HOW DO RIGHT-TO-CARRY LAWS AFFECT CRIME RATES?

Coping with Ambiguity using Bounded-Variation Assumptions

Charles F. Manski

Department of Economics and Institute for Policy Research, Northwestern University

and

John V. Pepper

Department of Economics, University of Virginia

forthcoming in *Review of Economics and Statistics*

Abstract: Despite dozens of studies, research on crime has struggled to reach consensus about the impact of right-to-carry (RTC) gun laws. With this in mind, we formalize and apply a class of *bounded variation* assumptions that flexibly restrict the degree to which outcomes may vary across time and space. Using these assumptions, we present empirical analysis of the effect of RTC laws on violent and property crimes in Virginia, Maryland and Illinois. Imposing specific assumptions that we believe worthy of consideration, we find that RTC laws increase some crimes, decrease other crimes, and have effects that vary over time for others.

Acknowledgments: We thank David Rodina for research assistance. We are grateful to Dan Black, Jennifer Doleac, Jeff Dominitz, John Donohue, Mark Duggan, Joel Horowitz, Brent Kreider, John Lott, Stephane Mechoulan, Daniel Nagin, Franco Peracchi, David Rivers, and participants at the 2015 Southern Economic Association Conference and the 2016 NBER Economics of Crime Working Group Summer Institute session for helpful comments. We are also grateful to three anonymous reviewers for their comments.

1. INTRODUCTION

Research on crime in the United States has commonly used data on county or state crime rates to evaluate the impact of laws allowing individuals to carry concealed handguns – so called right-to-carry (RTC) laws. Theory alone cannot predict even the direction of the impact. The knowledge or belief that potential victims may be carrying weapons may deter commission of some crimes but may escalate the severity of criminal encounters.[1] Ultimately, how allowing individuals to carry concealed weapons affects crime is an empirical question.

Lott (2010) describes this empirical research to a lay audience in a book with the provocative and unambiguous title *More Guns, Less Crime*. Yet, despite dozens of studies, the research provides no clear insight on whether "more guns" leads to less crime. Some studies find that RTC laws reduce crime, others find that the effects are negligible, and still others find that such laws increase violent and/or property crime. In a series of papers starting in 1997, Lott and coauthors have argued forcefully that RTC laws have important deterrent effects which can play a role in reducing violent crime. Lott and Mustard (1997) and Lott (2010), for example, found that RTC laws reduce crime rates in every violent crime category by between 5 and 8 percent. Using

[1] Donohue and Levitt (1998) formalize a simple stylized model showing that in illegal markets with scarce resources, firearms (or lethality more generally) do not have a clear impact on violence. This model implies that the impact of firearms might vary across markets and over time, a finding which is inconsistent with standard models used in empirical analysis. Duggan (2001) and Durlauf *et al*. (2016) make a similar point when noting that the impact of RTC laws may be sensitive to rates of gun ownership and the underlying environments and propensities for criminality. We elaborate in Section 3.

different models and revised/updated data, however, other researchers have found that RTC laws either have little impact or may increase violent crime rates.[2] Consider, for example, Aneja, Donohue and Zhang (2011), who make seemingly minor modifications to the basic model and data (e.g., reducing the number of demographic covariates) used by Lott (2010). Whereas Lott (2010) found that RTC laws decrease the different violent crime rates by between 5 and 8 percent, Aneja *et al.* (2011) found that RTC laws have a negligible impact on murder and increase other violent crime rates by 20 to 30 percent. Aneja *et al.* (2011) also report that many of the point estimates are statistically insignificant by conventional criteria.

This ambiguity may seem surprising. How can researchers using similar data draw such different conclusions? In fact, it has long been known that inferring the magnitude and direction of treatment effects is an inherently difficult undertaking. Suppose that one wants to learn how crime rates (an outcome of interest) would differ with and without a RTC law (a treatment) in a given place and time. Data cannot reveal counterfactual outcomes. That is, data cannot reveal what the crime rate in a RTC state would have been if the state had not enacted the law. Nor can data reveal what the crime rate in a non-RTC state would have been if a RTC law had been in effect. To identify the law's effect, one must somehow "fill in" the missing counterfactual observations. This requires making assumptions that cannot be tested empirically. Different assumptions may yield different inferences.

Yet the empirical research on RTC laws has struggled to find consensus on a set of credible assumptions. Reviewing the literature, the National Research Council (NRC) Committee to

[2] See, for example, Black and Nagin (1998), Duggan (2001), Aneja *et al.* (2011), and Durlauf *et al*. (2016).

Improve Research Information and Data on Firearms concluded that it is not possible to infer a credible causal link between RTC laws and crime using the current evidence (National Research Council, 2005). Indeed, the Committee concluded that (p. 150): “additional analysis along the lines of the current literature is unlikely to yield results that will persuasively demonstrate” this link. The Committee found that findings are highly sensitive to model specification. Yet there is no solid foundation for particular assumptions and, as a result, no obvious way to prefer specific results (also see Durlauf *et al.,* 2016). Hence, drawing credible precise findings that lead to consensus about the impact of RTC laws has thus far proven to be impossible.

How then should research proceed? We think that analysis of treatment response is most useful if researchers perform inference under a spectrum of assumptions of varying identifying power, recognizing the tension between the strength of assumptions and their credibility. Research on RTC laws has commonly made *invariance* assumptions asserting that specified features of treatment response are constant across space or time. These assumptions may be too strong, but weaker assumptions asserting bounded variation may be credible.

For example, an invariance assumption may assert that neighboring states such as Virginia and Maryland have identical environments and propensities for criminality. This assumption may seem attractive because it enables contemporaneous comparison of the two states, but it may not be credible. Yet it may be credible to assume that these states are similar to one another. Likewise, an invariance assumption may be that, in the absence of RTC statutes, Virginia and Maryland would have experienced the same change in murder rates between two periods, say 1988 and 1990. This assumption enables difference-in-differences analysis of treatment response but, again, it may not be credible. On the other hand, it may be credible to assume that the two states would have experienced similar changes.

With this in mind, we formalize and apply a class of *bounded-variation* assumptions that flexibly restrict the variation of treatment response across states and years. These assumptions generally do not point-identify the effects of RTC laws on crime rates. Instead, they partially identify them, yielding bounds rather than precise conclusions.[3]

Some readers may believe that partial identification analysis cannot lead to informative conclusions about the effects of RTC laws on crime. Under very weak assumptions, it is correct that the data cannot reveal whether RTC laws increase or decrease crime. However, under a specific set of bounded variation assumptions that we believe worthy of consideration, we draw substantive conclusions about the qualitative and quantitative impact of RTC laws (see Section 4). Moreover, our analysis makes transparent how these conclusions depend on the maintained assumptions.

The paper is organized as follows: After providing a brief overview of the data in Section 2, Section 3 formally defines the empirical question and the selection problem, and then introduces invariance and bounded variation assumptions. Throughout, we allow for the possibility that the

[3] Partial identification analysis of treatment effects from observational data was initiated in Manski (1990) and has developed subsequently. Manski (2007) provides a textbook exposition. The closest methodological precedent to the present study is the Manski and Pepper (2000) study of monotone instrumental variable assumptions. The closest applied precedent to the present paper is the Manski and Pepper (2013) analysis of the deterrent effect of the death penalty, which uses some simple bounded variation assumptions to examine data from two repeated cross sections. In this paper, we introduce a much broader class of bounded variation assumptions to evaluate annual state crime rates from multiple repeated cross sections.

effects of RTC laws vary across years, states, and crimes. To keep our task manageable, we focus on drawing inferences on the impact of RTC laws in a single state, Virginia. In an online appendix, we further illustrate the approach by examining what can be inferred about the impact of RTC laws in Maryland and Illinois, two states that did not adopt RTC statutes during the period we study.

In Section 3, we illustrate the sensitivity of inferences to different identifying restrictions but we do not argue in a favor of any particular assumption. Here, as in earlier research, we find it valuable to entertain a set of assumptions of varying strength and to determine how the strength of the assumptions imposed affects the strength of the conclusions drawn. See Manski and Pepper (2000, 2013) for other applications of this research philosophy.

In Section 4, we motivate particular bounded variation assumptions and present results on the impact of RTC laws under these assumptions. Here, we find that RTC laws appear to reduce some crimes, increase others, and for some the results vary over time. In Virginia (VA), for example, the data combined with these assumptions imply that the RTC law in VA reduced murder and larceny rates in nearly every year after 1989, but increased assaults after 1997. For other crimes, the sign of the bounds is generally negative in the 1990s but positive or indeterminate in the 2000s. Section 5 draws conclusions.

2. DATA

To evaluate the impact of RTC laws on crime, we use state level data on annual crime rates (per 100,000 residents) from 1970 to 2007. The focus on annual crime rates within states has been common in the literature.[4] However, we do not normalize crime rates to lie in the unit interval or take the natural log of the crime rate, as is done in much of the literature. Rather we are interested in directly studying state-year crime rates: 100,000 times the ratio of the number of reported crimes to the population of the state. We mostly focus on crime rates in Virginia. To further illustrate the bounded variation assumptions, we examine crime rates in Maryland and Illinois in an online appendix.

