

A PUBLIC CHOICE THEORY
OF CRIMINAL PROCEDURE

Keith N. Hylton
Vikramaditya S. Khanna

Revised Version of Discussion Paper No. 318

08/2004

Harvard Law School
Cambridge, MA 02138

The Center for Law, Economics, and Business is supported by
a grant from the John M. Olin Foundation.

The March 2001 version of this paper can be downloaded without charge from:

The Harvard John M. Olin Discussion Paper Series:
http://www.law.harvard.edu/programs/olin_center/

A PUBLIC CHOICE THEORY OF CRIMINAL PROCEDURE

By: Keith N. Hylton[†] & Vikramaditya S. Khanna^{††}

Abstract

We provide a more persuasive justification for the pro-defendant bias in Anglo-American criminal procedure than the most commonly forwarded justifications to date. The most commonly forwarded rationale for the pro-defendant bias is that the costs of false convictions – specifically, the sanctioning and deterrence costs associated with the erroneous imposition of criminal sanctions – are greater than the costs of false acquittals. We argue that this rationale provides at best a partial justification for the extent of pro-defendant procedural rules. Under our alternative justification, pro-defendant protections serve primarily as constraints on the costs associated with improper enforcement or rent seeking in the law enforcement process. The theory developed here explains key institutional features of Anglo-American criminal procedure and provides a positive theory of the case law as well. The theory is also corroborated by empirical evidence on corruption from several countries.

[†] Professor of Law, Boston University School of Law; J.D., Harvard Law School; Ph.D. (Economics), M.I.T. Email: knhylton@bu.edu.

^{††} Professor of Law, University of Michigan Law School; S.J.D. Harvard Law School. Email: vskhanna@umich.edu.

A PUBLIC CHOICE THEORY OF CRIMINAL PROCEDURE

By: Keith N. Hylton[†] & Vikramaditya S. Khanna^{††}

© 2004, Keith N. Hylton & Vikramaditya S. Khanna. All rights reserved.

I. INTRODUCTION

For many people who study Anglo-American criminal law the procedural protections offered to defendants are a bit of a puzzle. At the simplest level, these protections permit some factually guilty defendants to escape conviction, which should increase the incentives of wrongdoers to commit crimes.¹ This seems odd given that criminal wrongs are the most serious wrongs in society and hence the ones we should most wish to reduce.² What explains the willingness

[†] Professor of Law, Boston University School of Law; J.D., Harvard Law School; Ph.D. (Economics), M.I.T. Email: knhylton@bu.edu.

^{††} Professor of Law, University of Michigan Law School; S.J.D. Harvard Law School. Email: vskhanna@umich.edu. This paper has been a long time in gestation, and many have contributed to it with comments. The authors would like to thank Jack Beermann, Ron Cass, John Coates, Kevin Cotter, Dhammika Dharmapala, Ward Farnsworth, Jesse Fried, Ed Glaeser, Wendy Gordon, Mike Harper, Louis Kaplow, Santosh Khanna, Reinier Kraakman, James Krier, Joan Larsen, Fred Lawrence, James Liebman, Michael Meurer, Mitch Polinsky, Adam Pritchard, Mark Ramseyer, Mark Roe, Roberta Romano, David Rossman, Suzanne Scotchmer, Steven Shavell, Peter Siegelman, Stephen Spurr, William Stuntz, Kip Viscusi, and participants at the National Bureau of Economic Research, Summer Institute in Law & Economics, University of Michigan Law School, Law & Economics Workshop, Boston University School of Law, Faculty Workshop, Fordham University School of Law, Faculty Workshop, and Wayne State University, Microeconomics Workshop for their helpful discussions and suggestions. We would also like to thank Michelle Carlucci, Obert Chu, Barbara Coleman, Adam Forchheimer, Jessica Fritz, Angie Nguyen, Kendrick Nguyen, Nicholas Oldham, Seema Srinivasan, Paula Uscilla, and Andrew Yang for excellent research assistance. In addition, many thanks to the John M. Olin Center for Law, Economics, & Business at Harvard Law School for funding support while Professor Khanna was visiting at Harvard Law School in Spring 2001.

¹ See, e.g., Raymond A. Atkins & Paul H. Rubin, *Effects Of Criminal Procedure On Crime Rates: Mapping Out The Consequences Of the Exclusionary Rule*, (Oct. 23, 1998) (unpublished manuscript available on file with authors) (finding *Miranda* may have increased total crime rates by eleven percent and violent crimes rates by thirty-three percent); Paul G. Cassell, *The Guilty & The "Innocent": An Examination Of Alleged Cases Of Wrongful Conviction From False Confessions*, 22 HARV. J.L. & PUB. POL'Y 523 n.30 (1999)[hereinafter *Guilty & Innocent*]; Paul G. Cassell, *Miranda's Social Costs: An Empirical Reassessment*, 90 NW. U.L. REV. 387, 451 (1996)[hereinafter *Miranda's Social Costs*] (finding after *Miranda* criminal suspects are less willing to confess to their crimes). *But see*, John J. Donohue, III, *Did Miranda Diminish Police Effectiveness*, 50 STAN. L. REV. 1147 (1998) (dissecting the statistical analyses in Cassell's work and criticizing the use of statistics in measuring the import of Court decisions).

² See S.E. Marshall & R.A. Duff, *Criminalization and Sharing Wrongs*, 11 CAN. J.L. & JURIS. 7, 7 (1998) (noting that the term "criminal" indicates a serious condemnation of an activity or action); Susan Estrick, *Rape* 95 YALE L.J. 1087, 1183 (1986) (noting that "conduct is labeled 'criminal' in order to announce to society that these actions are not to be done and to secure that fewer of them are done"); Henry M. Hart, Jr., *The Aims of The Criminal Law*, 23 LAW & CONTEMP. PROBS. 401, 404-05 (stating that "[w]hat distinguishes a criminal from a civil sanction . . . is the judgement of community condemnation which accompanies and justifies its imposition"). See also John. C. Coffee, Jr., *Does "Unlawful" Mean "Criminal"? : Reflections on the Disappearing Tort/Crime Distinction in American Law*, 71 B.U.L. REV 193, 194 (1991) (noting that "the factor that most distinguishes the criminal law is its operation as a system of moral education and socialization . . . [a]s

of Anglo-American law to permit wrongdoers to escape conviction and thereby potentially increase the crime rate?

Commentators have offered a variety of justifications for “pro-defendant” procedural protections.³ The most commonly-forwarded justification is that in the criminal process we should be more concerned about the social costs generated by a false conviction than the social costs generated by a false acquittal.⁴ According to this view, false convictions are especially worrisome because they involve the state’s denial of liberty to an innocent individual, the infliction of irreparable harm to an individual’s reputation, and weaken the moral force or authority with which the criminal law speaks.⁵ In *In Re Winship*, the Supreme Court adopted this position as the primary rationale for viewing the high standard of proof used in criminal trials as constitutionally mandated.⁶ We will refer to this position as the “traditional error cost justification .”

However, the traditional error cost justification does not, we argue, provide a complete explanation for pro-defendant procedural protections. In this paper we provide and develop an alternative justification for the procedural protections: to constrain the costs associated with abuses of prosecutorial or punishment authority (or rent seeking more broadly). We call this a public choice theory of criminal procedure. This justification, we argue, provides a superior explanation for the procedural protections than the traditional error cost justification.

a result, the criminal law often and necessarily displays a deliberate disdain for the utility of the criminalized conduct to the defendant”); Sanford H. Kadish, *Excusing Crime*, in *BLAME AND PUNISHMENT: ESSAYS IN THE CRIMINAL LAW* 87 (1987) (concluding that “Criminal conviction charges a moral fault . . .”; HERBERT PACKER, *THE LIMITS OF THE CRIMINAL SANCTION* (1968).

³ See JOSHUA DRESSLER, *UNDERSTANDING CRIMINAL PROCEDURE*, (2d ed. 1997) (discussing various pro-defendant biases and their justifications); Donald J. Boudreaux & Adam C. Pritchard, *Civil Forfeiture and the War on Drugs: Lessons from Economics and History*, 33 *SAN DIEGO L. REV.* 79 (1996); Randolph N. Jonakait, *Biased Evidence Rules: A Framework for Judicial Analysis and Reform*, 1992 *UTAH L. REV.* 67, 68 (arguing that pro-defendant evidence rules are defensible as they promote justice).

⁴ See RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 604 – 05 (5th ed., 1998).

⁵ See *In Re Winship*, 397 U.S. 358, 363-364 (1970) (Brennan, J.) (“The accused during a criminal prosecution has at stake interests of immense importance, both because of the possibility that he may lose his liberty upon conviction and because of the certainty that he would be stigmatized by the conviction. Accordingly, a society that values the good name and freedom of every individual should not condemn a man for commission of a crime when there is a reasonable doubt about his guilt. . . . Moreover, use of the reasonable-doubt standard is indispensable to command the respect and confidence of the community in applications of the criminal law. It is critical that the moral force of the criminal law not be diluted.”), 372 – 74 (Harlan, J., concurring) (noting that the reasonable doubt standard is used in criminal trials because society views false convictions as being far worse than false acquittals largely due to liberty and reputation costs). See also Donald A. Dripps, *People v. Simpson: Perspectives On The Implications For The Criminal Justice System: Relevant But Prejudicial Exculpatory Evidence: Rationality Versus Jury Trial And The Right To Put On A Defense*, 69 *S. CAL. L. REV.* 1389, 1418 (1996) (noting that “[the reasonable doubt standard] strikes the balance very much in favor of [increasing] false acquittal[s]” as opposed to false convictions); Bruce H. Kobayashi & John R. Lott, Jr., *Low-Probability-High-Penalty Enforcement Strategies and the Efficient Operation of the Plea-Bargaining System*, 12 *INT’L REV. L. & ECON.* 69 (1992) (analyzing the costs of false convictions).

⁶ See *In Re Winship*, *supra* note 5, at 364.

To be sure, many courts and commentators have made the point that criminal procedural protections constrain the potential for abuse by government agents.⁷ However, the key contributions of this paper are (1) the development and elaboration of this observation in terms of an economic framework, (2) the use of this observation, as part of that framework, to form the basis of a positive theory of the case law and conventions of criminal procedure, and (3) an empirical analysis showing that procedural protections dampen corruption, a byproduct of rent-seeking. In summary, the theory we develop provides a better explanation, than the traditional story, for the existence and specific form of key rules and institutional features: the reasonable-doubt and double-jeopardy rules, restrictions on excessive and retroactive punishments, and features of the right to a jury, such as the unanimity requirement and peremptory challenges. The theory is also consistent with the historical development of these rules and is corroborated by empirical evidence on corruption from several countries.

We start with the claim that procedural protections make it more costly for self-interested actors, whether individuals or government enforcement agents, to use the criminal process to obtain their desired ends; which in turn saves resources and enhances deterrence. Absent some constraint, prosecutors and government agents would use the criminal process to benefit themselves or their constituents. History provides a number of examples of this. Procedural protections impose constraints and make the criminal process more costly to use, providing enforcement agents and those who would lobby them with a disincentive to use it for selfish ends. This saves resources that otherwise would be eaten up by lobbying and makes it more difficult for a variety of forms of corruption to flourish.⁸ In addition, procedural protections are likely to enhance deterrence. The reason is that when it is easy to enforce the law selectively, enforcement agents will come under pressure to sacrifice deterrence objectives in favor of other goals for which people may lobby.

Part II provides a description of some core criminal procedural protections – primarily the reasonable-doubt and double jeopardy rules. Part III inquires into the traditional rationales for these protections. Part IV presents the public choice rationale for procedural protections. In this part we explain the incentives

⁷ See e.g., YALE KAMISAR, WAYNE R. LAFAVE & JEROLD H. ISRAEL, *MODERN CRIMINAL PROCEDURE* 114 – 127 (8th ed., 1994); WAYNE R. LAFAVE & AUSTIN W. SCOTT, JR. *CRIMINAL LAW* 97 - 100 (2d. ed.); WAYNE R. LAFAVE, JEROLD ISRAEL & NANCY J. KING, *CRIMINAL PROCEDURE* 25 – 39 (3rd ed., 2000); *Ashe v. Swenson*, 397 U.S. 436, 447 (1970). See David Friedman, *Why Not Hang Them All: The Virtues Of Inefficient Punishment*, 107 J. POL. & ECON. 259, 262 - 63 (1999) (discussing how “inefficient” punishments, like prison, may also constrain government agents).

⁸ See, e.g., GORDON TULLOCK, *THE ECONOMICS OF SPECIAL PRIVILEGE AND RENT SEEKING* 96, (Gordon Tullock ed., Kluwer Academic Publishers 1989); CHARLES K. ROWLEY ET. AL., *THE POLITICAL ECONOMY OF RENT-SEEKING* 465-478, (Charles K. Rowley et al. eds., Kluwer Academic Publishers 1988).

for and various types of wealth extraction in the law enforcement process. Part V discusses the costs of rent seeking (or abuses of prosecutorial or punishment authority) in the criminal law enforcement process.

Part VI examines methods of constraining self-interested prosecutors, which include procedural protections and restrictions on penalties. In this part we present theories of the core functions of pro-defendant protections. We argue that their core function is to constrain the set of contractible bribes between a prosecutor and another party. We also argue that the reasonable-doubt and double jeopardy rules function as complements in suppressing corruption in enforcement. This works for both decisions to enforce *and* not to enforce the law. Lastly, the unanimity rule regarding jury convictions and the original function of peremptory challenges are both explained as features designed to dampen rent seeking in enforcement.

Part VII applies the core theories developed in Part VI to provide a positive theory of the case law on procedural protections. Part VIII provides empirical evidence that pro-defendant protections dampen corruption, which is consistent with our public choice theory.

II. SOME CORE PRO-DEFENDANT CRIMINAL PROCEDURAL PROTECTIONS

There is a vast panoply of procedural protections attached to the criminal process in the U.S. To focus our analysis we hone in on core protections that impose a significant pro-defendant bias in the criminal law process, by reducing either the probability of conviction or the severity of punishment.⁹ In particular, we focus on the reasonable-doubt standard of proof and double jeopardy protections.

The reasonable-doubt standard requires that the moving party (i.e., the prosecution) prove that the defendant is guilty, beyond a reasonable doubt, of the criminal offense(s) with which he is charged.¹⁰ Although the reasonable-

⁹ Rules that indirectly impose a bias by restricting the type of evidence that can be presented, such as the exclusionary rule or the right to silence, do not fall within our definition of core protections. This is because their impact on the probability of conviction is not as direct as the reasonable doubt standard and also because they do not have the same kind of historical pedigree that the reasonable doubt standard has. See Lawrence H. Tribe, *Constitutional Calculus: Equal Justice Or Economic Efficiency?*, 98 HARV. L. REV. 592, 607 (1985) (noting that “[E]xclusionary rule cases...are today treated as occasion for the assessment of the marginal deterrent effects of excluding particular categories of evidence”). Also note that it might be easier for police and other government agencies to satisfy some parts of the Fourth Amendment (e.g., giving a *Miranda* warning) as compared to satisfying the reasonable doubt standard. See Charles Weisselberg, *Saving Miranda*, 84 CORNELL L. REV. 109 (1998) (advocating a return to the “original” vision of *Miranda*). However, much of the analysis in the text can be applied to the “non-core” protections as well.

¹⁰ See *In Re Winship*, *supra* note 5, at 361.

doubt formulation seems to have first appeared in 1798,¹¹ the notion that the standard of proof in criminal trials should favor defendants has ancient origins. Blackstone, in his description of the criminal process, noted that “all presumptive evidence of felony should be admitted cautiously: for the law holds, that it is better that ten guilty persons escape, than one innocent suffer.”¹² Coke, considerably earlier, said that “the evidence against a prisoner should be so manifest, as it could not be contradicted.”¹³ In 1970, the Supreme Court “constitutionalized” this norm by holding in *In Re Winship* that the due process clause protects the defendant against conviction except upon proof beyond a reasonable doubt.¹⁴

The reasonable-doubt standard stands in contrast to the “preponderance of the evidence” standard, used most frequently in non-criminal cases and for sentencing issues in criminal proceedings.¹⁵ It requires that the moving party prove that the defendant is liable on the preponderance of the evidence or, put simply, is more likely liable than not.¹⁶ Sometimes the preponderance rule is assumed to require that the decision-maker be 51% certain that the defendant is liable before finding against him, whereas the reasonable doubt standard is assumed to require that the decision-maker be somewhere in the range of 90% to 95% certain before convicting the defendant.¹⁷

¹¹ See C. McCORMICK, EVIDENCE § 341, at 576 – 78 (1992).

¹² 4 WILLIAM BLACKSTONE, COMMENTARIES, at 358.

¹³ *Id.*, at 349 – 50.

¹⁴ See *In Re Winship*, *supra* note 5, at 364.

¹⁵ See *Concrete Pipe and Products Of California, Inc. v. Construction Laborer’s Pension Trust for Southern California*, 508 U.S. 602, 622 (1993) (noting that preponderance of the evidence is the “most common standard in the civil law”); *Apprendi v. New Jersey*, 530 U.S. 466 (2000) (noting that the preponderance of the evidence standard can be used for sentencing as long as the sentence is not more severe than the statutory maximum for the offense established by the jury’s verdict); *U.S. v. Lombard*, 72 F.3d 170, 176 (1st Cir. 1995) (noting that the preponderance of the evidence can be used in sentencing issues [enhancements in this case]).

¹⁶ See *U.S. v. Mandanici*, 205 F.3d 519, 532 (2d Cir. 2000) (Kearse, J. concurring) (noting that “a preponderance means more likely than not”).

¹⁷ See *Ethyl Corp. v. EPA*, 541 F.2d 1, 28 n. 58 (D.C. Cir. 1976) (stating that “It may be that the ‘beyond a reasonable doubt’ standard of criminal law demands 95% certainty [internal cites omitted]. But ..., a preponderance of the evidence demands only 51% certainty.”); *Brown v. Bowen*, 847 F.2d 342, 345-46 (7th Cir. 1988) (suggesting that certainty of 90% or more is sufficient to meet the reasonable doubt standard). There is yet another standard of proof that might be used in some instances. The clear and convincing evidence standard is used in some non-criminal contexts often as a method of determining if greater-than-compensatory damages should be awarded in a particular case. See *id.*, at 346. This standard occupies an intermediate position between the preponderance and the reasonable doubt standards. See, e.g., *Anderson v. Liberty Lobby, Inc.*, 477 U.S. 242, 271 (1986); *Cornell v. Nix*, 119 F.3d (8th Cir. 1997). Although these three standards are sufficient for purposes of our analysis it is worth noting that in the past there have been other standards. In fact, there was at one time a standard even higher than the reasonable doubt standard. See Anthony M. Morano, *A Reexamination of the Development of the Reasonable Doubt Rule*, 55 B.U.L. REV. 507, 511-19 (1975) (arguing that prior to the articulation of the reasonable doubt standard, a higher standard existed that required proof beyond any doubt). Indeed, there is, in theory, a continuum of standards of proof, but in practice only three. See V.S. Khanna, *Corporate Criminal Liability: What Purpose Does It Serve?*, 109 HARV. L. REV. 1477, 1516 n. 210; See also McCORMICK, *supra* note 11, at 378. See also, *Sullivan v. LTV Aerospace and Defense Co.*, 82 F.3d 1251, 1260 (2d Cir. 1996) (noting that “[t]he familiar burden of proof standards occur

Another procedural protection is the prohibition against Double Jeopardy, which stems from the 5th Amendment to the U.S. Constitution. The 5th Amendment states “nor shall any person be subject for the same offense to be twice put in jeopardy of life or limb”.¹⁸ In many respects this protection is similar to the doctrines of Res Judicata and Collateral Estoppel that are found in non criminal cases.¹⁹ However, there are some differences. In particular, one that has garnered much attention is the rule, unique to criminal proceedings, that normally prohibits prosecutorial appeals of initial trial acquittals, but permits defense appeals of initial trial convictions.²⁰

Other core protections that are relevant in the criminal context are the right to a jury trial,²¹ the ex post facto punishment rule,²² and the excessive punishments prohibition.²³ This does not exhaust the list of protections that impose a pro-defendant bias, but the crucial point for now is that many criminal procedural rules appear biased in favor of the defense. What justifies this bias?

