

# Michigan Law Review

---

Volume 95 | Issue 8

---

1997

## Deterrence's Difficulty

Neal Kumar Katyal

*Georgetown University Law Center*

Follow this and additional works at: <https://repository.law.umich.edu/mlr>



Part of the [Criminal Law Commons](#), [Law and Economics Commons](#), [Law and Psychology Commons](#), and the [Law and Society Commons](#)

---

### Recommended Citation

Neal K. Katyal, *Deterrence's Difficulty*, 95 MICH. L. REV. 2385 (1997).

Available at: <https://repository.law.umich.edu/mlr/vol95/iss8/3>

This Article is brought to you for free and open access by the Michigan Law Review at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Law Review by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact [mlaw.repository@umich.edu](mailto:mlaw.repository@umich.edu).

# DETERRENCE'S DIFFICULTY

Neal Kumar Katyal\*

## TABLE OF CONTENTS

INTRODUCTION .....	2386
I. THE ECONOMICS OF SUBSTITUTION .....	2389
A. <i>Conventional Deterrence and its Marginal Refinement</i> .....	2389
B. <i>Substitution</i> .....	2391
C. <i>The Example of Crack Cocaine</i> .....	2402
D. <i>Further Refinements</i> .....	2408
1. <i>A Model of Economic Substitution</i> .....	2408
2. <i>Probability Versus Sentence</i> .....	2410
3. <i>Two Types of Deterrence</i> .....	2416
II. SOME NORMATIVE RESULTS .....	2420
A. <i>The New Benefit of the Sentencing Guidelines</i> ..	2420
B. <i>Redefining Optimal Penalties</i> .....	2423
C. <i>The Income Effect and Y-Optimality</i> .....	2432
D. <i>Income Effects and Drugs</i> .....	2434
1. <i>Giffen Goods</i> .....	2434
2. <i>Price Discrimination</i> .....	2439
III. SHAPING SUBSTITUTION .....	2441
A. <i>The Preference-Shaping View of Criminal Law</i> .....	2442
B. <i>Sociological Norms and the Coercion of Preferences</i> .....	2447
1. <i>Know Your Law</i> .....	2447
2. <i>The Coercion of Punishment</i> .....	2455
3. <i>The Counterproductive Coercion of Stigma</i> .....	2457
C. <i>Preference and Reference</i> .....	2461
D. <i>Adaptive Preferences Revisited</i> .....	2467

---

\* Associate Professor of Law, Georgetown University Law Center. — Ed. For generous comments and criticism, I thank Bruce Ackerman, Akhil Amar, Antonia Apps, Ian Ayres, Lea Brilmayer, Dick Craswell, Steven Duke, Henry Hansmann, Melissa Hart, Julie Hilden, Christine Jolls, Sonia Katyal, Avery Katz, Roger Masters, Steve Morse, Kathy Ruemmler, Alan Schwartz, Michael Seidman, David Weisbach, and Kenji Yoshino.

I am also grateful for the insightful critical remarks Dan Kahan has provided in a response that accompanies this article. See Dan M. Kahan, *Between Economics and Sociology: The New Path of Deterrence*, 95 MICH. L. REV. 2477 (1997). This article is dedicated to the great teacher of my great teachers, Guido Calabresi.

IV. SYNTHESIZING THE EFFECTS OF DETERRENCE . . . .	2470
TO THE FUTURE . . . . .	2475

### INTRODUCTION

We all crave simple elegance. Physicists since Einstein have been searching for a grand unified theory that will tie everything together in a simple model. Law professors have their own grand theories — law and economics's Coase Theorem and constitutional law's Originalism immediately spring to mind. Criminal law is no different, for the analogue is our faith in deterrence — the belief that increasing the penalty on an activity will mean that fewer people will perform it. This theory has much to commend it. After all, economists and shoppers have known for ages that a price increase in a good means that people will consume less of it. But sometimes the consumption picture is more complicated than this simple economic account. Indeed, the leap from ordinary goods to criminal behavior is a large one, and one that presents complications of its own.

