it is a state level ruling, most states adopted some sort of law in the years following the decision similar to *Tarasoff*.

At the onset of *Tarasoff*, the duty owed to third parties became the subject of a "cottage industry of commentary" (Perlin 1992) in both the legal and mental health services communities. Since *Tarasoff*, both legal scholars and mental health professionals have argued that it, " . . . would lead to more danger by discouraging patients from seeking treatment and/or chilling patients' willingness to discuss issues of violence with their therapists" and that patients at most risk of dangerous activity will miss out on necessary counseling due to the costs mental health professionals incur while counseling risky patients (Stone 1976, Fliszar 2002, Ginsberg 2007). Ackerman (1976) predicted the result of *Tarasoff* to be the "end of effective psychotherapy". The question of *Tarasoff*'s effectiveness in deterring violence, and specifically homicides, remains unanswered empirically.

---

[1] *Tarasoff v. Board of Regents of the Universities of California*, 551 P.2d 334 (Cal. 1976).

This analysis contributes both to our general understanding of the role of confidentiality as well as the specific effect *Tarasoff* has had by codifying each state's *Tarasoff* law and employing a fixed effects (FE) model to estimate *Tarasoff*'s effect on homicides in the United States. Comparing states before and after the law suggests that the presence of duty to warn laws causes an increase in state homicide rates by about 5%.

2. Background

Crime and Mental Health

Economists and other researchers have tried to explain why crime rose steadily in the 1980's and abruptly fell in the early 1990's. Based on economic theory, the reasons for the sudden changes in crime are expansive and far reaching (Levitt 2004). Some research suggests that abortion (Donohue and Levitt 2001; Joyce 2009), gun control (Ayres and Donohue 2003; Black and Nagin 1998), and the death penalty (Donohue and Wolfers 2009; Katz et. al. 2003) may have contributed to the rise and fall of crime. A recently explored theory suggests a causal relationship between mental illness and crime. Either as the victim (Sliver et al. 2008; Teplin et al. 2005; Choe et al. 2008), or perpetrator (Link et al. 1995; Nestor 2002), previous literature suggests a link between sufferers of mental illness and crime.

Marcotte and Markowitz (2009) propose that the decline in crime was due in part to the widespread administration of psycho-pharmaceuticals. In fact, they report that 22% of inmates surveyed were found to suffer from some sort of serious mental illness. In an overview study, Choe et al. (2008) report that some studies have found that almost 50% of sampled mental health patients have a higher propensity towards violence. Swanson et al. (1990) finds that the mentally ill are 4 to 5 times more likely to be violent than the general population. The results of these studies suggest that violent acts in the United States, including homicide, might be disproportionately committed by the mentally ill. The findings presented in this analysis further strengthen the argument of a causal link between mental health and crime.

The History of Tarasoff

At the core of tort theory is the existence of duty of one person to another. In order to be found liable, a tortfeasor must have owed a duty of care to the victim and breached that duty. The duty doctrine inspires would-be tortfeasors to act in a manner that is better for society than the tortfeasor might have acted otherwise (Hylton 2006). Under classic tort theory, therapists counseling dangerous patients would owe no duty to potential victims or law enforcement agencies. However, warnings to potential victims based on information revealed during counseling could be socially beneficial. Thus, courts have ruled that an exception to the duty standard is the situation of a hostile patient. This exception of third party duty was first established in *Tarasoff*.

After unsuccessful attempts to court Tatiana Tarasoff, Prosenjit Poddar, a graduate student at the University of California at Berkeley in the late 1960's, sought professional help from a psychologist for depression. While receiving counseling, Poddar admitted desires to kill Tarasoff. Poddar's psychologist had Poddar detained temporarily, but at the discretion of the supervising psychologist, Poddar was released. Neither Tarasoff nor her family were ever made aware of Poddar's intentions.[2] Later, Poddar successfully carried out his plan and murdered Tarasoff. The family of Tarasoff

[2] Though this may have not been possible since Tarasoff was in Brazil at the time the threats were made.

sued the hospital, the psychologist, and the superior stating that a professional duty should exist to protect third parties from imminent harm.

In a landmark decision, the Supreme Court of California ruled that while traditionally no duty is afforded to a third party, in the case of mental healthcare professionals, a duty to warn a third party exists under certain circumstances, and the failure to warn is cause for suit. In its opinion, the Supreme Court of California stated, "When a therapist determines . . . that his patient presents a serious danger of violence to another, he incurs an obligation to use reasonable care to protect the intended victim against such danger."[3] Two years later in *Thompson v. County of Alameda*,[4] the Supreme Court of California determined that as long as the victim or class of victims is clearly defined, and the threat is substantial, the therapeutic professional holds a duty to, " . . . warn the intended victim or others likely to apprise the victim of the danger, to notify the police, or to take whatever other steps are reasonably necessary under the circumstances."[5]

In subsequent years, dramatic changes occurred in both the law associated with therapeutic professionals and the way they conducted business (Wise 1976; Givelber et al. 1980). Courts across the country used the ruling of *Tarasoff* as a basis for creating a duty by mental health professionals to warn third parties of imminent harm. In addition, a variety of states codified these case law rulings into statutory law defined in each state's legal code.

---

[3] *Tarasoff* 551 P.2d at 431.
[4] *Thompson v. County of Alameda*, 614 P.2d 728 (1980).
[5] *Tarasoff* 551 P.2d at 431.

Though almost all states took a stand with respect to third party duty in response to the *Tarasoff* ruling, not all states hold therapists[6] to the same standard. There are essentially four elements that play a role in each state's stance relative to *Tarasoff*. Those elements are: professionals named, standard of threat, standard of victim, and what party is entitled to be informed. Important distinctions exist between the types of professionals named in each state. While some states state specifically each professional potentially liable under *Tarasoff*, others define it more broadly. The standard of threat also varies by state, but in general, in order for a duty to exist, the threat made by the patient must be "clear and immediate"[7] and a "threat of serious physical harm"[8] and the victim must be readily identifiable. For example, the Arizona statute states that the health provider will be liable if:

> *The patient has communicated to the mental health provider an explicit threat of imminent serious physical harm or death to a clearly identified or identifiable victim or victims, and the patient has the apparent intent and ability to carry out such threat.*[9]

Likewise, the duty to warn statute for Utah states that a therapist will be held liable for the actions of a patient if the,

> . . . *client or patient communicated to the therapist an actual threat of physical violence against a clearly identified or reasonably identifiable victim.*[10]

[6] Throughout this paper I use words therapist, psychologist, mental health professional, therapeutic professional, etc interchangeably for stylistic purposes admitting that they are quite different in relation to the law (*See* Edwards 2010 for further explanation).
[7] Fla. Stat. § 491.0147
[8] Alaska Stat. § 08.86.200
[9] A.R.S. § 36-517.02
[10] Utah Code Ann. § 78B-3-502

Under state statutory duty to warn laws, the therapeutic professional is still only liable if the patient makes a credible threat and the professional does not take the proper action in providing warning to the appropriate persons. Of these states that have codified duty to warn laws, therapeutic practitioners can avoid liability by notifying proper authorities and the victim or victims named.

Tables 1 and 2 show that from 1981 to 2003 46% of the state\year observations have some sort of mandatory duty to warn law, 18% have discretionary duty to warn laws, and the remaining have no law (Edwards 2010). Five states have no case or statutory law on the duty to warn doctrine. Similarly, four states have suggested an adoption of *Tarasoff* through case law by expressing desire to rule in favor of *Tarasoff* when the correct fact pattern is presented.[11] The only state to reject outright the *Tarasoff* ruling is Virginia.[12] By 1986, about half of all states had passed some sort of *Tarasoff* ruling.

The remaining states constitute the large minority ruling which in Table 1 is classified as a discretionary duty to warn. These eleven states have adopted a policy that allows the therapeutic professional to use her best judgment in deciding to report threats of harm. These statutes are formed as part of legal bars to break patient/doctor confidentiality privileges. In general, therapeutic professionals cannot divulge conversations had with a patient. However, each state has written statutes to allow confidentiality breaches which govern the ethical code of each state's respective mental health professional associations. So while the applicable professional association may

---

[11] *See Morton v. Prescott* 564 So. 2d 1188 (1985) (Alabama); *Bradley Ctr. Inc. v. Wessner* 296 S.E.2d 693 (1982) (Georgia); *Lee v. Corregedore* 925 P.2d 324 (1996) (Hawaii); *Currie v. United States* 836 F.2d 209 (1997) (North Carolina).

[12] *Nasser v. Parker*, 455 S.E.2d 502 (1995). *See also* Edwards (2010) for a description of New Mexico's unclear stance on duty to warn laws.

allow for a breach in confidentiality, the stronger incentive will be to comply with the current state of law. In these states with discretionary duties to warn, one acceptable reason to break the patient/doctor confidentiality agreement is a serious threat to life. This, in result, is a much looser standard than a mandatory duty to warn. A mandatory duty law requires the professional to warn while the discretionary duty simply protects the professional from breach of confidentiality if she chooses to inform a third party. For example, Connecticut has established that:

> *" . . . all communications shall be privileged and a psychologist shall not disclose any such communications unless . . . the psychologist believes in good faith that there is risk of imminent personal injury to the person or to other individuals . . . "*[13]

Similarly, Florida's statute states:

> *"Any communication between any person licensed or certified under this chapter and her or his patient or client shall be confidential. This secrecy may be waived under the following conditions . . . When, in the clinical judgment of the person licensed or certified under this chapter, there is a clear and immediate probability of physical harm to the patient or client . . . "*[14]

In *Thepar v. Zuzuka*[15], the Supreme Court of Texas explained that, "The statute . . . permits . . . disclosures but does not require them," reinforcing the notion that a discretionary duty makes warning permissible, but not required.

Tarasoff *and Homicides*

---

[13] Conn. Gen. Stat. § 52-146c

[14] Fla. Stat. § 491.0147

[15] *Thepar v. Zuzuka*, 994 S.W.2d 635 (1999).

The remaining question then is the extent to which *Tarasoff* could affect homicides. As mentioned previously, there is a vast literature that suggests that crimes, including homicides, are disproportionately committed by the mentally ill (Swanson et al, 1990; Choe et al. 2008; Taylor & Gunn 1999). In fact, in the United Kingdom about 10% of all homicides are estimated to have been committed by somebody with recent contact to mental health services (Swinson et al. 2007). Given the effect many purport *Tarasoff* to have on the treatment of the mentally ill, it seems reasonable to assume that *Tarasoff* might affect homicides in ways other than deterring homicides by third party warnings.

There are a number of channels by which duty to warn laws might enable homicides, but in order for these laws to have an effect, they must be known. First, from the point of view of the mental health services provider, we would anticipate that therapists have the incentive to learn about the law. Survey evidence provided by Givelber, et al. (1980) found that 86% of psychologists surveyed were aware of the *Tarasoff* laws. In addition, it seems reasonable to think that psychologists will learn about laws that increase their personal liability, and there is also some evidence to suggest that mental health professionals invoke the *Tarasoff* duty to warn.[16] Also, psychologists will likely have incentive to convey the information about the laws to their patients (Klinka 2009[17]). In an attempt to avoid liability, a therapist will likely warn the patient ex ante, both verbally and by signed contract, of the law to allow the patient to monitor what is divulged. The evidence found in previous research suggests that this happens at least in part (Givelber et al. 1980; Rosenhan et al. 1993).

---

[16] *See* Roberts et al. (2002)
[17] *Spec*. note 213

Under this heightened state of awareness, many fear that patients will be more reluctant to divulge their most violent thoughts that then go untreated. Wise (1978) found that 25% of therapists reported greater patient reluctance to discuss violent thoughts, and more recently, Rosenhan et al. (1993) report that 60% of therapists felt that patients were at least somewhat more reluctant to discuss sensitive information. In addition, Givelber et al. (1980) report that psychologists were 30% more likely post *Tarasoff* to commit patients involuntarily to the hospital. The increased awareness of the law, coupled with the increased threat of involuntary hospitalization has likely discouraged patient discourse.

It is important to note, however, that while these studies are informative and give some insight into the mechanism by which *Tarasoff* laws might affect homicides, the sample selection may bias results, as psychologists may have incentive to overestimate *Tarasoff*'s negative effects in an effort to give evidence to overturn the law that heightens their own liability. Nonetheless, they do provide some interesting insight into the adverse effect *Tarasoff* might have.

Treatment of violent thoughts might also be discouraged through patient knowledge of the law and the choice to forgo mental health services all together (Wise 1976; Rosenhan et al. 1998; Klinka 2009). Another channel of this adverse effect could be in the selection of patients a therapist is willing to treat (Borum and Reddy 2001). The risk of liability might be sufficiently great as to deter doctors from even counseling potentially "risky" patients (Klinka 2009; Harmon 2008).

In sum, there is reason to believe that state duty warn laws are known by the mental health professional and that the information is passed to the patient. Many fear and there is some evidence to suggest that the patient response will be to suppress violent thoughts which leaves the homicidal tendency both untreated and unknown by the therapist. As a result, there are multiple channels by which homicidal tendencies could potentially go untreated.[18] The intent of the law is to stop patients from committing heinous acts of violence by warning the potential victim. If this were true, homicides should decrease. It could also be that duty to warn laws prohibit patients from getting necessary mental counseling which may cause homicides to increase. Given the ambiguity of the effects, the question is best answered empirically.

This analysis contributes to the related literature by offering some empirical evidence of *Tarasoff*'s effect and helps answer many of the theoretical questions raised by the *Tarasoff* duty over the past 20 years as well as explaining the rise and fall in crime over the last twenty years. In addition, it builds on recent work that links mental health conditions to crime in the United States (Marcotte and Markowitz 2009).

3. Model

A fixed effects (FE) model is used to estimate the effect of state *Tarasoff* laws on homicides. The panel nature of the data allows for the use of panel techniques that control for a lot of the unobserved time-invariant heterogeneity across states as well as unobserved national trends. This technique is particularly attractive in this setting as the laws vary both by time of adoption and style of law. Figures 1 and 2 display graphically both the variation in style of law and timing, respectively. In addition, the available window of data (1980-2003) captures some sort of law change in about 80% of the states.

[18] The question of which channel prevails (patients concealing homicidal tendencies versus doctors not treating at risk patients) is an interesting one, but not addressed in this paper. It suffices to say that as long as every channel does not systematically fail and at the same time, *Tarasoff* laws will, to some degree, encourage homicides.

Those states that do not vary within the sample window are included nonetheless in the analysis because while they do not contribute to the estimation of the coefficients of interest, they do still add value in explaining the overall variation in homicides. The most basic identification is:

$$h_{it} = \alpha m_{it} + \beta d_{it} + X_{it} + s_i + y_t + e_{it} \tag{1}$$

where *h* is the natural log of homicides per 100,000 that vary by state *i* and time *t*, *X* is a matrix of covariates,[19] *s* and *y* are state and year dummies, *e* is the error term, *m* is a binary variable that takes on the value one when state *i* in time *t* has a mandatory duty to warn, and zero otherwise. *d* is a binary variable that takes on the value one when state *i* in time *t* has a discretionary duty to warn and zero otherwise. Table 2 shows the summary statistics of the state\year cells. States with discretionary duty to warn laws encompass 17% of the sample while state\years with mandatory laws count for about 40% of the sample.

Identification Issues

There are several issues that require some thought while modeling the effect of *Tarasoff*. To start, court decisions and legislation do not usually pass right at the end or beginning of a year. This poses a problem because my preferred measure of homicides is reported yearly. This means I have to make a decision whether to count the law that passes in June of 1988 as beginning in 1988 (and thus overstating the life of the law) or 1989 (and understate the life of the law). [20] To reduce the chance of introducing bias into

[19] State controls are explained in the Data section.

[20] For instance, if Wisconsin passed its law in June of 1988, the law is counted as existing for the entire year of 1988. Because this tends to over state the length of the law, *t+1* law variables are included as alternate specifications. So in Wisconsin, the *t+1* law variable would count the law as starting in 1989. No significant difference results.

the model, I estimate equation (1), and all subsequent equations, both as explained in equation (1) and as:

$$h_{it} = \alpha_1 mc_{it+1} + \beta d_{it+1} + X_{it} + s_i + y_t + e_{it} \tag{2}$$

where *t+1* measures the second year since enactment, or the first full calendar year or enactment. This should create two types of estimates: equation (1) over counts the law's lifespan and equation (2) will under-count it. I can then compare the two estimates and see if they differ greatly. The results suggest that timing in this sense is not an issue.

There is another timing issue however. For law originating from judges and courts, it might be unclear at what point in the trial or appeal the law begins to be effective. It seems reasonable to think that people begin to react to the law after the highest court rules, but in case they do not, I estimate equations (1) through (4) with the timing of the laws adjusted to account for lower court rulings.

Another major methodological concern is how to correctly exploit the variation in state duty to warn laws. There might be some reason to believe that the origin of the law matters. For instance, laws created by legislatures might be perceived as more firm than those created by the court system. Conceivably, therapeutic professionals could be more aware, or respond more intently, to state laws passed by legislatures than decisions made in the state court system. To address this, models (1) and (2) are expanded to designate the source or type of law.

$$h_{it} = \alpha_1 mc_{it} + \alpha_2 ms_{it} + \alpha_3 ml_{it} + \beta d_{it} + X_{it} + s_i + y_t + e_{it} \tag{3}$$

$$h_{it} = \alpha_1 mc_{it+1} + \alpha_2 ms_{it+1} + \alpha_3 ml_{it+1} + \beta d_{it+1} + X_{it} + s_i + y_t + e_{it} \tag{4}$$

The variable *m* in equation (1) is expanded to three binary variables that incorporate the governing body deciding the law. The variable *mc* measures the effect of mandatory duty laws decided by the state court system. These are states that, when presented with evidence similar to *Tarasoff*, have ruled that a common law duty to warn exists. *ms* measures the effect of mandatory duty laws enacted by state statutory law, and *ml* measures the effect of state judicial ruling that has dictated a duty to warn will be enacted when the fact pattern is presented to the courts.

This situation occurs normally when the question presented before the court is something related to, but not exactly, the issue presented in *Tarasoff*. If, for example, a state court comes to a ruling about a psychologist duty to report child abuse, the court usually discusses a *Tarasoff* duty to warn.[21] An advantage to this model specification is that it allows states to switch from common law to statutory law as a state codifies existing common law doctrine.[22]

Perhaps the biggest threat to model validity is the possibility of some form of endogeneity biasing the results. There is little evidence to suggest that there might be an endogenous factor driving both homicide rates and the passing of these laws since they originated either as an exception to the rule of duty or as an issue of mental health. Even if these laws were created in response to trending state-levels of crime or the mental health status of the state, I can control for these trends by including multiple measures of crime and their lags. Another source of potential endogeneity is not only the mental health status of a state, which I hope to capture by including state suicide rates, but more precisely how legislatures might perceive the mental health status of a state. To address this, I include a set of variables that capture the uptake of state mandates for the

[21] *see* “Alabama” and “Georgia” in Edwards (2010)
[22] *Compare Tarasoff* 551 P.2d 334 *with* Cal. Civ. Code § 43.92 (California); *Naidu v. Laird*, 539 A.2d 1064 *with* 16 Del. C. § 5402.

availability of mental health insurance.[23] This should capture how legislatures view mental health. In addition, with the analysis of any law, there is some concern that the law is an artifact of the surrounding political environment. I can control for this potentially biasing effect by the inclusion of state political characteristics as well as state judge appointment and retention methods.[24]

Though unlikely, there might be reason to believe that these laws, especially the statutory laws, were passed in response to some underlying trend in crime. Thus the estimates I obtain are not predicting the effect of *Tarasoff* but rather just capturing some underlying trend in crime driven by the timing of the laws. To identify this possibility, I predict a series of models where the dependent variable is a common measure of crime that would be indicative to any underlying state trend in crime. If my estimates of the effect of these laws are merely capturing an underlying trend in crime, I should estimate a similar effect across different measures in crime. To test this, I run multiple regressions where the dependent variable is a unique measure of the level of crime in a state and report that duty to warn laws do not appear to have an effect on different measures of crime.