III. TRADITIONAL JUSTIFICATION FOR THE CORE CRIMINAL PROCEDURAL PROTECTIONS: THE REASONABLE-DOUBT STANDARD

The reasonable-doubt standard is the quintessential case of a pro-defendant protection. The traditional justification for it is that in the criminal process we are more concerned with false convictions than false acquittals and hence should prefer a pro-defendant bias.²⁴ Elaborating, the Supreme Court in

along a continuum with ‘preponderance of the evidence’ at one end and ‘beyond a reasonable doubt’ at the other. The ‘clear and convincing evidence’ standard falls somewhere in between.”).

¹⁸ U.S. CONST. amend.V.

¹⁹ See Robin W. Sardegna, *No Longer In Jeopardy: The Impact Of Hudson v. U.S. On The Constitutional Validity Of Civil Monetary Penalties For Violations Of The Securities Laws Under The Double Jeopardy Clause*, 33 VAL. U. L. REV. 115, 117 (1998) (noting that the civil procedure doctrines of res judicata and collateral estoppel are similar to prohibitions of the double jeopardy clause).

²⁰ See Vikramaditya S. Khanna, *Double Jeopardy’s Asymmetric Appeal Rights: What Purpose Do They Serve?*, 82 B.U.L.REV. 341 (2002); *U.S. v. Ball*, 163 U.S. 662, 671 (1896); Joshua Steinglass, *The Justice System in Jeopardy: The Prohibition on Government Appeals*, 31 IND. L. REV. 353 (1998); Kate Stith, *The Rise of Legal Error in Criminal Cases: Some Consequences of the Asymmetry in the Right to Appeal*, 57 U. CHI. L. REV. 1 (1990); Peter Westen & Richard Drubel, *Toward a General Theory of Double Jeopardy*, 1978 SUP. CT. REV. 81. In the non-criminal side the analogous doctrines (e.g., Collateral Estoppel) do not present such asymmetry in appeal rights. See *Cooper v. Federal Reserve Bank of Richmond*, 467 U.S. 867, 874 (1984) (noting that if the requirements are met both res judicata and collateral estoppel are available to plaintiffs and defendants).

²¹ U.S. CONST. AMEND. VI.

²² U.S. CONST. ART. I, § 9, CL. 3.

²³ U.S. CONST. AMEND. VIII.

²⁴ See 4 WILLIAM BLACKSTONE, COMMENTARIES *358; see also *In Re Winship*, *supra* note 5, at 372 (Harlan, J., concurring).

Winship identified three types of harm associated with false convictions: *loss of liberty stigma*, and *dilution of the moral force of the criminal law*.²⁵

In this part, we briefly reconsider the evidence supporting the traditional error cost argument of *Winship*. In order to address the argument directly, we will translate the Court's analysis into terms that can be identified as measurable or immeasurable. Specifically, we can classify the costs of false convictions and false acquittals as *sanctioning costs*, *deterrence costs*, and *disutility costs*.²⁶ Sanctioning costs are the social costs of punishment, which include the loss of liberty and stigma effects recognized in *Winship*. Deterrence costs refer to part of what the *Winship* Court intended when describing the dilution in law's moral force caused by false convictions. If the law punishes the innocent, then it weakens deterrence by giving people less incentive to comply and, similarly, if it fails to punish the guilty, then it weakens deterrence for obvious reasons.²⁷ Disutility costs refer to the other part of what the *Winship* Court intended when describing the dilution in the law's moral force: the disutility individuals suffer when they know that the law fails to punish the guilty or that it sometimes punishes the innocent.²⁸

Translating *Winship* into these terms, the Court has said in effect that the additional sanctioning, deterrence, and disutility costs of moving from the reasonable-doubt to the preponderance standard would outweigh the potential benefits. Is there empirical support for this claim?

²⁵ See *In Re Winship*, *supra* note 5, at 363-364. Although the Court did not conduct an examination of false acquittal costs, those costs presumably consist only of the third type of harm, dilution of the law's moral force.

²⁶ In theory one could consider the impact on the "expressive" effects of the law too. The expressive effect of the law is the effect it has on behavior without threatening a sanction. See Richard H. McAdams, *A Focal Point Theory of Expressive Law*, 86 VA. L. REV. 1649 (2000), Robert Cooter, *Expressive Law and Economics*, 27 J. LEGAL STUD. 585 (1998). However, the law's expressive effect is probably more closely connected to the perceived legitimacy of the law – that is whether it is corrupt or easy to corrupt – rather than a particular trade off between types of errors. See Jason Mazzone, *When Courts Speak: Social Capital and Law's Expressive Capital*, 49 SYRACUSE L. REV. 1039 (1999). The corruption concern is more closely tied to our rent-seeking/public choice theory than the traditional error cost account and as such is not discussed in this section. We discuss the expressive effect as impacted by our theory in Part V.B.2.

²⁷ On the deterrence cost, see A. Mitchell Polinsky & Steven Shavell, *Legal Error, Litigation, and the Incentive to Obey the Law*, 5 J. L. ECON. & ORG'N 99 (1989) (noting that if a potential defendant will not be prosecuted whether or not he obeys the law, he "will not obey the law, because there is a cost to obeying but no benefit (assuming, of course, that there is a benefit in disobeying)"). On the issue of those who go unpunished, see Don E. Scheid, *Constructing a Theory of Punishment, Desert, and the Distribution of Punishments*, 10 CAN. J.L. & JURIS. 441, 455 (1997) (noting that "[p]eople are willing to obey the law themselves so long as they can reasonably assume that those who break the law will not be able to do so with impunity, that they will not get away with it.>").

²⁸ Suppose people gain utility from knowing that the criminal justice system is fair. Then a system that has some parts perceived to be unfair will reduce the utility of those people who obtain utility from knowing the system is fair. This is clearly a harm that can be traced to any change that weakens the moral force of the criminal law.

To state a rough empirical test of the traditional error cost argument requires us to set aside disutility costs, for the moment, and focus only on the costs that are measurable: *sanctioning* and *deterrence* costs. Assuming that convictions drop under the reasonable doubt rule relative to the preponderance standard then we have fewer sanctioning costs. Further, under the reasonable doubt rule we would expect deterrence to worsen relative to the preponderance rule.²⁹ *The reasonable-doubt rule is thus preferable if the incremental sanctioning costs under the preponderance rule exceed the incremental (under-) deterrence costs under the reasonable-doubt rule.*

There are data that allow us to get a rough sense of whether this condition holds. The most direct sources are studies by David Anderson,³⁰ and by John Donohue and Peter Siegelman.³¹ If we use the sum of losses due to injuries and property theft (including fraud) as a conservative definition of the aggregate harm from crime, Anderson's study suggests the annual aggregate crime cost is \$1,031 billion.³² If we measure sanctioning costs by adding the opportunity costs

²⁹ In other words, we expect the incremental deterrence costs of false acquittal (D_{fa}) to exceed the incremental deterrence costs of a false conviction (D_{fc}). As a theoretical matter, this is demonstrable. First, both are positive because increases in either type of error reduce the difference in payoffs for complying and not complying with the law, *see, e.g.*, Louis Kaplow & Steve Shavell, *Accuracy in the Determination of Liability*, 37 J.L. ECON. 1, 5 (1994). Second, $D_{fa} > D_{fc}$, as a theoretical matter because in a well functioning enforcement system, enforcement efforts will tend to be targeted at guilty actors. Given this, changes in false acquittal probabilities will far outweigh changes in false conviction probabilities in terms of their influence on incentives. For a formal analysis reaching the same conclusion, *see* Henrik Lando, *The Optimal Standard of Proof in Criminal Law When Both Fairness and Deterrence Matter*, SSRN Working Paper available at http://papers.ssrn.com/paper.taf?abstract_id=238334. Both the argument here and Lando's model are easily reconcilable with Craswell and Calfee's article on the deterrence effects of error; *see* Richard Craswell & John E. Calfee, *Deterrence and Uncertain Legal Standards*, 2 J. L. ECON. & ORG. 279 (1986).

³⁰ *See* David Anderson, *The Aggregate Burden of Crime*, 42 J. LAW & ECON 611 (1999). The sanctioning costs of the death penalty are not discussed here because of the small number of executions each year. *See* Hashem Dezhbakhsh, Paul H. Rubin, & Joanna Mehlhop Shepherd, *Does Capital Punishment Have a Deterrent Effect? New Evidence from Post-moratorium Panel Data*, Draft 2001 (on file with authors), at 29 (providing a table of the number of executions over the last 23 years which appear to be around 40 to 50 a year).

³¹ *See* John J. Donohue III & Peter Siegelman, *Allocating Resources Among Prisons and Social Programs in the Battle Against Crime*, 27 J. LEG. STUD. 1 (1998).

³² *See* Anderson, *supra* note 30 at 629 (providing a figure of \$1,102 Billion which represents the sum of 'Risks to life and health', 'Crime-induced production', and 'Opportunity Costs' (for criminals and victims)). We have reduced this amount by \$71 Billion, which represents the sanctioning costs associated with using prison as a penalty as this is more appropriately accounted for as sanctioning costs for our analysis. *See infra* note 33. The "risks to life and health" reflect the value of crime-related lives lost (approximately 72,111 lives lost per year valued at about \$6.1 Million each) and the value of non-fatal injuries. *See* Anderson, *supra* note 30, at 624 – 626. The \$6.1 Million per life measure appears to be about the average value from many previous studies. *See id.*, at 626. The value of non-fatal injuries is also based on an average of prior studies valuing non-fatal injuries. *See id.*, at 626. The "crime-induced production" reflects the estimated amount of resources spent on items that result from or are associated with crime and hence cannot be spent on other items. *See id.*, at 616 – 617. "Examples include the production of personal protection devices, the trafficking of drugs, and the operation of correctional facilities." *Id.* at 616. Andersen provides a fairly detailed list of these expenditures. *See id.*, at 620. The "opportunity costs" associated with crime in the text represent the value of the days victims were unable to work due to the crime event, the time and effort spent by criminals in undertaking crime, and the time and effort of victims in attempting to prevent crime (e.g., locking things). *See id.*, at 623 – 624.

of the inmate's time while locked up plus the costs of maintaining inmates in prison, Anderson's study suggests that the annual sanctioning cost for *all convictions* is \$71 billion.³³

These numbers imply that the aggregate costs of crime, which can be treated as underdeterrence costs, are on the order of 15 times greater than the sanctioning costs for all convictions. In view of the magnitude of this differential, it seems improbable that the savings from a measure that reduces crime by improving deterrence, such as moving to the preponderance standard, would be swamped by a rise in sanctioning costs. Indeed, the empirical evidence on the responsiveness of crime to changes in prison population suggests that this is unlikely.

Surveying evidence on the responsiveness of crime to incarceration, Donohue and Siegelman estimate that the elasticity of crime with respect to incarceration is .15 and Steven Levitt finds an elasticity of .30.³⁴ Using the lower figure, an increase in sanctioning costs of 200% (i.e., tripling the prison population and increasing sanctioning costs from \$71 to \$213 billion – a net increase of \$142 billion), should be associated with a reduction in crime on the order of 30%, $((200)(.15) = 30)$.³⁵ This yields a deterrence benefit of \$309 billion. Since the sanctioning cost of tripling the prison population (\$142 billion) is less than the deterrence benefit (\$309 billion), a switch to the preponderance standard appears desirable on traditional error-cost grounds.

In sum, the readily-available empirical evidence does not suggest that increases in sanctioning costs would fully offset the crime-reduction benefits from switching to the preponderance rule.³⁶ Indeed, the evidence suggests that

³³ See *id.*, at 620, 624 (this figure represents the sum of "Crime-Induced Production: Corrections" and "Criminal lost workdays: in prison"). We recognize that these numbers miss some important costs. Indeed, they do not include the stigma costs, part of the liberty costs, and the moral disutility costs of diluting the moral force of the criminal law referred to by the Court in *Winship*. However, we are using the numbers only to determine whether the sanctioning and deterrence costs of errors, as commonly understood, would justify the reasonable doubt rule. An important part of our analysis is to show the rough magnitude of other costs that need to be considered to justify the reasonable-doubt rule if standard sanctioning and deterrence costs do not.

³⁴ See Donohue & Siegelman, *supra* note 31, at 13; Steven Levitt, *The Effect of Prison Population Size on Crime Rates: Evidence from Prison Overcrowding Litigation*, 111 Q. J. ECON. 319 (1996)(finding elasticity figure of .30).

³⁵ There is a net gain of \$167 Billion on these numbers by switching to the preponderance standard. Crime has dropped by \$309 Billion and sanctioning costs have risen to \$213 Billion from \$71 Billion (i.e., an increase of \$142 Billion). Thus, \$309 Billion less \$142 Billion is \$167 Billion, which makes switching to the preponderance rule desirable on traditional error-cost grounds.

³⁶ We should address a few issues raised by this discussion. First, the numbers we have chosen from the Anderson study are the most conservative figures that he uses, in the sense that they represent the lowest estimate of the cost of crime that Anderson provides. Thus, they should be the most supportive of the traditional error cost rationale - but even they do not support the traditional error cost rationale. Second, we have excluded costs associated with risk bearing. In particular, we have excluded the risk bearing costs

the preponderance rule's additional sanctioning costs are substantially less than the reasonable-doubt rule's under-deterrence costs.

This analysis suggests that for the traditional error-cost rationale to justify society's choice of the reasonable-doubt standard, other immeasurable costs (e.g., stigma or disutility costs) not explicitly considered here must be quite large. Our analysis allows one to see the size that must be assumed of these costs. For example, return to the scenario just considered: a tripling of the prison population produces a deterrence benefit of \$309 billion and an increased sanctioning cost of \$142 billion. To justify the reasonable-doubt standard, the immeasurable costs would have to add \$167 billion to the change in total sanctioning costs.

The required magnitude of the immeasurable costs is so large that one is forced to question it. Put another way, *Winship's* justification for the reasonable doubt rule does not appear to be strongly supported by the readily-available empirical evidence. In light of this, some other justification should be considered.

IV. A PUBLIC CHOICE JUSTIFICATION FOR PRO-DEFENDANT PROCEDURAL PROTECTIONS

We contend that the reasonable-doubt standard is designed primarily to make it harder for individuals and groups to use the criminal process as a mechanism for wealth extraction. This reduces the social costs generated by rent-seeking efforts to influence law enforcement and enhances the deterrent effect of the law.

Implicit in this justification for pro-defendant criminal procedure is the assumption that the state and its agents may have incentives that diverge from those of the social ideal. The social ideal is the outcome that a platonic philosopher-king would choose in order to maximize social welfare. For reasons that have been explained in the public choice literature, it is not necessarily the choice made by a democratically elected legislature,³⁷ much less that of a non-

of being a potential victim and of being falsely convicted. However, since the risk bearing costs associated with becoming a victim are probably far larger than those associated with a false conviction, this exclusion biases the example against our thesis. We thank Mitch Polinsky for the risk-bearing point.

³⁷ See, e.g., Jonathan R. Macey, *Promoting Public-Regarding Legislation Through Statutory Interpretation: An Interest Group Model*, 86 COLUM. L. REV. 223, 230-231 (1986) (expounding on the methods which interest groups can use to influence the political process, and noting that interest groups can distort legislators' thinking, which would then serve special interests rather than public interest). See also Gary Minda, *Interest Groups, Political Freedom, and Antitrust: A Modern Reassessment of the Noerr-Pennington Doctrine*, 41 HASTINGS L.J. 905, 945 (1990) (commenting that "[i]nterest groups are seen as 'rent-seekers' who influence the political process to achieve their selfish economic interests, and legislators are seen as motivated to advance their own self-interest"); Stephen J. Schulhofer, *Plea Bargaining as Disaster*, 101 YALE L.J. 1979, 1987 (1992)

democratic government. Moreover, prosecutors' interests may diverge from the social ideal as well. For example, the social ideal might be to maximize the number of correct convictions and minimize the number of false convictions subject to a budget constraint.³⁸ Prosecutors, however, might be interested in trying to maximize the number of convictions, to advance further in their careers, make money, or a variety of other things.³⁹ The incentive divergence can manifest itself in behaviors that induce lobbying of prosecutors or other state actors.⁴⁰

Rent seeking in the law enforcement process can occur in a variety of forms. However, we think two general types capture its observed forms. One is

(articulating that as an agent of the state, prosecutors' decisions "tend to diverge from those that would most efficiently serve the public interest" due to personal or professional reasons); Stephen J. Schulhofer, *Criminal Justice Discretion as a Regulatory System*, 17 J. LEGAL STUD. 43 (1988).

³⁸ See Khanna, *supra* note 20, at 361 – 62. For a similar approach, also see Edward Glaeser & Andrei Schleifer, *Incentives For Enforcement*, Draft 2000; Scott Baker & Claudio Mezzetti, *Prosecutorial Resources, Plea Bargaining, and the Decision to Go To Trial*, forthcoming 17 J. L. ECON. & ORG'N (2001) (using a model where prosecutors are concerned about correct convictions and false convictions amongst other things); Dirk G. Christensen, *Comments: Incentives vs. Nonpartisanship: The Prosecutorial Dilemma in an Adversary System*, 1981 Duke. L. J. 311 (1981); Darryl K. Brown, *Criminal Procedure Entitlements, Professionalism, and Lawyering Norms*, 61 OHIO ST. L.J. 801, 801 (2000).

³⁹ See Edward L. Glaeser et al., *What Do Prosecutors Maximize? An Analysis of the Federalization of Drug Crimes*, 2 J.L. & ECON. 259, 262 266 (2000); Gordon Van Kessel, *Adversary Excess in the America; Criminal Trial*, 67 NOTRE DAME L. REV. 403, 441 (1992); William M. Landes, *Economic Analysis of the Court*, 14 J. LAW & ECON. 61 (1971); Daniel C. Richman, *Essay Old Chief v. United States: Stipulating Away Prosecutorial Accountability?*, 83 VA. L. REV. 939, fn. 93 (1997); Christensen, *supra* note 38, at 321. Other reasons for a divergence between the social ideal and what prosecutors do may include that prosecutors and philosopher-kings consider the same things important, but weight them differently. For example, a philosopher-king might value the avoidance of a false conviction more than a prosecutor might. Society might also value certain convictions more than prosecutors might (e.g., society might value the conviction of one drug overlord more than 20 convictions of small time drug dealers, but prosecutors might view things differently). In addition, error by prosecutors in assessing the social value of certain cases may further widen the gap between prosecutorial and social welfare maximizing behavior.