This article sketches out several possible outcomes that arise from the criminalization of behavior. Incorporating recent work in economics, sociology, and psychology, it explains the ways in which the deterrence question is more difficult than many of us have assumed and illustrates how criminalization can create unintended, and sometimes perverse, incentives.

The first part of this article introduces (or reintroduces) criminal lawyers to the idea of substitutes and complements in economic theory. Briefly, two products are substitutes when they compete with each other and are complements when they "go together."<sup>1</sup> Consumers will tend to use more of a good — to substitute in favor of the good — when its relative price falls, and to use less of it — to substitute away from the good — when its relative price increases. If the price of tea increases, for example, substitution theory predicts that the demand for coffee would increase. But the demand for other products that go with tea, such as lemons, may drop because tea and lemons are complementary products. An increase in tea prices will not, however, directly affect the use of foot powder. Because tea and foot powder do not have much to do with

---

1. Put a different way, two goods are substitutes if an increase in the price of one causes an increase in demand for the other. They are complements if an increase in the price of one causes a decrease in demand for the other. For explanations of these concepts, see WALTER NICHOLSON, *INTERMEDIATE MICROECONOMICS AND ITS APPLICATION* 98 (4th ed. 1987); PAUL A. SAMUELSON & WILLIAM D. NORDHAUS, *ECONOMICS* 411 (12th ed. 1985).

each other — the consumption of one is not affected by the consumption of the other — they are considered “independent goods.”

Consumption is also affected by what economists term the “income effect.” The income effect predicts that an increase in the price of a good reduces the real income of a consumer of that good. This reduction in real income means that consumption of virtually all other goods may decrease. To continue our simple example, a drastic increase in the price of tea may indirectly decrease the consumption of foot powder because tea consumers will not have as much money to spend on foot powder. And here again, substitution is at work — a high price on tea will not have a strong income effect if consumers are willing to shift to coffee. The substitution effect, therefore, tempers the income effect of a price increase.

What do tea and lemons have to do with criminal law? The criminal law can be seen as setting prices for crimes, and these price effects may cause substitution. (Indeed, standard deterrence models implicitly assume substitution by holding as a central tenet that a penalty on activity *X* will lead people to substitute the legal behavior of refraining from *X*.) Just as a reduction in the price of tea may lead to an increase in the demand for coffee, an increase in the price of one crime may induce consumers and dealers to find a substitute. That is, a penalty on crime *X* may lead to behavior *Y* and *Z*. *Y* and *Z*, moreover, may be criminal acts. An assessment of the social utility of penalizing *X* should therefore discount the benefit of decreasing *X* by the corresponding harm of increasing *Y* and *Z*. This idea suggests that a penalty cannot be set at a level based simply on the harm that an activity causes. Rather, it must be set at a level that assimilates consumer reaction to the penalty.

Granted, this economic model is thin, and many qualifications and enhancements must be made. Begin by questioning whether criminals actually know the law; for if not, then how can this view of deterrence make any sense? A theory of behavior that understands the norm-creating component of law provides one answer: substitution-like effects will occur when the law influences the social understanding of the blameworthiness of particular acts, even when people lack knowledge about the law. To take another complication, consider whether criminals act rationally. Cognitive psychologists have demonstrated that people decide between various alternatives in ways that are often “irrational,” through framing effects, aversion to extremes, and so on. This work may be extrapolated to show how, even when strict assumptions about rational choice are relaxed, making very harmful substitutes look worse and

less harmful substitutes look better might minimize dangerous substitution effects. Other modifications that this article develops address the income and distributional effects of penalties and the issue of how preferences and tastes toward crime develop. This article is not centered around the old-fashioned question of whether deterrence works. Instead, it asks: Under what conditions does deterrence work, and what are some of its effects when it succeeds and fails?

Part I begins by sketching out a common understanding of deterrence and substitution. Part I also introduces, as one simple example of substitution, how consumption and dealing of heroin may have increased as a reaction to the new crack penalties. Part II then illustrates some of the normative conclusions that follow from this economic model, including a new argument in favor of the Sentencing Guidelines. This Part also illustrates how, once the income effect is understood, criminalization of acts may have perverse consequences, possibly increasing commission of those acts in certain circumstances.