An advantage I have in weeding out endogeneity is that I am able to distinguish the source of the law. The traditional notion of policy endogeneity comes from law created by legislatures in response to something they observe, but there is some evidence to suggest that these laws created by courts are not susceptible to the same type of

[23] *See* Klick and Markowitz (2007)
[24] The full list of covariates can be seen in the data section as well as in Table 2.

endogeneity (Shepherd 2009).[25] Though these laws could potentially have adverse consequences to criminal law, they are a matter of civil law. So a judge with a predisposition towards being "tough on crime" or "sympathetic to the victim" is going to have better, more direct avenues to affect crime rates through criminal law cases. Given the large quantity of cases presented to judges[26], it seems reasonable to think that appeals and state Supreme Court judges and justices who care a lot about affecting crime will simply substitute away from civil to criminal cases. This is evidenced by the opinions explaining the rationale behind the ruling published in each case. Of all the published opinions, the word *crime* is only mentioned four times—only two of which are original prose from the justice.[27]

Given that, it will be useful to compare the estimated effect of *Tarasoff* laws by the court of origination. If laws generated by courts are probably not endogenous, and I estimate similar results between statutory and common law, this might suggest exogenous state statutes. Any remaining unobserved factor that influences homicides and be correlated with the laws will hopefully get picked up by state effects, year effects, and state-specific time trends.[28]

With the inclusion of state effects, this makes each estimate a within-state estimation of the impact of *Tarasoff*—meaning the treated state is compared with its pre-

---

[25] Shepherd (2009) finds that state appointed judges tend to side more frequently with litigants from government branches around the time of reappointment. Even this type of bias would not be present in the context of *Tarasoff* because the state is never a litigant in a duty to warn case.

[26] *See* Huang (2010)

[27] *See* 77 Ohio St. 3d 284 at 1338. The legal realist might argue that what the published opinion says is not necessarily indicative of the courts motivation for ruling (i.e. a judge really wants to control crime through *Tarasoff* but publishes in the opinion the duty of care rationale). At least in the case of Ohio, Justice Stratton does not mask her acknowledgement that these might have an effect on crime. This suggests that if more judges considered homicides when making the ruling, they could just say so in the opinion. The fact that they don't suggests it wasn't a consideration.

[28] It could be the case that state policy makers react with legislation to a high profile murder where psychological counseling should have played a role. As long as legislators consistently respond it furthers the exogeneity of the law since high profile murders are probably random.

treated self. There is evidence to suggest that this provides a balanced comparison.[29] FE estimation does, however, have serious potential threats to unbiased estimation. Chief among those concerns is the possibility of underestimated standard errors caused by failure to identify serial correlation, non-random treatment (or non-random assignment of duty laws) and the usual binary nature of the treatment variable (Bertrand et al. 2004). Bertrand et al. (2004) show that clustered standard errors at the state level correct for serial correlation by allowing for correlation in the error terms within a state across time. This resolves the issue of understated error terms by, in many cases, conservatively overstating them in allowing the error terms to be robust and correlated. In this analysis, I use a test developed by Wooldridge (2002) to test for serial correlation in the error terms and correct as needed.

Data

Dependent Variables

The data on homicide rates comes from a variety of places. The preferred measure of homicides comes from the WISAQARS database compiled by the CDC from the National Center for Health Statistics (NCHS). It spans from 1981 to 2003 and captures the timing of the majority of legal changes in law. NCHS data comes from collection of death certificate information and contains information on nearly all deceased persons. Since this seems to be the most complete database of homicides, I use the NCHS measure of homicides as my preferred specification. To provide some robustness to the NCHS data set, I employ yearly homicide data from the Uniform Crime Report (UCR). The UCR provides a useful check to ensure the results are not an artifact

[29] Further discussion of balance in the data with a distributional analysis of the covariates is available in the Appendix.

of just one measure of homicides. In addition, the UCR reports more information about the nature of the relationship between the victim and perpetrator which will be useful in this analysis.

The UCR publishes a Supplemental Homicide Report (SHR) that consists of incident-level homicide reports. SHR data is collected from volunteer participation of over 17,000 law enforcement agencies across the United States but has several shortcomings (Levitt 1998; Marcotte and Markowitz 2009; Katz et al. 2003). Levitt (1998) outlines when self-reporting will lead to bias in the UCR and the example of how police force size will either encourage or discourage self-reporting. In addition, the UCR accuracy might suffer from heterogeneity across reporting agencies in reporting practices and technology. One problem with the UCR database is how to interpret a zero count of homicides. It is unclear in many cases whether a zero means no homicides or simply missing data. Stevenson and Wolfers (2006) point out that in at least a couple of cases, a zero count definitely means no reporting.[30] There is reason to believe though that a zero count on homicides could either mean no homicides or very little homicides or it could mean that resources are so scarce or homicides are so high, that allocation to reporting statistics is undesirable. Nonetheless, Joyce (2009) reports however that the SHR accounts for 90% of homicides. Given the ambiguity behind the rationale for zero counts of homicides, I omit those state/year observations from the analysis. Despite its shortcomings, the SHR database is particularly attractive to this study because information is reported on the relationship between the victim and perpetrator. Given the *Tarasoff* standard that the potential victim be readily identifiable, whatever effect we

[30] They refer to Illinois that for portions of the 1980's reported no homicides which is clearly false.

observe for the general population of homicides should be larger when stranger homicides are omitted.

The ideal pool of homicides would be only those affected by duty to warn laws. This however would be difficult to define because any relationship where the mental health patient could identify the victim would potentially be affected by *Tarasoff* laws. Because of this, I employ a strategy similar to Stevenson and Wolfers (2006) and narrow the pool of homicides by taking out murders by strangers. This will probably include some murders on which state duty to warn laws have no effect due to the imperfect nature of identifying the relationship between the two parties.[31] It will be an improvement on the entire sample of homicides and provide a useful comparison to the results found with the more complete NCHS database of homicides.

As discussed earlier, in addition to the NCHS measure of homicides and the SHR measure of non-stranger homicides, I include as dependent variables multiple measures of crime to see if state duty to warn laws explain their variation. To test this, I run multiple regressions where the dependent variable is a unique measure of the level of crime in a state. Those measures are the natural logs of the auto theft rate, larceny rate, robbery rate, and manslaughter rate. All these variables were collected from the UCR database and suffer from the same shortcomings inherent in the UCR but still provide some interesting evidence about the exogeneity of *Tarasoff* laws.

Independent Variables

One major difficulty in measuring the effect of any law passed in the 1980's and 90's is the task of correctly controlling for all the observed factors that might have an effect on homicides. The independent variables of interest are the coded law dummy

[31] *See* "Uniform Crime Reporting Handbook" available at: www.fbi.gov/ucr/ucr.htm

variables explained previously. The state duty to warn laws were compiled and coded by the author and a full description of each state and relevant court cases can be found in Edwards (2010).

It is unclear whether many demographic controls typically attributed as causing changes in crime actually predict levels of crime (Zimring 2007). Nevertheless, in an effort to be exhaustive in all possible predictors of homicides, I include state demographic characteristics that consist of median age, percent of the population that is black, percent of the population that is male, and real median income. These all come from census micro-level data compiled by the IPUMS website (Ruggles et al. 2009). Economic conditions have been linked to mortality (Ruhm 2000), so state unemployment rates collected from the Bureau of Labor Statistics are included.

Alcohol and drug use have long been associated with violence (Markowitz 2005). To control for the role alcohol plays on homicides, I include the real state beer excise tax reported by the *Brewers Almanac*. Another major concern is the prevalence of drugs, specifically crack cocaine. Unfortunately, good national measures of crack usage are scarce (Levitt 2004). Some of the effect of crack should be captured by the year fixed effects, but many inner-city homicides in the 1980's were likely caused by the prevalence of crack. To account for this, I use two data sources. First, I include the number of Drug Enforcement Agency (DEA) arrests made regarding cocaine in the model estimation.[32] Data on DEA arrests only exists between 1993 and 2006. For the remaining years of the sample, I use drug related deaths reported by the Morality Detail File and create per capita index of crack usage.

[32] Obtained from the Federal Justice Statistics Resource Center (fjsrc.urban.org)

With the creation of any law, politics plays a large role (Shepherd 2009). To capture that effect, the political controls I use are the share of democrats in the state senate and house, the political party of the governor, the selection and retention methods of state Supreme Court judges, and real per capita judicial expenditures. These political controls come from Shepherd (2009). Other good measures of the political atmosphere of a state are state enacted tort reform laws. With the majority of state tort reform laws passing within the window of this study (Avraham 2006), including tort reform laws will hopefully contribute in controlling for how a state perceives plaintiffs and defendants.[33]

As discussed earlier, it need be clear that these laws were not passed in response to trends in crime which would bias the results. To show this, I run multiple regressions with different measures of crime as the dependent variable. Another approach is to simply control for the underlying trends in crime. To do this I include in each model state rates of assault, auto theft, and robbery which all come from the UCR database. In addition, Levitt (1998) suggests that the size of the policing agency might explain homicides, especially among self-reported data. To control for this, I include per capita real state police expenditures collected from the Bureau of Justice Statistics.

These laws also might be endogenous if policy makers pass them in response to trends in the mental health status of a state. This is not likely the case because the mental health status of each state has experienced very little variation over time (Marcotte & Markowitz 2009). Nonetheless, I am able to control for the mental health status of a state by including the state suicide rate. Suicide rates have long been seen as the measure of the mental health status of a state (Klick & Markowitz 2006). In an effort to mitigate any endogeneity, I include the state suicide rate collected from the NCHS. There could be,

[33] All tort reform laws are included in every regression and can be reviewed in Avraham (2006).

however, a difference between the actual mental health status of a state, and how policy makers perceive the mental health of a state. To account for this, I include in each model a control for the adoption of state mandated employer offered mental health insurance coverage outlined in Klick and Markowitz (2006). This will hopefully capture any unobserved discrepancy between trends in mental health and perceived trends in mental health.

Another variable that might matter in these model specifications is the number of psychologists per state. Potentially, psychologists might migrate to states with more psychologist-friendly laws. Marcotte and Markowitz (2009) find this to not be the case, so it is not of great concern.[34] Nonetheless, I include the rate of psychologists per 100,000 of the state population gathered from the Area Resource File in all models.

4. Results

Estimation

Table 3 reports the estimations results of the effect of state *Tarasoff* laws on the natural log of NCHS homicide rates. Each model estimate is weighted by the square root of the state population, includes state and year fixed effects, state-specific time trends, demographic and economic controls, and robust standard errors clustered at the state level to allow for correlation of the error term within state across time.[35] Columns (1) and (3) of Table 3 estimate equations (1) and (2), respectively, while columns (2) and (4) are estimates of equations (3) and (4) respectively. Table 4 replicates Table 3 using UCR-SHR data on homicides by non-strangers.

---

[34] Even if there were psychologist migration, it would be captured, at least in part, by state-specific time trends.

[35] The test developed by Wooldridge (2002) suggests serial correlation in the error. State level clustering is used to control for this.

Columns (1) and (3) of Table 3 report coefficients in the case where each state is coded simply as having a duty law or discretionary law, regardless of the origination of the law. Looking at the results in columns (1) and (3), we see an effect of mandatory duty to warn laws of about 5%. It is also worth noting that these estimates are largely insensitive to how the law is timed (over- or understated). In addition, this magnitude is consistent with recent estimates of the disproportionate share of homicides committed by those receiving mental health services.[36]

The expanded specification that accounts for the origin of the law is quite telling. First we see that similar to the basic identification reported in columns (1) and (3), the results in columns (2) and (4) are very robust to over- and understating the lifetime of the law. In addition, we see that splitting up the law by its origin doesn't seem to make a difference to the sign and significance of each coefficient. Though the magnitudes fluctuate slightly, I fail to reject the null hypothesis that the mandatory duty laws split by origin are equal—the p-value is reported at the bottom of the table. As mentioned earlier, since each test of equality of effect fails to reject the null, this suggests that if laws created by courts are exogenous statutory law is likely exogenous as well.

Table 4 is a replication of Table 3 using the natural log of homicide by non-stranger rate measured by the UCR-SHR. Comparing Table 3 to Table 4, we see that in every case, the significance and sign of each law does not change. Given the "readily identifiable victim" standard required to impose a *Tarasoff* duty, I'd expect *Tarasoff* laws to have a larger effect in explaining the variation in non-stranger homicides, as these are

[36] Recall that Swinson et al. (2007) suggest that 10% of homicides are committed by people receiving mental health services. The number of homicides by those receiving mental health services *and* those who would have received mental health services were it not for *Tarasoff* laws (as has been suggested in previous literature, *see* Wise 1976; Rosenhan et al. 1998; Klinka 2009) is likely to be higher.

homicides more directly applicable to *Tarasoff*. Table 4 suggests an increase in the effect of around 7 percentage points (depending on the specification) when homicides are restricted to homicides by non-strangers. This is consistent with the notion that duty to warn laws have a greater effect among victims where the murderer is known.

The coefficient on the discretionary duty variable is occasionally negative across specifications but insignificant. The negative sign suggests that given the opportunity to decide when to report, therapists successfully distinguish between real threats and idle patient banter, but the large variance prohibits any sort of meaningful interpretation. In general, however, we see a positive and persistent relationship between duty to warn laws and homicides.

Robustness

Table (5) attempts to account for a possible delay in information dissemination. It is possible that newly enacted state laws are not immediately known by the professional counseling community, or that the effect of the law might be heterogeneous across years. Given the initial costs borne by the psychologist directly after a *Tarasoff* ruling, we might expect the initial effect to be smaller than the effect of subsequent years. It could also be that information traveled slower during the period in which most of these laws passed. Conversely, the law could have a large initial impact due to an uptake effect then cool as a new steady state is achieved. To model this, I re-estimate equations (1) through (4) partitioning the *m* and *d* variables to measure the potentially different effect the law has in its initial uptake to the subsequent years following. As example, equation (1) is defined as:

$$h_{it} = \alpha_1 m_{it}^{initial} + \alpha_2 m_{it}^{subsequent} + \beta_1 d_{it}^{initial} + \beta_1 d_{it}^{subsequent} + X_{it} + s_i + y_t + e_{it} \qquad (5)$$

where equation (5) is an expansion of equation (1), and *m* is divided to measure the effect of mandatory duty laws in the initial two years of the law passing, $m_{it}^{initial}$, and the effect the law has on the subsequent years, $m_{it}^{subsequent}$. While the results are robust to other cutoffs, two years seems like the appropriate amount of transition time to allow learning and bear the cost. Given the nature of the changing law, and the required change of habit and practice required by the mental health professional, we might expect to see that less psychologists are responding to a new duty law initially, and thus α_1 should be smaller in magnitude to α_2, or that as the years pass, and the more psychologists become familiar and comfortable with the law, the law should have a greater effect.

Table 5 replicates the results presented in Table 3 with each law indicator variable split to reflect heterogeneous treatments of each law. We see that in general, most law indicator variables estimate a smaller effect of the law in the initial two years than the subsequent years, but not in any statistically significant manner.[37]

There is still another model specification that provides additional evidence of the effect of duty to warn laws. The timing of the court decisions in Tables 3 and 4 represent the decision of the highest ruling court. It is possible that *Tarasoff* laws began to alter patient and professional incentives at the trial or appellate court level thus changing the year of uptake for states with court made law. To test this, I specify a different model where the variables *mc* and *ml* are timed to reflect the lower court ruling.[38] This only applies to states where the *Tarasoff* duty was discussed in the lower courts. In some

[37] Though not reported here, model (5) was estimated for each specification reported in the tables. In the majority of cases, those results report, like the ones in Table 5, and increased effect of law as time passes. Those results are available upon request.
[38] The majority of state rulings have been made at the state Supreme Court level in states with appellate courts, thus the lower court is the appellate court. Adjusting timing to the lowest court level didn't yield substantially different results from the appellate court.

cases where a *Tarasoff* duty is implied (coded as *mandatory duty likely* in the data), the lower court opinion does not discuss the *Tarasoff* duty thus the original coding remains unchanged. Tables 6 and 7 replicate Tables 3 and 4 with the *mc* and the *ml* variables recoded to reflect the timing of the lower court opinions. We see that generally signs and significances do not change, though magnitudes fluctuate slightly.

The results from Tables 3 through 5 suggest robustness across data sources (NCHS, UCR-SHR), coding of the independent variables of interest (over- , under-counting), and timing of court decisions. Another possible source of bias might be from an overly influential state that biases the results.[39] Initially, there is no reason to believe this is the case since the log transformation of the homicide rate makes in near a normal distribution. Nonetheless, I excluded the most influential states on each tail of the distribution and found no significant change to the results.[40]

Though unlikely, I can further test for endogeneity by checking for evidence of reverse causality. This is done by including a two year lead[41] of each law variable (Carvell et al. 2009). The lead law variables shouldn’t explain any of the variation in current homicides, and I find no significant leads.

To further the claim that these laws are not in response to an underlying trend in crime, as a falsification test I attempt to explain other measures of crime by these duty to warn laws. In general, every state sets a lofty requirement of bodily harm required to induce a therapist’s duty such as, “substantial risk of imminent and serious physical

---

[39] Bias could also be introduced into the model if these laws were enacted as part of a larger health care bill. This was the case with the state of Nebraska (R.R.S. Neb. § 38-3132), but dropping Nebraska doesn’t significantly change the results.
[40] These results are available upon request.
[41] In addition to the evidence presented here, one and three year leads of the duty to warn laws were included in each model and were found insignificant in almost every instance suggesting that the laws do not reflect some underlying trend in crime.

injury,"[42] "serious physical harm . . . causing death,"[43] or "explicit threat to kill or inflict serious bodily injury."[44] Additionally, by definition if the potential crime is discussed with a therapist, it would not be considered manslaughter, as homicides require some sort of premeditation. Thus, manslaughter rates offer an interesting counterfactual of a trend in crime that should have nearly no relationship with *Tarasoff* laws.

Tables 8 and 9 replicate the estimation of equations (1) – (4) where the dependent variable is the natural log of the manslaughter rate as reported by the UCR database. As seen in Tables 8 and 9, duty to warn laws do not explain in almost every specification any of the variation in manslaughter rates. The lack of a significant effect on manslaughter rates suggests that what is being captured by the duty to warn laws is not some spurious trend in crime.