Some of these issues are tied up with how prosecutors are rewarded. See Christensen, *supra* note 38, at 311; Daniel C. Richman, *Federal Criminal Law, Congressional Delegation and Enforcement Discretion*, 46 UCLA L. REV. 757, 818 n. 101 (1999); Tracey L. Meares, *Rewards for Good Behavior: Influencing Prosecutorial Discretion and Conduct with Financial Incentives*, 64 FORDHAM L. REV. 851 (1995). See Glaeser, et al., *supra*, at 261 (noting that decisions to prosecute are often influenced by prosecutors' interest in running for a higher office); Van Kessel, *supra*, at 442 (1992) (noting how win-loss records are very important to state prosecutors). This raises interesting questions about why we reward prosecutors in the way we do. This is the subject of a separate paper and outside the scope of our present inquiry. As a first cut, one suspects convictions rates are the more verifiable assessment criteria relative to others (much like profits are a bit easier to verify as assessment criteria relative to others in the corporate sphere when dealing with the agency problem in that context).

⁴⁰ Because prosecutorial or state interests may not match the social ideal we face what can be termed an "agency cost" problem. See Michael C. Jensen & W.H. Meckling, *Agency Costs and the Theory of the Firm*, 3 J. FIN. ECON. 305, 310 (1976) (discussing "agency cost" in the context of the agency theory of corporations). In other words, to the extent state actors take decisions that deviate from the social ideal, we can think of the social welfare losses associated with their decisions as "agency" or incentive-divergence costs. Admittedly, this is a different approach to agency costs. The standard account in Jensen & Meckling's article treats agency costs as the costs that result because the agent's incentives differ from those of the principal. See Jensen & Meckling, *supra*. In our description in the text, we are treating agency costs as the costs that result because state actors have incentives that deviate from those of a platonic philosopher-king.

inter-group wealth expropriation, which arises when one group attempts to gain some advantage from law enforcement agents at the expense of other members of society.⁴¹ Equivalently, we might describe this type of rent-seeking as an advanced version of tribalism. The other general type of rent-seeking is *simple corruption*, which occurs when an individual uses bribery or some other means, in connection with law enforcement, to extract wealth from society or to avoid punishment for wrongdoing.

A. *Inter Group Wealth Expropriation*

Rent-seeking behavior is a result of prosecutors acting in their unconstrained self-interest. If there are no constraints prosecutors may use the criminal process to benefit themselves by selling their power to enforce the criminal law to the highest bidder. What prosecutors receive could include direct or indirect monetary gain, enhanced chances for power and prestige, or anything else of value to the prosecutor.⁴² To simplify we can say that prosecutors receive a certain sum – say \$1 million – from the highest bidding group (A) to enforce the law in a particular manner. Enforcement of the law in this manner must benefit the highest bidding group by more than \$1 million – say \$1.5 million – and these gains would appear to come at the expense of the non-A groups (i.e., group B for simplicity). Thus, in a sense, A is using the prosecutor and the criminal process to extract \$1.5 million from B by paying the prosecutor \$1 million. The various groups realizing that the prosecutor is willing to sell his services will lobby to secure some of these gains from prosecutors and to prevent other groups from extracting wealth from them.⁴³

Inter-group wealth expropriation can be effected in a number of ways. For example, there is lobbying that results in targeted enforcement. This occurs when prosecutors disproportionately target certain groups for purposes of bringing prosecutions due to the lobbying efforts of other groups. The familiar case is where prosecutors disproportionately bring charges against members of group B because of the lobbying efforts, or simply to curry the favor, of the dominant group A. This permits group A members to shift the burden of criminal enforcement onto group B members or to impose costs on group B which result in some benefit to group A – e.g., the maintenance of a caste system.

⁴¹ See Einer R. Elhauge, *Does Interest Group Theory Justify More Intrusive Judicial Review?*, 101 YALE L.J. 31, 35 (1991); Richard H. McAdams, *Cooperation and Conflict: The Economics Of Group Status Production And Race Discrimination*, 108 HARV. L. REV. 1003, 1029-1030 (1995) (noting that “[W]hen interest groups pursue what economists call “rent-seeking” legislation, such as farm subsidies and tax “loopholes,” they seek merely to transfer resources from one group to another”); Angela O. Davis, *The American Prosecutor: Independence, Power, and the Threat of Tyranny*, 86 IOWA L. REV. 393, 459 (2001).

⁴² See, e.g., Richman, *supra* note 39, at 818, n. 102 (noting prosecutors’ eagerness for career advancement); Christensen, *supra* note 38, at 318 (noting that prosecutors are financially motivated).

⁴³ Cf. Paul H. Rubin & Martin J. Bailey, *The Role of Lawyers in Changing the Law*, 23 J. LEGAL STUD. 807, 822-23 (1994) (rent-seeking motives of lawyers result in social waste).

Indeed, in regimes in which prosecutors are elected, candidates for the position will have incentives to seek support from group A by promising to direct enforcement efforts against group B.⁴⁴ Perhaps the best known example of this in United States history is law enforcement in the South during the Jim Crow period, which involved numerous instances of prosecutors refusing to enforce the law against white citizens, while using the threat of criminal punishment to coerce black citizens.⁴⁵

Another example of wealth expropriation in the enforcement process is where prosecutors ask for some benefit to avoid bringing charges against members of politically marginal groups. This is akin to extortion or a protection racket. Knowing that politically powerful groups will have him removed from office if he threatens their interests, the self-interested prosecutor would focus his extraction efforts on the politically marginal.⁴⁶ The difference between this version of inter-group wealth extraction and the first is slight: in the first, the dominant group initiates the wealth extraction process and in the second, the prosecutor initiates the process. In the first case, the dominant group is likely to pay the prosecutor the smallest amount necessary to accomplish their ends, allocating the surplus to themselves. In the second, the prosecutor who initiates the wealth extraction process allocates the surplus to himself. One example of this occurs in rural areas of China, where local police officers have tried to enrich themselves by enforcing certain prohibitions, such as the one-child policy, against relatively poor farmers.⁴⁷

Yet another example of wealth expropriation is the passing of laws with disproportionate burdens on different groups.⁴⁸ This is analogous to targeted enforcement. In order to effect wealth expropriation, the laws need not apply directly to group B members. The dominant group (A) may find that certain activities are carried out only, or predominantly, by group B members, or that

⁴⁴ See Richman, *supra* note 39; William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 520 – 34 (2001) (discussing the incentives of various participants in the American Criminal Justice system); Tracey L. McCain, *The Interplay of Editorial and Prosecutorial Discretion in the Perpetuation of Racism in the Criminal Justice System*, 25 COLUM. J.L. & SOC. PROBS. 601, 648, n. 81 (1993) (noting that decisions to prosecute are susceptible to political influence because most prosecutors are elected); Dwight L. Greene, *Abusive Prosecutors: Gender, Race & Class Discretion and the Prosecution of Drug-Addicted Mothers*, 39 BUFF. L. REV. 737, 777 (1991) (noting that “Prosecutors ... are capable of conducting their offices in ways to advance their own political careers”).

⁴⁵ See RICHARD EPSTEIN, *FORBIDDEN GROUNDS* 91-97 (1992); William J. Stuntz, *Race, Class, and Drugs*, 98 COLUM. L. REV. 1795, 1839 (1998) (noting that robbery laws in the Jim Crow era were enforced against blacks more often than whites especially where white robbers stole from black victims).

⁴⁶ See William N. Eskridge, *Religion, Homosexuality, and Collisions of Liberty and Equality in American Public Law*, 106 YALE L.J. 2411, 2447 (1997); Greene, *supra* note 44, at 799.

⁴⁷ See Elisabeth Rosenthal, *Rural Flouting of China’s One-Child Policy Undercuts Census*, N.Y. Times, April 14, 2000.

⁴⁸ See, e.g., Stuntz, *supra* note 45, at 1795 (discussing the heightened police attention given to urban crack markets dominated by the lower economic class while law enforcement pays relatively little attention to the upscale powder-cocaine market).

group B members carry out these activities in a different manner from others. With this information, the dominant group may prohibit or place special burdens on the activity when carried out in a particular manner. For example, white majorities in the western United States enacted several facially neutral statutes in the late 1800s that had the effect of prohibiting Chinese laundries, both to limit competition from them and to limit the independent work options of Chinese laborers.⁴⁹

These cases are united by the common theme of one group benefiting at the expense of others through the use of the governmental process, whether law enforcement or legislation. As our concern is with law enforcement, we will not discuss in much depth the passing of laws with disproportionate burdens. However, the legislative and law enforcement processes provide alternative routes through which a predatory dominant group could extract wealth from others. In this sense, the legislative and law enforcement processes are substitutes in the eyes of the wealth extractor. The laws controlling rent seeking in these governmental processes can be understood as complements, in the sense that they prevent a predatory dominant group from shifting its expropriation efforts from one governmental process to another.⁵⁰

B. *Simple Corruption*

The other type of rent-seeking, simple corruption, involves the effort of a single individual or group to extract wealth from the general population. We have in mind two cases: that of an enforcement agent (police officer or prosecutor) who threatens to apprehend and charge an individual unless he pays the agent, and that of an enforcement agent who is merely willing to accept bribes from the general public or individuals to enforce or not enforce the law in a particular way.⁵¹ Thus, we consider instances where enforcement agents take bribes not to enforce the law or where they take bribes to enforce the law in a particular way. In general, these payoffs can take two forms. One, *ex ante* bribery, occurs when an individual bribes an enforcement agent before he commits a crime in exchange for an agreement by the agent not to enforce the

⁴⁹ See David E. Bernstein, *Lochner, Parity and the Chinese Laundry Cases*, 41 WM. & MARY L. REV. 211 (1999).

⁵⁰ For discussion of how the procedural protections may induce greater lobbying at the legislative level see Stuntz, *supra* note 44, *passim*. We discuss this matter at *infra* notes 102.

⁵¹ One could argue that the simple corruption category is the same as our second example of wealth expropriation, and we concede that the difference is more a matter of degree than of character. In the second example, the enforcement agent maintains his position through the support of local dominant groups. In the simple corruption story, the enforcement agent is either unconcerned with maintaining support from local dominant groups (in the case of the actively predatory enforcer), or passively accepts bribes in exchange for not enforcing the law.

law against him. In the other form, ex post bribery, the individual bribes the agent after he commits the crime.⁵²

There are many examples of simple corruption. A common example of ex post bribery is a police officer that accepts bribes in return for not issuing a ticket to a speeding motorist. Ex ante bribery appears to be less common, though there are many examples of it too. In most towns in the U.S., local government business is carried out by boards made up of residents with deep and strong connections to many of the parties who appear before them. In these settings, it is hard to distinguish the ordinary reciprocal exchanges that are part of normal social intercourse from ex ante bribery. The Supreme Court grappled with a routine example of this in *City of Columbia v. Omni Outdoor Advertising*,⁵³ which involved the efforts of a local billboard company to protect its incumbency advantage by encouraging the city council to prohibit the erection of new billboards.⁵⁴

Because the criminal law enforcement process can be used by groups or by individuals as a means to extract wealth, we should anticipate a steady stream of efforts to use it for that purpose. In light of this it becomes important to get a sense of the precise costs generated by such behavior and methods for constraining them.

V. COSTS ASSOCIATED WITH ABUSIVE OR “RENT-SEEKING” BEHAVIOR IN THE CRIMINAL PROCESS

We divide our discussion of the costs of rent-seeking in law enforcement into two parts: first, the costs related to the act of lobbying, and second, the costs related to the effect of lobbying on the deterrent force of the criminal law.

A. Direct Costs

Rent-seeking in the criminal process involves lobbying efforts by individuals or groups to influence the selection and prosecution of cases. The process of lobbying itself generates costs that are from a societal perspective often wasteful.

⁵² For an economic analysis of criminal law that incorporates these distinctions, see Mehmet Bac & Parimal Kanti Bag, *Law Enforcement Costs and Legal Presumptions* 5 - 6 (Draft, 2000)(on file with authors). See also Mehmet Bac, *Corruption, Supervision and the Structure of Hierarchies*, 12 J.L. ECON & ORG’N. 277 (1996).

⁵³ 499 U.S. 365 (1991).

⁵⁴ The excluded firm brought an unsuccessful antitrust lawsuit on the ground that the incumbent firm had colluded with city officials. The Supreme Court’s reluctance to apply the antitrust laws to this behavior is based in large part on the difficulty in distinguishing ex ante bribery from ordinary social intercourse. See *id.* at 379 – 80.

As a useful analogy, consider efforts to obtain a monopoly. Expenditures to obtain a monopoly may be desirable when a firm secures a dominant position through competition because a firm typically does this by improving its product, or reducing its costs; activities that increase the total surplus.⁵⁵ However, some expenditures to obtain a monopoly are socially wasteful; consider, for example, lobbying efforts to obtain a government privilege (e.g., an exclusive license or tariff protection) or to deny a deserved privilege to someone else.⁵⁶ This kind of naked wealth extraction creates no additional surplus and merely transfers an asset from one party to another at the social cost of the resources used to effectuate that transfer.⁵⁷

In the context of criminal law enforcement, efforts by group A to lobby the prosecutor to enforce selectively against group B are often wasteful in a sense similar to naked wealth extraction. Of course, the result is not an entire waste if targeting group B reduces the overall costs of crime.⁵⁸ However, there is little reason to believe that lobbying for selective enforcement will always bring about an efficient result. For example, group A will have no interest in inducing the prosecutor to go after cases of crime involving only members of group B as victims. Further, group A members may discourage the prosecutor from enforcing the law when members of their own group commit crimes against group B. The social waste from lobbying includes some portions of the lobbying efforts of politically dominant groups, the counter-lobbying efforts of politically non dominant groups, and some portion of the effort and time spent by government officials in addressing wealth transfers and maneuvering to obtain positions in which they can direct such transfers.⁵⁹

So far we have focused on the costs associated with inter-group expropriation efforts, but analogous arguments apply in the context of simple

⁵⁵ See PHILLIP AREEDA & LOUIS KAPLOW, *ANTITRUST ANALYSIS: PROBLEMS, TEXT, AND CASES* 7 (5th ed. 1997) (noting that “[c]ompetitive forces generate efficiency in two ways. Productive efficiency occurs as low cost producers undersell and thereby displace the less efficient. Allocative efficiency occurs as exchanges in the marketplace direct production away from goods and services that consumers value less and toward those they value more...”).

⁵⁶ See DENNIS C. MUELLER, *PUBLIC CHOICE* II 229 – 46 (rev. ed. 1989).

⁵⁷ See *id.*, at 229 – 246.

⁵⁸ It might be that lobbying coincides with what might be socially desirable. Also lobbying is not necessarily limited to inter-group lobbying. For example, if the victims of crimes by members of group B are other group B members then group B may lobby prosecutors to stop crime in their areas and hence lobby for prosecutions against other group B members. This sort of lobbying does not raise the kinds of concerns we are discussing in this paper. See Gordon Tullock, *Efficient Rent-Seeking*, in *TOWARD A THEORY OF THE RENT-SEEKING SOCIETY* 97 – 112 (James M. Buchanan, R.D. Tollison & Gordon Tullock eds., 1980).

⁵⁹ See, e.g., Gary S. Becker, *A Theory of Competition among Pressure Groups for Political Influence*, *QUARTERLY JOURNAL OF ECONOMICS* 371, 373-374 (August 1983). The easier it is to obtain governmental favors the more lobbying should appear. Also, as the value of the issues at stake increases one should expect greater expenditures as well.

corruption.⁶⁰ Corruption creates costs in terms of the resources spent in bribing the enforcement agent to enforce or not to enforce the law in some way and the efforts of the agent in positioning himself to take bribes. However, there is an important feature of the corruption model that suggests that rent-seeking costs can be much larger than appears initially. The enforcement process is vertically fragmented, in the sense that it moves through a chain beginning with a police officer (and next his superiors), moving to a prosecutor, and on to a magistrate or judge, and so on. If each one of these agents demands bribes to enforce or not enforce the law, then the total social waste will be considerably larger than in a vertically integrated enforcement regime in which a single agent controls the process from arrest to punishment.⁶¹

B. *Deterrence and Other Costs from Rent-Seeking*

In the context of the criminal process rent-seeking, if successful, has the effect of skewing the enforcement of the law. If members of society perceive (i) skewed enforcement and (ii) that such enforcement is related to rent-seeking efforts, then compliance with the law and the deterrent effect of the law are likely to be compromised.

1. *Some deterrent effects of selectively enforcing the law*

Selective enforcement of the law due to lobbying can have corrosive effects on deterrence. Consider, for example, where lobbying by group A

⁶⁰ Note that a bribe might be efficient in some cases, much as lobbying for targeted enforcement might be. See Gary S. Becker & George J. Stigler, *Law Enforcement, Malfeasance and Compensation of Enforcers*, 3 J. LEGAL STUD. 1 (1974); Padideh Ala'i, *The Legacy of Geographical Morality and Colonialism: A Historical Assessment of the Current Crusade Against Corruption*, 33 VAND. J. TRANSNAT'L 877, 899 (2000).

⁶¹ See Andrei Shleifer & Robert W. Vishny, *Corruption*, 108 Q. J. ECON. 599 (1993). The vertical fragmentation of law enforcement means that each individual agent is in a position similar to that of successive owners of the pieces of a toll road. One of the standard results is that the sum of the tolls charged by successive owners will be larger than the toll charged by a single owner of a road. This is known in the monopolization literature as the double-marginalization problem, see Joseph J. Spengler, *Vertical Integration and Antitrust Policy*, 58 J. POL. ECON. 347 (1950); Andy C.M. Chen & Keith N. Hylton, *Procompetitive Theories of Vertical Control*, 50 HASTINGS L.J. 573, 623 (1999). See also Howard A. Shelanski & J. Gregory Sidak, *Antitrust Divestiture in Network Industries*, 69 U. CHI. L. REV. 1, 23 (2001) (arguing that "under simple models of vertical control where the downstream firm is assumed to have market power, social welfare is unambiguously increased by the elimination of the double marginalization. One firm rather than two marks up the price of the upstream product, leading to a lower price and higher output."); John E. Lopatka & Andrew N. Kleit, *The Mystery of Lorain Journal and the Quest for Foreclosure in Antitrust*, 73 TEX. L. REV. 1255, 1297 (1995) (arguing that "if the products were complements, their provision by a single supplier could increase efficiency, for monopoly provision of complementary products can avoid the kind of double-marginalization problem well recognized in the context of vertically related, or successive, monopolies.") For discussion of a related problem – the problem of the anti-commons see Michael A. Heller, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621 (1998); James Buchanan & Yong J. Yoon, *Symmetric Tragedies: Commons and Anticommons Property*, JOURNAL OF LAW AND ECONOMICS, (forthcoming 2000); Francesco Parisi, Ben Depoorter, & Norbert Schulz, *Duality in Property: Commons and Anticommons*, University of Virginia School of Law, Law & Economics Research Papers Series, Research Paper No. 00-16 (Draft 2001) (on file with authors).

members results in selective law enforcement against group B members, with little regard to the actual guilt of the defendants, and disproportionately less enforcement against group A members.

In this scenario, deterrence is likely to drop for both groups A and B. In general, deterrence can be achieved through *substitution* effects or *scale* effects. Substitution effects occur when a change in the effective sanction leads potential offenders to substitute legitimate, law-complying conduct for illegitimate, undesirable conduct. Scale effects occur when enforcement causes potential offenders to stay out of certain areas, or off the streets at certain times. A selective enforcement policy in which group B is targeted implies, within a fixed budget setting, a diversion of resources from substitution-oriented policies to scale-oriented policies. This is analogous to shifting from a strategy of ticketing every motorist that speeds (inducing substitution toward slow driving) to a strategy of ticketing motorists who meet the profile of a speeder, *whether or not they are speeding* (inducing drivers who fit the profile to stay off the roads).