Parts I and II utilize an economic model of deterrence that assumes that preferences are fixed. This model yields a tidy set of prescriptions for policy, but it has the shortcomings noted above. It suffers, in short, from its own quest for simple elegance. Part III, therefore, posits what ultimately may be a more persuasive model of criminal punishment, one that relaxes the assumption of fixed preference. Instead of assuming that penalties act only to constrain opportunity, this Part examines how punishment may reduce the taste for an activity and explains how this preference-shaping conception of criminal law may either enhance or detract from the substitution effect. This Part goes on to develop a theory of the mechanism by which preferences adapt to punishment. Highlighting theories of social control from sociology and theories of choice from psychology, it shows how the criminal law exerts a strong, and sometimes unconscious, force on people's preferences.

The effects of criminalization are both varied and complicated, and it is only natural for readers to stumble on the contradictory effects and new vocabulary that appear in the pages that follow. Part IV tries to reduce this complexity by compiling and describing all the effects of a penalty and by providing examples of how penalties might work in practice. (While it might strike some as unnatural to save these illustrations until the end, doing so yields a fuller picture.) It must be stressed, however, that this piece is only a preliminary attempt to begin to peel away some of the assumptions

behind deterrence. Future work will have to examine substitution elasticities, consumer preferences, interactions between social attitudes and law, and a host of other details. This article is meant to raise more questions than it answers; what follows is therefore not a complete theory of criminal deterrence, but rather a first installment in an attempt toward one. In sum, this article illustrates the complexity of the deterrence question and shows that a simple and elegant answer to the deterrence question has not yet been found. As a result, much research and reflection is needed, particularly with regards to elasticities of substitution and the subterranean norm-creating role of the criminal law.

## I. THE ECONOMICS OF SUBSTITUTION

### A. *Conventional Deterrence and its Marginal Refinement*

To understand the implications of substitution, it is important to set the context by examining the conventional perspective on deterrence. Gary Becker pioneered modern economic analysis of criminal deterrence in his 1968 article.<sup>2</sup> Following Becker's lead, deterrence analysis has primarily focused on whether a particular penalty for a crime and the enforcement of the penalty will deter the commission of that crime. The analysis turns on whether the penalty is set at an appropriate level to optimize deterrence — balancing the cost of the activity against the cost of enforcement. It makes sense to think about many problems in substantive criminal law this way. The approach provides us, for example, with a straightforward way to examine whether a fifty-dollar fine for spitting gum out on the sidewalk will deter the optimal number of people from spitting their gum out and messing up sidewalks.

A few law and economics scholars, however, noted that Becker missed a crucial variable for optimality: marginal deterrence. The idea is essentially the problem of cliffs — exacting equal penalties for crimes of lesser and greater magnitude leads to crimes of greater magnitude. As its primary exponent, George Stigler, put it, “[i]f the thief has his hand cut off for taking five dollars, he had just as well take \$5,000.”<sup>3</sup> Stigler's insight tracked that of the

---

2. See Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169 (1968). The concept of price was first applied to criminal penalties, I believe, by the French criminologist Gabriel Tarde, who wrote: “The same thing applies, it seems to me, to the effect of penalties upon criminality as applies to the effect of prices on consumption.” GABRIEL TARDE, *PENAL PHILOSOPHY* 482 (Rapelje Howell trans., Little, Brown, and Co. 1912) (1890).

3. George J. Stigler, *The Optimum Enforcement of Laws*, 78 J. POL. ECON. 526, 527 (1970). For economic discussions of marginal deterrence, see Dilip Mookherjee & I.P.L. Png,

eighteenth-century Italian theorist Cesare Beccaria, who argued: "If an equal punishment is laid down for two crimes which damage society unequally, men will not have a stronger deterrent against committing the greater crime if they find it more advantageous to do so."<sup>4</sup> Jeremy Bentham made a similar move as well, arguing that the goal of a sanction is "to induce a man to choose always the least mischievous of two offences; therefore [w]here two offences come in competition, the punishment for the greater offence must be sufficient to induce a man to prefer the less."<sup>5</sup>