The crime data, obtained from the FBI's Uniform Crime Reports (UCR), were originally assembled by Lott and Mustard and have subsequently been modified, corrected, and updated several times. Our analysis uses the iteration assembled and evaluated by Aneja, Donohue and Zhang (2011).[5] For each state and year, we observe crime rates separately for murder, rape, assault, robbery, auto theft, burglary, and larceny. For each state-year, we observe whether a RTC statute is in place.[6]

[4] Some studies take the county rather than the state as the geographic unit of analysis, and most studies condition the analysis on characteristics of the state or county (e.g, demographics, number of police per capita, or median county income). It may be of interest to perform more disaggregated analyses that examine crime commission by particular categories of person or even by individuals. However, the requisite data to support such analyses are not available. Likewise, rather than evaluating the effect of a RTC statute, it may be of interest to study how the number of RTC permits impact crime. However, in general, data on the number of permits or the number of persons carrying concealed handguns are not available (Lott, 2010).

[5] The data were downloaded from http://works.bepress.com/john_donohue/89/ in June 2012.

[6] As of 2014, all 50 states have passed laws allowing citizens to carry concealed firearms. Not all

Figures 1A and 1B display the annual time series of murder and robbery rates in Virginia and Maryland over the period 1970–2007. The figures reveal several interesting characteristics of the crime rates. First, notice that except for murder rates in a few years during the 1970s, crime rates in Maryland exceed the analogous rate in Virginia, in many cases by a substantial margin. Second, the figures show well-known temporal patterns in crime: crime rates rose in the 1980s and then declined sharply beginning in the mid-1990s. Crime rates in Maryland, which did not adopt a RTC statute over this period, have more pronounced changes than those in Virginia, rising faster in the 1980s and dropping faster in the 1990s. Virginia enacted a RTC statute in 1989.

3. BASIC ISSUES IN INFERENCE ON TREATMENT EFFECTS, WITH AN ILLUSTRATIVE APPLICATION

---

states, however, have RTC laws as defined in the empirical research. The research has classified a state as having a RTC law when legal gun owners are allowed to carry concealed firearms, perhaps after certification and training. These are often referred to as "shall issue" provisions, in that the state shall issue a permit subject to the applicant meeting basic determinate criteria. Other states require applicants to demonstrate "good cause" before granting a concealed carry permit. These so called "may issue" states are not classified as having RTC laws. Ten states and the District of Columbia have "may issue" provisions and forty-one states have "shall issue" provisions. A recent wave of high-profile state legislation allows citizens to openly carry firearms in public, enacting so called "open carry" laws. While not considered in this analysis, the same identification problems creating uncertainty in the research on RTC laws apply to any evaluation of the impact of "open carry" laws on crime.

In this section, we formally define the empirical question and the selection problem, and introduce invariance and bounded variation assumptions. We begin by defining the average treatment effect and then assess the effect of a RTC law on the 1990 murder rate in Virginia. We focus on this somewhat narrow question to clearly illustrate the utility of bounded variation restrictions. In particular, we show how bounded variation restrictions provide an intuitive and simple way to improve the credibility of empirical research. We also illustrate the sensitivity of inferences to different identifying restrictions, without arguing in favor of any particular set of assumptions.[7] In Section 4, we use the observed data on crime rates to motivate a particular set of bounded variation assumptions and explore, in some detail, the inferences that arise from this set of assumptions.

Departing from conventional practice in applied econometric analysis, we do not refer to our empirical findings as "estimates" and do not provide measures of statistical precision when studying the impact of RTC laws. Instead, we perform a finite-population analysis that views states as the units of interest rather than as realizations from some sampling process. One reason is expositional, being that we want to focus attention on the identification problem arising from the unobservability of counterfactual outcomes. The bounded-variation assumptions that we consider place deterministic constraints on the values of counterfactual outcomes. Hence,

---

[7] Manski (2007, Chapers 7 through 10) provides a textbook exposition of the identifying power of alternative assumptions, ranging from weak assumptions yielding partial identification of average treatment effects to strong ones yielding point identification. Richardson *et al*. (2014) also reviews some of this literature and the related body of research on sensitivity analysis, which examines the implications of varying certain unknown parameters within some class of models.

imprecision in our findings results from the selection problem and is expressed through the width of the bounds that we report rather than through any measure of sampling uncertainty.

Considering the available data, a fundamental reason for not performing statistical inference is that measurement of statistical precision requires specification of a sampling process that generates the data. Yet we are unsure what type of sampling process would be reasonable to assume in this application. One would have to view the existing United States as the sampling realization of a random process defined on a super-population of alternative nations.[8] That is, one would have to pose a random process generating actual American history, with its division of the country into states with their populations of persons, as one among a set of possible histories that could have generated alternative state-year crime rates. But what random process should be assumed to have generated the existing United States, with its realized state-year crime rates?

There is no obvious answer and the literature analyzing RTC laws has not engaged the question. We do not argue that it would be impossible to specify a credible random process generating the United States and base statistical inference on this process, but doing so goes well beyond the scope of this paper. We leave this as an open question that may perhaps draw attention in future research by ourselves or others.

A reviewer has suggested that we compute confidence intervals using methods studied in the econometric literature on inference in settings with partial identification. Existing methods for computing such confidence intervals assume that the data are a random sample drawn from an infinite population; see, for example, Imbens and Manski (2004), Chernozhukov, Hong, and Tamer

[8] See Cochran (1977) and Deaton (1997), among others, for expositions of the distinction between finite-population and super-population analysis.

(2007), Chernozhukov, Lee and Rosen (2013), and Kaido, Molinari and Stoye (2016). Likewise, random sampling assumptions are implicit in the inferential methods used to estimate the standard errors of parameter estimates in the linear panel data models prevalent in the existing literature on RTC laws.[9] Random sampling assumptions, however, are not natural when considering states or counties as units of observation.

Similar concerns about the sampling process arise if one takes the unit of analysis to be an individual person and considers state crime rates to aggregate individual crime counts across the population of a state. To draw inferences in this setting, one must have prior information on the sampling process generating the individuals who populate a state. Random sampling from an infinite population, which is the underlying assumption used in the partial identification literature, applies if one is willing to assume that the realized population of a state in a given year is a random sample of persons drawn from a super-population of potential state residents and that crimes are committed independently by members of the population. However, this random sampling assumption may not be credible because commission of crimes may involve complex social

[9] There is an ongoing debate about the appropriate method of estimating standard errors for the linear models used in the RTC literature (see Aneja *et al.*, 2011, and National Research Council, 2005). Many researchers report standard errors that allow for arbitrary correlation within a state or county – so called state/county clustered standard errors – while others do not allow for such correlations. The NRC report explains that these clustered sampling standard errors are inappropriate in the standard linear panel data models (with fixed county effects) used in the literature. Our concerns are distinct from these issues. We find no basis for reporting standard errors using any of the conventional methods, allowing for clustering or not.

interactions that are not independent across individuals or over time. Moreover, it may not be credible to suppose that the observed population of a state in a given year is a random sample from some super-population.

Suppose that, despite the above concerns, one is willing to assume that a state population is a random sample drawn from a super-population. Given that state populations are so large, containing millions of persons, the sampling variability of counts would be a negligible consideration except possibly when considering relatively rare crimes such as murder.

In the setting of a randomized experimental design, Abadie *et al.* (2014) raise similar concerns about drawing inferences in state level analyses and propose an alternative approach for measuring uncertainty. In particular, they argue that if one observes the population, there is no sampling variability but there is statistical uncertainty if treatments are assigned at random. The fact that one cannot observe counterfactual outcomes creates "randomization uncertainty." Abadie *et al*. (2014) propose using randomization-based standard errors to express such uncertainty. They find that these standard errors are generally smaller than the conventional sampling-based standard errors made under a random sampling assumption.

This approach provides an alternative way to conceptualize uncertainty in cases where one observes a population rather than a sample. The proposed randomized-based standard errors, however, only apply when treatments are randomly assigned, in which case the average treatment effect is point identified. Their approach is not applicable in observational settings where the treatment, such as RTC laws, may be endogenous.

3.1 Treatment Effects and the Selection Problem

Consider the problem of inferring the average treatment effect (ATE) of a RTC statute on the rate of commission of a specified crime in a specified year in a group of states that share specified observed covariates:

$$(1) \qquad ATE_{dx} = E[Y_d(1)|X] - E[Y_d(0)|X].$$

Outcome $Y_d(1)$ denotes the crime rate if a state were to have a RTC statute in year d, $Y_d(0)$ denotes the analogous outcome if the state were not to have a RTC statute, X denotes the specified covariates, and d indicates the specified year. To define the ATE, we consider the fifty states plus DC to be the population and use the expectation operator $E[\cdot|X]$ to denote an average across states with covariates X. The unit of observation being a state, the average treatment effect measures how the crime rate would differ if all states with covariates X were to have a RTC statute in year d versus the rate if all such states did not have a RTC law. The ATE can vary with d and X. We follow the literature by assuming that treatment response is individualistic; that is, a RTC law in state j may impact crime in state j, but not elsewhere. Thus, there are no spillover effects to other states.

For each state j and year d, there are two potential outcomes, $Y_{jd}(1)$ and $Y_{jd}(0)$. The former outcome is counterfactual if state j did not have a RTC statute in year d while the latter is counterfactual if the state did have a RTC statute. The fact that the data only reveal one of the two mutually exclusive outcomes constitutes the selection problem. The observed crime rate is $Y_{jd}(Z_{jd}) = Y_{jd}(1) Z_{jd} + Y_{jd}(0) (1 - Z_{jd})$, where $Z_{jd} = 1$ if state j has a RTC statute in year d and $Z_{jd} = 0$ otherwise. We henceforth write $Y_{jd} = Y_{jd}(Z_{jd})$ for short.