To further this falsification test, I replicate equations (1) – (4) with various measures of crime as the dependent variable.[45] The dependent variables are the natural log of the auto theft rate, larceny rate, and robbery rate as measured by the UCR. In total, 30 coefficients of interest were estimated in 12 models and only three of the duty to warn coefficients were significant at the ten percent level. This is approximately what is expected and suggests that state duty to warn laws do not explain any of the variation in crime except for that of homicides.

5. Conclusion

The effect of state duty to warn laws on homicidal activity has been debated for decades. This paper shows that all else equal, mandatory duty to warn laws cause an

[42] D.C. Code § 7-1203.03(a)
[43] ORC Ann. 2305.51
[44] 59 Okl. St. § 1376. *See generally* Edwards (2010) for each state's specific standard of harm.
[45] While of course excluding each measure of crime as a control in its own regression.

increase in homicides of 5% or 4.35 people per 1,000,000.[46] This is consistent with previous literature suggesting that worsening mental health conditions lead to more crime. Duty to warn laws change the incentives of both the patient and the doctor. The original intention of the law was to deter dangerous patients from committing heinous crimes, but what may actually have happened was that the law changed the incentives to the patient and the doctor such that the patient has incentive to withhold homicidal tendencies, and the doctor has incentive to not explore homicidal tendencies. In addition, these laws increase the liability to health professionals and incentivize those professionals to not treat the most at risk patients, or at very least make the current state of the law abundantly clear to the patient as to suggest suppression of the most dangerous statements leaving the psychologist in liability-free ignorance to the true mental state of the patient. As a result the mental help needed to treat the patient is foregone, and all too often violence ensues.

I find these results to be robust across a multitude of specifications and falsification tests. The policy implications are simple and fairly easily employed. A change in law to no duty or discretionary duty should decrease in homicides.

[46] 39.3 homicides per state per year, on average.

6. References

Ackerman, Bruce. 1976. "Tarasoff v. Regents of the University of California: The duty to Warn: Common Law and Statutory Problems for California Psychotherapists". 14 *Cal. W.L. Rev.* 153-82.

Avraham, Ronen. 2006. "Database of State Tort Law Reforms." Working Paper No. 06-08.

Northwestern Law and Econ Research Paper. Available at: http://ssrn.com/abstract=902711.

Ayres, Ian, and John Donohue. 2003. "Shooting Down the 'More Guns, Less Crime' Hypothesis." 55 *Stanford Law Review* 1193–312.

Beail, N. 2001. "Recidivism Following Psychodynamic Psychotherapy Amongst Offenders With Intellectual Disabilities" 3 *British Journal of Forensic Practice* 33-37.

Black, Dan, and Daniel Nagin. 1998. "Do 'Right-to-Carry' Laws Deter Violent Crime?" 27 *Journal of Legal Studies* 209–19.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" 114 *Quarterly Journal of Economics* 249–275.

Borum, Randy and Marisa Reddy. 2001. "Assessing violence risk in Tarasoff situations: a fact-based model of inquiry". 19 *Behavioral Sciences and Law* 375-385.

Carvell, Daniel, Janet Currie and W. Bentley MacLeod. 2009. "Accidental Death and the Rule of Joint and Several Liability." NBER Working Paper No. 15412.

Centers for Disease Control and Prevention, National Centers for Injury Prevention and Control. Web-based Injury Statistics Query and Reporting System (WISQARS) [online]. 2005. Available from: www.cdc.gov/ncipc/wisqars

Choe, Jeanne Y., Linda A. Teplin, Karen M. Abram. 2008. “Perpetration of Violence, Violent Victimization, and Severe Mental Illness: Balancing Public Health Concerns.” 59 *Psychiatric Services* 153-164.

Donohue, John, and Steven Levitt. 2001. “The Impact of Legalized Abortion on Crime,” 116 *Quarterly Journal of Economics* 379-420.

Donohue, John, and Justin Wolfers. 2009. “Estimating the Impact of the Death Penalty on Murder.” 11 *American Law and Economics Review* 249-309.

Duggan, Mark. 2001. “More guns, more crime,” 109 *Journal of Political Economy* 1086-1114.

Edwards, Griffin Sims. 2010. “Database of State *Tarasoff* Laws.” Available from: http://userwww.service.emory.edu/~gsedwar/research.html

Fliszar, Gregory M. 2002. “Dangerousness and The Duty To Warn, Emerich V. Philadelphia Center For Human Development, Inc. Brings Tarasoff To Pennsylvania.” 62 *U. Pitt. L. Rev.* 201-22.

Givelber, Daniel, William Bowers and Carolyn Blitch. 1980. “Tarasoff, Myth and Reality: An Empirical Study of Private Law in Action.” 1984 *Wis. L. Rev.* 443-98.

Ginsberg, Brian. 2004. “Tarasoff at Thirty: Victim's Knowledge Shrinks the Psychotherapist's Duty to Warn and Protect”. 21 *Journal of Contemporary Health Law & Policy* 1-35.

Ginsberg, Brian. 2007. “Therapists Behaving Badly: Why The Tarasoff Duty Is Not Always Economically Efficient”. 43 *Willamette L. Rev*. 31-75.

Harmon, A. J. 2008. “Back from wonderland: A linguistic approach to duties arising from threats of physical violence.” 37 *Cap. U.L. Rev.* 27-91.

Huang. Bert I. (2010) “Lightened Scrutiny” Working paper.

Hylton, Keith. 2006 “Duty in Tort Law: An Economic Approach” 75 *Fordham L. Rev*. 1501-1533.

Joyce, Ted, 2009. “A Simple Test of Abortion and Crime”. *Review of Economics and Statistics*, forthcoming.

Katz, Lawrence, Steven Levitt and Ellen Shustorovich. 2003. “Prison Conditions, Capital Punishment, and Deterrence.” 5 *American Law and Economics Review* 318-43.

Klick, J. & Markowitz, S. 2006. “Are mental health insurance mandates effective? Evidence from suicides”. 15 *Health Economics* 83-97.

Klinka, Elisia. 2009. “It's Been a Privilege: Advising patients of the Tarasoff duty and its legal consequences for the federal psychotherapist-patient privilege” 78 *Fordham L. Rev*. 863-931.

Levitt, Steven D. 2004. “Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not.” 18 *Journal of Economic Perspectives* 163-90.

Levitt, Steven D. (1998) “The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports.” 14 *Journal of Quantitative Criminology* 61-81.

Link, Bruce G., and Ann Stueve. 1995. "Evidence Bearing on Mental Illness as a Possible Cause of Violent Behavior." 17 *Epidemiologic Reviews* 172-181.

Ludwig, Jens and Philip Cook. 2000. "Homicide and Suicide Rates Associated with Implementation of the Brady Handgun Violence Prevention Act." 284 *Journal of the American Medical Association* 585–91.

Markowitz, Sara. 2005. "Alcohol, Drugs, and Violent Crime." 25 *International Review of Law and Economics* 20-44.

Marcotte, David and Sara Markowitz. 2009 "A Cure for Crime? Psycho-Pharmaceuticals and Crime Trends." NBER working paper # 15354.

Meyer, Bruce. 1995. "Natural and Quasi-Natural Experiments in Economics," 12 *Journal of Business and Economic Statistics* 151–162.

Nestor, Paul G. 2002. "Mental disorder and violence: personality dimensions and clinical features." 159 *American Journal of Psychiatry* 1973–1978

Perlin, Michael. 1992. "Tarasoff and the Dangerous Patient: New Directions for the 1990's". 16 *Law & Psychology Review* 29-64.

Roberts, Laura W., Cynthia M. A. Geppert, and Robert Bailey. 2002. "Ethics in Psychiatric Practice: Essential Ethics Skills, Informed Consent, the Therapeutic Relationship, and Confidentiality," 8 *Journal of Psychiatric Practice,* 290–305.

Rosenhan, David L., Terri W. Teitelbaum, Kathi W. Teitelbaum, and Martin Davidson. 1993. "Warning Third Parties: The Ripple Effects of Tarasoff". 24 *Pacific Law Journal* 1165-231.

Ruggles, S., Sobek, M., Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. Integrated Public Use Microdata

Series: Version 4.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2009.

Ruhm, Christopher. 2000. "Are Recessions Good for Your Health?" 115 *Quarterly Journal of Economics* 617–50.

Shepherd, Joanna. 2009. "The Influence of Retention Politics on Judges' Voting" 38 *The Journal of Legal Studies* 169-204.

Silver, Eric, Richard B. Felson, and Matthew Vaneseltine. 2008. "The relationship between mental health problems and violence among criminal offenders." 35 *Criminal Justice and Behavior* 405-426.

Stevenson, Betsy, and Justin Wolfers. 2006. "Bargaining in the Shadow of the Law: Divorce

Laws and Family Distress." 121 *Quarterly Journal of Economics* 267–88.

Stone, Alan A. 1976. "The Tarasoff Decisions: Suing Psychotherapists to Safeguard Society". 90 *Harvard Law Review* 358-78.

Stone, Alan A. 1984. *Law, Psychiatry and Morality: Essays and Analysis.* American Psychiatric Press.

Swanson, Jeffrey W., Charles E. Holzer, III, Vijay K. Ganju, and Robert Tsutomu Jono. 1990. "Violence and Psychiatric Disorder in the Community: Evidence From the Epidemiologic Catchment Area Surveys." 41 *Hosp Community Psychiatry* 761-770.

Swinson, Nicola, Ashim Bettadapura, Kirsten Windfuhr, Navneet Kapur, Louis Appleby, and Jenny Shaw. 2007. "National Confidential Inquiry into Suicide and Homicide by People with Mental Illness: new directions." 31 *The Psychiatrist* 161-163.

Teplin, Linda A., Gary M. McClelland, Karen M. Abram, and Dan A. Weiner. 2005. “Crime Victimization in Adults with Severe Mental Illness, Comparison with the National Crime Victimization Survey.” 62 *Arch Gen Psychiatry* 911-21.

Wise, Toni Pryor. 1978. “Where the Public Peril Begins: A Survey of Psychotherapists to Determine the Effects of Tarasoff” 31 *Stanford Law Review* 165-190.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.

Zimring, Franklin E. 2007. *The Great American Crime Decline.* New York: Oxford University Press.

Table 1

Summary of State Duty to Warn Laws

State	*Duty Law*	*Date Passed*	*Deciding Body*
Alabama	mandatory likely	1985	Court
Alaska	discretion	1986	Legislation
Arizona	mandatory	1977	Court
Arkansas	no law	--	--
California	mandatory	1976	Court
Colorado	mandatory	1987	Legislation
Connecticut	discretion	1989	Legislation
Delaware	mandatory	1988	Court
District of Columbia	discretion	1979	Legislation
Florida	discretion	1987	Legislation
Georgia	mandatory likely	1982	Court
Hawaii	mandatory likely	1996	Court
Idaho	mandatory	1991	Legislation
Illinois	discretion	1990	Legislation
Indiana	mandatory	1998	Legislation
Iowa	no law	1981	Court
Kansas	no law	--	--
Kentucky	mandatory	1986	Legislation
Louisiana	mandatory	1986	Legislation
Maine	no law	--	--
Maryland	mandatory	1989	Legislation
Massachusetts	mandatory	1989	Legislation
Michigan	mandatory	1989	Legislation
Minnesota	mandatory	1986	Legislation
Mississippi	mandatory	1991	Legislation
Missouri	mandatory	1995	Court
Montana	mandatory	1987	Legislation
Nebraska	mandatory	1980	Court
Nevada	no law	--	--
New Hampshire	mandatory	1987	Legislation
New Jersey	mandatory	1979	Court
New Mexico	no law	1989	Court
New York	discretion	1984	Legislation
North Carolina	mandatory likely	1987	Court
North Dakota	no law	--	--
Ohio	mandatory	1997	Court
Oklahoma	mandatory	2009	Legislation
Oregon	discretion	1977	Legislation

Pennsylvania	mandatory	1998	Court
Rhode Island	discretion	1978	Legislation
South Carolina	mandatory	1998	Court
South Dakota	mandatory	1978	Court
Tennessee	mandatory	1989	Legislation
Texas	discretion	1979	Legislation
Utah	mandatory	1988	Legislation
Vermont	mandatory	1985	Court
Virginia	no duty	1995	Court
Washington	mandatory	1983	Legislation
West Virginia	discretion	1977	Legislation
Wisconsin	mandatory	1988	Court
Wyoming	discretion	1999	Legislation

Notes: For a complete database of each law, including references see Edwards (2010).

Table 2

Summary Statistics

	(1) Full Sample	(2) Manda-tory Duty State	(3) Discret-ionary Duty State	(4) State with No Law
Homicides per 100,000 of the population (NCHS)	8.92	6.54	10.44	17.74
	(12.76)	(3.90)	(11.85)	(28.12)
Non-Stranger Homicides per 100,000 of the population (UCR-SHR)	4.50	3.43	3.85	10.75
	(7.25)	(2.08)	(2.36)	(17.84)
Manslaughter per 100,000 of the population (UCR-SHR	0.57	0.47	0.50	1.16
	(0.86)	(0.36)	(0.43)	(2.04)
Suicides per 100,000 of the population (NCHS)	12.74	12.43	12.15	15.13
	(3.34)	(2.53)	(4.17)	(4.17)
Mandatory Duty	0.44	0.68	0.00	0.00
	(0.50)	(0.47)	(0.00)	(0.00)
Mandatory Duty (case law)	0.10	0.16	0.00	0.00
	(0.30)	(0.37)	(0.00)	(0.00)
Mandatory Duty (statutory law)	0.27	0.41	0.00	0.00
	(0.44)	(0.49)	(0.00)	(0.00)
Mandatory Duty Likely (case law)	0.07	0.11	0.00	0.00
	(0.25)	(0.31)	(0.00)	(0.00)
Discretionary Duty	0.17	0.00	0.81	0.00
	(0.38)	(0.00)	(0.40)	(0.00)
Demographic Characteristics	.	.	.	.
Percent Male	0.49	0.49	0.48	0.49
	(0.01)	(0.01)	(0.01)	(0.01)
Percent Black	0.10	0.10	0.13	0.07
	(0.12)	(0.10)	(0.17)	(0.07)
Median Age	33.51	33.33	33.84	33.83
	(3.17)	(3.13)	(3.36)	(3.05)
Unemployment	5.96	5.78	6.75	5.56
	(2.10)	(2.10)	(2.16)	(1.67)

Real Median Income	30578.30	30630.68	31819.96	28380.18
	(4276.10)	(4172.94)	(4489.44)	(3519.19)
Per Capita Real Police Expenditures	92.89	83.85	128.52	79.52
	(44.26)	(24.03)	(71.95)	(29.79)
Per Capita Real Judicial Expenditures	42.89	38.75	59.42	36.45
	(24.10)	(16.10)	(37.83)	(14.75)
Crack Index	2.41	1.66	2.50	5.78
	(7.08)	(2.58)	(4.48)	(17.03)
Real Beer Tax	0.16	0.18	0.13	0.16
	(0.13)	(0.15)	(0.10)	(0.07)
Psychiatrist per 100,000 of the population	21.66	21.31	23.67	20.15
	(9.83)	(9.16)	(13.03)	(5.77)
Crime Trends	.	.	.	.
Assaults per 100,000 of the population (SHR)	305.73	286.46	392.39	260.38
	(183.78)	(148.96)	(245.79)	(177.32)
Robberies per 100,000 of the population (SHR)	154.76	130.23	257.25	109.35
	(161.33)	(84.30)	(284.18)	(90.78)
Auto Theft per 100,000 of the population (SHR)	419.10	396.15	558.77	307.80
	(244.82)	(203.14)	(324.28)	(179.14)
Political Characteristics	.	.	.	.
Democratic Governor	0.52	0.51	0.56	0.49
	(0.50)	(0.50)	(0.50)	(0.50)
Republican Governor	0.01	0.01	0.02	0.05
	(0.12)	(0.07)	(0.12)	(0.22)
Independent Governor	0.47	0.49	0.42	0.46
	(0.50)	(0.50)	(0.50)	(0.50)
Proportion Democrat of State House	0.56	0.54	0.60	0.57
	(0.18)	(0.18)	(0.17)	(0.17)
Proportion Democrat of State Senate	0.56	0.56	0.58	0.56
	(0.18)	(0.18)	(0.19)	(0.17)
Selection Method of State Supreme Court Judges	.	.	.	.
Gubernatorial Appointee	0.08	0.09	0.00	0.14
	(0.27)	(0.29)	(0.00)	(0.35)
Legislative Appointment or	0.06	0.03	0.09	0.14

Election				
	(0.24)	(0.17)	(0.29)	(0.35)
Missouri Plan	0.41	0.45	0.45	0.14
	(0.49)	(0.50)	(0.50)	(0.35)
Nonpartisan Election	0.25	0.30	0.09	0.29
	(0.44)	(0.46)	(0.29)	(0.45)
Partisan Election	0.18	0.12	0.27	0.29
	(0.38)	(0.33)	(0.45)	(0.45)
Retention Method of State Supreme Court Judges	.	.	.	.
Gubernatorial Reappointment	0.08	0.06	0.09	0.14
	(0.27)	(0.24)	(0.29)	(0.35)
Judicial Nominating Committee Reappointment	0.02	0.03	0.00	0.00
	(0.14)	(0.17)	(0.00)	(0.00)
Legislative Reappointment or Reelection	0.08	0.06	0.09	0.14
	(0.27)	(0.24)	(0.29)	(0.35)
Life Tenure	0.06	0.06	0.09	0.00
	(0.24)	(0.24)	(0.29)	(0.00)
Nonpartisan Reelection	0.27	0.33	0.09	0.29
	(0.45)	(0.47)	(0.29)	(0.45)
Partisan Reelection	0.14	0.09	0.18	0.29
	(0.34)	(0.29)	(0.39)	(0.45)
Unopposed Retention Election	0.33	0.36	0.36	0.14
	(0.47)	(0.48)	(0.48)	(0.35)
Tort Reform Laws	.	.	.	.
Caps on Punitive Damages	0.31	0.30	0.29	0.39
	(0.46)	(0.46)	(0.46)	(0.49)
Caps on Total Damages	0.13	0.12	0.03	0.29
	(0.33)	(0.33)	(0.18)	(0.46)
Caps on Non-economic Damages	0.26	0.30	0.21	0.18
	(0.44)	(0.46)	(0.40)	(0.39)
Joint and Several Liability	0.53	0.53	0.62	0.39
	(0.50)	(0.50)	(0.49)	(0.49)
Periodic Payments	0.78	0.75	0.64	1.19
	(0.88)	(0.91)	(0.80)	(0.78)
Punitive Damages Evidence	0.47	0.55	0.23	0.42
	(0.50)	(0.50)	(0.42)	(0.50)
Split Recovery of Punitive	0.09	0.09	0.14	0.00