Now, in our scenario, group A members are under-deterred because they are facing low expected sanctions for engaging in undesirable activities.⁶² The deterrent effect on group B members would also be reduced because they are now punished whether they have acted “good” or “bad.”⁶³ In other words, the incentive to comply with the law is reduced for group B members because the payoffs from compliance and noncompliance have gotten closer.

An enforcement policy chosen by group A members that reduces deterrence may seem irrational, because it could lead to additional crimes being committed against group A members. However, there are scenarios in which such a policy could be chosen rationally by group A. For example, if groups A and B are geographically segregated, group A may choose to reduce potential crimes by Bs on As by apprehending all Bs who venture into their territory, whether or not the Bs are complying with the law.⁶⁴ Such a policy would make it costly for Bs to move among As, encouraging the Bs to stay in their own territory. In other words, group A members may choose to rely on scale effects

⁶² See Polinsky & Shavell, *supra* note 27.

⁶³ See *id.*, at 104 (stating that “A type II error (a truly innocent defendant is found liable) lowers the incentive to obey the law because he will face liability even if he obeys, thereby reducing the benefit to him of obeying the law”). See, e.g., Joel Schrag & Suzanne Scotchmer, *Crime and Prejudice: The Use of Character Evidence in Criminal Trials*, 10 J. L. ECON. & ORG. 319 (1994); Polinsky & Shavell, *supra* note 27.

⁶⁴ See, e.g., David A. Harris, *The Stories, The Statistics, and the Law: Why “Driving While Black” Matters*, 84 MINN. L. REV. 265, 271(1999) (discussing law enforcement’s tendency to subject black drivers in upscale neighborhoods to traffic stops). Note that one implication of the theory here is that “racial profiling” may be socially undesirable under some conditions. For some empirical evidence on racial profiling see John Knowles, Nicola Persico & Petra Todd, *Racial Bias in Motor Vehicle Searches: Theory and Evidence*, 109 J. POL. ECON. 203 (2001); John J. Donohue III & Steven Levitt, *The Impact of Race on Policing, Arrest Patterns, and Crime* (Draft 1998)(on file with authors).

to reduce the risks posed to them by group B offenders. Such a policy, in a fixed budget setting, could easily result in a weakening at the substitution-effect level in deterrence among Bs.⁶⁵

2. *Other costs: stigma effects, expressive effects, enforcement cost effects*

Rent seeking may also reduce the stigma associated with the criminal law. If the criminal label carries some stigma then that becomes part of the total sanction for criminal behavior and influences deterrence.⁶⁶ If something reduces the stigma without a countervailing increase in the official sanction then deterrence is reduced.⁶⁷

The stigma from being labeled a criminal stems in part from a belief that the person so labeled has violated a societal norm meriting condemnation and has been adjudicated in a “fair and impartial” manner.⁶⁸ If, however, being labeled a criminal is perceived by members of one’s social circle as indicative of a political or biased use of the law, which itself may be the result of rent-seeking in the law enforcement process, then the stigma from the criminal label diminishes.⁶⁹

For similar reasons, rent seeking is likely to reduce any “expressive” or symbolic features of the criminal law. The criminal law may express to society

⁶⁵ Another scenario might be where lobbying by group A members leads to selective enforcement against group B members focusing on those who are guilty, but again with disproportionately less enforcement against group A members. In this case the same potential under-deterrence problem for group A members exists because again they face low expected sanctions. However, for the group B members the situation has changed. The deterrent effect on group B members may still remain or be enhanced because in this case if a group B member complies with the law he is probably not going to be prosecuted and if he fails to comply he is more likely to be prosecuted. The overall impact on deterrence is ambiguous in this case. However, given that A members are assumed to control the enforcement process, one probable scenario is where their law enforcement agent (the prosecutor) is unable, because of unfamiliarity or indifference, to distinguish complying from non-complying members of group B. With a prosecutor unable or unwilling to distinguish the “good” from the “bad” among group B, a policy of targeting only the bad in group B would be infeasible. This suggests that the outcome discussed in the text, in which deterrence falls unambiguously, is a plausible outcome of selective enforcement.

⁶⁶ See Khanna, *supra* note 17, at 1508 – 9; David Cole, *The Paradox of Race and Crime: A Comment on Randall Kennedy’s “Politics of Distinction”*, 83 GEO. L. J. 2547, 2561 (1995).

⁶⁷ See Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 VA. L. REV. 349, 350 (1997); Richard Moran, *Home Sweet Home: Given a Choice, Many Convicts are Now Opting for Jail Instead of Probation: Why?*, BOSTON GLOBE, October 29, 2000, at E1.

⁶⁸ See generally Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW.U. L. REV. 453 (1997).

⁶⁹ See ERIC POSNER, *LAW AND SOCIAL NORMS* 97 – 100 (2000) (noting that criminal offenders can signal loyalty to a subcommunity by violating the law and being punished by the dominant group. The subcommunity is more likely to view criminal punishment as a signal of loyalty to the subcommunity the more the subcommunity believes the criminal justice system is “infected with a political agenda”); Dan Kahan, *supra* note 67, at 357 – 58; Janice Nadler, *The Effects of Perceived Injustice on Deference to the Law*, at 9-10, Draft 2000 (on file with authors)(suggesting that, if the law is seen to be imposed in an irrelevant or immoral manner, it will not be deferred to).

what is undesirable behavior and this encourages individuals not to engage in this activity because they respect or otherwise value the law (i.e., legal norm internalization).⁷⁰ If so, then a perceived increase in rent-seeking can only serve to weaken this respect because the message that the activity is undesirable is clouded by the message that law enforcement is selective and biased. This reduces whatever ability the criminal law has to express condemnation, shape preferences or influence behavior outside of pure deterrence based on expected sanctions.⁷¹ Second, members of society may derive some utility from expressing condemnation.⁷² Although this may be true in some instances the benefits inuring from this must be weighed against the costs of rent-seeking and how much the perception of rent-seeking reduces the utility from expressing condemnation.⁷³

In addition to stigma and expressive effects, rent seeking may have enforcement-cost effects. If law enforcement is perceived to be biased then it is likely that some people will refuse to assist law enforcement.⁷⁴ This would increase the difficulty and costs associated with apprehensions and prosecutions, which in turn reduces the likelihood that wrongdoers would be sanctioned.⁷⁵

VI. METHODS OF CONSTRAINING RENT-SEEKING

In this Part we examine two methods of constraining law enforcement agents in order to dampen rent seeking in the law enforcement process: procedural protections and penalty restrictions.⁷⁶ We also provide a positive theory for the basic procedural rules in this area.

⁷⁰ See, e.g., Robert Cooter, *Do Laws Make Good Citizens? An Economic Analysis of Internalizing Legal Values*, U.C. Berkeley Law and Economics, Working Paper Series, p. 16; Dan Kahan, *What do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591, 603 – 04 (1996); Richard McAdams, *Legal Construction of Norms*, 86 VA. L. REV. 1649 (2000); Cass Sunstein, *Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2022, 2025-29 (1996).

⁷¹ See Kahan, *supra* note 70; Sunstein, *supra* note 70.

⁷² See Dan M. Kahan, *Punishment Incommensurability*, 1 BUFF. CRIM. L.R. 691, 693 – 94; see also Interdisciplinary Program Series, *The New Chicago School Myth of Reality?*, 5 U. CHI. L. SCH. ROUNDTABLE 1, 3 (1998); Robinson & Darley, *supra* note 68, at 472.

⁷³ See generally Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 HARV. L. REV. 961, 1244 – 1245 (2001).

⁷⁴ Cf. Nadler, *supra* note 69, at 32 – 41 (discussing results of certain experiments which suggest that group identity does matter when law enforcement is perceived to be “unjust” reflected in willingness to commit unrelated wrongs and mock juror verdicts).

⁷⁵ One possibility worth noting is that if a group of people are reluctant to provide information to law enforcement about other members in their group (absent selective enforcement issues) then that might be a reason itself to proceed with selective enforcement in that group to increase the probability of being sanctioned. See Kay B. Perry, *Fighting Corruption At The Local Level: The Federal Government’s Reach Has Been Broadened*, 64 MO. L. REV. 157, 162 (1999). Of course, such a policy might only exacerbate the “us” and “them” mentality.

⁷⁶ Constraining prosecutorial behavior to reduce rent-seeking costs is, in a rough sense, analogous to the efforts of corporate law in constraining the agency costs arising from the separation of ownership and control. This analogy is far from perfect. We want enforcement agents to protect society’s interests, even if

A. Procedural Protections

Procedural protections constrain rent-seeking in the criminal process by making the process costly for both law enforcement agents and for society. These costly protections deter those who would seek to use the criminal process as a means to transfer wealth. Procedural protections increase costs to society because they require greater effort by the prosecutor to obtain convictions, which increases the number of false acquittals and the number of meritorious cases not brought.⁷⁷ Prosecutors, who are rewarded when conviction rates are high, bear much of the brunt of the costs of these procedures.⁷⁸

A key function of pro-defendant procedural protections is to increase the probability that a prosecutor and a bribing party will be unable to find a mutually-acceptable bribe, thus making the set of contractible bribes zero or close to it. This works as follows.

On the prosecutor's side, procedural protections raise the cost of targeting innocent parties *as well as* the costs of letting the guilty go free compared to where these protections are absent. If the prosecutor targets innocent group B

they diverge from those of the government. In the corporate context, we usually ask corporate agents to act on behalf of shareholders, whether or not shareholder interests diverge from the social ideal. *See generally* ROBERT C. CLARK, *CORPORATE LAW* (1986); WILLIAM A. KLEIN & J. MARK RAMSEYER, *BUSINESS ASSOCIATIONS – AGENCY, PARTNERSHIPS, AND CORPORATIONS* (4th ed. 2000). The prosecutor can be viewed as an agent for society in a manner similar to how a manager or employee is often viewed as an agent for a corporation. *See* Rebecca Hollander-Blumoff, *Getting to “Guilty”: Plea Bargaining as Negotiation*, 2 HARV. NEGOTIATION L. REV. 115 (1997) (noting that “[a] prosecutor... is the agent of the people whom the office purports to protect.”). However, unlike the corporate context, environmental factors that constrain the agency costs of private firms are not present in the case of governments. Since governments do not issue stock, we do not observe discounts in their share prices due to agency costs, nor do governments face the same risk of losing out to competitors as business enterprises do. *See* CLARK, *supra*, Ch. 4 (1986); Tamar Frankel, *Fiduciary Layoff* 71 CALIF. L. REV. 795, 811 (1983) (arguing that different compensation schemes and techniques should have the effect of reducing conflicts of interest with fiduciaries). Hence, explicit constraints on rent-seeking play a relatively important role in the public sector.

⁷⁷ *See generally* Atkins & Rubin, *supra* note 1; Cassell, *Guilty & Innocent*, *supra* note 1; Cassell, *Miranda’s Social Costs*, *supra* note 1; Donohue, *supra* note 1; Kahan, *supra* note 1, at 389.

⁷⁸ *See* J. Mark Ramseyer & Eric B. Rasmussen, *Why is the Japanese Conviction Rate So High*, 30 J. LEGAL STUD. 53 (2001) (arguing that because of the stigma in Japan of acquitting defendants and the corresponding detrimental career effects it may have for prosecutors and judges, they prosecute only strong cases); Fred C. Zacharias, *Structuring the Ethics of Prosecutorial Trial Practice: Can Prosecutors Do Justice*, 44 VAND. L. REV. 45, n. 264 (arguing that “[t]o the extent a prosecutor’s conviction rate is all that counts, the institutional incentives point toward minimizing the responsibility to ‘do justice.’”) *See also* Tracey L. Meares, *Rewards for Good Behavior: Influencing Prosecutorial Discretion and Conduct with Financial Incentives*, 64 FORDHAM L. REV. 851 (proposing a system of incentives to effect prosecutors’ conduct).

The difficulty of using the process and the attendant cost involved makes it less useful as a wealth-extraction tool than perhaps other options and should shift lobbying and rent-seeking away, to some extent, from the criminal process to other methods of influencing government behavior. *See* Paul H. Rubin, Christopher Curran, & John Curran, *Litigation Versus Legislation: Forum Shopping by Rent-Seekers*, Draft 1999 (on file with authors).

members, he is unlikely to be successful given all the pro-defendant procedures. If he maintains his promise to target group B, he will have very few successful prosecutions, and will probably lose his job (as high conviction rates are important to prosecutors). This suggests that the prosecutor will demand a very high bribe in order to adopt a selective enforcement policy. Moreover, given the risk of losing his job, potential bribers should doubt the credibility of the prosecutor's promise to selectively enforce. Given the difficulty of implementing a successful selective enforcement policy, the doubtful credibility of the prosecutor, and the negligible benefits to both parties, the potential briber's willingness-to-pay should fall substantially.⁷⁹

Procedural protections also make it more difficult for corruption to flourish. A prosecutor who threatens to arrest individuals on false charges would find it considerably more difficult to mount a credible threat against his victims in the presence of pro-defendant procedural protections. Moreover, the prosecutor's ability to credibly promise *not to enforce* the law against a particular defendant (in exchange for a bribe) should fall. For if the prosecutor charges the wrong person or no one at all, he most likely will be unsuccessful in obtaining a conviction for the crime. Given his difficulty in charging and convicting an alternate candidate, the cost to the prosecutor of promising not to enforce against a particular defendant is relatively high, and the promise probably cannot be considered fully credible. These factors suggest the prosecutor will demand a large bribe. Hence, the prosecutor's power to shake down individuals for money in exchange for a promise not to bring charges is diminished in the regime with procedural protections. This does not mean that prosecutors will never take bribes not to enforce the law, but rather than the procedural protections increase the costs and reduce the benefits to the prosecutors of attempting this strategy compared to if these protections were absent. From the perspective of the potential defendant, his willingness to pay a bribe falls since he is less likely to be convicted in the first place, and any promise by the prosecutor not to enforce cannot be regarded as fully credible in light of the protections.⁸⁰

⁷⁹ See A. Mitchell Polinsky & Steven Shavell, *Corruption and Optimal Law Enforcement*, J. PUB. ECON. (forthcoming 2001); A. Mitchell Polinsky and Steven Shavell, *On the Disutility and Discounting of Imprisonment and the Theory of Deterrence*, 28 J. LEGAL STUD. 1 (1999). See also Kaplow & Shavell, *supra* note 73, at 961 n. 637. On the social costs of corruption, see ROBERT KLITGAARD, *CONTROLLING CORRUPTION* (1988); Susan Rose-Ackerman, *CORRUPTION: A STUDY IN POLITICAL ECONOMY* (1978); Pranab Bardhan, *Corruption and Development: A Review of Issues*, 35 J. ECON. LITERATURE 1320 (1997); and Andrei Schleifer & Robert W. Vishny, *Corruption*, 108 Q.J. ECON. 599 (1993).

⁸⁰ We do not discuss what might happen if the defendant suffered a large stigma simply from being charged or indicted for certain kinds of wrongdoing. See *In re Fried*, 161 F.2d 453, 458 (2nd Cir. 1947) (noting that "[f]or a wrongful indictment is no laughing matter; often it works a grievous, irreparable injury to the person indicted. The stigma cannot be easily erased."). In such cases the potential for corruption and wealth extraction are greater because the prosecutor can gain or impose costs without actually having to win at a trial (i.e., without having to obtain a conviction). These instances are simply outside the scope of this paper.

We can see how this works in the context of the two major types of procedural protections we have emphasized: the reasonable doubt standard and the double jeopardy rule. Both reduce the prosecutor's power to selectively enforce and hence help to constrain rent-seeking costs. In this sense, they clearly fall within the analysis of this section because they simultaneously raise the cost to the prosecutor of implementing a selective policy and lower the value to the potential beneficiary of seeking such a policy. The reasonable-doubt rule accomplishes this task by directly reducing the probability of a guilty verdict and increasing the amount of evidence necessary for conviction. The double jeopardy rule aids in this task by preventing the prosecutor from bringing successive prosecutions against the same defendant, with the hope of eventually learning how to convict the defendant on weak evidence.

This theory suggests that the reasonable doubt and double jeopardy rules have complementary functions. No matter how low the probability of success is reduced as a result of the reasonable doubt rule, the prosecutor may still have an incentive to adopt a selective enforcement policy if he can bring successive actions against a defendant. In the extreme case in which the prosecutor can bring an infinite number of successive actions against the defendant, he is likely to eventually get a conviction, no matter how small the probability of conviction in the individual trial.⁸¹ The more worrisome case, however, is where the prosecutor learns from a previous mistake and uses the information from a "test trial" to boost the probability of conviction to a near certainty in the second trial.⁸² The double-jeopardy-like rule emerges in this framework as a rule designed to prevent enforcement agents from substituting toward a successive prosecution strategy in order to avoid the constraint imposed by the reasonable doubt rule.⁸³

B. *Penalty Restrictions or "Inefficient Punishments"*

Another way to constrain the costs associated with abuses of prosecutorial or governmental authority is to put restrictions on the size of penalties or the process by which they are levied. David Friedman has described the size and

⁸¹ Suppose the probability of conviction in one trial is p . If the prosecutor can bring an infinite number of successive actions, each with the same probability of conviction, the likelihood of eventual conviction is $p + (1-p)p + (1-p)^2p + \dots + (1-p)^Np$, which approaches 1 as N approaches infinity. For example, if p is 30% against an innocent individual then by the fourth trial the cumulative probability of conviction has risen to approximately 75%.

⁸² See *Developments in the Law—Corporate Crime: Regulating Corporate Behavior Through Criminal Sanctions*, 92 HARV. L. REV. 1341-49 (1979).

⁸³ This statement is not meant to exclude alternate means of perhaps restraining prosecutorial retrials. For example, if in the initial trial the reasonable doubt standard required 95% certainty of guilt before conviction then we could require a 96% certainty of guilt in retrial number 1. If there was a second retrial then we could require a 97% likelihood of guilt and so forth. Even then, given the argument in *supra* note 81 it is doubtful that much would be gained through this approach.

process restrictions as “inefficient” punishments.⁸⁴ Friedman distinguishes inefficient punishments, like prison, from efficient punishments, like the death penalty administered quickly or a large monetary penalty equal to the defendant’s wealth.⁸⁵ The argument is that efficient punishments do not impose large direct costs on the state and hence prosecutors may have an incentive to use them to extract wealth from defendants.⁸⁶ However, inefficient punishments impose substantial costs on the state and hence reduce the incentive of prosecutors and the state to use the criminal process to extract wealth.⁸⁷

At least two types of penalty restrictions are relevant to this analysis: the prohibition of retroactive punishments and of cruel and unusual punishments.⁸⁸ The latter restriction fits comfortably with the analysis here as well as Friedman’s. As Friedman notes, the state could easily adopt a low-cost system of punishment; defendants could be executed, enslaved, or put into laboratories for scientific experimentation and the harvesting of organs and tissue.⁸⁹ Instead, we observe a system in which the state forgoes the opportunity to extract all of the defendant’s wealth, and prison terms force the state to forgo the full value of the convict’s labor. The constitutional prohibition on cruel and unusual punishments is in part responsible for this choice, although the choice seems to have been made in some countries where there is no such prohibition.⁹⁰

The historical evidence is consistent with this view of the purpose of the cruel and unusual punishment clause.⁹¹ The first restrictions on excessive penalties in English law appeared in the Magna Carta, in chapters regulating discretionary fines.⁹² The discretionary fines, or “ameracements,”⁹³ obviously had the potential to be used as a source of revenue for the state, and as a source of private income in a period in which much of criminal prosecution was undertaken by private parties.⁹⁴ It was their abuse that led the authors of the

⁸⁴ See Friedman, *supra* note 7, at 259.

⁸⁵ See *id.*, at 260 – 61.

⁸⁶ See *id.*, at 261.