3.2. Invariance Assumptions

The conventional practice used to address the selection problem has been to invoke assumptions that are strong enough to point-identity counterfactual mean outcomes and, hence, the average treatment effect. These assumptions typically assert invariance of some kind. One might, for example, assume that the ATE is constant across X and/or d.

To illustrate in a simple setting, consider inference on the impact of adoption of a RTC law on the 1990 murder rate in Virginia (VA). Virginia enacted a RTC statute in 1989. In this case, d = 1990, X = VA, and $Z_{VA,1990} = 1$. There exists only one state with X = VA, so $P(Z_{1990} = 1|X = VA) = 1$ and $E[Y_{1990}(t) \mid X = VA] = Y_{1990}(t)$. Hence, (1) reduces to

$$
\begin{aligned}
(2) \qquad ATE_{1990,VA} &= E[Y_{1990}(1) \mid X = VA] - E[Y_{1990}(0) \mid X = VA] \\
&= Y_{VA,1990}(1) - Y_{VA,1990}(0).
\end{aligned}
$$

The available data reveal that $Y_{VA,1990}(1) = Y_{VA,1990}$, but they do not reveal the counterfactual quantity $Y_{VA,1990}(0)$.

Throughout this paper we focus attention on treatment response in a single state, such as Virginia, rather than on a group of states. This being so, we henceforth omit the expectation notation. We also omit the adjective "average" when discussing treatment effects and use the abbreviation TE_{jd} to denote the treatment effect in state j in year d.

Table 1 displays the murder rates per 100,000 residents in Virginia and Maryland in 1988 and 1990. Neither state had an RTC statute in 1988. Virginia enacted one in 1989 but Maryland did not. Thus, $Y_{VA,1990}(1) = Y_{VA,1990} = 8.81$.

These data may be used to compute three simple but different findings on the counterfactual 1990 murder rate in Virginia, $Y_{VA,1990}(0)$, under alternative invariance assumptions:

(3) ***Time invariance***: $Y_{VA,1990}(0) = Y_{VA,1988}(0) = Y_{VA,1988} = 7.75$;

(4) ***Interstate invariance***: $Y_{VA,1990}(0) = Y_{MD,1990}(0) = Y_{MD,1990} = 11.55$;

(5) ***Difference-in-difference invariance***: $Y_{VA,1990}(0) =$

$$[Y_{MD,1990}(0) - Y_{MD,1988}(0)] + Y_{VA,1988}(0) = (Y_{MD,1990} - Y_{MD,1988}) + Y_{VA,1988} = 9.67.$$

These, in turn, imply treatment effects of 1.06, -2.74, and -0.86 respectively. Thus, the three empirical findings differ in direction and magnitude.

Given certain invariance assumptions, each finding appropriately measures the effect of the RTC law on the 1990 murder rate in Virginia. The one assuming time invariance is correct under the assumption that no determinant of criminal behavior changed in Virginia between 1988 and 1990 except for enactment of the RTC statute. Interstate invariance of the 1990 murder rates is correct under the assumption that the populations of Virginia and Maryland had the same propensities for criminal behavior and faced the same environments except for the presence of the RTC statute in Virginia. The difference-in-difference (DID) finding is correct under the assumption that, in the absence of RTC statutes, Virginia and Maryland would have experienced the same change in murder rates between 1988 and 1990. Thus, each of the three findings can be justified by specific invariance assumptions. However, the variation in empirical findings shows that these invariance assumptions cannot hold jointly. Indeed, it may be that none of the assumptions holds. Thus, the TE may equal none of the values 1.06, -2.74, and -0.86.

The literature evaluating RTC laws has analyzed crime data across many states and years rather than two states and two years. Having more data, however, does not reduce the dependence of empirical findings on the assumptions that researchers maintain. It has been common, for example, to assume a linear model with a homogeneous treatment effect and state-year fixed effects.[10] This model has the form

$$(6) \quad Y_{jd}(t) = \theta \cdot t + X_{jd}\beta + \alpha_j + \gamma_d + \varepsilon_{jd}.$$

Here treatment t = 1 denotes the presence of an RTC law and t = 0 otherwise. The parameter θ is the treatment effect, which does not vary with j and d. Thus, the model assumes that right-to-carry laws have the same effect on crime rates, θ, in all states and years.[11] Model (6) permits variation in crime rates across states and years only through the composite additive intercept $X_{jd}\beta + \alpha_j + \gamma_d + \varepsilon_{jd}$. Here β is a parameter vector, while α_j and γ_d are state and year fixed effects. The unobserved

[10] The existing research on RTC laws does not formally model the selection process by which states adopt RTC statutes or, in general, apply an invariance assumption that the response functions are mean independent of observed instrumental variables. Notable exceptions are Lott and Mustard (1997) and Lott (2010), who present results under a variety of instrumental variable assumptions.

[11] One may allow the treatment effect θ to vary across states and years if one assumes that passage of RTC laws is exogenous; that is, statistically unrelated to the state-date specific treatment effect. However, researchers typically do not want to make this assumption.

variable ε_{jd} is a random state-year interaction assumed to have mean zero conditional on each realized value of t and X.

This model relies on the invariance assumption that the effect of RTC laws, θ, is the same for all states and all years. While this homogeneity assumption is convenient and has substantial identifying power when combined with certain other assumptions, there is little support for the notion that the effects of RTC laws are identical across states and time. In fact, the empirical literature provides some evidence to the contrary. Some researchers have concluded that the effects of RTC laws vary over time and across states (see Black and Nagin, 1988; National Research Council, 2005; Aneja *et al.*, 2011; Durlauf *et al*., 2016).

The VA and MD crime rates displayed in Figure 1 provide direct evidence that the invariance assumptions in (3), (4) and (5) do not hold across all years and states. Figure 1 displays the annual murder rates in Maryland and Virginia, from 1970 to 2007. While neither state had a RTC statute prior to 1989, crime rates vary over time and across states from 1970 to 1988. For example, in 1988, the murder rate in Virginia was 1.89 less than the rate in Maryland in that year and 0.39 greater than the Virginia rate in 1987.[12] Thus, the invariance assumptions are violated in the years before VA adopted a RTC statute, and the signs and magnitudes of these violations differ over time and across assumptions.

While all three of these invariance assumptions are rejected in the pre-1989 period, the existing literature on the effects of RTC laws on crime consistently applies these types of restrictions, especially variations of the DID model in (5) or the related model in (6). In this setting,

---

[12] Also notice that the time invariance assumption does not hold in Maryland, where the murder rate increased from 9.63 in 1988 to 11.55 in 1990.

researchers implicitly assume that the invariance restrictions apply when outcomes are counterfactual even though they are rejected in periods where outcomes are observed. This is hard to motivate and the literature using these invariance assumptions fails to do so. Likewise, the variety of alternative but related difference-in-difference type approaches that point identify the average treatment effect in data from repeated cross sections -- for example, the propensity score and the synthetic control (see Abadie, Diamond and Hainmueller, 2010) approaches -- rely on invariance assumptions that are difficult to motivate in this application.

3.3. Bounded Variation Assumptions

3.3.1. The Assumptions

Invariance assumptions have a sharpness that often makes them suspect. Consider each of the three equalities (3), (4), and (5) that point-identify the impact of the RCT law in Virginia in 1990. Why should any of them hold exactly? Why should the treatment effect be exactly constant across states and years as assumed in the linear model (6)? Empirical researchers often say that such assumptions are "approximations," but they do not formalize what this means. Instead, they perform analyses that use some invariance assumptions as if they are truth and that entirely dismiss other assumptions. For example, DID estimation maintains assumption (5) but places no restrictions on response levels.

A simple way to improve the credibility of empirical research is to weaken invariance assumptions to assumptions of bounded variation. In this paper we report empirical findings under bounded-variation assumptions that weaken assumptions (3), (4), and (5) for a specified treatment t as follows:

(7) ***Bounded Time Variation***: $|Y_{jd}(t) - Y_{je}(t)| \leq \delta_{j(de)}$,

(8) ***Bounded Interstate Variation***: $|Y_{jd}(t) - Y_{kd}(t)| \leq \delta_{(jk)d}$,

(9) ***Bounded DID Variation***: $|[Y_{jd}(t) - Y_{je}(t)] - [Y_{kd}(t) - Y_{ke}(t)]| \leq \delta_{(jk)(de)}$,

where (j, k) are specified states, (d, e) are specified years, and ($\delta_{j(de)}$, $\delta_{(jk)d}$, $\delta_{(jk)(de)}$) are specified positive constants. Manski and Pepper (2013) observed that the bounds on counterfactual quantities implied by such bounded-variation assumptions are solutions to linear programming problems and, hence, are relatively easy to compute. See also Lafférs (2015).

To simplify the notation, we suppress the dependence of the δ values on the treatment t. We also suppress for simplicity the logical requirement that crime rates must take non-negative values. This requirement does not bind in the empirical analyses that we report.

The bounded time variation assumption restricts the absolute difference in treatment response in state j between two years, say 1988 and 1990, to be less than $\delta_{j(de)}$. Letting $\delta_{j(de)} \geq 0$ weakens the traditional time invariance assumption by supposing that response may differ over time by no more than $\delta_{j(de)}$. The larger the selected value of $\delta_{j(de)}$, the weaker the assumption. Similarly, the bounded interstate variation assumption restricts the contemporaneous absolute difference at year d in response between two states, say Virginia and Maryland, to be less than $\delta_{(jk)d}$. Letting $\delta_{(jk)d} \geq 0$ weakens the traditional invariance assumption by supposing that response across the two states may differ by no more than $\delta_{(jk)d}$. The larger the selected value of $\delta_{(jk)d}$, the weaker the assumption.