Damages				
	(0.28)	(0.28)	(0.35)	(0.00)
Collateral Source	0.52	0.53	0.57	0.40
	(0.50)	(0.50)	(0.50)	(0.49)
Contingency Fee	0.26	0.25	0.40	0.09
	(0.44)	(0.43)	(0.49)	(0.29)
Patient Compensation Fund	0.19	0.18	0.17	0.29
	(0.40)	(0.39)	(0.38)	(0.45)
Mental Health Insurance Mandates	.	.	.	.
Benefits mandated for either all plans or other plans	0.19	0.21	0.09	0.24
	(0.39)	(0.40)	(0.28)	(0.42)
Benefits mandated for all plans	0.16	0.17	0.09	0.24
	(0.36)	(0.37)	(0.28)	(0.42)
Benefits mandated with some parity	0.11	0.12	0.09	0.12
	(0.31)	(0.32)	(0.28)	(0.31)
Benefits mandated will complete parity	0.09	0.09	0.08	0.12
	(0.29)	(0.29)	(0.27)	(0.31)
Sample Size	1173.00	759.00	299.00	115.00

Table 3

FE estimation of the effect of state duty to warn laws on the log of state NCHS homicide rates

dependent variable: ln(NCHS homicide rate)	(1)	(2)	(3)	(4)
Mandatory Duty	0.058†	.	0.045‡	.
	(0.030)	.	(0.019)	.
Mandatory Duty (case law)	.	0.100^	.	0.067‡
	.	(0.055)	.	(0.027)
Mandatory Duty (statutory law)	.	0.058†	.	0.055†
	.	(0.028)	.	(0.026)
Mandatory Duty Likely (case law)	.	-0.027	.	-0.028
	.	(0.047)	.	(0.045)
Discretionary Duty	0.004	-0.001	0.013	0.011
	(0.053)	(0.051)	(0.054)	(0.052)
P-value of case and statutory law equivalence test	.	0.394	.	0.723
Sample Size	1114	1114	1114	1114
Adjusted R-squared	0.975	0.975	0.975	0.975
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of homicides per 100,000 of the population as measured by the NCHS. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. Each model contains state and year fixed effects, state specific time trends, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 4

FE estimation of the effect of state duty to warn laws on the log of state non-stranger UCR-SHR homicide rates

dependent variable: ln(UCR-SHR non-stranger homicide rate)	(1)	(2)	(3)	(4)
Mandatory Duty	0.123‡	.	0.095‡	.
	(0.039)	.	(0.037)	.
Mandatory Duty (case law)	.	0.214‡	.	0.184‡
	.	(0.062)	.	(0.050)
Mandatory Duty (statutory law)	.	0.124‡	.	0.104†
	.	(0.050)	.	(0.050)
Mandatory Duty Likely (case law)	.	-0.177	.	-0.195
	.	(0.075)	.	(0.082)
Discretionary Duty	0.062	0.055	0.020	0.015
	(0.114)	(0.118)	(0.108)	(0.109)
P-value of case and statutory law equivalence test	.	0.246	.	0.216
Sample Size	1081	1081	1081	1081
Adjusted R-squared	0.903	0.905	0.903	0.905
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of non-stranger homicides per 100,000 of the population as measured by the UCR-SHR. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. Each model contains state and year fixed effects, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ p<0.10 † p<0.05 ‡ p<0.001

Table 5

FE estimation of the heterogeneous effect of state duty to warn laws on the log of state NCHS homicide rates

dependent variable: ln(NCHS homicide rate)	(1)	(2)	(3)	(4)
Mandatory Duty--First Two Years	0.053†	.	0.041†	.
	(0.025)	.	(0.020)	.
Mandatory Duty--Subsequent Years	0.056^	.	0.050^	.
	(0.030)	.	(0.030)	.
Mandatory Duty (case law)--First Two Years	.	0.040	.	0.017
	.	(0.056)	.	(0.031)
Mandatory Duty (case law)--Subsequent Years	.	0.055^	.	0.038
	.	(0.031)	.	(0.033)
Mandatory Duty (statutory law)--First Two Years	.	0.061‡	.	0.064‡
	.	(0.025)	.	(0.027)
Mandatory Duty (statutory law)--Subsequent Years	.	0.101‡	.	0.109‡
	.	(0.043)	.	(0.044)
Mandatory Duty Likely (case law)--First Two Years	.	-0.048	.	-0.072
	.	(0.036)	.	(0.023)
Mandatory Duty Likely (case law)--Subsequent Years	.	-0.059	.	-0.055
	.	(0.035)	.	(0.031)
Discretionary Duty--First Two Years	0.023	0.029	0.016	0.020
	(0.054)	(0.055)	(0.050)	(0.050)
Discretionary Duty--Subsequent Years	0.003	0.053	0.002	0.058
	(0.072)	(0.073)	(0.066)	(0.068)
Sample Size	1114	1114	1114	1114
Adjusted R-squared	0.969	0.970	0.969	0.970
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of homicides per 100,000 of the population as measured by the NCHS. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. Each model contains state and year fixed effects, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 6

FE estimation of the effect of state duty to warn laws on the log of state NCHS homicide rates with lower court timing

dependent variable: ln(NCHS homicide rate)	(1)	(2)	(3)	(4)
Mandatory Duty	0.069‡	.	0.046†	.
	(0.026)	.	(0.023)	.
Mandatory Duty (case law)	.	0.095^	.	0.032
	.	(0.051)	.	(0.027)
Mandatory Duty (statutory law)	.	0.072‡	.	0.087‡
	.	(0.030)	.	(0.034)
Mandatory Duty Likely (case law)	.	-0.114	.	-0.058
	.	(0.042)	.	(0.027)
Discretionary Duty	0.044	0.046	0.027	0.043
	(0.062)	(0.061)	(0.055)	(0.055)
P-value of case and statutory law equivalence test	.	0.706	.	0.165
Sample Size	1114	1114	1114	1114
Adjusted R-squared	0.969	0.969	0.969	0.970
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of homicides per 100,000 of the population as measured by the NCHS. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. In addition, these models account for the possibility the behavior started to change at the trial or appeals level. Each model contains state and year fixed effects, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 7

FE estimation of the effect of state duty to warn laws on the log of state non-stranger UCR-SHR homicide rates with lower court timing

dependent variable: ln(UCR-SHR non-stranger homicide rate)	(1)	(2)	(3)	(4)
Mandatory Duty	0.076	.	0.084†	.
	(0.051)	.	(0.038)	.
Mandatory Duty (case law)	.	0.191‡	.	0.172‡
	.	(0.081)	.	(0.056)
Mandatory Duty (statutory law)	.	0.066	.	0.100†
	.	(0.057)	.	(0.050)
Mandatory Duty Likely (case law)	.	-0.118	.	-0.200
	.	(0.105)	.	(0.079)
Discretionary Duty	0.067	0.065	0.024	0.019
	(0.110)	(0.110)	(0.107)	(0.109)
P-value of case and statutory law equivalence test	.	0.706	.	0.165
Sample Size	1114	1114	1114	1114
Adjusted R-squared	0.969	0.969	0.969	0.970
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of non-stranger homicides per 100,000 of the population as measured by the UCR-SHR. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. In addition, these models account for the possibility the behavior started to change at the trial or appeals level. Each model contains state and year fixed effects, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 8

FE estimation of the effect of state duty to warn laws on the log of UCR manslaughter rates

dependent variable: ln(NCHS homicide rate)	(1)	(2)	(3)	(4)
Mandatory Duty	0.088	.	0.045	.
	(0.030)	.	(0.019)	.
Mandatory Duty (case law)	.	0.100	.	0.067
	.	(0.055)	.	(0.027)
Mandatory Duty (statutory law)	.	0.058	.	0.055
	.	(0.028)	.	(0.026)
Mandatory Duty Likely (case law)	.	-0.027	.	-0.028
	.	(0.047)	.	(0.045)
Discretionary Duty	0.004	-0.001	0.013	0.011
	(0.053)	(0.051)	(0.054)	(0.052)
P-value of case and statutory law equivalence test	.	0.944	.	0.915
Sample Size	1018	1018	1018	1018
Adjusted R-squared	0.805	0.806	0.805	0.805
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of manslaughters per 100,000 of the population as measured by the UCR. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. Each model contains state and year fixed effects, state specific time trends, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ p<0.10 † p<0.05 ‡ p<0.001

Table 9

The effect of state duty to warn laws on the log of UCR manslaughter rates with lower court timing

dependent variable: ln(UCR manslaughter rate)	(1)	(2)	(3)	(4)
Mandatory Duty	-0.045	.	0.035	.
	(0.079)	.	(0.098)	.
Mandatory Duty (case law)	.	-0.117	.	0.121
	.	(0.131)	.	(0.132)
Mandatory Duty (statutory law)	.	-0.051	.	-0.042
	.	(0.085)	.	(0.102)
Mandatory Duty Likely (case law)	.	0.353^	.	0.067
	.	(0.212)	.	(0.184)
Discretionary Duty	-0.639	-0.644	-0.665	-0.702
	(0.537)	(0.535)	(0.550)	(0.546)
P-value of case and statutory law equivalence test	.	0.605	.	0.229
Sample Size	1018	1018	1018	1018
Adjusted R-squared	0.694	0.695	0.697	0.698
Law Lifespan Over- or Under-counted	O	O	U	U

Notes: State-level clustered standard errors are in parenthesis and all models are weighted by the square root of the state population. The dependent variable is the natural log of manslaughters per 100,000 of the population as measured by the UCR. The values in models (3) and (4) represent the specification where each respective law is not counted until its first full year of enactment. Each model contains state and year fixed effects, political controls, crime trend controls, and lagged demographic and crime trend characteristics. A detailed outline of mandatory and discretionary duty to warn laws can be found in Edwards (2010). ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Figure 1

Current Duty to Warn Laws

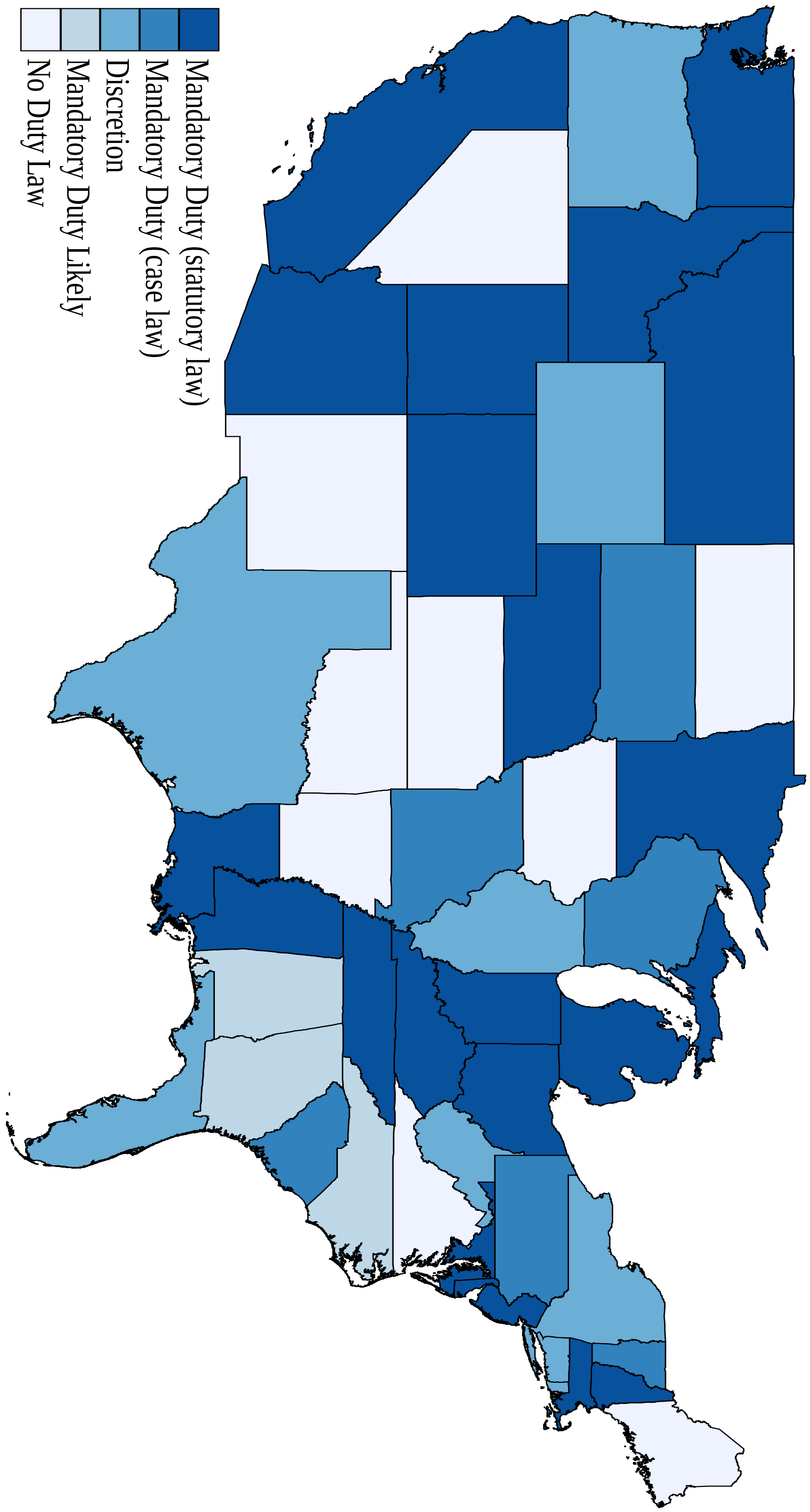
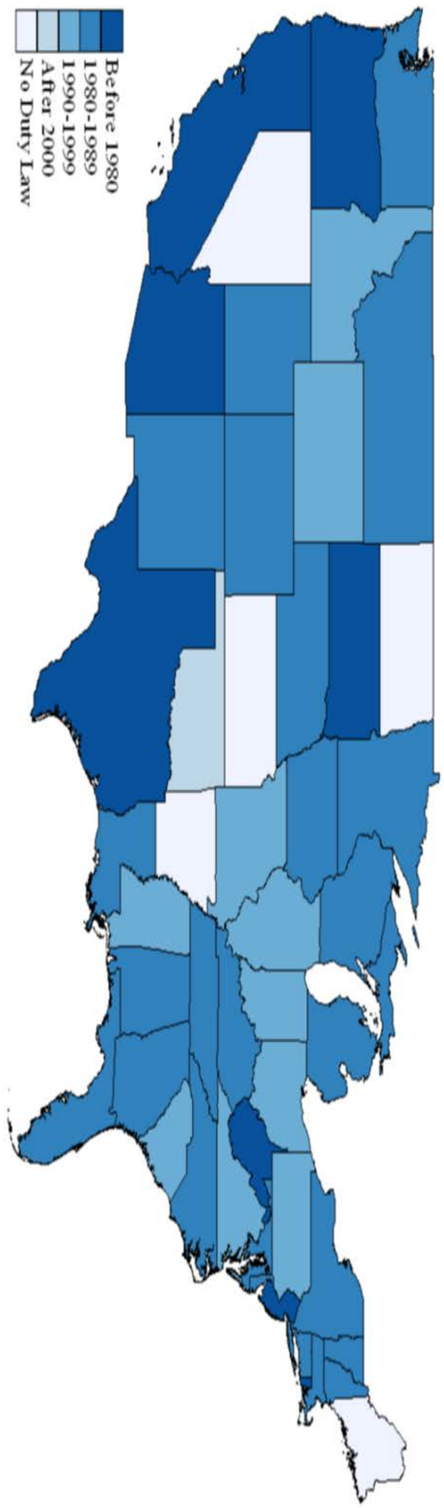


Figure 2

Duty to Warn Laws Year of First Enactment



A Selection-Corrected Estimate of *Chevron*'s Impact on Agency Deference

1.A Chevron

Congress has longed passed laws and appointed federal agencies to carry out the text of the statute. Often agency interpretation and actions surrounding the governing statute is disputed. The landscape surrounding how agencies interpret these statutes and the judicial decision to defer to the agency interpretation when challenged changed greatly with the passing of *Chevron U.S.A., Inc v Natural Resources Defense Council, Inc*[47] on June 24th, 1984. Prior to *Chevron*, adjudication surrounding agency interpretation of statutes was ambiguous and inconsistent (Miles & Sunstein 2006), but the ruling in *Chevron* created a test aimed to reduce ambiguity in the law. This test, also known as the *Chevron* two-step, is:

1- Has "Congress . . . directly spoken to the precise question at issue,"

2- Is the agencies interpretation of the statute "permissible."[48]

The *Chevron* two-step aimed to harmonize agency adjudication by creating a simple rule that first confirms that the governing statute does not explicitly state anything contrary to the agencies action, and second that the interpretation of the statute by the agency is "permissible". While rule 1 is arguably just a special case of rule 2 (Stephenson and Vermeule 2009) it remains that the *Chevron* two-step implies a large degree of agency deference.

1.B Incentives

Changing incentives for both the challenging party and the agency pre- and post-*Chevron* play a crucial role in measuring *Chevron*'s impact. In the days prior to

[47] 467 US 837 (1984).
[48] *Id*. at 842–44

Chevron, the degree to which judges deferred to agency decisions varied greatly by judge and case (Miles & Sunstein 2006). The ambiguity in the expectation of judicial agency deference was surely calculated in a challenger's decision to challenge (Manning & Stephenson 2009).[49] This same ambiguous threat of litigation would likely affect how agencies operate.

From the viewpoint of the challenger, it is unclear a priori whether ambiguous expected outcomes would create an incentive to challenge more or less often. In one scenario, the challenger might be more inclined to challenge (compared to a post-*Chevron* world) due to private information held that signals to the strength of the case. On the other hand, if there is no clear rubric in agency adjudication, a would-be challenger would likely not be able to accurately calculate the probability of success, and this might discourage challenging. The same ambiguity would likely affect how agencies make decisions.

After the implementation of the *Chevron* two-step, the incentives to both the would-be challengers and the agencies likely change. Post-*Chevron* both the challenger and the agency should be able to more precisely predict the expected outcome conditional on an agency challenge. That being known, we are likely to see a decrease in the number of challenges but an increase in the quality of those challenges.

1.C Judicial Decisions

The implementation of *Chevron* affected not only the incentives to challengers and agencies, but also judges. Prior to *Chevron*, judges considered factors such as statute\interpretation contemporaneity and the longevity and consistency of the

[49] In addition, Manning and Stephenson (2009) state the reasonable argument that *Chevron* might not have had the practical impact that many originally thought. The hope with this paper is to provide further evidence that it did.

interpretation (Manning and Stephenson 2009). These rules left considerable room for judges to interject their own policy preferences in each ruling (Miles and Sunstein 2006). One potential draw of the *Chevron* rule is its potential to standardize how courts rule on administrative interpretations of statutes (Strauss 1987). A clear standard for agency deference should reduce the amount judges can interject their own biases into rulings. Miles and Sunstein (2006) found strong evidence suggesting that judges rule according to their political leanings despite *Chevron*.

This bias introduced by judges' policy preferences as well as the bias introduced by the change in incentives to challenge muddies our ability to see if *Chevron* increased agency deference and by what magnitude. A clear understanding of *Chevron*'s effect is important to determine how much a Supreme Court ruling can actually change the judicial landscape with specific focus on how incentives are altered. It is also important to those who see agency deference as an unconstitutional delegation of power and a smearing of the checks and balances system. Additionally it should help to correctly identify what influence major judicial decisions have on the individual incentives to litigate.