⁸⁷ See *id.*, at 263 – 64.

⁸⁸ See U.S. CONST. ART I, § 9, CL. 3 (stating “No Bill of Attainder or ex post facto Law shall be passed.”) A similar prohibition applies to the states: “No State shall . . . pass any Bill of Attainder . . .” *Id.*, ART. I, § 10, CL. 1; U.S. CONST. AMEND. VIII (stating “Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.”).

⁸⁹ See Friedman, *supra* note 7.

⁹⁰ For example, Malaysia, Morocco, Senegal, and the Ivory Coast do not have a rule prohibiting cruel and unusual punishment. Yet, the state in those countries has apparently not gone to the extreme of trying to profit from punishing the guilty. For information on Transparency International’s Corruption Perceptions Index, see <<http://www.worldbank.org/html/prddr/trans/>>.

⁹¹ See Anthony F. Grannuci, “Nor Cruel and Unusual Punishments Inflicted:” *The Original Meaning*, 57 CAL. L. REV. 839, 844 – 848 (1969).

⁹² See *id.*, at 845.

⁹³ *Id.*

⁹⁴ On privately-initiated enforcement, see, e.g., Friedman, *supra* note 7, at 264.

Magna Carta to devote three chapters to their control.⁹⁵ These provisions later evolved into the more modern prohibitions of excessive and disproportionate punishments.⁹⁶

Consistent with Friedman's argument, our analysis suggests that penalty restrictions increase the cost of punishment to the state, dampening incentives for wealth extraction, and at the same time reduce the amount a potential defendant would be willing to pay in order to avoid being charged with a crime. The prohibition on cruel and unusual punishments reduces the potential for the state or prosecutor to punish innocent individuals in order to profit from their punishment. By raising the cost of punishment, the prohibition enhances the likelihood that the state will punish only the guilty.⁹⁷ It also dampens incentives individuals have to become prosecutors in order to enrich themselves.

The prohibition of retroactive punishments – a restriction on the penalty imposition process – constrains rent-seeking at the legislative level. In the absence of such a restriction, interest groups could use the criminal process to confiscate the wealth of other groups or particular individuals. A predatory enforcement regime could retroactively impose a criminal penalty on the activity of a particular group and use the new law as leverage to expropriate their wealth. The ex post facto and bill of attainder clauses in the Constitution both apply to this type of activity. The ex post facto clause applies specifically to legislative attempts to punish retroactively.⁹⁸ The bill of attainder clause applies to legislative attempts to punish particular individuals without a trial.⁹⁹

The distinction between penalty restrictions and procedural protections suggests that these types of rules, both designed to dampen rent-seeking, serve as complements: procedural protections work better in the short run while

⁹⁵ See Grannuci, *supra* note 91, at 845 – 846.

⁹⁶ See *id.*

⁹⁷ The reason for why the incentive to go after the truly guilty increases because of inefficient punishments requires some explanation. An inefficient punishment increases the costs of convicting anyone (guilty or innocent), and hence increases the costs of convicting both types of defendants. This should result in the state (and the prosecutor) shifting more resources towards targeting the truly guilty rather than bringing cases against the innocent as the expected payoffs from convicting the truly guilty are probably higher than the expected payoffs from convicting the innocent as a general matter.

⁹⁸ See Harold J. Krent, *The Puzzling Boundary Between Criminal and Civil Retroactive Lawmaking*, 84 GEO. L.J. 2143 (1996); Wayne A. Logan, *The Ex Post Facto Clause and The Jurisprudence of Punishment*, 35 AM. CRIM. L. REV. 1261 (1998).

⁹⁹ See Jane Welsh, *The Bill of Attainder Clause: An Unqualified Guarantee of Process*, 50 BROOK. L. REV. 77 (1983); Thomas B. Griffin, *Beyond Process: A Substantive Rationale for the Bill of Attainder Clause*, 70 VA. L. REV. 475 (1984). We will focus our discussion *infra* on the ex post facto clause for the sake of brevity. Note that both clauses induce a reactive type of rent seeking, as they seem to invite defendants to challenge virtually every effort to punish on the ground that it is either a disguised bill of attainder or retroactive penalty.

penalty restrictions work better in the long run.¹⁰⁰ The procedural protections described earlier – the reasonable-doubt rule and the double-jeopardy rule – operate in the short run to remove incentives for the prosecutor to selectively enforce the law. However, if the background institutional structure is one that allows the state to profit from the punishment of individuals, we should worry about how long the prosecutor will be able to stay out of the predatory enforcement game. Just as the potential for profit induces entry of new businesses in the private sector, the potential for profit in enforcement should induce entry of a similar sort in the public sector. Creative prosecutors would find ways to modify the procedural rules, plea bargain around them, or to lobby the legislature until the desired changes were enacted.¹⁰¹ Penalty restrictions constrain this long-term erosion process by reducing the potential for profit in criminal law enforcement.¹⁰²

C. *Some Implications for the Jury*

Although we have focused on rent-seeking in the enforcement process rather than in the legislative process, the jury serves as an important constraint against both types of rent-seeking.¹⁰³ A prosecutor who brings politically motivated charges against members of politically weak groups faces the risk, under the jury system, of being unable to gain a unanimous verdict from a jury consisting of some members from the weak group. Indeed, the theory of this paper suggests an important rationale for the requirement of unanimity among

¹⁰⁰ A key difference between procedural protections and penalty restrictions has to do with how much of the cost of each method the prosecutor bears. The costs of penalty restrictions or inefficient punishments are borne largely by the state and society, but not particularly by the prosecutor. They impose only an indirect cost on the prosecution. In contrast, pro-defendant procedures impose direct costs on prosecutors because they make obtaining convictions more difficult. Given that high conviction rates are important to many prosecutors for career reasons, we would expect prosecutors to respond more to these direct measures than to indirect concerns about how much a punishment is costing the state. This is especially so when many prosecutors are paid by local authorities and prisons are funded by state authorities. See Stuntz, *supra* note 44, at 520 – 33 (discussing the incentives of the different actors in law enforcement). In some respects one can view the procedural protections as alignment measures – trying to align prosecutors’ interests with social welfare (or the benevolent state), whereas the inefficient punishments might be seen as constraints on a non-benevolent state.

¹⁰¹ See William J. Stuntz, *The Uneasy Relationship Between Criminal Procedure and Criminal Justice*, 107 YALE L.J. 1, 26 (1997) (arguing that despite pro-defendant protections, prosecutors have held conviction rates constant and lowered the average cost of prosecution by prosecuting “winning” cases). See Stuntz, *supra* note 44, at 509 – 515 (arguing that because criminal codes are so broad it gives a prosecutor the ability to selectively enforce).

¹⁰² The procedural constraints we have discussed might reduce the incentive to lobby law enforcers, but might lead parties to more lobbying on the legislative side. Penalty restrictions, to the extent they are part of constitutional law, reduce incentives for legislative lobbying. In addition, any set of restraints that forces rent-seeking groups to use more costly methods of extracting wealth should have the desired effect of dampening rent-seeking incentives. At a minimum, the restraints considered here have that effect.

¹⁰³ See *Duncan v. Louisiana*, 391 U.S. 145, 156 (1968) (“[p]roviding an accused with the right to be tried by a jury of his peers gave him an inestimable safeguard against the corrupt or overzealous prosecutor and against the compliant, biased, or eccentric judge.”)

jurors in criminal trials.¹⁰⁴ The need to obtain a unanimous verdict makes it more difficult for the prosecutor to selectively target politically marginal groups or individuals in the law enforcement process. The need to obtain a unanimous verdict from the jury also gives the jury the power to nullify statutes designed to expropriate wealth from politically marginal groups.

Our theory provides some insight into the original function of challenges to the jury's composition, including the controversial problem of peremptory challenges. Blackstone said that challenges could be put to either the whole array of jurors or to individual jurors, and that

challenges to the array are at once an exception to the whole panel, ... and they may be made upon account of partiality or some default in the sheriff, or his under-officer who arrayed the panel. . . Also, though there be no personal objection against the sheriff, yet if he arrays the panel at the nomination, or under the direction of either party, this is good cause of challenge to the array.¹⁰⁵

In other words, the fundamental common law rationale for permitting challenges to the whole jury is the suspicion, no doubt grounded on evidence, that the sheriff chose the jurors in order to maximize his chances of obtaining a conviction. Challenges to the *whole array* were apparently permitted to prevent the sheriff from implementing a selective enforcement policy.

Challenges to individual jurors could be based on cause, or could be peremptory, in the sense of not being based on any of the accepted grounds.¹⁰⁶ Peremptory challenges were granted only to the defendant.¹⁰⁷ Although peremptory challenges have come under attack more recently as a form of invidious discrimination,¹⁰⁸ the original purpose is somewhat easier to see in the context of a rent-seeking model. One could view the peremptory, in this analysis, as giving the defendant a zone of unquestioned authority in the choice of jurors, so long as he did not use it to an excessive degree. If a wily predatory sheriff had managed to choose conviction-prone jurors in a way that would be

¹⁰⁴ An alternative explanation for the unanimity requirement – that it reduces false convictions may not be terribly convincing. See Timothy Feddersen & Wolfgang Pesendorfer, *Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts* (Draft, 1997)(on file with authors) (arguing that under plausible assumptions the unanimity requirement may result in an increase in false convictions relative to a supermajority vote requirement).

¹⁰⁵ See BLACKSTONE, COMMENTARIES, *supra* note 12, at 359.

¹⁰⁶ See *id.*, at 361 – 63.

¹⁰⁷ See *id.*, at 362.

¹⁰⁸ See *J.E.B. v. Alabama ex rel. T.B.*, 511 U.S. 127, 140 (1994) (holding that the Equal Protection Clause prohibits gender discrimination in the use of peremptory challenges); *Batson v. Kentucky*, 476 U.S. 79, 89 (1986) (holding the prosecutor's peremptory challenges based solely on race were unconstitutional under the Equal Protection Clause.)

difficult to challenge on the accepted grounds, the defendant could always fall back on his peremptory challenges. To the extent that this obstruction stood in the way of any effort to selectively enforce the law, the sheriff would have a much smaller incentive to try to control the composition of the jury.¹⁰⁹

VII. APPLICATIONS OF POSITIVE THEORY

We have been concerned so far with explaining broad institutional features. In this Part we extend the argument by taking a look at the case law associated with pro-defendant protections. Since criminal procedure is a vast area, we can provide only a sketch here. We claim that this paper's framework

¹⁰⁹ In addition, *when* the right to a jury trial is available may be consistent with a rent-seeking approach as well. Jury trials are available as of right for most criminal cases except those that carry trivial or fairly small penalties. *See* *Duncan v. Louisiana*, 391 U.S. 145, 160 (1968) (holding that "there is a category of petty crimes or offenses which is not subject to the 6th Amendment jury trial provisions"). This is consistent with rent-seeking because trials that carry very small or trivial penalties may not be particularly attractive means with which to extract wealth for prosecutors. *See* *Friedman*, *supra* note 7, at 268. For such small sanctions the costs of the jury trial are probably not justified by any reduction in rent-seeking (which is probably small in this context).

Further, alternative explanations for the right to a jury trial do not appear to provide as complete a picture as they might if they considered concerns with rent-seeking. One potential explanation for the right to a jury trial is that society values the expression of the popular will as reflected in a jury decision. *See* *Spaziano v. Florida*, 468 U.S. 447, (1984) (Stevens, J. dissenting in part) (arguing that the right to have an authentic representation of the community's views on the determination that must precede a deprivation of liberty supports the constitutional entitlement to a trial by jury). If this were the only purpose behind the right to a jury trial we would expect all jury decisions to be unreviewable. However, this is not the case because the law permits jury convictions to be appealed but not jury acquittals. *See* *Kepner v. United States*, 195 U.S. 100 (1903). *See also* *Steinglass*, *supra* note 20, at 354 – 55.

Another potential explanation for the right to a jury trial is that it either reduces erroneous decisions relative to bench trials or is less likely to falsely convict relative to bench trials. *See* HARRY KALVEN, JR. & HANS ZEISEL, *THE AMERICAN JURY* 6 – 7 (1971). This argument is not particularly convincing because it is a little difficult to believe that jury trials are likely to be more accurate (i.e., less error prone) than bench trials. One doubts there is any empirical evidence to support this result and our legal system also seems to suggest that jury trials may be more prone to errors than bench trials. *See* VALERIE P. HANS & NEIL VIDMAR, *JUDGING THE JURY* 126 – 27 (1986) (stating that "[w]e are thus led to the conclusion that jurors may not always be able to follow the law as it is intended to be."). Much of the law of evidence seems to try to protect the jury from misperceptions and bias, whereas we seem less concerned with these matters for bench trials. *See, e.g.,* FED. R. EVID. 103(c), 403. This suggests bench trials are probably more accurate than jury trials overall or at least not less accurate as a general matter.

It may be, however, that we believe juries are less likely to falsely convict compared to bench trials. It is not entirely clear why we would believe this if we think bench trials are generally more accurate. Perhaps the argument is that judges are more biased against defendants than a jury of the defendant's peers as judges tend to be in quite a different socio-economic strata as compared to most defendants. *See* HANS & VIDMAR, *supra* (noting that "[f]or criminal trials the pattern disagreement shows that the jury was usually more lenient toward the defendant than was the judge."); Pnina Lahav, *The Chicago Conspiracy Trial: Character and Judicial Discretion*, 71 U. COLO. L. REV. 1327, 1340 (2000). There may be some empirical evidence supporting the differing rates of false convictions (or maybe tied to it). *See generally* KALVEN & ZEISEL, *supra*. Perhaps this is true, but when phrased this way it appears more consistent with concerns about rent-seeking. This is because this suggests that judges as a group may discriminate (i.e., use the criminal process) against criminal defendants as a group. If so, then this justification squares well with a rent-seeking approach. Also even if jury trials result in fewer false convictions and more false acquittals than bench trials the issue is raised about whether the asymmetry is desirable for the same reason that the reasonable doubt standard may be too severe on traditional error cost grounds.

provides a good positive theory of substantial parts of criminal procedure doctrine.

A. *Double Jeopardy*

As a general matter, Double Jeopardy reduces the prosecutor's power to selectively enforce or abuse his discretion in a manner complementary to the reasonable doubt standard. Double Jeopardy complements the reasonable doubt rule by preventing the prosecutor from bringing successive prosecutions against the same defendant with the hope of eventually learning how to convict the defendant on weak evidence.¹¹⁰ In brief, our claim is that *the essential purpose of Double Jeopardy doctrine is to prevent prosecutors from substituting toward successive prosecutions in order to avoid fundamental single-trial procedural constraints such as the reasonable-doubt rule.* This proposition explains puzzles in Double Jeopardy doctrine as well as its overall structure.

The "puzzles" in Double Jeopardy doctrine have been set out with remarkable clarity by Akhil Amar.¹¹¹ One of the key doctrinal tensions he notes is between the decisions in *Blockburger v. United States*¹¹² and *Diaz v. United States*.¹¹³ *Blockburger* established the principle that, for Double Jeopardy purposes, a greater offense is treated as the same as each of its lesser-included offenses. Thus, under the *Blockburger* test it would be a violation of the Double Jeopardy rule to retry an individual for murder after that individual is acquitted on a charge of attempted murder (arising from the same set of facts).¹¹⁴ However, in *Diaz* the Court permitted an individual to be retried for murder, after having been convicted of attempted murder, in a case in which the victim died after the first trial from injuries received in the initial attack.¹¹⁵ Under a strict application of the *Blockburger* test, *Diaz* would have to be considered a mistake.

The tension between *Blockburger* and *Diaz* (and between *Blockburger* and several other decisions) appears to be a puzzle under the approach urged by Amar, which would apply the rule of *Blockburger* strictly to the facts of *Diaz*. However, under the anti-substitution thesis implied by our framework, there is no tension between *Blockburger* and *Diaz*. *Blockburger* prohibits a particular

¹¹⁰ See *Benton v. Maryland*, 395 U.S. 784, 796 (1969) (noting that the rationale for double jeopardy includes the policy against allowing multiple prosecutions that will enhance the possibility that the innocent may be found guilty).

¹¹¹ See Akhil Reed Amar, *Double Jeopardy Law Made Simple*, 106 YALE L. J. 1807 (1997).

¹¹² 284 U.S. 299 (1932).

¹¹³ 223 U.S. 442 (1912). As Amar notes, the principle of *Diaz* was affirmed later by the Supreme Court in *Garret v. United States*, 471 U.S. 773 (1985). See Amar, *supra* note 111, at 1813.

¹¹⁴ See Amar, *supra* note 111, at 1813.

¹¹⁵ The state was prepared to offset against the murder sentence time served under the attempted murder conviction. See *id.*

substitution strategy: the use of a later trial on either a greater or lesser-included offense in order to take two shots at convicting a defendant on one particular set of facts. The “*Diaz* exception” (if one wishes to call it that) applies to the case in which there is clearly no evidence that the prosecutor has adopted such a strategy. *Blockburger* and *Diaz* are easily reconciled under the anti-substitution principle.

Consider another puzzle under *Blockburger*. Suppose one charge consists of circumstances *A*, *B*, and *C*, and another of the circumstances *A*, *B*, and *E*. To use Amar’s example, suppose one charge is robbery and the other is murder, arising from the same set of events, and element *B* requires proof of the defendant’s presence at the scene of the crime.¹¹⁶ The defendant is tried under the *ABC* charge first, and the government loses because it cannot prove element *B*. Can the government turn around and bring the *ABE* charge? *Blockburger* would seem to imply that the answer is yes, because the charges do not involve the same elements. However, the Supreme Court held in *Ashe v. Swenson* that, under the Double Jeopardy Clause, the government was collaterally estopped from attempting to prove element *B* in the second prosecution.¹¹⁷

The tension between *Blockburger* and *Ashe* disappears under this paper’s anti-substitution principle. If prosecutors were free to split up their charges into separate yet overlapping bundles – {*A*, *B*, *C*}; {*A*, *B*, *E*}; {*B*, *C*, *D*} – without having to fear that they would be barred by the Double Jeopardy rule, then this would clearly be a desirable way of skirting the single-trial constraints imposed by the reasonable-doubt rule. It follows that Double Jeopardy doctrine should be understood to prevent this kind of prosecution stratagem. If the *Blockburger* doctrine is read to incorporate the anti-substitution rule proposed here, then there is no apparent conflict between *Blockburger* and the collateral estoppel rule of *Ashe*.

Our claim that the overall structure of Double Jeopardy doctrine is consistent with the model of this paper is supported by three other aspects of the doctrine: the asymmetry of appeal rights, the treatment of mistrials, and the application of Double Jeopardy to civil suits by government agencies.

Consider, first, the treatment of appeal rights. In *Kepner v. U.S.*, the Supreme Court held that appeal rights are asymmetric,¹¹⁸ in the sense that the defense generally can appeal any conviction, but the prosecution’s right to

¹¹⁶ See Amar, *supra* note 111, at 1827.

¹¹⁷ 397 U.S. 436 (1970); for discussion see Amar, *supra* note 111, at 1827 – 28.