These bounded variation assumptions have identifying power because they imply that counterfactual state-year crime rates are similar to observed rates in other states and years. The degree of similarity is determined by the bound parameters ($\delta_{j(de)}$, $\delta_{(jk)d}$, $\delta_{(jk)(de)}$). These assumptions resemble the smoothing assumptions made in kernel nonparametric regression analysis, with δ acting as the bandwidth (aka tuning parameter). A fundamental difference is that the asymptotic theory for nonparametric regression makes smoothing go to zero with sample size, so smoothing is performed to achieve desired asymptotic statistical properties rather than identification. In contrast, we impose assumptions that relate counterfactual quantities to observed ones and thus provide identifying power.

It is of interest to compare the bounded variation assumptions to those made in the linear homogeneous model. These assumptions are not nested. The bounded variation assumptions are weaker in two respects. First, they permit the treatment effect to vary across states and years, whereas model (6) restricts the treatment effect to have the constant value θ across states and years. Second, assumptions (7)-(9) do not impose a condition on mean variation in treatment response akin to the restriction of model (6) that ε_{jd} has mean zero conditional on each realized value of t and X. On the other hand, assumptions (7)-(9) restrict the potential magnitudes of time, interstate, and DID variation in outcomes under a given treatment, whereas model (6) does not.

3.3.2. Findings for Virginia in 1990

Figure 2 shows bounds on $TE_{1990,VA}$ for different values of $\delta_{j(de)}$, where d = 1990, e = 1988, and j = VA. Our objective is to illustrate the sensitivity of inferences to different identifying restrictions, without arguing in favor of any particular set of assumptions. The traditional time invariance assumption ($\delta_{j(de)} = 0$) point identifies the TE, revealing that enactment of the RTC statute increases the murder rate by 1.06. Ambiguity about the TE increases with $\delta_{j(de)}$. The figure shows that the bound on the TE is entirely positive when $\delta_{j(de)}$ is less than one, but any value of $\delta_{j(de)}$ larger than one renders it impossible to sign the TE. For example, when $\delta_{j(de)} = 2$, we know that $Y_{VA,1990}(1) = 8.81$ and $Y_{VA,1990}(0) \in [5.75, 9.75]$. Hence, $TE_{1990,VA} \in [-0.94, 3.06]$.

Figure 2 also shows bounds on $TE_{1990,VA}$ for different values of $\delta_{(jk)d}$, where k = Maryland. The traditional interstate invariance assumption ($\delta_{(jk)d} = 0$) point identifies the TE, revealing that the RTC statute decreases the murder rate by -2.74. Ambiguity about the TE increases with $\delta_{(jk)d}$. The figure shows that the bound is entirely negative when $\delta_{(jk)d} < 2.74$, but values in excess of 2.74 make it impossible to sign the TE. For example, when $\delta_{(jk)d} = 3$, we know that $Y_{VA,1990}(1) = 8.81$ and $Y_{VA,1990}(0) \in [8.55, 14.55]$. Hence, $TE_{1990,VA} \in [-5.74, 0.26]$.

Bounded variation assumptions can be easily adapted to fit particular features of the application. For example, Figure 1 reveals that crime rates in MD generally exceed the contemporaneous crime rates in VA, even in the years before 1989 when neither state had adopted a RTC statute. One might therefore think it reasonable to make the bounded variation assumption one sided; that is,

$$(10) \qquad 0 \leq Y_{MD,1990}(0) - Y_{VA,1990}(0) \leq \delta_{(MD,VA)1990}$$

Assumption (10) makes the lower bound on the TE equal -2.74, the finding under invariance assumption (4). The upper bound is identical to the one displayed in Figure 2.

Finally, letting $\delta_{(jk)(de)} \geq 0$ weakens the traditional DID invariance assumption. The traditional assumption ($\delta_{(jk)(de)} = 0$) point identifies the TE, revealing that the RTC statute decreases the mean murder rate by -0.86. Ambiguity about the TE increases with $\delta_{(jk)(de)}$, and for values in excess of 1 it is impossible to sign the TE.

3.4. Joint Bounded Geographic and Time Variation Assumptions

Rather than use a single invariance assumption in isolation, one may want to combine assumptions. For example, one might simultaneously assume that Virginia and Maryland are similar to one another and that Virginia had similar characteristics in 1988 and 1990. Here we evaluate bounds on$TE_{1990,VA}$ under joint interstate and time bounded variation assumptions.

Under a joint interstate and time variation assumption, the counterfactual murder rate $Y_{VA,1990}(0)$ is bounded as follows:

$$\begin{aligned} &\max(Y_{MD,1990} - \delta_{(MD,VA)1990},\ Y_{VA,1988} - \delta_{VA(1990,1988)}) \\ &\leq\ Y_{VA,1990}(0)\ \leq \\ &\min(Y_{MD,1990} + \delta_{(MD,VA)1990},\ Y_{VA,1988} + \delta_{VA(1990,1988)}). \end{aligned} \tag{11}$$

A necessary condition for this assumption to be valid is that $\delta_{VA(1990,1988)} + \delta_{(MD,VA)1990} \geq |Y_{MD,1990} - Y_{VA,1988}|$. Otherwise, the lower bound exceeds the upper bound.

Table 2 displays the upper and lower bound on the TE under joint bounded interstate and time variation assumptions. Findings highlighted in dark grey are not valid; the lower bound exceeds the upper bounds. Findings in light grey identify the sign of the treatment effect, with those in bold text as negative and those in italicized text as positive. Findings that are not highlighted are informative, but do not reveal whether adoption of the RTC law in 1989 increased or decreased the 1990 murder rate in VA.

There are several interesting findings. First, the necessary condition for validity of assumptions rules out a range of small values of $\delta_{VA(1990,1988)}$ and $\delta_{(MD,VA)1990}$ as infeasible. Thus, it cannot jointly be true that (Virginia, Maryland) are highly similar in 1990 and that (1988, 1990) are highly similar years in Virginia.

Second, the assumptions point identify the TE for a variety of feasible values (i.e., $\delta_{VA(1990,1988)} + \delta_{(MD,VA)1990} = 4$), and they identify the sign of the TE for others. For example, when $\delta_{VA(1990,1988)} = \delta_{(MD,VA)1990} = 2$, the TE is identified to equal -1.2. When $\delta_{va(1990,1988)} = 0.5$ and $\delta_{(MD,VA)1988} = 3.5$, the TE is identified to equal 0.3. The sign of the TE is identified to be negative for all feasible $\delta_{(MD,VA)1990} \leq 3$, and greater than zero for all feasible $\delta_{VA(1990,1988)} \leq 0.5$. For larger values of $\delta_{VA(1990,1988)}$ and $\delta_{(MD,VA)1990}$, the assumptions do not identify the sign of the TE.

To summarize, in this section we have examined the sensitivity of inferences to different bounded variation assumptions, focusing on the effect of the 1989 adoption of an RTC law on the 1990 murder rate in Virginia. We find that the sign of the TE is identified for some assumptions, but not others. Especially interesting findings emerge when different bounded variation assumptions are combined. In particular, the strongest assumptions are ruled out, including the traditional strict invariance assumptions. Less strong assumptions either point identify the TE or identify the sign of the TE, while weaker ones do not identify whether adoption of an RTC law increased or decreased murder in Virginia in 1990.

The type of sensitivity analysis performed here, which fully maps out the ambiguity resulting from the selection problem, may prompt one to seek a purely data-driven approach to determine the validity of different assumptions – a "stopping rule." In general, however, no such rule exists: the data alone cannot identify the counterfactual outcomes.

Alternatively, one might rely on an understanding of the institutional setting, the historical record, theory or other sources of information to develop a set of credible assumptions on treatment response. For example, in most markets economists are comfortable assuming demand is downward sloping even though the market data cannot reveal the demand function (see Manski, 1997). When evaluating the impact of RTC laws, finding common ground based on theory is difficult. Instead, in Section 4, we use data on crime rates prior to Virginia's adoption of a RTC law to provide guidance on what we believe to be a set of bounded variation assumptions worthy of consideration.

4. EVALUATING THE IMPACT OF RTC LAWS USING BOUNDED VARIATION ASSUMPTIONS

Whereas Section 3 assessed the sensitivity of inference to different assumptions, in this section we evaluate the impact of RTC laws under particular bounded variation assumptions. We begin in Section 4.1 by developing an approach for selecting the parameters δ based on observed data. Then, using these parameters, we present findings for Virginia in Section 4.2 and for Maryland and Illinois in an online appendix. The latter states did not adopt RTC statutes prior to 2007. In Section 4.4 we discuss and summarize the results.

4.1. Selecting the Bound Parameters

In Section 3, we assessed how inferences vary across different values of the bound parameters δ. In this section we use observed data in VA and MD to develop sensible data-based parameter values. In particular, we use data prior to 1989, when VA did not have a RTC statute, to determine the minimum parameter values that would be required to make bounded variation assumptions consistent with the observed data. While this approach does not ensure the validity of the bounded variation assumptions, it provides what we believe to be a reasonable starting point for our analysis.

To illustrate the idea, consider the murder rates for Maryland and Virginia in 1987 and 1988. The observed murder rates vary over time and across states: the 1988 murder rate in VA is 7.75 and the rate in MD is 9.63, while the analogous rates for 1987 are 7.36 and 9.55, respectively. Neither state had a RTC statute in these years, so these data can be used to test whether invariance and bounded variation assumptions are valid. In fact, the data are inconsistent with the time invariance assumption (3), the interstate invariance assumption (4), and the DID assumption (5).