Identifying the effect of *Chevron* is in practice quite difficult with so many individual decisions factoring into whether or not an agency ruling is challenged and upheld. Previous studies have found both theoretically and empirically that *Chevron* favors agencies and their interpretation of statutes (Miles & Sustein 2006; Czarnezki 2009). In addition, early studies found that *Chevron* is associated with an increase in agency deference. Schuck and Elliot (1990), using comparison of means, found *Chevron* caused about a 14% increase in agency deference. Some years later Richards et al.

(2006) used a logit model to estimate an 8 percentage point increase in deference. The difficulty in using a comparison of means approach is that many factors not observed by the means affects the level of means. Richards et al. does a good job of controlling for observable case characteristics that affect the estimated impact of *Chevron*, but other unobservable factors like post-*Chevron* incentives not to challenge will likely bias the estimated effect.

The preferred research design would be to observe 50 years of agency deference then implement *Chevron* and go back in time to redo those 50 years under the *Chevron* regime. This is unfortunately not possible. A good alternative would be to analyze the cases that were pending when *Chevron* was decided since those cases would not be biased by the post-*Chevron* incentives to challenge. Unfortunately, only 58 votes were cast between 1984 and 1985 which raises the question of model validity when less than 60 observations are identified in the data.[50]

Using the data collected and generously shared by Richards et al. (2006), I attempt to do a hybrid of the two approaches. To address the issue of sample selection after *Chevron* passed, I estimate the jump in the trend of agency deference at the date *Chevron* was decided. This allows me to use the full sample of observed agency deference while placing greatest weight on the cases in the pipeline when *Chevron* was decided.

2. Estimation Strategies

One possible design to measure the jump in agency deference was first established in the education literature by Thistlethwaite and Campbell (1960). They

[50] It is important to note that balance in the data has been considered rigorously. Distribution analysis of the covariates suggests some imbalance globally which is consistent with the need for a selection-corrected estimate. Unfortunately, there is not enough data locally comment on balance around the cutoff date.

assumed that comparing student academic performance right above and below a cut off for merit-based awards provided a near randomized experiment of the effect of the awards. Essentially Thistlethwaite and Campbell (1960) compare two groups, one that received the treatment and the other that did not because randomization is not possible.

This method essentially tests for a jump in the data right at the cut off value by defining the estimated function in such a way as to assign highest weight on the data right near the cut off and decreasing as the data gets farther away from the cut off. Intuitively, it makes sense to compare, say, two students who scored similarly but who received different treatments. The idea being that a student earning 70 points on an exam might score slightly different if we were to rewind time and allow the student to take the test again—not because she is any more or less prepared, just because there is some randomness associated with how well somebody scores on a test on any given day. Clearly, the best students will be at the top of the distribution, and the lowest on the bottom, but where each stands locally is close to random. I, in this paper, extend this idea of randomness right around the treatment, but instead of a little randomness locally around any given test score, I exploit the randomness associated around the decision date of a case.

Similar estimation strategies extend from government assistance programs (Ludwig & Miller 2007) to program evaluation of mandatory summer school (Jacob & Lefgren 2004) and optional full-day kindergarten (Edwards & Frisvold 2010).

Recent work by Lee (2008) has proven theoretically that even if subjects have some control over the timing (in the context of *Chevron*) variable, some control is as good as randomization around the forcing variable.

In the case of *Chevron*, the concern would be that the Supreme Court decision might not be exogenously determined in the model. That is, the Supreme Court might be ruling based on the observed trend in agency deference. If this is the case, traditional estimation techniques would likely suffer from bias caused by endogeneity of the passing of *Chevron*. Measuring the jump in the time trend, however, would not suffer from the same bias as long as the forcing variable, or the uptake of *Chevron*, is not precisely determined by some unobserved attitude, say in the Supreme Court.

The conclusions of Lee (2008) in the context of *Chevron* suggest that as long as *Chevron* was not determined specifically to manipulate the current cases filed but rather get at a overall trend in agency deference, the timing of *Chevron* is as good as random.

There will be however, selection on the cases that get challenged in the post-*Chevron* world which are not the same as the pre-*Chevron* world. Table 1 shows summary statistics of the total sample and the pre- and post-*Chevron* world. While it appears upon comparison of means that challenging went up among some challenging groups after *Chevron*, these means fail to account for the selection on the unobservable changes in incentives after *Chevron*. In addition, agencies might have changed how they interpret statutes post-*Chevron*. This is precisely the issue that muddles the estimation of *Chevron*.

A number of pending cases that were filed pre-*Chevron* and decided post-*Chevron* simulates a randomized trial. This estimation design allows me to place greater weight on these cases decided right around *Chevron* and lesser weight out on the tails of the time distribution. What I hope to answer is what the world would be like if incentives to challenge agencies hadn't changed post-*Chevron* and everybody acted the same before and after.

2.A Parametric v. Nonparametric Model Specification

I employ two strategies to estimate any discontinuity in agency deference, and both are well-outlined in Lee and Lemieux (2010). The first is to specify a flexible parametric model, and the second is to estimate the trend using nonparametric techniques. In the context of *Chevron*, I use both methods.

The choice between parametric and nonparametric specification is important if the underlying data follows a nonlinear trend as is the case with agency deference. To allow for function form flexibility, polynomials of varying order can be added to parametric specifications. This approach has been criticized for allowing too much weight on the entire sample of data while not placing enough weight on the local data around the cut off point (Lee and Lemieux 2010). In the context of *Chevron*, this is not a concern of particular severity because the lack of a large mass of local data.

The alternative approach is to specify the trend using nonparametric regression techniques. Nadaraya (1964) and Watson (1964) outline the procedure which sorts the data into bins based on proximity of the variable containing the trend (in this context, time until/after the *Chevron* ruling) then calculates, in each bin, a smoothed local mean. This allows the data to "speak for themselves" in such a way that is not bound by parametric specification. This idea has been extended by many, including Fan (1992) to specifying local polynomial regressions within each bin weighted by a kernel function.

Two potential shortfalls to nonparametric regression are that the results are often influence heavily by the choice of bandwidth (bin size) and that nonparametric regressions have poor boundary properties. There are available many options for kernel choice that improves the boundary performance of nonparametric regressions (Ludwig and Miller 2007), and the issue of bandwidth choice is discussed in great detail later.

Since it is unclear, a priori, if the gains from one technique outweigh the costs of another, I report both results.

3. Model

To measure the jump in deference, I estimate:

$$Y_{is} = \alpha R_{is} + m(T_{is}) + \varepsilon_{is} \tag{1}$$

where Y_{is} is an indicator variable for whether justice i in time s voted in favor of agency deference, R_{is} is an indicator variable taking on the value 1 if the vote cast by justice i at time s was decided after *Chevron* was decided, T_{is} is a measure of the number of days between the vote i at time s and the ruling in *Chevron*, and ε is random error. T_{is} is essentially a time trend centered around June 24th, 1984—the day *Chevron* was decided. The estimate of the impact of *Chevron* on agency deference is then the estimated value of α. I identify $m(\cdot)$ both parametrically and nonparametrically.

3.A Parametric Specification

Identified parametrically, $m(T_{is})$ becomes

$$m(T_{is}) = \beta_1 T_{is} + \beta_2 T_{is}^2 + \beta_3 T_{is}^3 + \ldots + \beta_k T_{is}^k \tag{2}$$

where polynomials of T_{is} are included until

$$t_{\hat{\beta}_{k+1}} < c_\alpha$$

where $t_{\hat{\beta}_{k+1}}$ is the t statistic on the *k+1*th order polynomial and *c* is the critical value at rejection level α.[51] To provide the most flexibility in this parametric specification, the exercise of adding additionally higher order polynomials until insignificant is completed separately on each side of the cutoff, so equation (1) becomes

$$\lim_{T_{is}\uparrow\tilde{T}} m(T_{is}) = \beta_1^L T_{is} + \beta_2^L T_{is}^2 + \beta_3^L T_{is}^3 + \ldots + \beta_k^L T_{is}^{k_L} \text{ for } T_{is} < \tilde{T} \tag{3.a}$$

and

$$\lim_{T_{is}\downarrow\tilde{T}} m(T_{is}) = \beta_1^R T_{is} + \beta_2^R T_{is}^2 + \beta_3^R T_{is}^3 + \ldots + \beta_k^R T_{is}^{k_R} \text{ for } T_{is} > \tilde{T} \tag{3.b}$$

where $\tilde{T}$ represented the day *Chevron* was decided and k_L and k_R represent the “optimal” polynomial order for the time trend leading into and out of *Chevron*, respectively.[52]

3.B Nonparametric Specification

Perhaps the chief concern in nonparametric regression is the choice of bandwidth (Racine 2008). While there are many methods to choose bandwidth selection, there yet remains a universally “optimal” bandwidth selection technique (Lee and Lemeuix 2010). The tradeoff that exists in bandwidth selection is that larger bins provide more accurate local estimates within each bin but may introduce bias from over-smoothing. In this paper, I use a common bandwidth choice technique, the Rule of Thumb (ROT) technique to calculate the center point of a range of possible bandwidths. The ROT bandwidth is defined as

---

[51] Throughout this paper, α=*.05*

[52] In the context of this paper, k_L=k_R, so both sides of the *Chevron* decision can be included in the same regression.

$$h_{ROT} = \kappa \left[\frac{\tilde{\sigma}^2 R}{\sum_{i=1}^{n} \{\tilde{m}''(T_{is})\}^2} \right]^{\frac{1}{5}} \tag{4}$$

where κ is a kernel-specific constant, *R*, is the range of the bin, $\tilde{m}''(\cdot)$ is the measure of the curvature of the local regression, and $\tilde{\sigma}^2$ is the estimated local variance. The ROT bandwidth provides a starting point bandwidth. I then estimate nonparametric regressions from the lower bound bandwidth of $bw_0 = h_{ROT} - \varepsilon$ to the upper bound bandwidth of $bw_X = h_{ROT} + \varepsilon$ at increments ϕ where ε is a number sufficiently large to cover all probable bandwidths.

Formally, the set of bandwidths used, *H*, is defined as

$$H \in [h_{ROT} - \varepsilon, h_{ROT} + \varepsilon] \tag{5}$$

with $\frac{2\varepsilon}{\phi} + 1$ number of unique bandwidths. [53]

4. Data

The data used in this analysis was originally collected and coded by Richards et al. (2006). They describe in detail how the data was collected. Essentially, they ran a legal search of all Supreme Court administrative law cases decided between 1969 and 2000 where there was a clear majority and *Chevron* was sited.[54] This excludes administrative cases to which *Chevron* does not directly apply. In addition, Richards et al. excluded tax and criminal cases, as they seem to represent a different aspect of agency adjudication.

[53] The “+1” comes from including both the upper and lower bound.
[54] Superior to this dataset would be one that contains appellate court rulings, at least in the DC Circuit. Unfortunately, this data is not currently available.

The coded variable of whether or not a judge voted for deference is the dependent variable, and an indicator variable stating whether or not the case was decided post-*Chevron* is our independent variable of interest. As controls, the authors employed methods common in the Political Science literature to account for political ideology of the judge and of the 'liberalness' or 'conservativeness' of the agency ruling. In addition, controls were included to account for the number of *amicus curiae* briefs filed, the challenging party, advocating party, whether or not the president can fire the head of the agency, whether or not the case is a rulemaking case, and the length of the underlying statute.

5. Results

Table 2 replicates the results reported in Richards et al. (2006). Column (1) replicates the logit coefficients. Column (2) reports the marginal effect of each logit coefficient in column (1), and column (3) reports the results of the analogous linear probability model (LPM) estimation. Richards et al. estimated the effect of *Chevron* to be 8.3 percentage points. The LPM estimation measures a similar post-*Chevron* effect, and in no covariate confuses the sign or significance found in the logit model.

Given the closeness of the LPM with the logit seen in Table 2, and the ability to control for heteroskedasticity using robust standard errors, there is no reason to believe the LPM will bias the results.

5.A Parametric Results

Table 3 reports the results of multiple parametric specifications of varying order polynomials. The "optimal" polynomial based on the criteria of equations 3.a and 3.b is a third order polynomial for both the pre- and the post-*Chevron* periods. Those results are reported in rows one and two of Table 3 with the estimated effect of 21 percentage points. Each even numbered row reports the results of the previous row including the controls outlined in Richards et al. to show the insensitivity of the results to the baseline covariates. In addition to the "optimal" order polynomial specification, Table 3 includes higher order polynomials as the parametric specification should be insensitive to higher order polynomials (Lee and Lemieux 2010).[55] Table 3 reports this to generally be the case, though the magnitude increases quite a bit when we move from a 4th to a 5th order polynomial. The parametric specification suggests that the enactment of the *Chevron* doctrine increased agency deference by 21 percentage points. The parametric estimate of *Chevron*'s effect can also be seen in Figure 1. The discontinuity surrounding *Chevron* can also be seen in Figure 1.

To generate the scatter points in Figure 1, each vote is ranked chronologically centered on the day *Chevron* was passed. Votes are then bunched into bins. Thus the vertical distance represents the proportion of votes for deference within each bin, and the horizontal distance represents the average number of days until or after the *Chevron* ruling. The size of each dot represents the mass of votes in each bin. For example, a dot with the coordinates (-365,1) would represent cases that happened on average a year before *Chevron* was decided and that all judges voted for deference. Each line represents the values of $\hat{Y}_{is}$, and the dashed vertical line represents the *Chevron* decision day.

5.B Nonparametric results

[55] In an alternative scheme to choose polynomial order, the "optimal" polynomial was chosen locally using only data contained in one bin pre- and post-*Chevron*, then repeated for two bins pre- and post-*Chevron* and so on. The idea was to find a good local fit to then apply globally. In general, the results of this exercise suggested a polynomial order someone between 3 and 6 which can be seen in Table 3. These results are available upon request.

The results specifying $m(T_{is})$ nonparametrically are reported in Table 4. At the ROT bandwidth of 148, the nonparametric estimate of *Chevron* is 18 percentage points. Standard errors of the jump are calculated using the calculations documented by McCrary (2008).[56] Previous research has found the choice of weighting scheme within each bin, or kernel, to not substantially affect the results (Lee and Lemieux 2010, Ludwig and Miller 2007). Throughout this paper, I use the Epanechnikov kernel to calculate local mean smoothed averages within each bin.[57]

Figure 2 displays the nonparametric equivalent of Figure 1 for measuring the effect of *Chevron*. Visual inspection of both Figures suggests that the nonparametric specification captures much more accurately the raw data represented by the binned scatter points although the nonparametric estimate coincides closely with the 21 percentage point effect estimated by the optimal parametric specification.

Both specifications suggest an effect much larger than previously estimated. In fact, off of a mean of 62% deference, we would expect to see judges, under the *Chevron* two-step, defer to agency rulings around 80% of the time.

As mentioned previously, the choice of bandwidth can greatly influence the results. To address this, Table 4 reports the nonparametric results of the effect of

[56] McCrary (2008) uses the following formula to calculate the standard error of the discontinuity:

$$\hat{\sigma}_\theta = \sqrt{\frac{1}{nh}\frac{24}{5}\left(\frac{1}{\hat{f}^+} + \frac{1}{\hat{f}^-}\right)}$$

where n is the number of observations, h is the bandwidth size, $\hat{f}^+$ is the estimated effect on one side of the break and $\hat{f}^-$ is the estimated effect on the other side (thus the jump is measured as $\theta = \hat{f}^- - \hat{f}^+$). $\frac{24}{5}$ is a constant derived in part from the specific kernel used (triangle) which changes slightly to $\frac{108}{25}$ when using the Epanechnikov kernel.

[57] *See* DesJardins and McCall (2008) for an example of using the Epanechnikov kernel. Local mean smoothing is used in place of local linear or local polynomial regressions because it provides a better fit to the raw data in this context.

Chevron using a wide range of bandwidths. Anchoring off the ROT bandwidth of 148 as the center point, I re-estimate the model using increasingly larger and increasingly smaller bandwidths as described previously. Creating a sufficiently large window of bandwidths around the ROT bandwidth should give us an idea of the sensitivity of the results to bandwidth selection. Table 4 reports a sample of bandwidths used that includes the minimum and maximum bandwidths used. A fuller picture of the effect of bandwidth selection can be seen in Figure 3.

A total of 101 different bandwidths were used. The horizontal axis of Figure 3 represents the value of each bandwidth while the vertical axis represents the estimated effect at each bandwidth size. The dashed line in the center is the ROT bandwidth. While there is some variation, the results are remarkably insensitive to bandwidth size even for small bandwidths. Given the concordance of the two specifications and the insensitivity to bandwidth size, it seems reasonable to conclude that the nonparametric specification is probably accurate.

6. Robustness

Inherent in the idea of capturing the effect of a change in law like *Chevron* is that the change is exogenous to outside factors. As mentioned previously, it is reasonable to believe that the *Chevron* ruling was made exogenously to the pending cases at the time, and as long as this is the case, there is no theoretical basis for a jump in any other observable characteristic. One way to test for this is, however, is to include as the dependent variable placebo outcomes and test for a jump at the onset of *Chevron*. Table 5 reports those results.

Each row represents a different nonparametric regression with a unique dependent variable. In each case, the ROT bandwidth was used, and the standard errors were calculated using the formula found in McCrary (2008). The majority of placebo regressions yield statistically insignificant jumps at the onset of *Chevron*. There is, however, a nontrivial amount of significant jumps reported in Table 5. Closer inspection would suggest that these jumps are a product of noise in the data as opposed to some sort of systematic pattern that would warrant concern.[58] The nonparametric specification for each placebo effect is displayed graphically in Figure 4 in the same manner as the previous figures.

Of particular interest is the nonparametric specification of judge ideology found in Figure 4b. Similar to previous figures, the vertical axis measures the average political leaning of each point in time. One potential benefit of *Chevron* is its ability to steer judges away from deciding cases according to their personal policy preferences (Strauss 1987). If this were in fact a problem, and *Chevron* fixed it, we should see an obvious trend in the pre-*Chevron* years and no discernable trend after. There appears to be, however a significant trend that persists through the cutoff suggesting first that political biases at the time of *Chevron*'s ruling is exogenous and second that the policy pushing framework proposed by Miles and Sunstein (2004) persists post-*Chevron*.

7. Conclusion

[58] For instance, the coefficients associated with the number of amicus briefs filed in favor and opposing the agency decision seem to contradict one another in that they are both positive. In addition, the estimated jump in the proportion of cases where the president can fire the agency head and the proportion of cases that are rulemaking cases report implausibly large coefficients suggesting probabilities close to or even greater than one. After discounting the jumps in these trends, only a few significant jumps in the placebo dependent variables remain—which is to be expected.

There are many likely reasons to believe that incentives play a role in the decision to challenge an agency ruling and that those incentives likely were altered with the passing of the *Chevron* two-step. This requires an estimator that corrects for the selection issues that occurred in the post-*Chevron* world. The simplest approach would be to exploit the cases that were filed prior to the *Chevron* ruling but decided after. Unfortunately, a small sample size around the *Chevron* ruling makes statistical precision difficult. The approach here is to use a large dataset of Supreme Court agency rulings with a weighting scheme that places greatest emphasis on the plausibly exogenous cases surrounding the day *Chevron* was decided. The jump in agency deference is then measured using two estimation techniques.