¹¹⁸ See *Kepner v. United States*, 195 U.S. 100, 105 (1904).

appeal acquittals is severely limited.¹¹⁹ On its face, this rule seems like it might reduce the number of false convictions and increase the number of false acquittals relative to symmetric appeal rights because it denies the prosecution the ability to correct false acquittals at the trial level, while also denying the prosecution the ability to turn correct acquittals into false convictions through the appeals and retrial process.¹²⁰ However, on closer inspection, this is not necessarily the case. Not only are the above mentioned effects possible, but so are countervailing effects. For example, by giving the prosecution only one shot at obtaining a conviction we provide the prosecution with an incentive to increase spending in the initial trial, which could increase the number of false convictions in the initial trial.¹²¹ The net effect is most likely ambiguous and small.¹²² Thus, the effects of asymmetric appeal rights on false convictions are ambiguous as a theoretical matter.¹²³ Similar arguments suggest that false acquittal effects are likely to be small and ambiguous.¹²⁴

In light of this, it seems doubtful that traditional error-cost arguments provide a solid basis for asymmetric appeal rights. However, if we incorporate concerns over rent seeking then a stronger rationale for asymmetric appeal rights emerges.¹²⁵ The asymmetric appeal rights rule of *Kepner* has the effect of making the jury's initial determination of acquittal final. By denying prosecutors the option to have a jury's acquittal determination reviewed by an appellate court, the *Kepner* asymmetry rule enhances the power of the jury relative to that of the prosecutor.¹²⁶ Given the unanimity requirement and the jury's composition after the defendant's challenges, the *Kepner* rule increases the difficulty facing any prosecutor who mounts a selective enforcement campaign.

Justice Holmes's dissent in *Kepner* is instructive largely because it focused on the wrong theory for asymmetric appeal rights. Holmes argued that the majority's decision in *Kepner* made no sense if understood as a rule preventing

¹¹⁹ See *Sanabria v. United States*, 437 U.S. 54, 64 (1978) (holding that even if legal rulings on the exclusion of evidence leading to acquittals were erroneous the prosecution could not appeal); *Fong Foo v. United States*, 369 U.S. 141, 143 (1962) (holding that the prosecution could not appeal an acquittal where the judge, who lacked the authority to do so, directed a verdict of acquittal before the prosecutor has rested his case); *Carroll v. United States*, 354 U.S. 394, 400 (1956); *United States v. Ball*, 163 U.S. 662 (1896) (holding that a defendant may not be prosecuted more than once for an offense).

¹²⁰ See *Khanna*, *supra* note 20, at 360 – 63.

¹²¹ See *id.*, at 374 – 83.

¹²² See *id.*

¹²³ Note that given that there are few acquittals (both in symmetric and asymmetric appeal rights jurisdictions) and that in those countries where the prosecutor is allowed to appeal they do so infrequently, one suspects that the false convictions reducing or increasing effects are likely to be small because few resources are being saved by prohibiting the few prosecutorial appeals that might arise under symmetric appeal rights. See *id.*, at 33 n. 120, 39 – 40.

¹²⁴ See *id.*, at 47 – 48.

¹²⁵ See *id.*, at 68 – 70.

¹²⁶ Cf. *Westen & Drubel*, *supra* note 20, at 122 – 55 (discussing the role of jury nullification in the context of asymmetric appeal rights).

retrials, because some retrials would occur following a successful *defense* appeal of a conviction.¹²⁷ However, the desirability of preventing retrials is an incomplete and inadequate rationale for the decision in *Kepner*. The theory we advance provides a superior rationale: the Double Jeopardy rule is not designed for the sole purpose of controlling or preventing retrials; its purpose is to prevent prosecutors from successfully implementing a selective enforcement policy. This is a greater concern when prosecutors appeal as compared to when defendants appeal. A regime in which prosecutors could appeal acquittals *ad infinitum* would be much more vulnerable to selective enforcement pressures than one where they could not.¹²⁸

The treatment of mistrials is another area of Double Jeopardy doctrine that shows a concern for constraining prosecutorial abuse. The kind of abuse we are concerned with here is that the prosecution may think, at some point in the initial trial, that a conviction is not likely and may then try to have a mistrial declared by the court to try to get another shot at the defendant.¹²⁹ If we permitted the prosecution to do this and bring another trial then the prosecution would have a tremendous incentive to have mistrials declared whenever it thought it might not win the initial trial. This potential would, in turn, increase the incentive to engage in selective enforcement, which induces rent seeking.¹³⁰

The law appears to reflect these concerns in the way it addresses whether another trial will be permitted following a mistrial. One could characterize the law's approach to this problem as one that depends greatly on the defense's attitude towards a mistrial. Thus, if the defense seeks or does not oppose a motion for a mistrial then the prosecution will normally be permitted to bring another suit.¹³¹ This is consistent with our substitution thesis because if the *defense* is seeking a mistrial, the prosecution is not likely to be using the mistrial process to seek another trial to go after the defendant.¹³² An exception to this is observed where the defense seeks a mistrial based on something the prosecution did that appears deliberately calculated by the prosecution to induce the defense to seek a mistrial.¹³³ This is also consistent with the rent-seeking framework

¹²⁷ See *Kepner v. United States*, 195 U.S. 100, 134-35 (Holmes, J., dissenting) (1903).

¹²⁸ See *Lockhart v. Nelsen*, 488 U.S. 33, 42 (1988); *Ashe v. Swenson*, 397 U.S. 82, 91 (1978); *United States v. Scott*, 437 U.S. 82, 91 (1978). This is similar to the argument in *supra* note 81.

¹²⁹ See *Ashe v. Swenson*, 397 U.S. 436, 447 (1970); Stephen Schulhofer, *Jeopardy and Mistrials*, 125 U. PA. L. REV. 449, 468-69 (1977).

¹³⁰ This concern is very similar to that discussed in the context of asymmetric appeal rights. At the same time if we never permitted the prosecution to bring another trial after a mistrial we would give the defendant a great deal of strategic power to inject matters that might lead to a mistrial when a conviction appears likely. See generally Vikramaditya S. Khanna, *The Mystery of Mistrials* (Draft 2002).

¹³¹ See *United States v. DiFrancesco*, 449 U.S. 117, 130 (1980); see also *Oregon v. Kennedy*, 456 U.S. 667, 676 (1982).

¹³² See Schulhofer, *supra* note 129, at 468-69.

¹³³ See *Arizona v. Washington*, 434 U.S. 497, 508 (1978); *United States v. Dinitz*, 424 U.S. 600, 611 (1976) (stating that "[the Double Jeopardy law] bar retrials where 'bad-faith' conduct by judge or prosecutor

because in such cases there is a clear risk that the prosecutor could have induced the defense's motion for a mistrial in order to take advantage of information gleaned from the initial trial.¹³⁴

On the other hand, when the defense opposes a mistrial motion the courts have adopted a more cautious stance to permitting another trial – the prosecution must prove a “manifest necessity” for the next trial.¹³⁵ In addition, the factors that go to showing whether “manifest necessity” is present appear to be designed to ascertain whether the prosecutor was trying to abuse the criminal process – e.g., by getting a mistrial in order to avoid a loss in the initial trial.¹³⁶ For example, a hung jury leading to a mistrial does not present the same specter of potential abuse as does the injection of prejudicial error by the prosecutor to obtain a mistrial.¹³⁷ In the former case the prosecution is often granted another trial while the latter case will normally not result in another trial.¹³⁸ Further, when the reason for the mistrial was a move by the defense, without prosecutorial provocation, then the scope for prosecutorial abuse is also low and another trial is usually granted.¹³⁹ These factors are all consistent with an approach that seeks to constrain the prosecutor's ability to substitute successive prosecutions in order to avoid the effect of pro-defendant procedural protections.

140

threatens the harassment of an accused by successive prosecutions [or] a more favorable opportunity to convict the defendant; [W]here a defendant's mistrial motion is necessitated by judicial or prosecutorial impropriety designed to avoid an acquittal, reprosecution might well be barred”); *United States v. Jorn*, 400 U.S. 470, 485 (1971) (stating that “[t]hus, where circumstances develop not attributable to prosecutorial or judicial overreaching, a motion by the defendant for mistrial is ordinarily assumed to remove any barrier to reprosecution, even if the defendant's motion is necessitated by prosecutorial or judicial error”).

¹³⁴ See *Commonwealth v. Starks*, 416 A.2d 498, 500 (1980) (arguing that prosecutorial overreaching “... signals the breakdown of the integrity of the judicial proceeding, and represents the type of prosecutorial tactic which the double jeopardy clause was designed to protect against.”). This exception to the bar on the application of double jeopardy where a defendant seeks a mistrial is a very narrow one. See *Green v. United States*, 451 U.S. 929, 931 (1981) (Marshall, J., dissenting) (stating that “I suspect that a defendant seeking to prevent a retrial will seldom be able to prove the Government's actual motivation.”). Few courts have found intentional prosecutorial inducement. See *Commonwealth v. Warfield*, 227 A.2d 177 (Pa. 1967) (defendant was indicted for murder and voluntary manslaughter and the trial judge suppressed the defendant's confession, yet the District Attorney revealed in his opening statement that the defendant had made a confession to the police. After the defendant moved for a mistrial the District Attorney admitted that he sought to induce the defendant to seek a mistrial).

¹³⁵ See *Arizona v. Washington*, 434 U.S. 497, 505 (1978); WAYNE R. LAFAVE, ET AL., *CRIMINAL PROCEDURE* 1176-80 (3d ed. 2000).

¹³⁶ See Schulhofer, *supra* note 129, at 468-69.

¹³⁷ See Steinglass, *supra* note 20, at 361.

¹³⁸ *Id.* at 487.

¹³⁹ See DRESSLER, *supra* note 3, at 607.

¹⁴⁰ See Schulhofer, *supra* note 129, at 454 (noting that “...reprosecution may be barred even though no adjudication results from the first proceeding. The doctrine thus provides more meaningful protection against the danger of governmental harassment and the burden of repeated trials...”). Schulhofer also makes the point that “A number of courts have barred retrial even when mistrial was triggered by absence of the defendant, impermissible cross-examination, or persistently objectionable behavior by defense counsel.” See Schulhofer, *supra* note 129, at 484.

Finally, consider the on-going debate about whether Double Jeopardy protections should apply to nominally “civil” suits brought by government agencies that otherwise appear “punitive”.¹⁴¹ The courts have generally not permitted Double Jeopardy protections to apply to nominally “civil” suits, however, if it can be shown that the “civil” suit is in reality a form of “punishment” then Double Jeopardy protections may apply.¹⁴² The courts seem to rely on the following factors to determine if a “civil” suit is in reality “punishment”:

- (1) whether the sanction involves an affirmative disability or restraint;
- (2) whether it has historically been regarded as a punishment;
- (3) whether it comes into play only on a finding of scienter;
- (4) whether its operation will promote the traditional aims of punishment-retribution and deterrence;
- (5) whether the behavior to which it applies is already a crime;
- (6) whether an alternative purpose to which it may rationally be connected is assignable to it; and
- (7) whether it appears excessive in relation to the alternative purpose assigned to it.¹⁴³

Consistent with our hypothesis, these seven factors appear to be designed to prevent law enforcement agents from *substituting* civil for criminal prosecution as a means of avoiding the Double Jeopardy rule.¹⁴⁴ Indeed, rather than announcing a seven factor test, courts could clarify the doctrine in this area by saying that its purpose is to prevent the prosecution from gaining through the adoption of such a substitution strategy.

B. *Ex Post Facto Clause*

The ex post facto clause bars retroactive application of certain changes in

¹⁴¹ See generally Mary M. Cheh, *Constitutional limits on Using Civil Remedies To Achieve Criminal Law Objectives: Understanding and Transcending the Criminal-Civil Law Distinction*, 42 HASTINGS L.J. 1325 (1997). See Kenneth Mann, *Punitive Civil Sanctions: The Middleground Between Criminal and Civil Law*, 101 YALE L.J. 1795, 1797, 1869-73 (1992).

¹⁴² See *United States v. Halper*, 490 U.S. 435, 449 (1989); *U.S. v. Ward*, 448 U.S. 242, 248-249 (1980).

¹⁴³ *Hudson v. United States*, 522 U.S. 93, 99 (1997); *Kennedy v. Mendoza-Martinez*, 372 U.S. 144, 168 - 169 (1963).

¹⁴⁴ Some of the other factors (e.g., scienter, historically regarded as punishment) may work as proxies for the kind of stigma associated with the particular civil wrong. See Cheh, *supra* note 141, at 1352-54. If the wrong appears like what most people consider criminal then the stigma may be quite high thereby giving the prosecutor greater ability to extract wealth. The fourth and fifth factors suggest that one important purpose for the doctrine in this area is to discourage enforcement agents from substituting civil enforcement for criminal enforcement in order to evade the procedural protections that come along with criminal enforcement. See *id.* at 1345, 1354-57, 1394.

the criminal law.¹⁴⁵ The standard justifications are that it provides *notice* to defendants of the conduct that is illegal and the sanction for it, as well as *constraining the government* from passing arbitrary or vindictive legislation against a particular defendant.¹⁴⁶ The prohibition is only concerned with matters that amount to “punishment”¹⁴⁷ and applies more frequently in the context of legislative decisions as compared to judicial decisions.¹⁴⁸

Given that a concern for potential government oppression has always been one of the justifications for the *ex post facto* clause, the key contribution of the rent-seeking framework is its identification of a particular justification as central to efforts to understand case outcomes. The two standard justifications, notice and prevention of abuse, provide alternate approaches to explaining the case law. Under the notice approach, just about every change in the criminal punishment process, whether in the law or in evidentiary requirements, could violate the prohibition against *ex post facto* rules if applied retroactively. *Under the “abuse” or rent-seeking approach, only those changes in the punishment process that create a potential for abuse, by increasing the risk of targeted enforcement, should be deemed violations of the ex post facto clause.* The case law, as we shall see, is more consistent with the latter view.

Traditionally the *ex post facto* rule applies in the four contexts set out in *Calder v. Bull*.¹⁴⁹ These are: (1) when the legislature creates a new criminal law it may not be applied retroactively to behavior that was not criminal at the time the behavior occurred; (2) when the legislature removes elements from the definition of a crime or otherwise increases the severity of a crime (making it a more serious crime than when it occurred) these changes may not be applied retroactively; (3) when the legislature increases the punishment for a particular crime then this change may not be applied retroactively; and (4) when the legislature changes the rules of evidence or changes the requirements for testimony relative to what they were when the act occurred then the changes may not be applied retroactively.¹⁵⁰ Contexts (2) and (3) are often considered as one.¹⁵¹ Contexts (1)

¹⁴⁵ See JOHN E. NOWAK & RONALD D. ROTUNDA, CONSTITUTIONAL LAW 428 (5th ed. 1995); *Lindsey v. Washington*, 301 U.S. 397 (1937).

¹⁴⁶ See WAYNE R. LAFAYE & AUSTIN W. SCOTT, JR. CRIMINAL LAW 97 - 100 (2d. ed).

¹⁴⁷ Sometimes “civil” sanctions may amount to “punishment” for purposes of *ex post facto* inquiry. See *Landgraf v. U.S.I. Film Products*, 511 U.S. 244 (1994) (finding that the punitive damage awards of the 1991 Civil Rights Act were similar enough to criminal sanctions to apply the *ex post facto* clause). However, this is not a frequent occurrence so that normally the legislative label is determinative (i.e., if it is labeled “civil” by the legislature then it will most likely not amount to “punishment”). See *Kansas v. Hendricks*, 521 U.S. 346 (1997).

¹⁴⁸ This is because the prohibition appears in Art. I of the U.S. Constitution, which deals with legislative power and not in Art. III, which deals with judicial power. See LAFAYE & SCOTT, *supra* note 146, at 97 - 100.

¹⁴⁹ See *Calder v. Bull*, 3 U.S. 386 (1778).

¹⁵⁰ See *id.* at 390.

through (4) represent fairly common methods in which prosecutors and the government could engage in selective or targeted enforcement or legislation.¹⁵²

The modern case law presents more difficult issues. For example, what happens when the government passes a law that changes parole requirements or privileges? In *California Dept. of Corrections v. Morales*¹⁵³ the California State government changed the rule (contained in a statute) on reconsideration hearings for prisoners who had their first attempt at obtaining parole rejected.¹⁵⁴ The old rules provided for annual reconsideration for such inmates and the new rules provided for the parole board to defer hearings for up to 3 years if the board “finds that it is not reasonable to expect that parole would be granted... during the following two years and states the basis for its findings.”¹⁵⁵ The new rule in effect changed the statutory scheme and granted additional power and discretion to the parole board. The Court held that this rule, when applied retroactively, did not violate *ex post facto* because there was only a remote likelihood of obtaining parole in those cases where the new rule applied.¹⁵⁶

¹⁵¹ See LAFAVE & SCOTT, *supra* note 146, at 97 n.3; NOWAK & ROTUNDA, *supra* note 145, at 428. The *ex post facto* prohibition will also be violated in some other situations. See *Beazell v. Ohio*, 269 U.S. 167 (1925); *Kring v. Missouri*, 107 U.S. 221 (1882) (noting that “the new Constitution of Missouri does take away what, [by] the law of the State when the crime was committed, was a good defence to the charge of murder in the first degree.”); *Thompson v. Utah*, 170 U.S. 343 (1898) (arguing that “the provision in the constitution of Utah providing for the trial in courts of general jurisdiction of criminal cases, not capital, by a jury composed of eight persons, is *ex post facto* in its application to felonies committed before the Territory became a State, because, in respect of such crimes, the Constitution of the United States gave the accused, at the time of the commission of his offence, the right to be tried by a jury of twelve persons, and made it impossible to deprive him of his liberty except by the unanimous verdict of such a jury.”).

¹⁵² Context (1) is seldom brought into question in modern times, but it still represents the quintessential instance of the *ex post facto* prohibition – where the government and the prosecutor could target a particular defendant in response to some political, or other gain, they may receive. See LAFAVE & SCOTT, *supra* note 146, at 98. Contexts (2) and (3) still arise in some form in modern times. See *Stogner v. California*, 539 U.S. 607 (2003); *Dobbert v. Florida*, 432 U.S. 282 (1977); *Hernandez v. State*, 43 Ariz. 424 (1934); *Lindsey v. Washington*, 301 U.S. 397 (1937). They also represent instances, like Context (1), where prosecutors could engage in selective or targeted enforcement. Context (4) involves instances where changes in the standard of proof and rules of evidence are applied retroactively. See *Walker v. State*, 433 So. 2d 469 (Ala. 1983); *Plachy v. State*, 239 S.W. 979 (1922); *Thompson v. Missouri*, 171 U.S. 380 (1898); LAFAVE & SCOTT, *supra* note 146, at 97 – 101. One of the old English cases establishing this point involved a defendant that had convinced a co-conspirator in treason to flee the country. See 9 T. MACAULAY, HISTORY OF ENGLAND, 31, 171-173 (1899). The English Parliament responded by reducing the number of witnesses needed to convict for treason, which was held to violate the *ex post facto* prohibition. See *id.* at 270. This case fits into our approach as it involves an instance where the legislature changed the law in order to target a particular defendant.

In the older cases any change (whether to increase or decrease the punishment) might have violated *ex post facto* if applied retroactively. See *In Re Tyson*, 13 Colo. 482 (1889); *Commonwealth v. McDonough*, 95 Mass 581 (1866). However, now that has changed so that a decrease in punishment is unlikely to violate *ex post facto*. See *U.S. v. Stewart*, 1993 U.S. App. Lexis 17634, *10, 1993 WL 265147, **3 (10th Cir. 1993) (noting that “the framers of the *Ex Post Facto* Clause intended it to preclude only increased punishment for preexisting criminal conduct”). See also *Miller v. Florida*, 428 U.S. 423, 430 (1987).

¹⁵³ 514 U.S. 499 (1995).

¹⁵⁴ See *Morales*, *supra* note 153, at 501.

¹⁵⁵ *Morales*, 514 U.S., at 503.

¹⁵⁶ See *id.* at 513.