While the invariance assumptions are inconsistent with the observed pre-1989 data, the bound parameters in assumptions (7), (8) and (9) can be chosen to ensure internal consistency. The bounded interstate variation assumption in (8), for example, is valid in 1988 if $\delta_{(VA,MD)1988} \geq 1.88$ (= 9.63 − 7.75). Likewise, the bounded time variation assumption is consistent with the VA 1988 murder rate if $\delta_{VA(1988,1987)} \geq 0.39$ (= 7.75 − 7.36), and the DID assumption is valid if $\delta_{(VA,MD)(1988,1987)} \geq 0.31$. These restrictions on the bound parameters make the bounded variation assumptions consistent with the 1987 and 1988 data.

Extending this approach to perform comparisons from 1970 to 1988, we find the minimum values of the bound parameters required for the assumptions to be consistent with the data prior to 1989. To do this, we compute the minimum valid parameter value for each pair of adjacent years from 1970 to 1988 (i.e., 1970 to 1971, 1971 to 1972, and so forth), and then select the maximum of these values.[13] Table 3 displays the minimum parameters that ensure the bounded variation

[13] There are many other ways to generate plausible restrictions on the bounded parameters using the pre-1989 data. For example, rather than finding parameters based on the observed absolute differences, one could compute separate bounds on the maximum positive difference and the maximum negative difference. Assumption (10), where the difference in the crime rates between Maryland and Virginia is assumed to be non-negative, is a variant on this idea. This approach weakly narrows one side of the bounded parameters and thus will generally give tighter bounds on the average treatment effect. An alternative approach would be to compare observed rates in both adjacent and non-adjacent years. This will weakly increase the bound parameters, and thus generally give wider bounds on the average treatment effect. For example, if we compare crime rates separated by two-years, the difference-in-difference parameter for murder remains at 2.3 but the parameter on robbery increases from 55 to 75. If we compare crime rates separated by 10 years, the parameter for murder increases to 3 and the parameter on robbery increases to 79.

Moreover, one can refine and/or expand the analysis to introduce additional bounded variation restrictions and refined measures of δ. For example, one might include other states or evaluate county crime rates. Likewise, one might consider the crime rate in Maryland after excluding Baltimore. These and other variations would allow one to apply additional moment restrictions on the counterfactual outcomes and develop refined measures of closeness, δ. For this assumptions are consistent with the observed pre-1989 data on various crime rates. For murder, the interstate variation parameter at least 2.7, the time variation parameter needs to be at least 2.0, and the DID variation parameter at least 2.3. For robbery, the analogous parameters are 292, 30 and 55.

We use the parameter values displayed in Table 3 as a starting point for implementing the bounded variation assumptions in (7), (8), and (9). Of course, whether these particular assumptions are valid is unknown, and one must invariably make judgements about the tradeoff between the strength and credibility of assumptions and findings. Using these parameters as an anchor, one might assess the sensitivity of findings by using some multiple of the maximum value or some quantile of the distribution. Table 3, for example, shows the 0.75-quantile of the minimum valid parameters for the DID assumption (DID_0.75). In this case, the DID parameter for murder falls from 2.3 to 1.2, and for robbery falls from 55 to 27.

4.2. The Impact of a RTC Statute in VA

In this section, we use bounded variation assumptions with the parameter values displayed in Table 3 as a starting point for drawing inferences on the impact of an RTC law on crime in VA. Tables 4 and 5 display findings on the TE of a RTC statute in VA by year, from 1990 to 2006, and for seven different crimes. Each set of results applies a different assumption. Recall, as discussed in Section 3, that we do not report measures of sampling variability. Valid standard

---

paper, we focus our analysis two states, VA and MD, over multiple years. This focus keeps our task manageable and results in transparent findings.

errors reflecting uncertainty due to the sampling process would likely impact some of the qualitative conclusions about the effect of RTC laws on crime, especially for highlighted findings that are close to zero and for rare crimes such as murder.

Table 4 presents results for all seven crimes and all years under the DID invariance assumption. We use 1988 as the anchor year for the DID invariance assumption. Findings in bold text identify the TE to be negative while those in italicized text are positive. For murder, the DID assumption findings are strictly negative. For the other crimes, the findings are negative in earlier years, but tend to be positive in later years. For example, the RTC statute is found to have increased the 2005 burglary rate by 112.4.

These results, however, rely on a strict invariance assumption. If we use the parameter values in Table 3 to relax this invariance restriction to a bounded variation DID assumption, nearly all of the bounds except those for larceny include zero. That is, under the bounded DID variation assumption, we cannot generally identify the sign of the TE. To see this, add and subtract the DID parameters in Table 3 to the findings in Table 4. For example, the impact of RTC laws on the 2000 murder rate is bounded between [-0.6 − 2.3, -0.6 + 2.3] = [-2.9, 1.7]. The bound on the impact on the 2005 burglary rate is [-24, 248].

The results from these two DID assumptions provide a stark illustration of the tradeoff between the strength of assumptions and conclusions. The strong invariance form of the DID assumption point-identifies the TE but the weaker bounded variation form of this assumption generally does not identify even the sign of the TE.

Partial identification analysis jointly using variations of assumptions (7), (8), and (9) allows us to bridge the gap between these two extremes by considering middle ground assumptions that yield information on the sign of the TE. To do this, we continue to apply DID bounded variation assumptions and also use bounded time and interstate variation assumptions that seem well suited for this application. In particular, we impose the following three assumptions:

A. $|[Y_{VA,d}(0) - Y_{VA,e}(0)] - [Y_{MD,d}(0) - Y_{MD,e}(0)]| \leq \delta_{(VA,MD)(de)}$, where $\delta_{(VA,MD)(de)}$ = DID_0.75 in Table 3 for $d \in [1990, 2006]$ and e = 1988;

B. $0 \leq Y_{MD,1990}(0) - Y_{VA,1990}(0) \leq \delta_{(MD,VA)1990,}$ where $\delta_{(MD,VA)1990}$ = interstate in Table 3 for $d \in [1990, 2006]$;

C. $Y_{VA,d}(0) \geq Y_{VA,88}(0)$ for $d \in [1990, 1996]$;

$Y_{VA,d}(0) \leq Y_{VA,88}(0)$ for $d \in [1999, 2006]$.

Assumption (A) is the DID bounded variation assumption (i.e., assumption (9)) where the bound parameter equals the 0.75 quantile of the distribution of the minimum bound parameters required for the assumptions to be consistent with the data in each year prior to 1989 (see the DID_0.75 column in Table 3). This assumption relaxes the strict invariance assumption where $\delta = 0$ but strengthens the bounded variation assumption using the DID parameters in Table 3. Assumption (B) is the one-sided bounded interstate variation assumption (i.e., assumption (10)) which restricts the crime rate in VA to be no greater than the crime rate in MD. Finally, assumption (C) is a bounded time variation assumption that operationalizes the idea that crime rates rose in the early 1990s and fell in the 2000s.[14]

[14] In general, the bounds are not sensitive to the cutoff dates used for the bounded time variation Assumption C. For example, changing the cutoff years from 1996 to 1995 and 1999 to 2000 does not impact any of the qualitative conclusions. If we drop Assumption C altogether, 10% of the bounds are impacted but only a handful of the qualitative results change. In particular, without

We do not apply the bounded time variation assumption (7). This assumption is not flexible enough to account for the large temporal changes in crime rates over the period 1970 to 2006. To identify the counterfactual crime rate that would have occurred had VA not adopted a RTC law, the time variation assumption in (7) adds and subtracts $\delta_{VA}(d,88)$ to the observed crime rate in 1988. This bound on the counterfactual crime rate does not depend on the year under consideration, d – the bound is the same for 1989 and 2006 – even though there were significant changes in observed crime rates over this period. Thus, while the bounded variation parameter $\delta_{VA}(d,88)$ in Table 3 may be sensible when evaluating the counterfactual crime rate in 1989, it is hard to defend when evaluating the rate in 2006. In fact, if we add-in the standard bounded time variation assumption in (7), just over 2 out of every 5 of the bounds are infeasible with the lower bound being above the upper bound.

Table 5 displays the findings using these bounded variation assumptions. Findings highlighted in in light grey identify the sign of the treatment effect, with those in bold text as negative and those in italicized text as positive. Findings highlighted in dark grey are infeasible, the lower bound being above the upper bound. Findings that are not highlighted do not identify the sign of the TE. The results vary by crime and year. We find that the RTC law in VA reduced murder and larceny rates in nearly every year, but increased assaults after 1997. For other crimes, the sign of the bounds is generally negative in the 1990s but positive or indeterminate in the 2000s.

Interestingly, these results are generally inconsistent with the standard assumption that the TE does not vary over time (see model 6). To see this, note that under the assumption that the TE

Assumption C, the bounds for assault in 2003 and 2004 are not identified to be positive, and the bounds for burglary from 1990 to 1993 are not identified to be negative.

is the same in every year, one can take the intersection of the bounds from each year to derive a refined tighter bound. This is an instrumental variable bound, as defined in Manski (1990). Yet, the intersection bounds are empty for all seven crimes; the lower bound exceeds the upper bound. For example, the lower and upper intersection bounds for robbery are 23.2 and -78.2. Thus, this bounded variation assumption is inconsistent with the homogeneous treatment effect assumption in model (6).