The first, a polynomial parametric specification, finds the *Chevron* effect to be a 21 percentage point increase in agency deference. The second technique is to specify the trend in agency deference nonparametrically. Nonparametric results suggest an 18 percentage point increase in agency deference. Previous studies have found a similar, albeit much smaller effect. The nonparametric results are quite robust to bandwidth specification, and the parametric results are robust to higher order polynomials. The insensitivity of each specification and their concordance in effect should give some insight to the effect of *Chevron.*

8. References

Czarnezki, J. (2009). "An Empirical Investigation of Judicial Decision-making, Statutory Interpretation & the Chevron Doctrine in Environmental Law." Marquette University Law School Legal Studies Research Paper Series Research Paper No. 07‐05.

Edwards, G., David Frisvold. (2010). "Full-Day Kindergarten and Student Achievement". Working Paper.

Fan, J. 1992. Design-adaptive nonparametric regression. *Journal of the American Statistical Association* 87: 998–1004.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001) "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201-209.

Jacob, B. & Lefgren, L. (2004). "Remedial Education And Student Achievement: A Regression Discontinuity Analysis." *The Review of Economics and Statistics*. 86(I): 226-244.

Lee, D. (2008) "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics*, 142(2), 675–697.

Lee. D. S., & T. Lemieux (2010) "Regression Discontinuity Designs in Economics," 48 *Journal of Economic Literature*.

Ludwig, J. & Miller, D. (2007). "Does Head Start Improve Children's Life Chances? Evidence From A Regression Discontinuity Design" *The Quarterly Journal of Economics*, February 122(1).

Manning, J. F., & M.C. Stephenson (2009), *Legislation and Regulation*. Foundation Press.

McCrary, J. (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142(2):698–714.

Miles, J., Sunstein, C. (2006). "Do Judges Make Regulatory Policy? An Empirical Investigation of *Chevron*" 73 *University of Chicago Law Review.*

Nadaraya, E. A. 1964. On estimating regression. *Theory of Probability and Its Application* 9: 141–142.

Racine, J.S. (2008), "Nonparametric Econometrics: A Primer," 3 *Foundations and Trends in Econometrics*: Vol. 3: No 1, pp 1-88.

Richards, M., Smith, J., & Kritzer, H. (2006). "Does *Chevron* Matter?" *Law & Policy*, 28(4).

Schuck P., & Elliott, E. (1990). "To the Chevron Station: An Empirical Study of Federal Administrative Law" *DUKE L.J*. 984.

Stephenson, M. C., & A. Vermeule (2009). "*Chevron* Has Only One Step." 95 Virginia Law Review 597.

Strauss, P. L. (1987). "One Hundred Fifty Cases Per Year: Some Implications of the Supreme Court's Limited Resources for Judicial Review of Agency Action" 87 *Columbia Law Review* 1093.

Thistlethwaite, Donald L. and Donald T. Campbell, "Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment," *Journal of Educational Psychology*, December 1960, 51, 309–317.

Matsudaira, J. (2008). “Mandatory Summer School and Student Achievement,” *Journal of Econometrics*, 142, 829-850.

Watson, G. S. 1964. Smooth regression analysis. Sankhya Series A 26: 359–372.

Table 1

Summary Statistics

	Full Sample		**Pre-*Chevron***		**Post-*Chevron***	
	mean	sd	mean	sd	mean	sd
Deference Vote	0.623	0.485	0.558	0.497	0.681	0.467
After Chevron	0.534	0.499	0.000	0.000	1.000	0.000
Judge Ideology	-0.196	0.674	-0.079	0.734	-0.298	0.599
Policy Direction of Agency Decision	0.060	0.954	0.122	0.957	0.006	0.949
Judge Ideology\Policy Direction Interaction	-0.017	0.669	-0.005	0.710	-0.027	0.632
Count of Amicus Briefs						
To Reverse Deference	1.520	2.499	1.164	1.871	1.819	2.892
To Support Deference	0.890	1.774	0.779	1.754	0.983	1.786
With Unclear Motives	0.242	0.690	0.187	0.438	0.289	0.849
Challenging Party						
Government	0.240	0.427	0.172	0.378	0.298	0.458
Corporation	0.393	0.489	0.405	0.491	0.382	0.486
Non-Corporation Interest Group	0.107	0.309	0.094	0.292	0.118	0.323
Individual	0.261	0.439	0.329	0.470	0.202	0.402
Advocating Party						
Agency Only	0.553	0.497	0.547	0.498	0.559	0.497
Agency with Co-Parties	0.206	0.405	0.224	0.418	0.190	0.393
Non-Agency	0.241	0.428	0.229	0.420	0.251	0.434

President Can Fire Agency Head	0.503	0.500	0.364	0.482	0.624	0.485
Rulemaking Case	0.604	0.489	0.508	0.500	0.688	0.464
Statute Length	72.285	89.897	61.790	58.395	81.443	109.508

Note: Data is at the judge/case level

Table 2

Replication of Richards et al. (2006) measures of the *Chevron*'s Effect on Judicial Voting for Agency Deference

Model **Coefficients Reported**	**Logit** **log odds** (1)	**Logit** **margins** (2)	**LPM** **margins** (3)
After Chevron	0.410‡	0.084‡	0.086‡
	(0.140)	(0.028)	(0.029)
Attitude	-0.111	-0.023	-0.021
	(0.098)	(0.020)	(0.020)
Agency Policy Direction	0.066	0.013	0.014
	(0.081)	(0.017)	(0.017)
Agency Policy Direction * Attitude	0.600‡	0.122‡	0.126‡
	(0.101)	(0.019)	(0.021)
Count of Amicus Briefs			
To Reverse Deference	-0.025	-0.005	-0.004
	(0.027)	(0.006)	(0.006)
To Support Deference	0.162‡	0.033‡	0.030‡
	(0.050)	(0.010)	(0.008)
With Unclear Motives	-0.174‡	-0.035‡	-0.035‡
	(0.062)	(0.013)	(0.013)
Type of Party Challenging Deference			
Corporation	0.909‡	0.186‡	0.197‡
	(0.188)	(0.037)	(0.039)
Non-corporation Interest Group	0.873‡	0.178‡	0.191‡
	(0.258)	(0.052)	(0.053)
Individual	0.960‡	0.196‡	0.206‡
	(0.205)	(0.041)	(0.042)
Type of Party			

Advocating Deference			
Agency	0.282	0.058	0.052
	(0.188)	(0.038)	(0.039)
Agency with Co-party	0.205	0.042	0.042
	(0.201)	(0.041)	(0.043)
President Can Fire Head	0.914‡	0.186‡	0.192‡
	(0.154)	(0.030)	(0.032)
Rulemaking Case	-0.512‡	-0.104‡	-0.106‡
	(0.137)	(0.028)	(0.028)
Statute Length	-0.003‡	-0.001‡	-0.001‡
	(0.001)	(0.000)	(0.000)
Sample Size	1174	1174	1174
Adjusted (or Pseudo) R-squared	0.1037	0.1037	0.119

Notes: Judge-Case clustered standard errors in parenthesis for columns (1) and (3) and delta-method computed standard errors in column (2). The dependent variable is an indicator variable signaling whether or not a judge voted for agency deference. Column (1) represents a replication of the findings in Richards et al. (2006), column (2) the marginal effects of each logit coefficient, and column (3) reports the analogous linear probability model. ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 3

Parametric Estimates of *Chevron*'s Effect on Judicial Voting for Agency Deference

	After Chevron	Std. Error	Poly Order	"Optimal" Poly	Includes Controls	N	Adjusted R^2
1	0.217†	(0.097)	3rd	X	N	985	0.055
2	0.210†	(0.100)	3rd	X	Y	952	0.143
3	0.289†	(0.128)	4th		N	985	0.054
4	0.283†	(0.127)	4th		Y	952	0.142
5	0.491‡	(0.164)	5th		N	985	0.055
6	0.426‡	(0.168)	5th		Y	952	0.142
7	0.350	(0.217)	6th		N	985	0.060
8	0.219	(0.223)	6th		Y	952	0.156

Notes: Judge-Case clustered standard errors in parenthesis. The dependent variable is an indicator variable signaling whether or not a judge voted for agency deference. Each row represents a different regression including all polynomials of the order less than equal to the value specified in column 4. The even numbered rows represent each order polynomial specification repeated to include the controls used in Richards et al. (2006). The "optimal" polynomial order is discussed in the Model section. ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 4

	Range of Bandwidths						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
After Chevron	0.130‡	0.149‡	0.161‡	0.184‡	0.199‡	0.209‡	0.182‡
	(0.0012)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Bandwidth Size	98.88	114.88	131.88	148.88	164.88	181.88	198.88
Rule of Thumb Bandwidth				X			

Notes: Standard errors in parenthesis. Each estimate is derived from a local mean-smoothing regression at varying bandwidths. The dependent variable is an indicator variable signaling whether or not a judge voted for agency deference. Column (4) represents the Rule of Thumb bandwidth. Each estimate uses the Epanechnikov kernel. ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 5

Non-Parametric Estimates of *Chevron*'s Effect on Placebo Outcomes

	After Chevron	Std. Error	Bandwidth
Judge Ideology	0.011	(0.024)	177.77
Policy Direction of Agency Decision	-0.752	(0.025)	107.58
Judge Ideology\Policy Direction Interaction	0.121‡	(0.047)	179.61
Count of Amicus Briefs			
To Reverse Deference	0.753‡	(0.012)	84.18
To Support Deference	1.230‡	(0.022)	108.46
With Unclear Motives	-0.042	(0.020)	209.91
Challenging Party			
Government	0.404‡	(0.034)	93.26
Corporation	-0.644	(2.798)	120.25
Non-Corporation Interest Group	0.000	(4.116)	111.10
Individual	0.262‡	(0.022)	131.10
Advocating Party			
Agency Only	-0.083	(0.014)	185.10
Agency with Co-Parties	-0.124	(3.026)	102.81
Non-Agency	0.278‡	(0.023)	120.13
President Can Fire Agency Head	0.786‡	(0.022)	106.54
Rulemaking Case	0.327‡	(0.017)	114.80
Statute Length	-55.779	(0.002)	116.83

Notes: Standard errors in parenthesis. Each estimate is derived from a local mean-smoothing regression at varying bandwidths. Each row represents a different regression where the first column is the dependent variable. Bandwidth was chosen using the Rule of Thumb, and each estimate uses the Epanechnikov kernel. ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Figure 1