There was only a speculative possibility of really extending the prison term through this new rule.¹⁵⁷

The result in *Morales* is consistent with the rent-seeking model of procedural constraints. Given the tiny risk that the change in *Morales* could have resulted from or facilitated successful targeting, the rent-seeking concerns behind the ex post facto clause are not seriously triggered. Although there appeared to be an increase in discretion and in the expected sanction, it was well constrained because the board could only defer hearings for those prisoners that had no reasonable chance of parole. If a prisoner had little chance of parole anyway then the scope for the new rule to encourage selective enforcement was limited.¹⁵⁸

More recent case law distinguishes between rules that affect witness competency and rules that change the quantum of evidence needed to convict defendants. In *Carmell v. Texas*,¹⁵⁹ the Supreme Court held that a change in the law regarding when uncorroborated testimony of sexual assault complainants would be admitted violates the ex post facto clause if applied retroactively.¹⁶⁰ Prior to the change in the law, Texas required that the testimony of sexual assault complainants be corroborated unless the complainant was under the age of 14.¹⁶¹ After the defendant's alleged wrongdoing the law changed so that uncorroborated testimony of any complainant below 18 years of age would be acceptable.¹⁶² The prosecution attempted to use this law because during the relevant time the complainant was between the ages of 14 and 18.¹⁶³ The Court held that retroactively using the new law violated the ex post facto prohibition because it reduced the quantum of evidence necessary to convict the defendant.¹⁶⁴ The decision ran counter to *Hopt v. Territory of Utah*,¹⁶⁵ where the Court held that a change in the law that permitted the prosecution to present the evidence of witnesses convicted of felonies was not a violation of ex post facto even if applied retroactively.¹⁶⁶

¹⁵⁷ *Id.* at 509, 514. The Parole board was required to provide reasons for its decisions too. *Id.* at 511.

¹⁵⁸ The US Supreme Court has recently upheld the *Morales* decision in *Garner v. Jones*, 120 S.Ct. 1362 (2000). The facts in *Garner* are similar to *Morales* in that the parole board was able to delay hearings on parole as a result of a regulatory amendment (not a statutory change as in *Morales*). See *Garner, supra*, at 1366 – 69. The Court remanded and required a finding on whether there was a “significant risk” under *Morales*. See *id.*, at 1370.

¹⁵⁹ See *Carmell v. Texas*, 120 S. Ct. 1620 (2000).

¹⁶⁰ See *id.* at 1643.

¹⁶¹ *Id.* at 1624.

¹⁶² *Id.* at 1625.

¹⁶³ See *Carmell, supra* note 159, at 1625.

¹⁶⁴ See *id.* at 1633.

¹⁶⁵ See *Hopt v. Territory of Utah*, 110 U.S. 574 (1884).

¹⁶⁶ See *id.* at 589.

Although *Carmell* and *Hopt* are quite difficult to reconcile under the notice theory of the ex post facto clause, they can be reconciled under the rent-seeking model. If our concern is with the ability of the prosecutor to use certain changes in the law of evidence to increase the chance of abuse of the criminal process then the *Hopt* and *Carmell* contexts present differing risks of abuse. In *Hopt* the prosecution is being permitted to use testimony of those who have committed felonies. Given that a witness's prior record may be used in court to challenge his testimony it is not clear that permitting felons to testify greatly increases prosecutorial discretion. This evidence may be important, but the jury or the judge will be inclined to discount a felon's testimony.¹⁶⁷ In the *Carmell* context, obtaining the testimony of the complainant is obviously important to the prosecutor's case. Increasing the prosecutor's discretion with regard to this type of testimony gives the prosecutor and other parties considerably greater room to use the system to their advantage.¹⁶⁸ Thus, the change in *Carmell* makes it significantly easier to credibly threaten to convict a defendant, whereas the change in *Hopt* may only increase this risk slightly because of discounting.¹⁶⁹ The potential for abuse is greater in *Carmell* than in the *Hopt* context.¹⁷⁰

C. Some Other Measures That Constrain Rent-Seeking

In this Part we address two areas of recent controversy, void-for-vagueness and entrapment doctrine.

1. Void-for-vagueness

The void-for-vagueness doctrine serves as a constraint on legislators and law enforcement agents that curtails their discretion. The early cases struck down laws that were deemed vague, in the sense that they granted law

¹⁶⁷ One could, of course, posit instances where such abuse may occur (e.g., the prosecution offering a felon, who is currently in jail, a reduced sanction in some manner for fabricating testimony), but one suspects that the risk is either not great or that the testimony would not be generally believed. See Joshua M. Levinson & Brian Lambert, *Twenty-Ninth Annual Review of Criminal Procedure*, 88 GEO. L.J. 1175 (2000) (discussing government's disclosure obligations). Note that the *Hopt* decision concerned a law that related to permitting all felons' testimony (even those not in jail and hence not subject to as much prosecutorial arm-twisting) so that the threat of prosecutor's using their power against this general group to extract false testimony is not entirely persuasive. See *Hopt*, *supra* note 165, at 588. The facts of *Hopt* involved a felon in prison at the time, but the rule was not limited to those instances. See *id.* at 589.

¹⁶⁸ See *Carmell*, *supra* note 159, at 1640.

¹⁶⁹ See *id.*

¹⁷⁰ Note that we have only argued that the *potential* for abuse is greater in *Carmell* than *Hopt* not necessarily that *Hopt* was correctly decided – for that to be the case we would need to believe that the potential for abuse in *Hopt* was below the threshold, whatever it might be, that is needed to trigger ex post facto prohibitions. We make no comment on that at this stage except to argue that there is a difference in abuse potentials between the cases. Also we make no comment on where the threshold for triggering ex post facto prohibitions should/might be.

enforcement agents broad discretion in deciding what is legal and what is not.¹⁷¹ Such discretion gives enforcement agents wide power to extract wealth through the criminal law enforcement process. This raises the specter of rent seeking.

More recently a new twist to the issue of vague statutes has arisen: what happens when the group against which selective enforcement may be used agrees to it or supports it? This is the situation that gave rise to the Supreme Court's decision in *Chicago v. Morales*.¹⁷² The case developed after Chicago passed an ordinance that prevented "criminal street gang members from loitering with one another or with other persons in any public place."¹⁷³ The definition of loitering was quite broad – "[remaining] in any one place with no apparent purpose,"¹⁷⁴ and police were required to ascertain whether some (at least one of two) persons who were loitering were gang members. The ordinance, it was argued, had the support of the communities in which it was most likely to be enforced, the high-crime urban neighborhoods of Chicago.¹⁷⁵ In spite of the claimed community support, the Supreme Court held the ordinance unenforceable on vagueness grounds.¹⁷⁶

In light of the broad grant of discretion to the police under Chicago's anti-gang ordinance it is obviously possible that the police could engage in selective enforcement.¹⁷⁷ The community may have consented to this risk in order to enhance its security.¹⁷⁸ The general question raised by *Morales* is whether a community should be allowed to make this trade off.

Some scholars have argued for an exception to the void-for-vagueness doctrine on the ground that selective enforcement or targeting is extremely unlikely in the *Morales* context.¹⁷⁹ Specifically, the use of an anti-gang ordinance to oppress a particular group, or transfer wealth from one group to another, is extremely unlikely in the *Morales* setting for two reasons. First, the ordinance, arguably, had a high degree of community support even within the

¹⁷¹ *Papachristou v. City of Jacksonville*, 405 U.S. 156 (1972); *Fields v. City of Omaha*, 810 F.2d 830 (8th Cir. 1987).

¹⁷² 527 U.S. 41 (1999).

¹⁷³ *Id.* at 45 – 46.

¹⁷⁴ *Id.* at 47 & n.2.

¹⁷⁵ *Id.* at 74 (Scalia, J., dissenting) ("[m]any residents of the inner city felt that they were prisoners in their own homes...Chicagoans decided that to eliminate the problem it was worth restricting some of the freedom they once enjoyed.").

¹⁷⁶ See *Morales*, *supra* note 172, at 51.

¹⁷⁷ See *id.*, at 58 – 59.

¹⁷⁸ See *id.*, at 74 (Scalia, J., dissenting) (stating that "[t]he minor limitation upon the free state of nature that is the prophylactic arrangement imposed upon all Chicagoans seemed to them (and it seems to me) a small price to pay for liberation of their streets").

¹⁷⁹ See Tracey L. Meares & Dan M. Kahan, *The Wages of Antiquated Procedural Thinking: A Critique of Chicago v. Morales*, 1998 U. CHI. LEGAL F. 197.

neighborhoods most likely to be burdened by its enforcement.¹⁸⁰ Second, the costs of selective enforcement probably would have been internalized within the relevant communities.¹⁸¹ Put another way, this case is unlike the example of selective law enforcement in the Jim-Crow South, which involved the use (or non use) of force by a politically dominant group to oppress a politically marginal group. The communities that supported the anti-gang ordinance made an apparently conscious decision to trade off protection from police harassment in order to reduce the crime rate in their neighborhoods. The general implication of this argument is that constitutional restraints on criminal law enforcement should be relaxed in settings where the risk of selective enforcement is minimal. For simplicity, we will refer to this as the internalization critique.

The framework of this paper provides an alternative to the internalization critique, as well as a justification for *Morales*. The internalization critique misses an important feature of pro-defendant procedural protections. Not only do they dampen temptation for selective enforcement, or inter-group wealth expropriation, they also cut down the prospects for what we have termed “simple corruption.” That is, procedural protections also have the function of reducing the opportunities for an individual enforcement agent to enrich himself by, for example, using his position to bully individuals who can be threatened with arbitrary arrest and punishment. This is potentially just as harmful as inter-group wealth extraction, because as long as it is possible for individuals to enrich themselves through the enforcement process, people will devote resources to acquiring positions as enforcement agents. Each position along the chain of enforcement (from the officer on the street to his immediate superiors, to prosecutors, to parole officers) could become a source of monopoly profits for the individuals who occupy them.¹⁸² When this occurs on a large scale, consistency and impartiality in enforcement are unlikely to be observed.¹⁸³

¹⁸⁰ See *id.*; *Morales*, *supra* note 172, at 74 (Scalia, J., dissenting) (stating that “[t]he minor limitation upon the free state of nature that is the prophylactic arrangement imposed upon all Chicagoans seemed to them (and it seems to me) a small price to pay for liberation of their streets”).

¹⁸¹ This seems implicit in Justice Scalia’s approach. See *Morales*, *supra* note 172, at 74.

¹⁸² This is essentially the argument made in the context of a vertically fragmented enforcement scheme. See *supra* text accompanying note 61.

¹⁸³ In addition to this argument it should be noted that the factual predicates of the case – that the minority dominated community supported this measure and hence was willing to trade off civil rights for enhanced safety – is a fairly contentious matter. First, there is significant debate over whether the community did actually support these measures. See Richard R.W. Brooks, *Fear and Fairness in the City: Criminal Enforcement and Perceptions of Fairness in Minority Communities*, 73 S. CAL. L. REV. 1219 (2000). The support seems to be equivocal. See *id.*, at 1233 – 35. Second, even if the community did support this measure it does not tell us too much about how much the community is willing to trade rights for safety. See *id.*, at 1262 (noting that “...if poor blacks are more supportive of the American legal system because they are less aware of the existence of race-based unfairness..., then a desire or willingness on their part to expand legal enforcement in poor urban communities is not a fully informed position for lawmakers to follow.”). Allit says is (assuming the community did support the measure) that the community preferred this mix of safety and civil rights over the current one. It does not tell us that this option would have been preferred over others that were not offered to the community. Indeed, it is possible that other alternatives could have

2. Entrapment

The fact that the entrapment defense is a relatively new common law doctrine probably has a lot to do with the expanding scope of criminal prohibitions.¹⁸⁴ The defense does not exist for common law crimes, such as murder.¹⁸⁵ This is a sensible result because even those who commit murder as a result of inducement, cajolery, or solicitation are still likely to present a danger to society.¹⁸⁶ The defense is connected today largely with drug prosecutions and other victimless crimes.¹⁸⁷

Our theory provides a rationale for the entrapment defense and for its relatively recent appearance in the law. Entrapment's recent appearance in connection with new prohibitions can be understood as a reaction to the rent-seeking hazards associated with expanding criminal prohibitions. The short list of common law crimes encompasses conduct that is uniformly considered undesirable.¹⁸⁸ It is possible, as we have argued, for enforcement agents to enforce them selectively, but there are many procedural protections in existence to constrain this incentive. Relatively new criminal prohibitions, on the other hand, often encompass conduct that is not uniformly considered undesirable, and may quite easily be made the basis for selective enforcement.¹⁸⁹ Consider, for example, the debates concerning the more severe punishments for crack cocaine, heavily used in minority neighborhoods, relative to powdered

been preferred by the community to the ordinance or the current state of affairs. Community support therefore only tells us so much. *See id.* at 1271. Finally, even if the community did support this ordinance over all others that does little to address concerns with the prosecutors' now enhanced power to extract wealth.

¹⁸⁴ *See Sorrells v. United States*, 287 U.S. 435, 453 (1932) (Roberts, J., concurring) (noting that "[t]he increasing frequency of the assertion that defendant was trapped is doubtless due to the creation by statute of many new crimes, (e.g., sale and transportation of ...narcotics) and the correlative establishment of special enforcement bodies for the detection and punishment of offenders").

¹⁸⁵ *See LAFAVE & SCOTT, supra* note 146, at 421 – 22; *See also* MODEL PENAL CODE §2.13 (noting that the defense of entrapment is unavailable when "causing or threatening bodily injury is an element of the offense").

¹⁸⁶ *See* MODEL PENAL CODE §2.13 Comment at 420 (1985) (noting that "one who can be persuaded to cause such injury presents a danger that the public cannot safely disregard").

¹⁸⁷ *See* Dana M. Todd, *In Defense of the Outrageous Government Conduct Defense in the Federal Courts*, 84 KY. L. REV. 415, 419 (1995); John F. Pries, *Witch Doctors and Battleship Stalkers: The Edges of Exculpation in Entrapment Cases*, 52 VAND. L. REV. 1869, 1872 (1999).

¹⁸⁸ These common-law crimes are *mala in se*, or morally wrong acts. These crimes are distinguished from *mala prohibita* crimes, which are acts made criminal by statute, but are not of themselves considered criminal. *See* BLACK'S LAW DICTIONARY 956 (6th ed. 1990). One argument in support of excluding the entrapment defense from conduct uniformly considered criminal is that, "from a moral perspective, it is wrong to punish those...who lack an opportunity to know and adhere to the law due to government conduct." John T. Parry, *Culpability, Mistake and Official Interpretations of Law*, 25 AM. J. CRIM. L. 1, 5 – 6 (1997).

¹⁸⁹ *See Sorrells, supra* note 184, at 453 (Roberts, J., concurring) (noting that "efforts...to obtain arrests and convictions (of these crimes) have too often been marked by reprehensible methods."). *See also* Sanford H. Kadish, *Fifty Years of Criminal Law: An Opinionated Review*, 87 CALIF. L. REV. 943, 970 (1999).

cocaine.¹⁹⁰ As a positive matter, then, we should expect to observe, and we have observed, the entrapment defense expanding in scope and gaining a stronger footing in criminal law doctrine as the scope of criminal prohibitions extends beyond basic common law crimes.

One rationale provided for the entrapment defense is that it enables courts to avoid becoming tainted by condoning inappropriate conduct, or “abhorrent transaction(s)”.¹⁹¹ The rationales could be expanded to include the claim that it discourages police officers from engaging in inappropriate conduct because it effectively denies them the reward (in terms of prosecutions) for engaging in such conduct.¹⁹² This rationale has been criticized as inadequate on the ground that the purity of the courts, or of enforcement agents, has no particular value in itself.¹⁹³ If some impurity enhances deterrence, why not allow it?

The justification suggested by this paper’s framework avoids this problem. The function of the entrapment defense is not simply to protect the purity of enforcement agents, but to dampen rent-seeking incentives at lower levels of the enforcement process. If enforcement agents are denied the fruit of entrapment efforts, the rewards from using the law enforcement process to target specific individuals or groups fall.

VIII. EMPIRICAL EVIDENCE

We ran a regression of Transparency International’s corruption index on several variables, including measures of key pro-defendant criminal procedural rules. The reasoning behind this exercise is that if pro-defendant criminal procedural rules reduce the incentives to use the criminal laws for inter-group wealth extraction and for personal enrichment, the degree of corruption should be lower in countries that have such procedural rules.

The key measures of pro-defendant rules used in the regression analysis are the existence of a rule prohibiting cruel and unusual punishment and the existence of a common law system. The common law system represents a greater degree of pro-defendant bias in its criminal procedures than civil law systems (e.g. the reasonable-doubt standard, certain aspects of Double Jeopardy, certain

¹⁹⁰ See *United States v. Anderson*, 82 F.3d 436 (D.C. Cir 1996); *United States v. Sanchez*, 81 F. 3d 9 (1st Cir. 1996) (holding that increased sentencing for possession of crack rather than powdered cocaine was not unconstitutionally void for vagueness).

¹⁹¹ *Sorrells*, *supra* note 184, at 459 (Roberts, J., concurring).

¹⁹² See *id.*, at 448.

¹⁹³ The majorities in *Sorrells* and *Sherman* used a subjective test in applying the entrapment defense (focusing on the defendant’s predisposition for crime and mental state), rather than focusing on the action of government and law enforcement officials. See *Sorrells*, *supra* note 184, at 451; *Sherman v. United States*, 356 U.S. 369, 375 – 76 (1958). See also Jason R. Schulze, *United States v. Tucker: Can the Sixth Circuit Really Abolish the Outrageous Government Conduct Defense?*, 45 DEPAUL L. REV. 943 (1996).

rules on Jury Trials). The results suggest that both types of pro-defendant protection are strongly negatively correlated with the degree of corruption. In other words, the presence of these protections is strongly correlated with low corruption.

Transparency International's 1997 "Corruption Perceptions Index" provides a score ranging from 10 (least corrupt) to 1 (most corrupt) for roughly 80 countries. The index, which measures international perceptions of corruption (bribe-taking and bribe-paying), is based on a survey of business people and analysts.¹⁹⁴ Our measure for the existence of a pro-defendant bias in criminal procedure is simply coding for whether the country has a common law system. In general, common law legal regimes appear to have a stronger pro-defendant bias than civil law regimes due in part to the presence of the reasonable-doubt standard, right to a jury trial, and certain aspects of Double Jeopardy amongst others.¹⁹⁵ The cruel and unusual punishment measurement is reflected by two "dummy variables." One variable, Crupun3, takes the value one if the country either does not have a rule prohibiting cruel and unusual punishment (e.g., Malaysia), or does not abide by the rule if it has one (e.g., Cameroon).¹⁹⁶ The other variable, Crupun2, takes the value one if the country has a rule prohibiting

¹⁹⁴ Our corruption index data are from 1997. The data are from surveys of business people and may capture general levels of corruption in the country. Our interest is in corruption in law enforcement, but we assume that corruption in law enforcement is probably correlated with general levels of corruption in a positive manner. For information on the Corruption Perceptions Index, see <http://www.worldbank.org/html/prddr/trans/>. See also Johann Graf Lambsdorff, *Background Paper to the 2000 Corruption Perceptions Index* (September 2000).