4.4 Discussion of Results

Our findings about the impact of RTC laws on crime are nuanced and not amenable to a simple punch line conclusion. Inferences are sensitive to assumptions. Under the weakest assumptions, we cannot identify whether RTC laws increase or decrease crime. Under stronger assumptions, we find that adoption of the RTC statute in VA has reduced some crimes (e.g., murder), increased others, and has had ambiguous sign effects on others. For some crimes the findings vary across years, indicating that the RTC law reduced the crime rate in the 1990s but increased it in the 2000s.

When the signs of the TE are identified for VA, the magnitudes are generally large. For example, using the middle ground DID assumption (see Table 5), RTC laws are found to reduce the VA murder rate in 1995 by at least 16% (from 9.1 to 7.6), the rape rate by at least 6%, the robbery rate by at least 37%, and the larceny rate by at least 13%. In contrast, in 2005 RTC laws are found to reduce the murder rate by at least 15%, but increase the rape rate by at least 47%, the assault rate by at least 8%, the robbery rate by at least 3%, and the burglary rate by at least 13%.

These findings of heterogeneous effects across years may partially explain why the findings obtained with homogenous response assumptions like (6) are sensitive to the years of data included in the sample. Analyses performed using more recent data tend to find that RTC laws have negligible or even positive effects on violent crime, whereas the same analyses performed using data through the early 1990s tend to find that RTC laws decrease violent crime rates. See Aneja *et al.*, (2011) and National Research Council (2005, Chapter 6, Tables 6-5 and 6-6).

Given that the effects vary over time and across crimes, and in many cases do not reveal the sign of the TE, it is difficult to provide a simple assessment of the overall efficacy of the RTC law in VA. One's perspective will necessarily depend both on the assumption that seems most plausible, if any, as well as how to weigh results that vary over time and across crimes. Under the weaker assumptions, the bounds generally do not identify the sign of the TE and thus provide little guidance about whether RTC laws increase or decrease crime. Under the stronger DID invariance assumption (Table 4), RTC laws are found to have reduced crime in the early 1990s and have mixed effects a decade later, in the 2000s. Similar but less definitive results, are found under the middle ground assumption used to generate the results displayed in Table 5.

To assess the overall benefit of a RTC statute, one must somehow aggregate the effects across the different crimes. Obviously, the decision of how to weight the effects for different crimes is a complex and subjective undertaking. A simple metric might be to count the aggregate change in crimes. In this case, using the results from Table 5, one would conclude that RTC laws tended to decrease the number of crimes in the 1990s but increase the total number of crimes in the 2000s.

Instead, one might adjust the raw counts for the fact that some crimes are more costly to society than others. A number of researchers have sought to estimate the average costs of different crimes.[15] By combining the effects of RTC laws on crime rates with a set of cost estimates, one can compute how the average costs of crime would change if a RTC statute were to be adopted in VA. While admittedly a tenuous exercise, this might provide a rough snapshot of the aggregate impact of RTC laws on crime.

A final complication in evaluating the impact of RTC laws arises because the results reported are retrospective, not prospective. The two forms of analysis differ in their objectives. Researchers performing retrospective analysis aim to learn past treatment effects in a study population, asking questions such as: What do we know about the effects of RTC laws in Virginia from 1990 to 2006? A central objective of prospective analysis is to inform treatment choice in a future population, asking questions of the form: What would happen if a state were to adopt a RTC statue in 2017?

Prospective analysis is more difficult than retrospective analysis. Empirical evidence on treatment response is entirely absent prospectively, but it is partially present retrospectively after the outcomes of realized treatments are observed. Under a time invariance assumption, as applied in Assumption (6), the answer is straightforward. The impact of RTC laws does not vary over time. Yet, this assumption is not tenable. One must address the forecasting problem inherent in prospective inference.

[15] As might be expected, there is a great deal of variation in these cost of crime estimates, but they all imply that murder is many times more costly than any other crime. For example, McCollister *et al*. (2010) estimate that the average costs (in 2008 dollars) of a murder is nearly $9 million, whereas the average costs of a rape, the next most costly offense, is $241,000.

In this paper, we have not considered prospective analysis. The literature on RTC laws, including this paper, may provide policymakers useful retrospective information. It is not, per se, informative about prospective questions, and in particular about what would happen if a RTC statute were to be adopted in a state that does not currently have such a law.

5. CONCLUSION: INCREDIBLE CERTITUDE

Given that research on RTC laws is often inconclusive, contradictory, and confusing, the research community has been largely marginalized in this important policy debate. Why? Researchers are rewarded for producing simple findings leading to definitive policy prescriptions – e.g., more guns leads to less crime – yet generating such findings requires strong assumptions that cannot be persuasively defended. In this setting, researchers report findings with “incredible certitude” rather than expressing due caution (Manski, 2013). Drawing inferences about the effects of RTC laws, or gun policy more generally, is an inherently difficult undertaking: conclusions are highly sensitive to the data and assumptions, the available data are limited, and a wide range of assumptions, and thus conclusions, are consistent with the data. Researchers combining data with different maintained assumptions reach different logically valid conclusions, yet fail to adequately express sensitivity of the findings to untestable assumptions.

For empirical research on complex policies such as gun laws, and RTC laws in particular, to be informative to the policy debate, we believe it is critical that the discussion move away from this paradigm of incredible certitude towards an honest portrayal of partial knowledge (Manski, 2013). Adequate expression of caution goes beyond using temperate language, replicating results under marginally different assumptions, or reporting confidence intervals. Although helpful, these means of expressing caution do not go nearly far enough. Drawing inferences under a variety of assumptions that are not credible does not resolve the problem. Adequate expression of caution requires formal methods to perform empirical research under assumptions that are weak enough to be credible.

In this paper, we develop and apply one such set of assumptions, namely bounded variation assumptions. These assumptions, which relax the traditional invariance assumptions applied in the literature, provide an intuitive, simple, and flexible way to improve the credibility of empirical research, assess the sensitivity of inferences to different identifying restrictions, and apply middle ground assumptions that sometimes identify the sign of the TE.

The results reveal the inherent tradeoff between the strength of assumptions and findings. Even under the strongest invariance models, the TE of a RTC statute varies across crimes, states, and time. Findings under the more credible bounded variation models suggest even greater degrees of ambiguity with many results not identifying the sign of the TE, and others varying over time and across states. Still, under these assumptions, we find that the RTC law in VA reduced murder and larceny rates in nearly every year, but increased assaults after 1997. For other crimes, the sign of the bounds is generally negative in the 1990s but positive or indeterminate in the 2000s.

These results are informative but they are inconsistent with the notion that RTC laws uniformly increase or decrease crime. In this light, we do not report findings with incredible certitude: there are no simple conclusions. Still, our findings may inform the policy debate by providing credible information that constrains the resulting discussion to lie within a set of bounds.

References

Abadie, Alberto, Susan Athey, Guido Imbens and Jeffrey Wooldridge (2014), "Finite Population Causal Standard Errors," NBER Working Paper No. 20325.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. (2010), "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program", *Journal of the American Statistical Association*, 105, 493-505.

Aneja, Abhay, John J. Donohue and Alexandria Zhang (2011), "The Impact of Right-to-Carry Laws and the NRC Report: Lessons for the Empirical Evaluation of Law and Policy," *American Law and Economic Review*, 13, 565-632.

Black, Dan A. and Daniel S. Nagin. (1998), "Do right-to-carry law deter violent crime?" *Journal of Legal Studies* 27, 209-19.

Chernozhukov, Victor, Han Hong, and Elie Tamer (2007), "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243-1284.

Chernozhukov, Victor, Simon Lee, and Adam Rosen. (2013), "Intersection Bounds: Estimation and Inference," *Econometrica*, 81, 668-737.

Cochran, William G.. (1977), *Sampling Techniques*, New York: Wiley.

Deaton, Angus. (1997), *The Analysis of Household Surveys*, Baltimore: The Johns Hopkins University Press.

Donohue, John J. and Steven Levitt. (1998), "Guns, Violence, and the Efficiency of Illegal Markets," *American Economic Review*, 88, 463-67.

Duggan, Mark. (2001). "More Guns, More Crime," *Journal of Political Economy*, 109, 1086-1114.

Durlauf, Steven. N., Salvador Navarro, and David A. Rivers. (2016), "Model Uncertainty and the

Effect of Shall-Issue Right-to-Carry Laws on Crime," *European Economic Review*, 81, 32-67.

Imbens, Guido and Charles F. Manski (2004), “Confidence Intervals for Partially Identified Parameters.” *Econometrica*, 72, 1845-1857.

Kaido, Hiroaki, Francesca Molinari and Jorg Stoye, (2016), “Confidence Intervals for Projections of Identified Parameters,” Department of Economics, Cornell University.

Lafférs, Lukas. (2015), “Bounding Average Treatment Effects Using Linear Programming,” cemmap Working Paper CWP70/15.

Lott, John R. (2010), *More Guns, Less Crime: Understanding Crime and Gun-Control Laws*, Chicago: University of Chicago Press.

Lott, John R. and David Mustard (1997), ‘‘Crime, Deterrence and Right-to-Carry Concealed Handguns,’’ Journal of Legal Studies, 26, 1–68.

Manski, Charles F. (1990), “Nonparametric Bounds on Treatment Effects,” *American Economic Review Papers and Proceedings*, 80, 319–323.

Manski, Charles F. (1997), “Monotone Treatment Response,” *Econometrica*, 65, 1311–1334.

Manski, Charles F. (2007), *Identification for Prediction and Decision*, Cambridge: Harvard University Press.