At its best, the marginal deterrence argument is one about creating incentives for individuals to refrain from committing the same crime on a greater scale.<sup>6</sup> As such, it is a much-needed refinement of the traditional deterrence question. While the traditional question asks whether a penalty for  $X$  deters  $X$ , the marginal deterrence theorist asks whether a penalty for  $X$  may prompt commission of the marginally more severe crime  $X + 1$  because that crime receives the same magnitude of punishment as  $X$ . For that reason, Stigler's solution to the marginal deterrence problem was to state that "[e]xpected penalties [should] increase with expected gains so there is no marginal net gain from larger offenses."<sup>7</sup> This is where substitution comes in.

---

*Marginal Deterrence in Enforcement of Law*, 102 J. POL. ECON. 1039 (1994); Louis L. Wilde, *Criminal Choice, Nonmonetary Sanctions, and Marginal Deterrence: A Normative Analysis*, 12 INTL. REV. L. & ECON. 333 (1992). A philosophical treatment of marginal deterrence, bristling with general insights into the nature of criminal law, is found in Dan M. Kahan & Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269 (1996).

4. CESARE BECCARIA, *On Crimes and Punishments*, in *ON CRIMES AND PUNISHMENTS AND OTHER WRITINGS* 1, 21 (Richard Bellamy ed. & Richard Davies et al. trans., Cambridge Univ. Press 1995) (1764). Or, as Chabroud complained 200 years ago, "if I have betrayed my country, I go to prison; if I have killed my father, I go to prison; every imaginable offence is punished in the same uniform way. One might as well see a physician who has the same remedy for all ills." MICHEL FOUCAULT, *DISCIPLINE AND PUNISH: THE BIRTH OF THE PRISON* 117 (Alan Sheridan trans., 2d ed. Vintage Books 1995) (1975) (internal quotation marks omitted) (quoting 26 ARCHIVES PARLEMENTAIRES DE 1787 A 1860 (1st ser.) 618 (1887) (statement of Chabroud)).

5. JEREMY BENTHAM, *AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION* 168 (J.H. Burns & H.L.A. Hart eds., Athlone Press 1970) (1789) (emphasis omitted); see also JEREMY BENTHAM, *THE THEORY OF LEGISLATION* 201 (Tripathi Private Ltd. 1975) (1802) ("Where two offences are in conjunction, the greater offence ought to be subjected to severer punishment, in order that the delinquent may have a motive to stop at the lesser." (emphasis omitted)).

6. This may explain why Stigler's attempt to rehabilitate marginal deterrence went over like a lead balloon with most people, with the important exception of a few academics. See, e.g., Richard Craswell, *Damage Multipliers in Market Relationships*, 25 J. LEGAL STUD. 463 (1996) (discussing the applicability of marginal deterrence in torts and contracts).

7. Stigler, *supra* note 3, at 531.

### B. *Substitution*

This article argues that Beccaria and Stigler's insight about marginal deterrence reflects only one instance of substitution. By pointing out that consumers (criminals) will base their choices between  $X$  and  $X + 1$  on the price (expected penalty) of each, marginal deterrence demonstrates a relationship between price and conduct. But criminals often have choices beyond  $X$  and  $X + 1$ ; thus an increase in the price of  $X$  may increase the commission of non- $X$  activities. The substitution perspective therefore expands the conventional deterrence question by asking whether a penalty for  $X$  will distort behavior and lead people to commit an altogether different crime ( $Y$ ,  $Z$ , or some combination of the two). These other acts may be other crimes, or they may be lawful endeavors. Substitution's chief insight is that it shows that the focus of marginal deterrence — one example of substitution — is too narrow.