¹⁹⁵ See CRAIG M. BRADLEY, *CRIMINAL PROCEDURE: A WORLDWIDE STUDY* xv-xxii (1999), for a discussion of the major differences in criminal procedure between common law and civil law systems. In civil law, or inquisitorial, systems (found in a majority of continental European countries) a "theoretically neutral judicial officer conducts the criminal investigation and a judge...determines guilt or innocence." *Id.* at xv. Common law systems (found in the United States, Great Britain and its former colonies) are based on a mistrust of the government, and "the defendant is endowed with a quiver of rights that he may launch against the government at various stages of the proceeding." *Id.* at xvi. The reasonable-doubt standard is one of the defendant's weapons against common law criminal systems. England, Wales, South Africa, and the United States all require that the defendant's guilt be shown beyond a reasonable doubt. See *id.* at 122, 349. We tried to determine whether countries differ with respect to the double jeopardy rule. We found that every country in our sample either had a double jeopardy rule or had signed a treaty requiring compliance with such a rule.

¹⁹⁶ In our sample, 31 countries fall in this category (i.e., have a rule prohibiting cruel punishment but practice it in spite of the rule). Four countries in our sample do not prohibit cruel and unusual punishment (Malaysia, Morocco, Senegal, and Ivory Coast). We have no fixed definition of cruel and unusual punishment. For those states that ban cruel and unusual punishment, we have taken evidence of torture and serious physical abuse as a strong sign of the existence of such punishment. For example, the State Department says the following of Cameroon: "The Penal Code proscribes torture, renders inadmissible in court evidence obtained thereby, and prohibits public servants from using undue force against any person; however, although President Biya also promulgated a new law in 1997 that bans torture by government officials, there were credible reports that security forces continued to torture, beat and otherwise abuse prisoners and detainees. In New Bell and other nonmaximum-security penal detention centers, beatings are common and prisoners reportedly are chained and flogged at times in their cells." See http://www.state.gov/www/global/human_rights/1999_hrp_report/99hrp_toc.html

cruel and unusual punishment, but there are some concerns expressed about the state's compliance with its own rule (e.g., El Salvador).¹⁹⁷

We also included in the regression a measure of the ratio of government spending to gross domestic product (GDP). The purpose of this measure, labeled Ratio, is to capture the relative number of opportunities for corruption in a country. Presumably as this ratio increases, the number of enforcement agents increases relative to the size of the economy. For example, a country that has one licensing agent for every business will presumably have a large ratio of government spending to GDP, and a correspondingly large number of opportunities for bribe-paying. It happens, however, that this argument is inadequate because a country may choose to pay its licensing agents nothing (allowing them to make up the shortfall in bribes) and then the ratio of government spending to GDP may be relatively small. Regardless, we have included the Ratio measure as it is likely to have a substantial effect, though the direction of the effect is unclear *a priori*.

A. *A First Cut: Ordinary Least Squares Regression*

The results in Table 1 are for an ordinary-least squares regression of the Corruption Perceptions Index (labeled CPI) on Ratio and the pro-defendant procedure variables. The results indicate that corruption is significantly lower where a common law system is in place.¹⁹⁸ Moreover, both measures of cruel punishment constraints indicate that the failure to prohibit such punishment is positively correlated with corruption. The results indicate that moving from a regime in which there is a prohibition of cruel and unusual punishment (that is complied with) to one in which there is no such prohibition (Crupun3) reduces the Corruption index (i.e., increases corruption) by 3 points. This is quite a substantial drop given that the maximum score is 10. The existence of a common law system raises the Corruption index (i.e., reduces corruption) by 1.5 points.

¹⁹⁷ Countries in this category have been described by the U.S. State Department as applying punishment methods that vary from sleep deprivation to beatings. However, the State Department's report indicates that the abuses in these countries are either uncommon or not serious. For the State Department's analysis, see http://www.state.gov/www/global/human_rights/1999_hrp_report/99hrp_toc.html. Thus, the difference between Crunpun 2 and Crupun 3 is largely a matter of degree.

¹⁹⁸ The results in Table 1 were largely replicated in a second regression that includes a variable measuring the ratio of public sector wages to financial sector wages. In the second regression, the COMMLAW and CRUPUN2 coefficients remained roughly the same. The CRUPUN3 variable dropped to statistical insignificance, but this may largely be a byproduct of the sharp drop in observations because of missing wage data. We had only 42 observations for the second regression. The new variable PAFIN, which measures the ratio of public sector to financial sector wages, came in highly significant with a coefficient of 3.395 (t-statistic = 2.3), suggesting that the ratio of public sector to financial sector wages is negatively associated with corruption.

Table 1: OLS regression of corruption index on procedural constraints

CPI	Coef.	Std. Err.	T	P> t
RATIO	.039	.022	1.800	0.076
CRUPUN3	-2.997	.513	-5.840	0.000
CRUPUN2	-2.104	.604	-3.482	0.001
COMMLAW	1.455	.471	3.090	0.003
CONS	4.986	.730	6.832	0.000

Number of obs = 75

R-squared = 0.458

Adj R-squared = 0.427

The substantial impact of the two variables measuring cruel punishment constraints and the common law variable were replicated in expanded regression models (see Appendix) controlling for educational levels (percentage at primary level), religion (percent Catholic, Muslim), and economy type (socialist, mixed). Although the coefficient for Crupun3 fell in absolute value from 3 to 2, it remained statistically significant and increased in proportion to the common law variable.

Interestingly, the results suggest that the cruel and unusual punishment measures have a larger impact on corruption than the common-law measure (which proxies for a stronger pro-defendant bias in criminal procedure). This has interesting implications for the recent literature on common law protections and economic growth.¹⁹⁹ The results suggest that hard constraints on the state's freedom to profit through punishment may be a more important restriction on corruption than the existence of common law rules. The results can also be taken as empirical verification of our claim in Part VI.B that procedural protections and penalty restrictions (e.g., the prohibition against cruel and unusual punishment) work as complements with penalty restrictions having the more powerful long run effect.

B. Problems and Extensions

The ordinary-least squares approach is arguably inappropriate for this study, and, if so, the regression results reported in Table 1 may be biased in the sense that they fail to capture the true relationship between the corruption index

¹⁹⁹ For greater discussion see, e.g., Rafael LaPorta, Florencio Lopez de-Silanes, Andrei Schleifer & Robert W. Vishny, *Law and Finance*, 106 J. POL. ECON. 1113, 1151 – 52 (1998); Paul G. Mahoney, *The Common Law and Economic Growth: Hayek Might be Right*, University of Virginia School of Law, Legal Studies Working Paper Series, Working Paper 00-8, January 2000, available at <http://papers.ssrn.com/paper.taf?abstract_id=206809> (finding results consistent with the notion that the common law leads to greater economic growth relative to civil law systems in the period 1960 – 1992).

and the procedural constraints. The most important problem for this analysis is endogeneity.²⁰⁰ The ordinary-least squares regression carried out in Table 1 assumes that the explanatory variables (i.e., procedural constraints) are exogenous or predetermined with respect to the dependent variable (i.e., corruption). This may not be true, however, if procedural restraints are adopted as a response to corruption because then they are clearly not predetermined with respect to corruption. The procedural restraint variables would then be endogenous rather than exogenous variables.

This argument presents a potentially serious issue in the case of the cruel and unusual punishment measure used in the Table 1 regression. Some of the countries in our sample may have adopted rules prohibiting cruel and unusual punishment in order to dampen corruption. A parliament, for example, might pass such a law in order to constrain rampant corruption in some executive branch office. This concern is probably not serious in the case of the common law measure. It is unlikely that any of the countries in our sample would have adopted the common law in response to corruption. Common law regimes have appeared as the result of colonization instead of problem-focused legislation.

The standard solution to an endogenous explanatory variable problem is to find an “instrumental” variable, i.e., a variable that is correlated with the (potentially) endogenous explanatory variable and that is not itself endogenous.²⁰¹ Our solution to the instrument-choice problem is to follow the approach of Paulo Mauro in his article “Corruption and Growth.”²⁰² Mauro found that an index measuring the degree of ethnolinguistic fractionalization served as a good instrument for corruption. Mauro’s ethnolinguistic fractionalization instrument measures the probability that two persons drawn at random from a country’s population will not belong to the same ethnolinguistic group.²⁰³ Following Mauro’s lead, if the ethnolinguistic fractionalization measure serves as a good instrument for corruption and if corruption does cause some countries to adopt procedural restraints (e.g., banning cruel and unusual punishment), then the ethnolinguistic fractionalization instrument should also serve as a good instrument for a potentially endogenous procedural restraint measure. We therefore collapsed our cruel and unusual punishment indexes into one index (labeled Crupun) taking the values 1, 2, and 3, where the value 2 represents the level of punishment reflected in the Crupun2 index and the value 3 represents the level of punishment reflected in the Crupun3 index. Table 2 presents results of ordinary-leasts quares and instrumental-variables regressions

²⁰⁰ On endogeneity, or simultaneous-equations bias, *see, e.g.*, WILLIAM H. GREENE, *ECONOMETRIC ANALYSIS*, 652 – 711 (4th ed. 2000).

²⁰¹ *See* GREENE, *supra* note 200, at 370 – 375 (discussing instrumental variable regressions).

²⁰² 110 Q.J. ECON. 681 (1995).

²⁰³ *See id.*, at 682 – 83.

of the corruption index on Crupun and other variables shown in Table 1 (full regression results are provided in the appendix). We used the ethnolinguistic fractionalization index for the year 1961 as the instrument.

Table 2: Comparison of OLS and IV regressions of corruption index on procedural constraints

OLS Regression				
CPI	<i>Coef.</i>	<i>Std. Err.</i>	<i>T</i>	<i>P> t </i>
RATIO2	.112	.037	3.00	0.004
CRUPUN	-1.177	.259	-4.53	0.000
COMMLAW	.882	.471	1.80	0.077

Source: Table A.4 of appendix

Instrumental Variables Regression				
CPI	<i>Coef.</i>	<i>Std. Err.</i>	<i>T</i>	<i>P> t </i>
RATIO2	.010	.097	0.10	0.918
CRUPUN	-4.660	2.329	-2.00	0.050
COMMLAW	2.281	1.351	1.69	0.097

Source: Table A.3 of appendix

The new results show that the instrumental variables estimate of the coefficient for the cruel and unusual punishment index is negative, as in the original ordinary-least squares regression. Moreover, the new Crupun coefficient estimate is even larger in absolute value relative to the common law coefficient estimate. The results in the bottom part of Table 2 suggest that the corruption index falls roughly 4 points for every increase (from 1 to 2, or from 2 to 3) in the cruel and unusual punishment index.

These results are sufficient for our purposes. A key claim of this paper is that procedural restraints dampen corruption, and that is confirmed by all of the regression results. Of course, to the extent endogeneity is a serious issue, which seems to be confirmed by the results in Table 2, the instrumental variables estimates should be viewed as the reliable ones. Moreover, the instrumental variable results are consistent in a deeper sense with the theory of this paper. The theory views procedural restraints as a price or tax that constrains the demand for corruption, and also as a product of or supply-side response to corruption. The instrumental variable results are consistent with and effectively

reconcile supply and demand side theories of the relationship between procedural constraints and corruption.²⁰⁴

One possible criticism of our interpretation of the regression results is that the Crupun measure may simply be a reflection of the stability or effectiveness of government. An ineffective government is likely to be corrupt. If our cruel and unusual punishment index is simply a proxy for ineffectiveness, the correlation between the index and government corruption may not tell us much about the relationship between criminal procedural constraints and corruption. Though we cannot be confident that this argument is wrong, there are two reasons to doubt it. First, as a conceptual matter, it is difficult to distinguish government ineffectiveness from corruption. We are inclined to think ineffectiveness results from corruption rather than the other way around. Second, in expanded regression models shown in the appendix we included several variables that could be interpreted as proxies for government effectiveness. Our central results are replicated in the expanded regressions.

IX. CONCLUSION

The strong pro-defendant bias in criminal procedure is a stalwart of Anglo-American law. This has seemed perplexing because it could encourage criminal wrongdoing, whereas one of the primary reasons for declaring something criminal is to try to severely reduce its incidence.²⁰⁵ This apparent contradiction has led to several attempts to justify the pro-defendant bias in criminal procedure. Our paper develops a simple positive theory for a pro-defendant bias: to constrain the costs associated with rent seeking in the criminal law enforcement process. The key institutional features of the punishment process (reasonable-doubt rule, penalty restrictions, the jury), the case law on

²⁰⁴ The size of the “endogeneity bias” in the estimate of the Crupun coefficient is simply the difference between the ordinary-least-squares and the instrumental-variables estimate. Looking at Table 2 we know that the ordinary-least-squares estimate is -1.177 and the instrumental-variables estimate is -4.660, and thus the suggested endogeneity bias is 3.483. The positive sign for the bias term is consistent with what we would expect under the traditional instrumental-variables framework. Our theory treats procedural constraints as a “price” reducing the demand for corruption. The alternative part of the model treats past corruption as a signal or call that leads to the production of more procedural restraints – the equivalent of a supply relationship. In the standard demand-supply model, any estimate of the slope of the demand curve by ordinary-least-squares will be biased upward due to movements along the supply curve caused by exogenous disturbances (or shocks) to demand. *See, e.g.,* G.S. MADDALA, *INTRODUCTION TO ECONOMETRICS* (2d ed. 1992); A. M. McGahan, *Cooperation in Process and Capacities: Trade Associations in Brewing After Repeal*, 38 J. LAW & ECON. 521, 531–532 (1995) (providing a mathematical analysis on the effect of exogenous factors in the demand-supply function). We seem to have an equivalent result here. Exogenous disturbances in the “demand” schedule for corruption generate an upward bias in the ordinary-least-squares estimate of the relationship between procedural restraints (as measured by Crupun) and the corruption level.

²⁰⁵ *See* Marshall & Duff, *supra* note 2; Estrick, *supra* note 2; Hart, *supra* note 2. *See also* Coffee, *supra* note 2; Kadish, *supra* note 2; PACKER, *supra* note 2.

criminal procedure, and empirical evidence on international corruption appear to be consistent with this theory.

We began by sketching some of the more common criminal procedures and examining the traditional justifications given for them. In particular, we focused on the reasonable doubt standard and the justification most commonly given for it – that we are more concerned with the sanctioning and deterrence costs associated with false convictions than with false acquittals. We found the traditional rationale makes implicit assumptions about the relative costs of false convictions and false acquittals that are inadequately unsupported by the empirical evidence.

In light of this, we argued that the pro-defendant bias can be better explained as a means to constrain costs associated with rent seeking. Absent some constraint, prosecutors and other agents would use the criminal process to benefit themselves. This prospect, in turn, encourages groups to lobby law enforcement agents for selective enforcement. The costs of unfettered prosecutorial behavior could very plausibly be large enough to justify pro-defendant procedural protections.

Approaching criminal procedure from the perspective of constraining self-interested actors is likely to provide important insights into the current scope of criminal procedure as well as insights about whether certain doctrines should be extended or not. The analysis here could be applied to other current topics, including the extension of criminal procedural protections to civil suits brought by government agencies and reform proposals to abolish Double Jeopardy and the right to a jury trial for serious crimes in the UK.²⁰⁶

²⁰⁶ See Andra Varin, *In U.K., Double Jeopardy's in Jeopardy*, http://abcnews.go.com/sections/world/DailyNews/britain_justice020801.html.

Appendix

1. Tobit Regression Analysis

Below we report the results of expanded corruption index regressions. The new variables below are EDU1 = percentage of population (25 and older in 1991) that has something less than a primary education only. This includes those who have no education at all or have completed primary school but have gone no further. (Source: Statistical Abstract of the World, 3d. ed., Annmarie Muth, ed., Gale Research (1997); RELC = percentage of the population who are Catholic, RELM = percentage of the population who are Muslim or Islamic (Source: Statistical Abstract, and where necessary supplemented from <http://www.adherents.com/>). ETM = dummy equal to one if economy is classified as mixed socialist-capitalist, ETS = dummy equal to one if economy is classified as socialist (Source: <http://www.cia.gov/cia/publications/factbook/indexgeo.html>).

Since the dependant (CPI) cannot be greater than 10 or less than 1, we used the "Tobit" regression model. The coefficient estimates in the Tobit regression were virtually indistinguishable from the ordinary least squares coefficient estimates. This is not surprising, since almost all of our dependent variable observations were within the limits of 1 and 10.

Table A.1: Tobit regressions

CPI	Coef.	t	Coef.	t
RATIO	.044	1.69	.047	2.05
CRUPUN3	-1.889	-3.04	-2.515	-4.61
CRUPUN2	-1.761	-2.80	-2.146	-3.85
COMMLAW	1.167	2.17	.904	1.83
EDU1	-.028	-2.26		
RELC	-.009	-1.30	-.014	-2.19
RELM	-.006	-0.61	-.012	-1.48
ETM	-.812	-1.45	-.408	-0.81
ETS	-1.668	-2.36	-1.619	-2.56
CONS	7.091	7.19	5.868	7.80
Number	61		75	
Pseudo R-squared	.16		.16	

To facilitate comparisons among the different regression results reported in this section, the next table shows a tobit regression using a few new variables. We also replaced RATIO with RATIO2, which is the World Bank's measure of the general government consumption as a percentage of GDP in 1997. The variable GNPCAP measure per capita GNP in 1995.

Table A.2: Tobit regression

CPI	Coef.	St. Err.	T	P> t
RATIO2	.116	.035	3.30	.002
GNPCAP	.000	.000	1.78	.080
CRUPUN3	-2.309	.489	-4.72	.000
CRUPUN2	-1.689	.549	-3.07	-3.85
COMMLAW	.801	.465	1.72	.090
RELC	-.012	.006	-1.96	.055
RELM	-.009	.007	-1.19	.239
ETM	.028	.435	-0.06	.949
ETS	-1.370	.584	-2.34	.022
CONS	4.607	.758	6.08	.000
Number	74			
Pseudo R-squared	.19			

2. Instrumental Variables Regressions

We collapsed the two cruel and unusual punishment indexes (Crupun2 and Crupun3) into one (Crupun), which takes the value 1 (prohibition of cruel punishment), 2 (prohibition but imperfect compliance), or 3 (no prohibition). As reported in the text, Table A.3 uses Mauro's ethno-linguistic fractionalization index from 1961 as an instrument for Crupun. Table A.4 presents ordinary-least-squares regressions for comparison.

Table A.3: Instrumental variables regressions

CPI	<i>Coef.</i>	<i>t</i>	<i>Coef.</i>	<i>t</i>
RATIO2	.092	0.90	.010	0.10
GNPCAP	-.000	-0.43	-.000	-0.41
CRUPUN	-4.219	-1.76	-4.660	-2.00
COMMLAW	2.012	1.64	2.281	1.69
EDU1	-.005	-0.16		
RELC	.005	0.30	.003	0.16
RELM	.043	1.21	.037	1.10
ETM	.195	0.17	.274	0.28
ETS	-.381	-0.24	-.724	-0.55
CONS	10.987	3.56	12.768	2.61
Number	53		66	
F statistic for regression	2.37		2.67	
Prob > F	0.0283		0.0146	

Table A.4: Ordinary least squares regressions

CPI	<i>Coef.</i>	<i>t</i>	<i>Coef.</i>	<i>t</i>
RATIO2	.108	2.23	.112	3.00
GNPCAP	.000	1.51	.000	1.70
CRUPUN	-.832	-2.74	-1.177	-4.53
COMMLAW	1.070	2.03	.882	1.80
EDU1	-.029	-2.33		
RELC	-.005	-0.83	-.011	-1.71
RELM	.000	0.03	-.007	-0.93
ETM	-.368	-0.70	.084	0.18
ETS	-1.389	-2.01	-1.303	-2.11
CONS	6.641	6.22	5.656	6.17
Number	60		74	
R-squared	0.61		0.58	
F statistic for regression	8.72		11.36	
Prob > F	0.0000		0.0000	