Manski, Charles F. (2013). *Public Policy in an Uncertain World: Analysis and Decisions*, Cambridge: Harvard University Press.

Manski, Charles F. and John V. Pepper (2000), “Monotone Instrumental Variables: with an Application to the Returns to Schooling,” *Econometrica*, 68, 997–1010.

Manski, Charles F. and John V. Pepper (2013), “Deterrence and the Death Penalty: Partial

Identification Analysis Using Repeated Cross Sections," *Journal of Quantitative Criminology*, 29, 123-141.

McCollister, Kathryn E., Michael T. French, and Hai Fang (2010), "The cost of crime to society: New crime -specific estimates for policy and program evaluation," *Drug and Alcohol Dependence*, 102, 98-109.

National Research Council, (2005), *Firearms and Violence: A Critical Review,* ed., Wellford, C., J. Pepper, and C. Petrie, National Academy Press, Washington, D.C.

Richardson, Amy, Michael G. Hudgens, Peter B. Gilbert, and Jason P. Fine (2014), "Nonparametric Bounds and Sensitivity Analysis of Treatment Effects," *Statistical Science*, 29, 596-618.

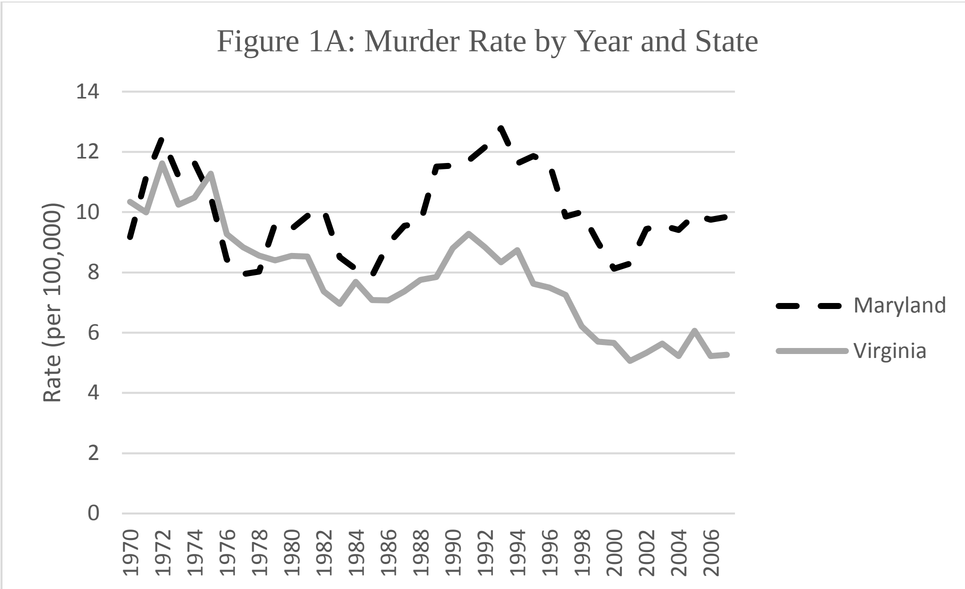
Figure 1A: Murder Rate by Year and State
Rate (per 100,000)
14
12
10
8
6
4
2
0
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
Maryland
Virginia

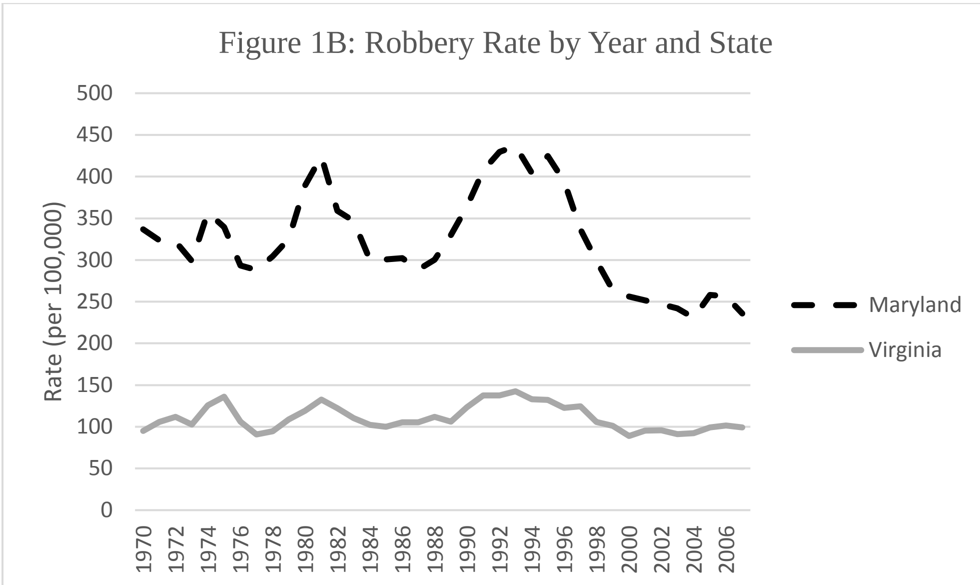

Figure 1B: Robbery Rate by Year and State
Rate (per 100,000)
500
450
400
350
300
250
200
150
100
50
0
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
Maryland
Virginia

Figure 2: Bounds on the TE as a Function of δ

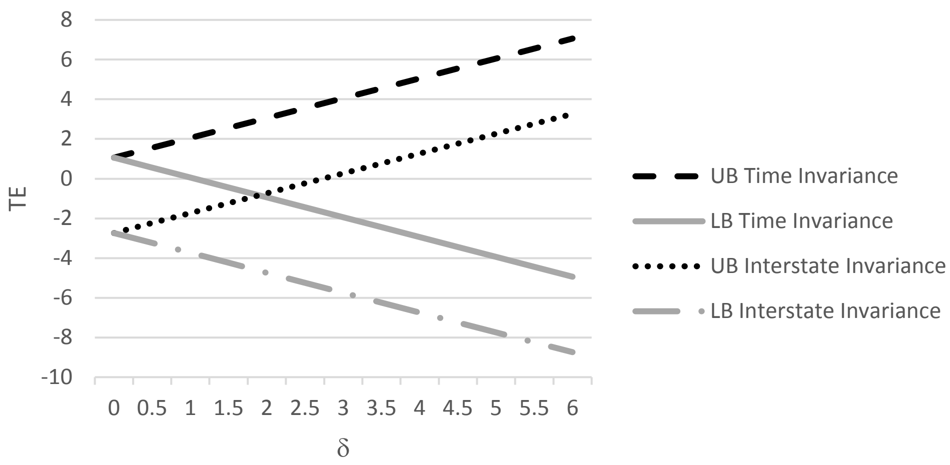

Table 1: Murder Rates per 100,000 Residents by Year and State

Year	Maryland	Virginia
1988	9.63	7.75
1990	11.55	8.81

Table 2: Bounds on the Treatment Effect Given Bounded Interstate and Time Variation Assumptions, Virginia, 1990

δ_VA(1990,1988) \ δ_(VA,MD)1990	0		0.5		1		1.5		2		2.5		3		3.5		4		4.5		5		5.5		6	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
0	0.8	-3.2	0.8	-2.7	0.8	-2.2	0.8	-1.7	0.8	-1.2	0.8	-0.7	0.8	-0.2	0.8	0.3	*0.8*	*0.8*	*0.8*	*0.8*	*0.8*	*0.8*	*0.8*	*0.8*	*0.8*	*0.8*
0.5	0.3	-3.2	0.3	-2.7	0.3	-2.2	0.3	-1.7	0.3	-1.2	0.3	-0.7	0.3	-0.2	*0.3*	*0.3*	*0.3*	*0.8*	*0.3*	*1.3*	*0.3*	*1.3*	*0.3*	*1.3*	*0.3*	*1.3*
1	-0.2	-3.2	-0.2	-2.7	-0.2	-2.2	-0.2	-1.7	-0.2	-1.2	-0.2	-0.7	**-0.2**	**-0.2**	-0.2	0.3	-0.2	0.8	-0.2	1.3	-0.2	1.8	-0.2	1.8	-0.2	1.8
1.5	-0.7	-3.2	-0.7	-2.7	-0.7	-2.2	-0.7	-1.7	-0.7	-1.2	**-0.7**	**-0.7**	**-0.7**	**-0.2**	-0.7	0.3	-0.7	0.8	-0.7	1.3	-0.7	1.8	-0.7	2.3	-0.7	2.3
2	-1.2	-3.2	-1.2	-2.7	-1.2	-2.2	-1.2	-1.7	**-1.2**	**-1.2**	**-1.2**	**-0.7**	**-1.2**	**-0.2**	-1.2	0.3	-1.2	0.8	-1.2	1.3	-1.2	1.8	-1.2	2.3	-1.2	2.8
2.5	-1.7	-3.2	-1.7	-2.7	-1.7	-2.2	**-1.7**	**-1.7**	**-1.7**	**-1.2**	**-1.7**	**-0.7**	**-1.7**	**-0.2**	-1.7	0.3	-1.7	0.8	-1.7	1.3	-1.7	1.8	-1.7	2.3	-1.7	2.8
3	-2.2	-3.2	-2.2	-2.7	**-2.2**	**-2.2**	**-2.2**	**-1.7**	**-2.2**	**-1.2**	**-2.2**	**-0.7**	**-2.2**	**-0.2**	-2.2	0.3	-2.2	0.8	-2.2	1.3	-2.2	1.8	-2.2	2.3	-2.2	2.8
3.5	-2.7	-3.2	**-2.7**	**-2.7**	**-2.7**	**-2.2**	**-2.7**	**-1.7**	**-2.7**	**-1.2**	**-2.7**	**-0.7**	**-2.7**	**-0.2**	-2.7	0.3	-2.7	0.8	-2.7	1.3	-2.7	1.8	-2.7	2.3	-2.7	2.8
4	**-3.2**	**-3.2**	**-3.2**	**-2.7**	**-3.2**	**-2.2**	**-3.2**	**-1.7**	**-3.2**	**-1.2**	**-3.2**	**-0.7**	**-3.2**	**-0.2**	-3.2	0.3	-3.2	0.8	-3.2	1.3	-3.2	1.8	-3.2	2.3	-3.2	2.8
4.5	**-3.2**	**-3.2**	**-3.7**	**-2.7**	**-3.7**	**-2.2**	**-3.7**	**-1.7**	**-3.7**	**-1.2**	**-3.7**	**-0.7**	**-3.7**	**-0.2**	-3.7	0.3	-3.7	0.8	-3.7	1.3	-3.7	1.8	-3.7	2.3	-3.7	2.8
5	**-3.2**	**-3.2**	**-3.7**	**-2.7**	**-4.2**	**-2.2**	**-4.2**	**-1.7**	**-4.2**	**-1.2**	**-4.2**	**-0.7**	**-4.2**	**-0.2**	-4.2	0.3	-4.2	0.8	-4.2	1.3	-4.2	1.8	-4.2	2.3	-4.2	2.8
5.5	**-3.2**	**-3.2**	**-3.7**	**-2.7**	**-4.2**	**-2.2**	**-4.7**	**-1.7**	**-4.7**	**-1.2**	**-4.7**	**-0.7**	**-4.7**	**-0.2**	-4.7	0.3	-4.7	0.8	-4.7	1.3	-4.7	1.8	-4.7	2.3	-4.7	2.8
6	**-3.2**	**-3.2**	**-3.7**	**-2.7**	**-4.2**	**-2.2**	**-4.7**	**-1.7**	**-5.2**	**-1.2**	**-5.2**	**-0.7**	**-5.2**	**-0.2**	-5.2	0.3	-5.2	0.8	-5.2	1.3	-5.2	1.8	-5.2	2.3	-5.2	2.8