Parametric Specification of *Chevron*'s Effect on Agency Deference

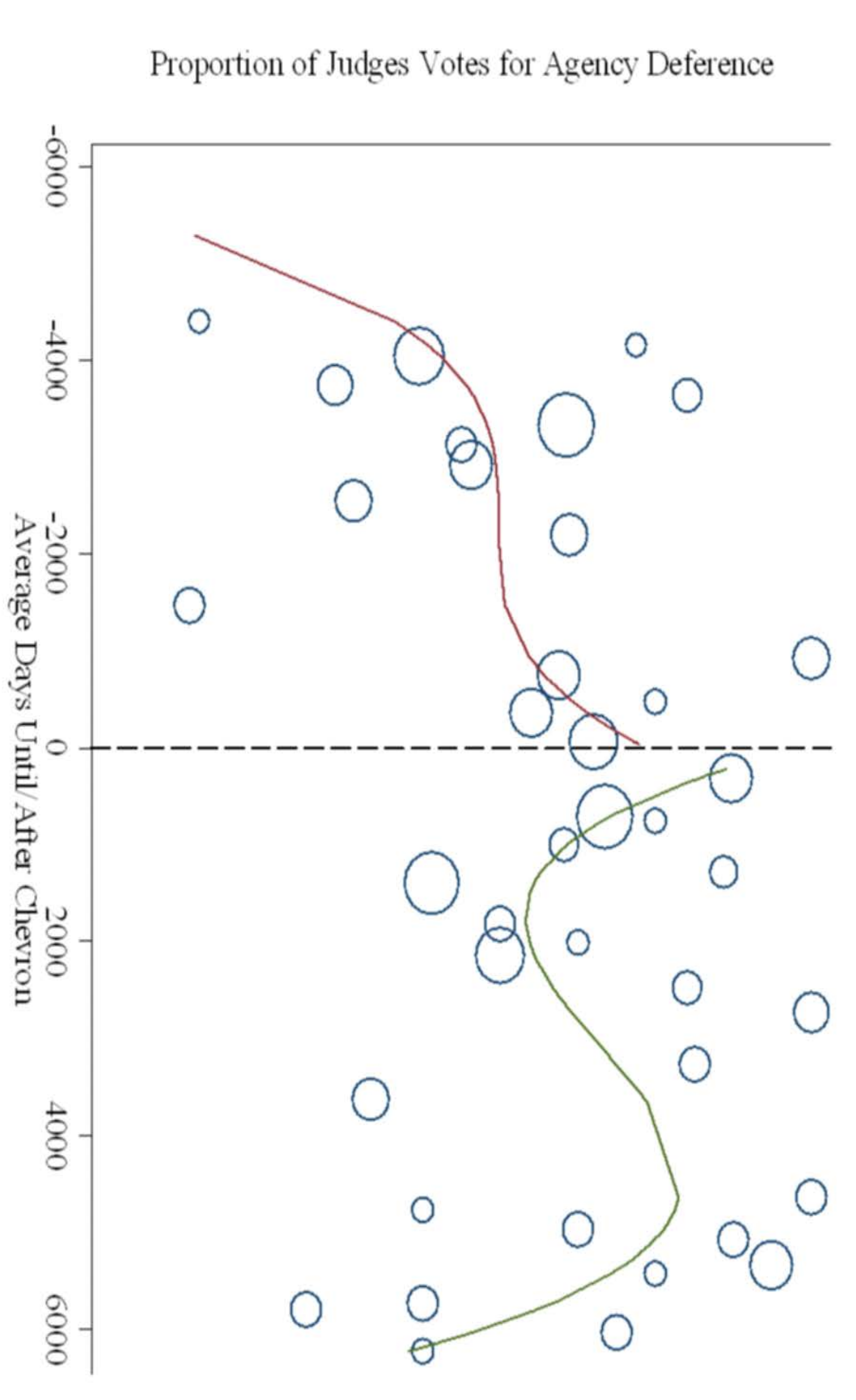


Figure 2

Nonparametric Specification of *Chevron*'s Effect on Agency Deference

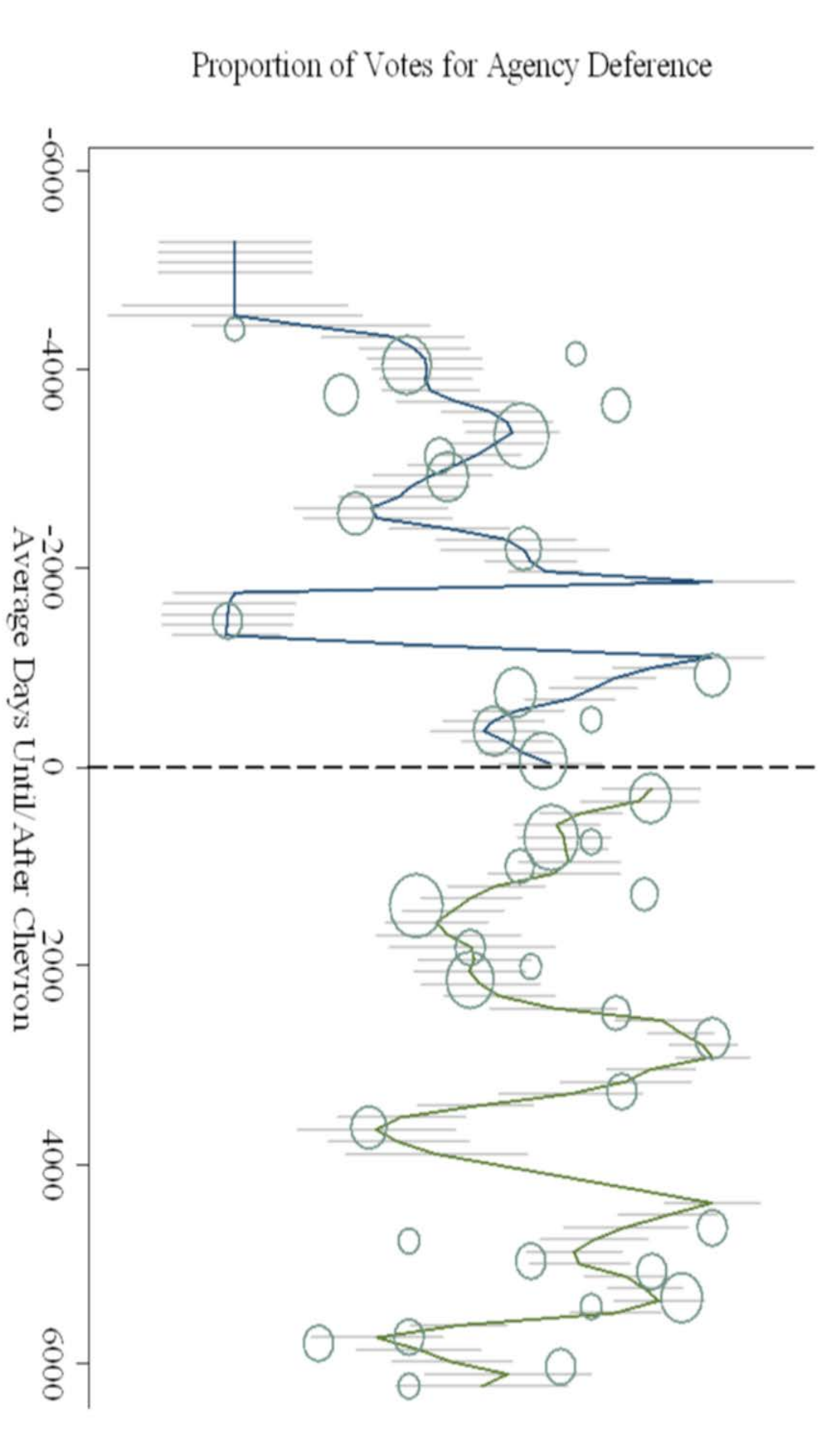


Figure 3

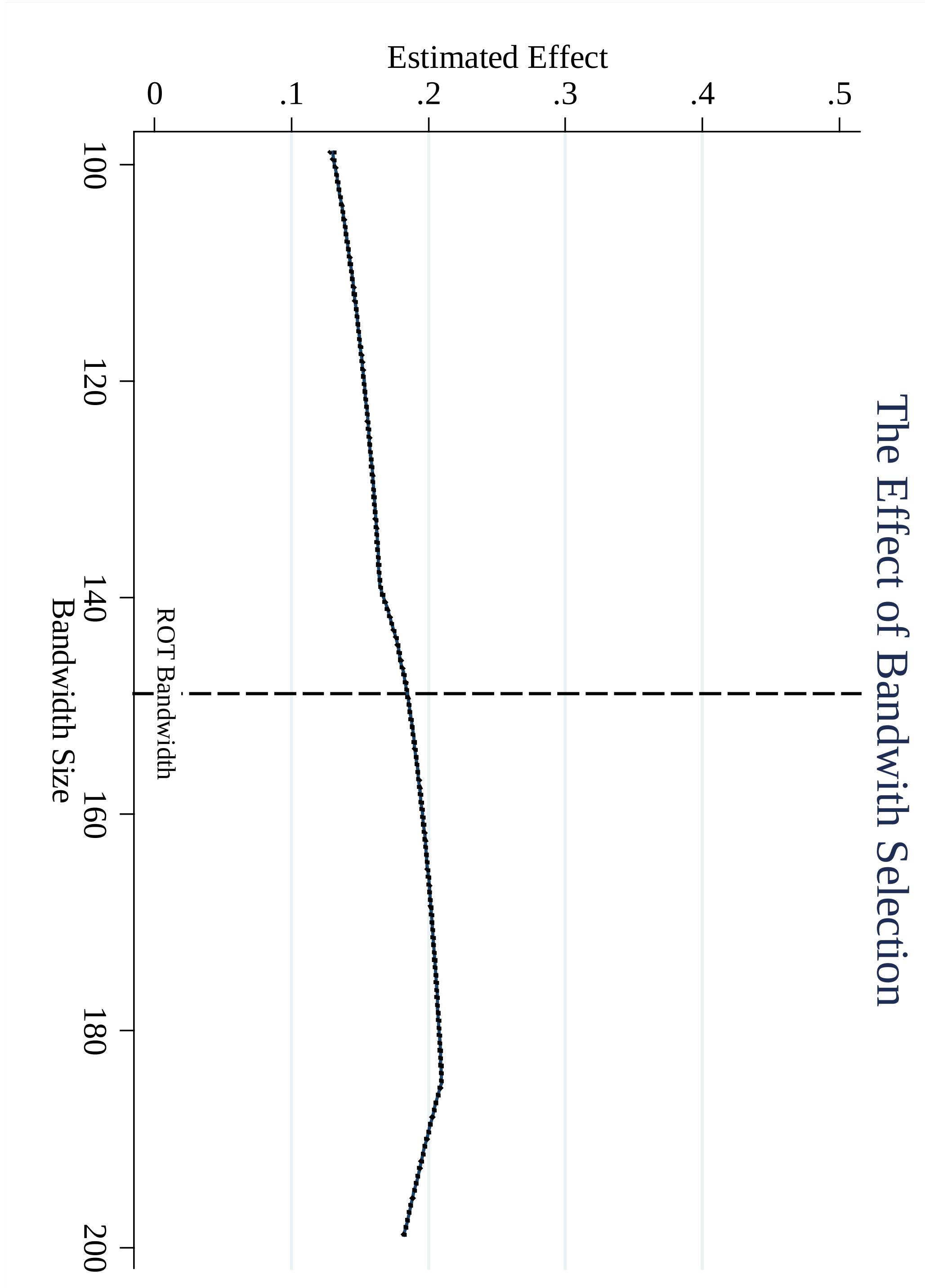

The Effect of Bandwith Selection
Estimated Effect
0
.1
.2
.3
.4
.5
Bandwidth Size
100
120
140
160
180
200
ROT Bandwidth

Figure 4.a

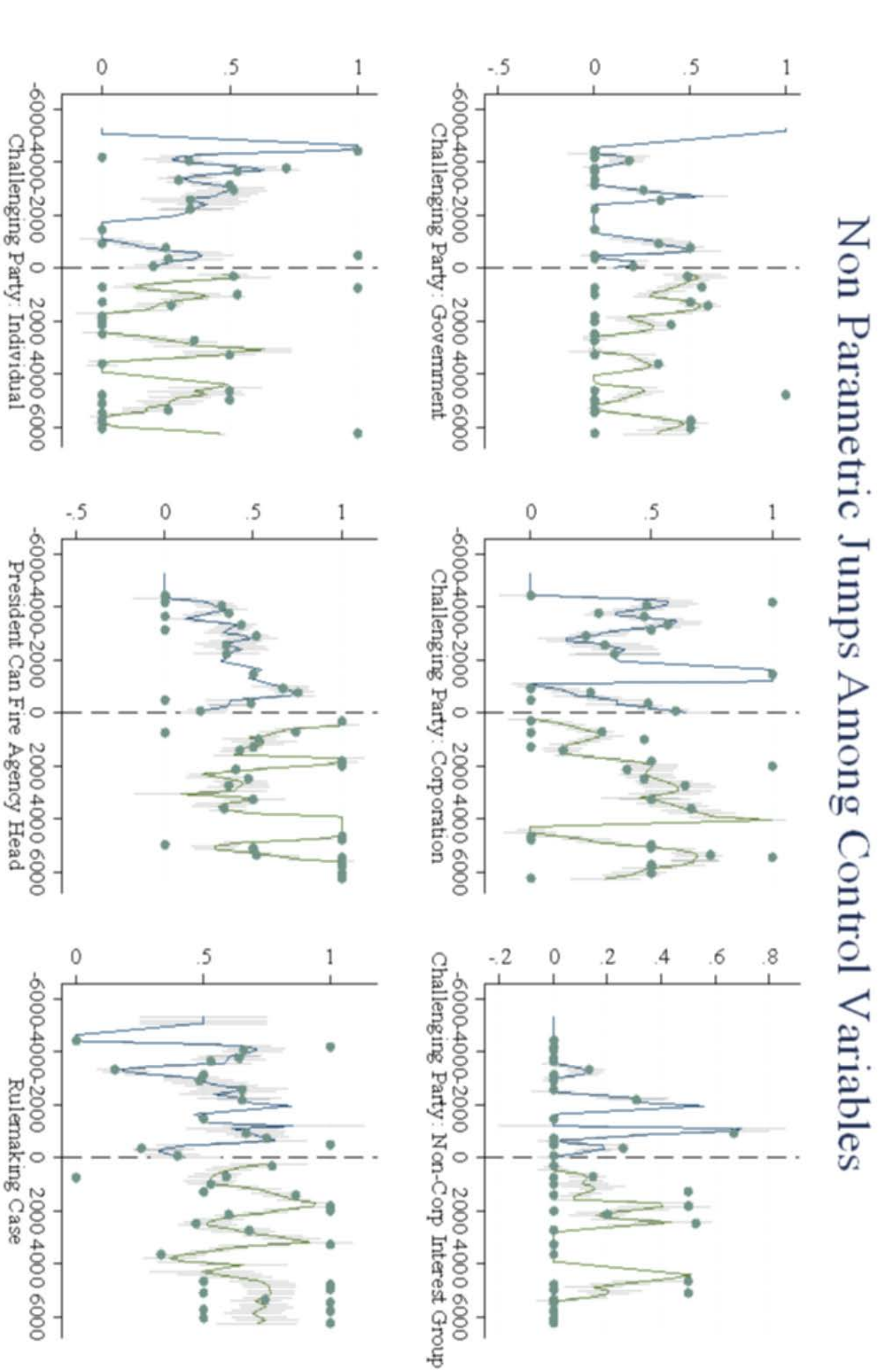

Non Parametric Jumps Among Control Variables
Challenging Party: Government
Challenging Party: Individual
Challenging Party: Corporation
President Can Fire Agency Head
Challenging Party: Non-Corp Interest Group
Rulemaking Case

Figure 4.b

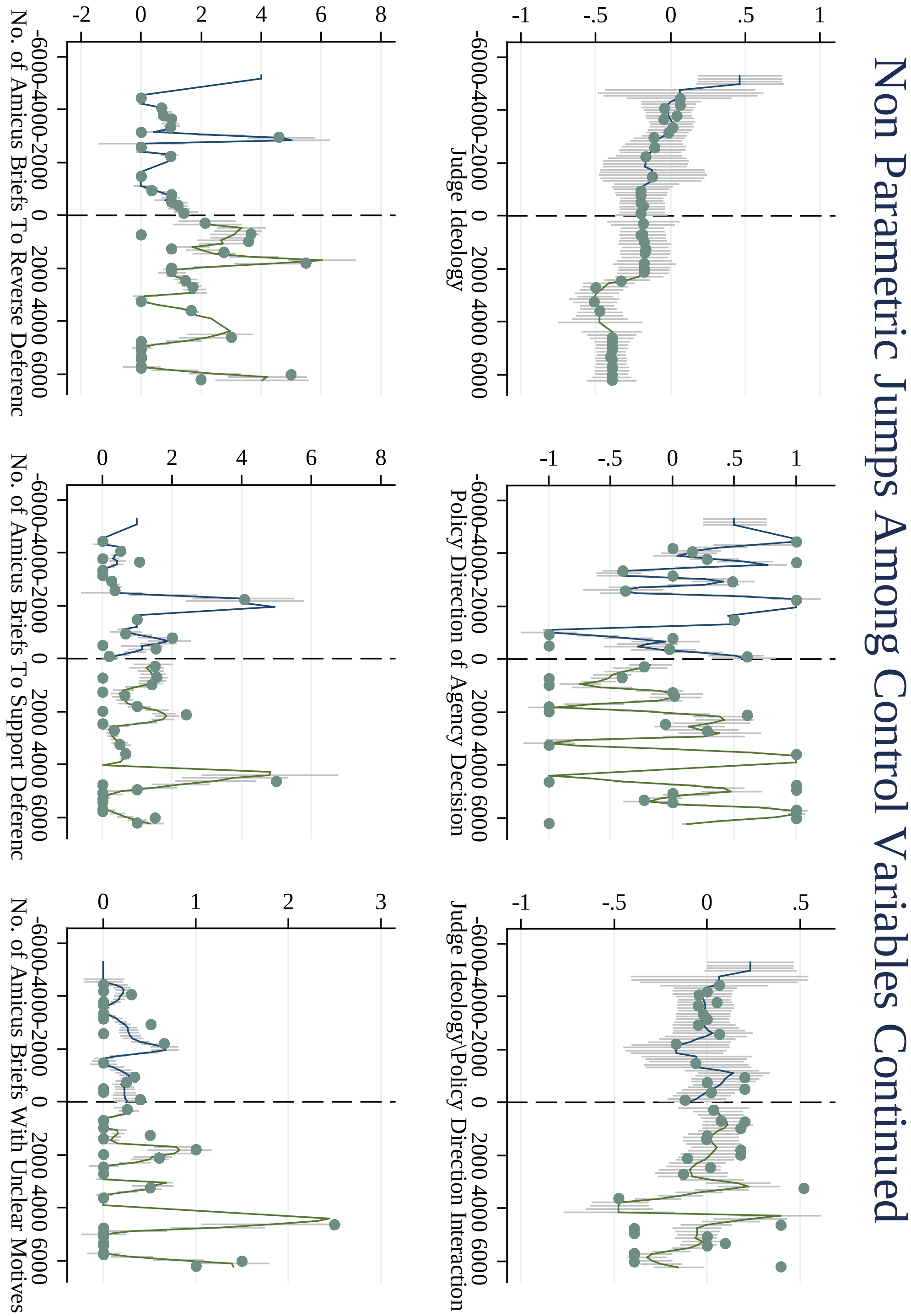
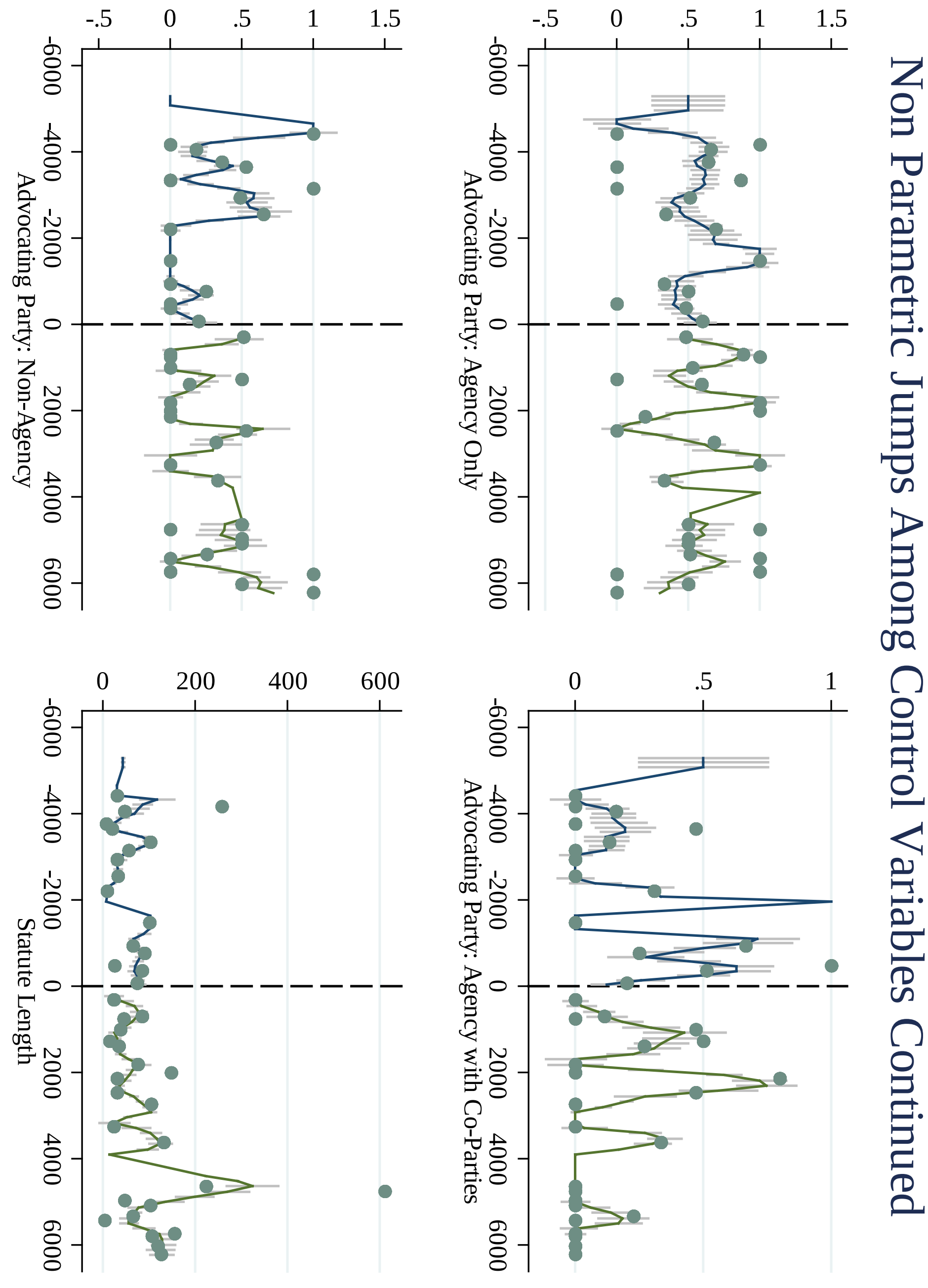

Non Parametric Jumps Among Control Variables Continued
Judge Ideology
Policy Direction of Agency Decision
Judge IdeologyPolicy Direction Interaction
No. of Amicus Briefs To Reverse Deferenc
No. of Amicus Briefs To Support Deferenc
No. of Amicus Briefs With Unclear Motives

Figure 4.c


Non Parametric Jumps Among Control Variables Continued
Advocating Party: Agency Only
Advocating Party: Agency with Co-Parties
Advocating Party: Non-Agency
Statute Length
-.5
0
.5
1
1.5
200
400
600
-6000
-4000
-2000
0
2000
4000
6000

The Power of the Racketeer: An empirical approach

Recently economists have attempted to explain the nature of incentives aligned with corruption. Much of the understanding of corruption in business and government comes form survey data (LaPorta et. al 1999) and suffers the standard shortcomings. In particular, surveys collecting data on illegal activities are especially plagued with bias. In general, people engaging in illegal activity go to great lengths to ensure the privacy of their activities. Thus, it is unclear what incentives would arise to motivate honest participation in a survey. Glaeser and Saks (2006) attempted to circumvent the plaguing issues of surveys by employing a state panel data set on the number of corruption convictions against government employees and found that in general corruption in state governments is more likely to occur in poor, less educated states. Here, I use a similar approach to measure the effect of corruption in a different realm.

The literature on corruption dates back to Becker's (1968) model of crime. Ehrlich (1972) pioneered the use of empirical methods in modeling illegitimate activities by explaining specific crime rates as a two step function of the probability of getting caught, and recent work in the field has attempted to model the criminal firm, their costs and choice of "goods" (Dick 1995), structure (Cheng et. al. 2005) , and payoffs. A recent working paper by Skarbek (2009) attempts to model the ability of incarcerated individuals to extort non-incarcerated criminals using information gathered about the Mexican Mafia, but to date, there has been very little work done empirically to test any of the theories proposed. One of the earliest works in the field by Rubin (1973) motivates the analysis of this paper. Many scholars have modeled corruption by organized criminal behavior as a monopoly, or monopolistic competition (see Schelling 1967, Buchanan 1973 & Garoupa 2000). Rubin uses this framework and extends it to the criminal act of extortion and racketeering.

Alexander (1997) gives an example of rackets that took place in the depression era Chicago market for pasta. Organized criminal firms would "sell" their labels to be put on the pasta packaging. The cost of the labels (25 cents per pound) acted as a per unit tax, and if the labels were not purchased, the pasta companies suffered violently.

Racketeering is thus defined as the use of illegal activities[59] to receive income[60]. In essence, racketeering is an attempt at a cartel, or loose combination, in which the racketeer earns the monopoly profit. The work of racketeering has been mainly the work of organized criminal units, as they are able to exploit scale economies in areas such as risk and enforcement of collections (Dick 1995).

Rubin (1973) theorized that organized criminal firms are able to extort rents through rackets. If a legitimate firm operates in a competitive market, any additional "tax", or fee the legitimate firm must pay to the criminal firm, will cause the victim firm's prices to rise above marginal cost and thus force the firm out of business. The criminal firm has two choices for extracting monopoly rents: It must charge this "tax" only on firms with market power, or charge the tax on every firm in the industry. As victims are usually few, the likely case is that of individual firms with some degree of market power. The racketeering firm will extract monopoly rents from the firm up to the point of zero normal profits. Rubin (1973) theorizes that monopolistic firms targeted by racketeering organizations are concentrated in areas with high populations of non-native speakers and have a natural monopoly created by the language of the goods and services offered.

A store offering service in Estonian and products from Estonia will have monopoly power over immigrants from Estonia, for example. Advances in transportation should allow for competitive pricing among English speaking firms, but have no effect on firms specializing in non-English language services. Thus, we should expect that a higher percentage of non-English speakers in a state should cause a greater amount of racketeering activity. Figure 1 shows the flow of events. Empirically I am able to measure the first step, the number of non-English speakers, and I am able to approximate the last step by collecting data on the count of racketeering indictments. The middle step, however, is more difficult to estimate. Without knowing the cost structure of the victim firm, I am unable to distinguish monopoly rents from normal profits. I can, however, measure the effect the presence of small businesses has on racketeering activity. Fewer small businesses in a state would suggest less competition and a higher probability of

---

[59] *See* 18 U.S.C. § 1961 for a list of what constitutes illegal activities.
[60] *See* 18 U.S.C. § 1962 for a more detailed description of what it means to receive income.

obtaining monopoly rents. Rubin (1973) suggests that where monopoly rents are greater, there should be an increase in racketeering activity to extract those rents.