Note: To simplify the presentation, the murder rates reported in Table 2 are rounded to the nearest whole number (except for the 1990 murder rate in VA). Bounds highlighted in dark grey are infeasible. Bounds highlighted in light grey and bold text identify the sign of the TE to be negative, and those in light grey and italicized text are identified to be positive.

Table 3: Bounded Variation Parameters

	Interstate	Time	DID	DID_0.75
Murder	2.7	2.0	2.3	1.2
Rape	17	4	6	3
Aggravated Assault	324	39	60	34
Robbery	292	30	55	27
Auto Theft	371	46	61	43
Burglary	496	206	136	67
Larceny	775	410	175	93

Note: These parameters are found by taking the maximum of the minimum parameter value for the models to be consistent with the data in each year prior to 1989. The DID_0.75 parameter is the 0.75 quantile.

Table 4: The TE of a RTC Law in VA Given the DID Invariance Model, by Year and Crime Type

	Murder	Rape	Assault	Robbery	Auto Theft	Burglary	Larceny
1990	**-0.9**	**-4.7**	**-3.9**	**-52.0**	**-10.9**	**-26.9**	*246.3*
1991	**-0.6**	**-5.9**	*11.1*	**-81.1**	**-22.1**	**-12.9**	*45.4*
1992	**-1.4**	**-4.9**	**-10.0**	**-103.1**	**-51.4**	**-61.8**	**-165.3**
1993	**-2.6**	**-1.9**	**-11.5**	**-105.4**	**-29.3**	**-105.4**	**-206.9**
1994	**-1.0**	**-2.2**	**-0.7**	**-82.3**	**-115.8**	**-45.2**	**-304.4**
1995	**-2.4**	**-5.1**	**-9.5**	**-104.0**	**-56.3**	**-102.4**	**-499.7**
1996	**-2.2**	**-0.8**	*1.1*	**-83.0**	**-63.8**	**-44.0**	**-373.8**
1997	**-0.7**	*1.5*	*28.6*	**-23.7**	*49.0*	**-7.1**	**-285.8**
1998	**-1.9**	*3.3*	*38.9*	**-4.7**	*91.0*	**-0.5**	**-296.4**
1999	**-1.4**	*5.1*	*48.6*	*26.1*	*139.7*	**-2.0**	**-222.5**
2000	**-0.6**	*3.8*	**-22.5**	*21.6*	*82.7*	*47.9*	**-381.7**
2001	**-1.3**	*7.7*	**-23.5**	*32.4*	*36.8*	*28.7*	**-231.6**
2002	**-2.2**	*10.1*	**-19.4**	*37.3*	**-1.8**	*66.6*	**-174.8**
2003	**-2.0**	*9.9*	*33.9*	*37.6*	**-48.2**	*53.9*	**-65.8**
2004	**-2.3**	*10.6*	*20.9*	*50.4*	**-43.6**	*84.8*	*18.9*
2005	**-2.0**	*10.7*	*45.8*	*29.8*	**-29.2**	*112.4*	*39.1*
2006	**-2.6**	*12.5*	*65.8*	*33.5*	*19.6*	*111.3*	**-109.3**

Note: Treatment effects in bold text identify the sign to be negative and in italicized text are identified to be positive.

Table 5: Bounds on the TE of a RTC Law in VA Given the Joint DID, Interstate and Time Bounded Variation Models (Assumptions A, B and C), by Year and Crime Type

	Murder		Rape		Assault		Robbery		Auto Theft		Burglary		Larceny	
	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
1990	**-2.0**	**0.0**	**-7.9**	**-1.4**	-38.0	13.1	**-79.2**	**-24.7**	**-53.6**	**-10.9**	**-94.3**	**-81.9**	*153.2*	*291.6*
1991	-1.7	0.3	**-9.2**	**-2.7**	-23.0	28.1	**-108.3**	**-53.8**	**-64.9**	**-22.1**	**-80.3**	**-29.5**	-47.8	138.6
1992	**-2.6**	**-0.6**	**-8.2**	**-1.7**	-44.1	7.0	**-130.4**	**-75.9**	**-94.1**	**-51.4**	**-129.2**	**-104.5**	**-258.5**	**-72.1**
1993	**-3.7**	**-1.7**	-5.2	1.4	-45.6	5.5	**-132.7**	**-78.2**	**-72.0**	**-29.3**	**-172.8**	**-142.4**	**-300.1**	**-113.7**
1994	**-2.1**	**-0.2**	-5.4	1.1	-34.8	16.3	**-109.6**	**-55.1**	**-158.6**	**-115.8**	-112.6	-172.5	**-397.6**	**-211.3**
1995	**-3.5**	**-1.5**	**-8.3**	**-1.8**	-43.6	7.5	**-131.2**	**-76.7**	**-99.0**	**-56.3**	-169.8	-216.2	**-592.9**	**-406.5**
1996	**-3.4**	**-1.4**	**-4.1**	**-0.1**	-33.0	18.1	**-110.2**	**-55.7**	**-106.5**	**-63.8**	-111.4	-223.9	**-467.0**	**-280.7**
1997	-1.9	0.1	-1.8	4.7	-5.4	45.6	-51.0	3.6	*6.2*	*49.0*	-74.5	60.3	**-379.0**	**-192.6**
1998	**-3.0**	**-1.1**	*0.1*	*6.6*	*4.8*	*55.9*	-32.0	22.5	*48.2*	*91.0*	-67.9	67.0	**-389.6**	**-203.3**
1999	**-2.0**	**-0.6**	*1.9*	*8.4*	*32.1*	*65.7*	*-1.2*	53.3	*97.0*	*139.7*	-69.5	65.4	**-315.7**	**-129.4**
2000	-1.7	0.3	*0.5*	*7.0*	13.5	-5.5	*-5.7*	48.8	*40.0*	*82.7*	-19.6	115.3	**-474.8**	**-288.5**
2001	**-2.5**	**-0.5**	*4.5*	*11.0*	15.4	-6.5	*5.2*	*59.7*	-6.0	36.8	-38.7	96.1	**-324.8**	**-138.5**

2002	**-2.4**	**-1.4**	*6.9*	*13.4*	15.0	-2.3	*10.1*	*64.6*	**-44.5**	**-1.8**	-0.8	134.0	**-267.9**	**-81.6**
2003	**-2.1**	**-1.2**	*6.6*	*13.1*	*6.0*	*50.9*	*10.4*	*64.9*	**-55.7**	**-48.2**	-13.6	121.3	-158.9	27.4
2004	**-2.5**	**-1.5**	*7.4*	*13.9*	*3.3*	*38.0*	*23.2*	*77.7*	**-65.7**	**-43.6**	*17.4*	*152.2*	-74.2	112.1
2005	**-1.7**	**-1.1**	*7.4*	*14.0*	*11.7*	*62.8*	*2.6*	*57.1*	**-71.9**	**-29.2**	*45.0*	*179.8*	-54.1	132.2
2006	**-2.5**	**-1.8**	*9.2*	*15.8*	*31.8*	*82.9*	*6.2*	*60.7*	-23.1	19.6	*43.9*	*178.7*	**-202.4**	**-16.1**

Note: Bounds highlighted in dark grey are infeasible. Bounds highlighted in light grey and bold text identify the sign of the TE to be negative, and those in light grey and italicized text are identified to be positive.