This paper attempts to add to the literature of corruption and organized crime by using state level panel data to explain the variation in racketeering charges and convictions in the United States as a function of a state's language ability. The findings confirm the theory of Rubin (1973) that an increase in non-English speakers in a state will cause an increase in the level of racketeering activity and are supported by the results that find that increases in the amount of small businesses per state lead to decreases in monopolist rents and racketeering activity.

II. Model

The model I estimate reflects the count nature of the response variable. For number of rackets, *y*, in state *i* and year *t*, we estimate the following count model in (A1):

$$E[y_{it} \mid p_{it}, sb_{it}, D_{it}, S_i, EI_{it}, d_t, c_i] = m(\alpha + \delta p_{it} + \sigma sb_{it} + \sum_{q \in \Phi} \beta^q D_{it}^q + \varphi S_i + \sum_{p \in \Theta} \psi^p EI_{it}^p + n_t + c_i)$$

where *p* is the percent of state *i* in time *t* that self-reports not being able to speak any English, *sb* is the per capita amount of small businesses per state per year, *D* is *q* demographic variables indexed by state and year, *EI* is *p* economic indicators by state and year, *S* is a variable indicating whether or not state *i* has state racketeering laws, *c* represents a time-invariant fixed effect for state *i* as outlined in Hausman, Hall, & Griliches (HHG) (1984), and *n* a time dummy to allow for flexibility in *m(·)*. The model proposed by HHG (1984) suits this issue well as it allows for dependence between the error term, *u*, and *c*. Following the pattern set by HHG (1984), we will allow *m(·)* to take the form *exp(·)* and use quasi-maximum likelihood (QML) estimation.

Gourieroux, Monfort, and Trognon (GMT) (1984) prove that the QML estimator is consistent and asymptotically normal when the Poisson model is correctly assumed, and Wooldridge (1999) shows that only correct specification of the condition mean, (A1) is necessary for consistency. This allows the model to be distributionally misspecified, without regard for over or under dispersion, and still get consistent and efficient results. This method is especially useful in this study, as it allows for robust standard errors clustered at the state level (Wooldridge 1999).

Additionally, the count model specified in (A1) might be misspecified if current racketeering activity is actually a function of previous year rates of non-English speakers possibly caused by delays required for authorities to build a case. To address this, non-English speaker rates are included in the models reported in Table 2 both as levels and lags.

III. Data

Racketeering activity is measured by the number of racketeering charges per state per year from the Executive Office for U.S. Attorney's and covers the years 1994 to 2007. This data represents the number of federal cases filed under the *Racketeering Influenced and Corrupt Organizations* Act (here on out referred to as RICO and unless otherwise noted, RICO implies federal RICO) by a federal prosecutor. As is the case in most empirical legal work, there is an issue of case selection. In every instance of federal indictment, there are many possible avenues of selection that could potentially bias the results. Decisions made federal prosecutor could introduce case selection bias. In this analysis, case selection will likely not play a biasing role as long the selection issues are random across states.

Using RICO charges might be problematic because of case selection issues, but also because it is unclear whether charges data over- or underestimates actual racketeering activity. It could be that when building a case against a suspect, a prosecutor might string together a plethora of charges against the suspect with the intention of using the large quantity of charges as a way to force a plea bargain. Prosecutors might also charge a suspect with much more heinous crimes than actually committed as a mechanism to force a plea bargain. If either of these mechanisms is used in practice, RICO charges will be an overestimate of the level of racketeering activity in a state.[61]

Charges data could also actually be an underestimate as the best criminals select out of the sample by simply not getting caught. It is likely the case that the most elite of criminal firms run successful rackets without getting caught. This would suggest that the count of racket charges is indeed an understatement of racketeering activity. While

[61] The extent to which a measure of RICO charges over estimates the actual level of racketeering activity in state depends on the degree by which they plea mechanisms are employed.

charges data might be suspect to bias, there is reason to believe the bias pulls in both directions.

Table 1 contains the summary statistics for RICO charges and all other variables included in the model. Other variables included to explain the variation in RICO charges included state demographic variables that were complied from Census data, with linear interpolation for years that fall between Census years. Other explanatory variables, including the variable of interest comes from IPUMS (Ruggles et al. 2008). Demographic controls include race, sex, real median income, and age. In addition GDP and unemployment rates of each state come from the Bureau of Labor Statistics and the Bureau of Economic Analysis. Additionally, in an effort to explain criminal activity, each model includes homicide rates from the Uniform Crime Report, and state police, judicial and corrections expenditures from the FBI, and an index of crack cocaine activity adopted from Levitt (2004).

To capture the political atmosphere of each state that might influence RICO charges, each model includes the share of the state house and senate that is democrat, governor's political party, judicial selection methods and judicial retention methods which all come from Shepherd (2006). Data on the number of small businesses per state per year comes from the Census Statistics of U.S. Businesses database. This data represents the per capita count of businesses with less than 10 employees on payroll on March 12th of each year. Much of the RICO law deals with issues surrounding immigration. To control for that, I include in each model the per capita amount of immigrants obtaining legal permanent residence by state as reported by the Department of Homeland Security and a dummy variable indicating the post-September 11th era of the dataset.

The explanatory variable of interest was constructed from a census question asking the individual their English speaking ability. This question is a follow up to a question about the native tongue spoken at home. Since this analysis aims to measure the effect people who don't speak the language of the land has on racketeering activity, and though there are some good proxies for somebody who doesn't speak English (such as foreign born, or language spoken at home) the best has to be a self-reported measure of English speaking ability. The choices for the question addressing English speaking ability are: *Very well; Well; Not well; Not at all*. Due to years of observing people over estimating their foreign language skills, I include in the variable that measures the percent of the population of a state that does not speak English those who report *Not at all* and also those who report *Not well*. This again, accounts for how we generally over estimate our ability to speak a non-native language. As a robustness check, however, I report in Table 3 the effect of the proportion of a state that reports only *Not at all*. This serves to show that the results are insensitive to small changes in the variable of interest, and allows for comparison of the magnitude.

The QML Poisson estimator allows for the inclusion of state and year fixed effects which hopefully captures any unobserved heterogeneity. Figure 2 displays the heterogeneity of charges across states. Certain costal states act as a large hubs for inflowing non-English speakers (California, New York, Florida), while other costal states (Oregon, Maine, Delaware) are not. These effects should get picked up by the state effects in the model. Also, the fixed effect design will nicely control for unobservable heterogeneity caused by courts interpretation of RICO statutes, and heterogeneity of juries.

Another time-invariant factor that will influence the amount of racketeering activity in a state is the presence of state RICO laws. In the seventies and eighties many states adopted state statutes modeled after federal RICO laws. These state laws, however, are much less frequently used and generate few cases compared to their federal counterpart (Floyd 1998). This may be, in part, due to the lack of state level policing agencies that act in similar manner to the FBI.

IV. Results

Table 2 displays the results with the dependent variable the count of RICO charges per state per year. Column (1) reports the coefficient estimate of the QML regression and column (2) reports the marginal effect. To account for the possibility that racketeering charging in year *t* are a function of foreign language speakers in year *t-1*, columns (3) and (4) repeat the models estimated in columns (1) and (2) respectively but with the foreign language variable lagged. The results in Table 2 suggest that a 0.1 percentage point increase in the proportion of the state that speaks little to no English causes an increase in racketeering cases charge of between 0.8 (level) and 1.2 (lagged) cases per state per year.

The marginal effect of non-English speakers on racketeering activity speaks the effect of step 1 on step 3 of Figure 1. To further the claim, an increase in the presence of small businesses decreases the amount of racketeering activity though statistically insignificant. This is consistent with the notion that greater amounts of small businesses leads to more competition among small businesses and thus fewer monopoly rents to be exploited.

An interesting and significant result is the relationship between legalized immigrants and racketeering activity. Through all specifications, the relationship is significant and negative. One possible explanation is that of human capital accumulation. It could be that legalized immigrants add to the pool of bilingual residents of a state (since reaching permanent residence status implies some mastery of the native tongue) that could potentially aid non-English speakers in translating business transactions thus making monopoly profits in this specific way smaller and less attractive to criminal firms.

Table 3 replicates Table 2 except that the independent variable of concern measures the proportion of each state that reports not speaking English at all, as opposed to that of Table 2 that reports the proportion of each state that either speaks English poorly or not at all. The results in Table 3 are similar to those in Table 2 except that the effect is much bigger in Table 3. This result provides further evidence of the relationship between foreign language speakers and racketeering activity because Table 3 captures the pool of people that would be most dependent on goods and services offered in a foreign language. [62]

V. Conclusion

This paper finds that all else equal, a 0.1 percentage point increase in the non-English speaking population will cause and increase of about 1 racket per state per year. This supports the theory that criminal firms are able to extract monopoly rents from businesses offering goods and services in foreign languages. Additionally, this analysis reports an interesting negative relationship between permanent residence granted and

[62] In addition, the results are not sensitive to outlying states. Dropping the states with the largest amount of RICO activity—New York and California—does not significantly alter the results.

racketeering. These results are fairly robust to model specifications and variations in data and are supported by the notion that the states with greater amounts of small businesses, and consequently less monopolies, experience less racketeering activity. This contributes to the existing understanding of corruption and organized crime in explaining in part how a criminal firm might extract rents from a victim firm. This happens when a victim firm gains a monopoly by offering goods and services in a language that is uncommon to the land. Criminal extortion can be maintained because the victim firm can extract monopoly rents which will be paid as "taxes" to the criminal firm.

References

Alexander, B. "The Rational Raketeer: Pasta Protection in Depression Era Chicago," XL *Journal of Law and Economics* 175-202, 1997.

Becker, G. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (March/April 1968): 169–217.

Buchanan, J.M. ''A Defense of Organized Crime?'' in The Economics of Crime and Punishment, edited by S. Rottenberg. Washington, DC: American Enterprise Institute, 1973, 119–32.

Chang, J. Lu, H., Chen, M. 2005. "Organized Crime or Individual Crime: Endogenous Size of a Criminal Organization and the Optimal Law Enforcement" *Economic Inquiry* 43(3): 661-675.

Dick, A. R. ''When Does Organized Crime Pay? A Transaction Cost Analysis.'' *International Review of Law and Economics*, 15(1), 1995, 25–45.

Floyd, John E. *RICO State by State: A Guide to Litigation Under the State Racketeering Statutes* (American Bar Association, Section of Antitrust Law, 1998)

Garoupa, N. ''The Economics of Organized Crime and Optimal Law Enforcement.'' Economic Inquiry, 38(2), 2000, 278–88.

Glaeser, Edward L., Saks, Raven E. 2006. "Corruption in America." *The Journal of Public Econoimcs.* 90: 1053– 1072.

Grossman, Gene M., Katz, Michael L. 1983. "Plea Bargaining and Social Welfare." *The American Economic Review.* 73(4): 749-757.

Gourieroux, C., Monfort, A., Trognon A., "Pseudo Maximum Likelihood Methods: Theory," Econometrica, 52(1984), 681-700.

Hausman, J.A., Hall, B.H., Griliches, Z., 1984. “Econometric models for count data with an application to the patents-R&D relationship.” *Econometrica* 52, 909-938.

LaPorta, Raphael, Lopes-de-Silanes, Florencio, Shleifer, Andrei, Vishny, Robert, 1999. “The Quality of Government.” *Journal of Law, Economics, and Organization.* 15 (1), 222–279.

Levitt, Steven D. 2004. “Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not.” 18 *Journal of Economic Perspectives* 163-90.

McCoy, Thomas R., Mirra, Michael J. 1980. "Plea Bargaining as Due Process in Determining Guilt," *Stanford Law Review*. 32: 887-941.

Reignanum, Jennifer F. 1988. “Plea Bargaining and Prosecutorial Discretion”. *The American Economic Review*. 78(4): 713-728.

Ruggles, S., Sobek, M., Alexander, T., Fitch, C., Goeken, R., Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series: Version 4.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2008.

Rubin, P. H. ‘‘The Economic Theory of the Criminal Firm,’’ in *The Economics of Crime and Punishment*, edited by S. Rottenberg. Washington, DC: American Enterprise Institute, 1973, 155–66.

Shepherd, Joanna. 2009. “The Influence of Retention Politics on Judges’ Voting” 38 *The Journal of Legal Studies* 169-204.

Schelling, T. S. ‘‘Economics and the Criminal Enterprise.’’ *Public Interest*, 7, 1967, 61–78.

Skarbek, D. “Criminal Extortion.” George Mason working paper. 2009.

Wooldridge, Jeffrey, “Distribution-Free Estimation of Some Nonlinear Panel Data Models,” *Journal of Econometrics*, Vol. XC, 1999

Table 1

Summary Statistics

	Mean	Min	Max
Count of Rackets	18.85	0.00	1052.00
	(74.26)	.	.
Percent of State that Speaks Little to No English	0.022	0.00	0.11
	(0.019)	.	.
Number of Small Businesses Per Capita	16.92	12.32	27.20
	(3.10)	.	.
Number of Legal Immigrants Per Capita	2.15	0.21	9.34
	(1.73)	.	.
Post-September 11th	0.43	0.00	1.00
	(0.50)	.	.
Homicide Rate	7.74	0.28	86.59
	(9.22)	.	.
Crack Index	4.94	-1.90	76.90
	(7.06)	.	.
Democrat Share of State House	0.52	0.23	0.89
	(0.11)	.	.
Democrat Share of State Senate	0.52	0.20	0.91
	(0.11)	.	.
Real Median Income	31207.90	21387.14	44946.74
	(4440.59)	.	.
Unemployment	4.81	2.30	8.70
	(1.20)	.	.
Real Police Expenditures Per Capita	121.81	42.51	396.74
	(65.38)	.	.
Real Judicial Expenditures Per Capita	51.44	22.27	179.49
	(15.53)	.	.
Real Corrections Expenditures Per Capita	98.37	-0.99	364.37
	(49.67)	.	.
Real GDP	107072.90	9556.43	757031.80
	(126290.70)	.	.

Median Age	37.03	26.40	45.00
	(3.08)	.	.
Percent Hispanic	0.07	0.00	0.44
	(0.09)	.	.
Percent Black	0.10	0.00	0.65
	(0.11)	.	.
Percent Male	0.48	0.45	0.52
	(0.01)	.	.
State RICO Law	0.59	0.00	1.00
	(0.49)	.	.
Year	.	1994	2007

Table 2

QML-FE Estimation of the Effect of Little to No English Speakers on the Count of RICO Charges

dependent variable: Count of RICO Charges	(1) Coefficient	(2) Margin	(3) Coefficient	(4) Margin
Little to No English	42.442^	848.120^	59.853‡	1219.626‡
	(23.364)	(466.892)	(22.491)	(458.297)
Small Businesses Per Capita	-0.210	-4.192	-0.171	-3.480
	(0.185)	(3.692)	(0.195)	(3.966)
Immigrants Per Capita	-0.112^	-2.240^	-0.144†	-2.927†
	(0.067)	(1.332)	(0.072)	(1.465)
Log Psuedolikelihood	-5219.336	.	-4906.244	.
Sample Size	653.000	653.000	602.000	602.000
Level or Lag	Level	Level	Lag	Lag

Notes: Robust standard errors are in parenthesis and clustered at the state level. All models include population with the coefficient constrained to 1. Columns (2) and (4) report the marginal effects of the models estimated in columns (1) and (3) respectively. Columns (3) and (4) account for the possibility that current racketeering counts are a function of the previous years non-native speaker rates. The dependent variable is the count of RICO charges per state as measured by the EOUSA. Each model contains state and year fixed effects, and political and demographic controls. ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Table 3

QML-FE Estimation of the Effect of Non-English Speakers on the Count of RICO Charges

dependent variable: Count of RICO Charges	(1) Coefficient	(2) Margin	(3) Coefficient	(4) Margin
No English Spoken	150.306‡	3003.586‡	162.950‡	3320.436‡
	(58.096)	(1160.941)	(45.684)	(930.913)
Small Businesses Per Capita	-0.150	-2.989	-0.097	-1.974
	(0.177)	(3.530)	(0.185)	(3.776)
Immigrants Per Capita	-0.137^	-2.747^	-0.168†	-3.415†
	(0.075)	(1.506)	(0.074)	(1.498)
Log Psuedolikelihood	-5153.914	.	-4874.045	.
Sample Size	653	653	602	602
Level or Lag	Level	Level	Lag	Lag

Notes: Robust standard errors are in parenthesis and clustered at the state level. All models include population with the coefficient constrained to 1. Columns (2) and (4) report the marginal effects of the models estimated in columns (1) and (3) respectively. Columns (3) and (4) account for the possibility that current racketeering counts are a function of the previous years non-native speaker rates. Table 3 differs from Table 2 in the exclusion of those who report speaking English "not well". The dependent variable is the count of RICO charges per state as measured by the EOUSA. Each model contains state and year fixed effects, and political and demographic controls. ^ $p<0.10$ † $p<0.05$ ‡ $p<0.001$

Figure 1

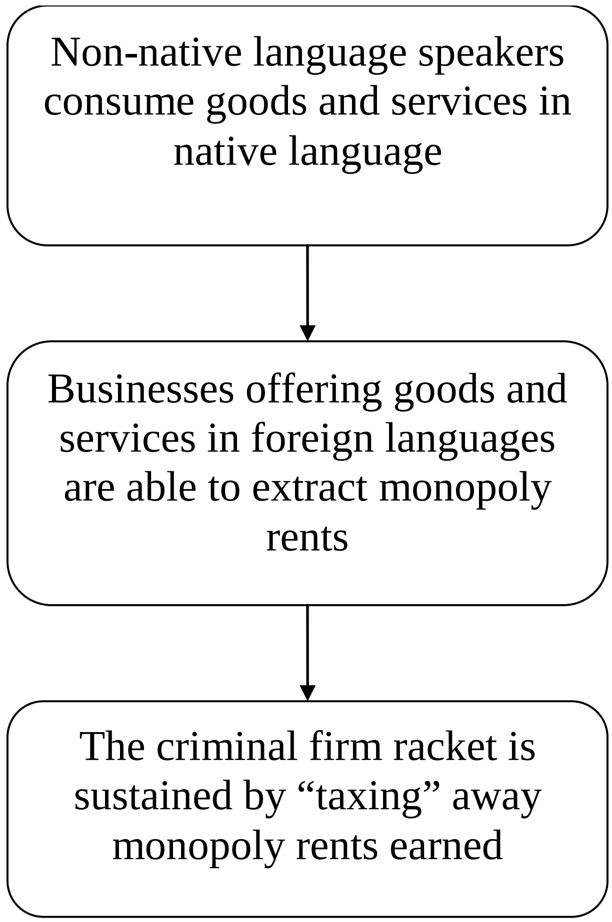

Non-native language speakers consume goods and services in native language
Businesses offering goods and services in foreign languages are able to extract monopoly rents
The criminal firm racket is sustained by “taxing” away monopoly rents earned

Figure 2