The move I am making away from marginal deterrence is one familiar to many readers in public economics. Essentially, I am arguing that marginal deterrence utilizes a partial equilibrium model that does not yield a complete account of the effect of a particular penalty, and that general equilibrium analysis is, in general, better suited to analyzing the problem due to the numerous distortions in behavior created by a penalty.<sup>8</sup> In part, this is because at a high enough price, virtually anything can become a substitute. Even a monopolist who can charge a higher price because of the absence of competition is constrained by substitution from charging a price that is sky-high. In the days when trains were effectively the only means of long-distance transport, for example, many believed that train companies had a monopoly and that they could charge whatever they wanted. But a sky-high price for a train ticket meant that most people would stay home, or that they would use their bicycles.<sup>9</sup> The point is that at high prices, even the things that aren't ordinarily considered substitutes, such as bicycles and train rides, function that way. Indeed, even exact opposites — a train ride and forgoing a train ride — are substitutes when the price of a train ride is high enough. Or, to take another example, many politicians used to think that salt was a perfect item to tax because it had no substi-

---

8. Cf. RICHARD W. TRESCH, *PUBLIC FINANCE: A NORMATIVE THEORY* 14 (1981) (describing faults of partial equilibrium analysis in public finance context).

9. The point at which consumers switch to alternate goods (or to no good) depends on the price elasticity of demand. Price elasticity measures the change in demand for a given change in price.



tutes. But if salt taxes were raised, consumers would have switched to different spices or would not have used salt at all.<sup>10</sup>

Criminal law has unconsciously relied on such concepts. Viewed one way, the whole point of deterrence is to make the price of a crime high enough so that a criminal will “substitute” forgoing the crime. Just as a high price on train rides meant that people would not take them, a high price on a crime means people will not commit it. But, as the bicycle option in the train example shows, a gradation of potential options is open to people. When it comes to crime, however, most of us don’t take the economics seriously enough to examine whether an analogue to the bicycle exists — we assume that crimes do not have much to do with each other. Even marginal deterrence is generally concerned with the problem of incentives to commit the same crime on a greater scale. The paradigm is generally (but not exclusively) that criminals have four choices: commit the crime, do not commit the crime, commit the crime on a greater scale, or commit the crime on a lesser scale. The general idea is that crimes are not substitutes but “independent goods.” This underlying assumption that crimes are independent may explain why the conventional deterrence perspective has endured. As we shall see, this assumption appears to have some merit, particularly for what I call nonmarket crimes.<sup>11</sup> But many crimes are not independent. Their substitutability will be determined by the “elasticity of substitution” — the ease with which the demand for one crime may be substituted for the demand for another.<sup>12</sup>

To illustrate this concept, imagine that the demand for two products is perfectly elastic. Under this condition, a small change in

---

10. Price is set by both demand and supply. As Marshall noted, in the short run, the price of a good is set by demand because supply is fixed. At some point, however, the supply will clear the market. If the price of a good increases beyond a certain point, consumers will then substitute other goods for the product. The price will be set by the price at which a marginal user will buy the good. Nevertheless, as time elapses, the cost of producing the good will set the price because supply will increase. At that point, the marginal cost of production will set the price because a lower price will induce people to substitute the product for the others that they had previously used. But as the demand for the good at the lower price increases, the supply will again clear the market. People will then invest to increase the supply, and the circle will repeat itself. This led Marshall to note that the question of whether price was set by demand or supply was like asking whether one end of a scissors did the cutting or the other. See 1 ALFRED MARSHALL, *PRINCIPLES OF ECONOMICS* 348 (9th ed. 1961); see also NICHOLSON, *supra* note 1, at 10-11.

11. See *infra* text accompanying notes 145-50.

12. See J.R. HICKS, *THE THEORY OF WAGES* 117 (2d ed. 1963); see also Charles Blackorby & R. Robert Russell, *Will the Real Elasticity of Substitution Please Stand Up?*, 79 *AM. ECON. REV.* 882 (1989); Y. Mundlak, *Elasticities of Substitution and the Theory of Derived Demand*, 35 *REV. OF ECON. STUD.* 225 (1968).

the price of one product will lead to a large change in demand for that product because consumers will substitute the other. This is the flipside of the marginal deterrence problem. Marginal deterrence states that when the price of two goods is equal, consumption will shift to the product that provides the user with greater utility. Conversely, when the utility of two goods is equal (that is, the goods are perfectly elastic) and their prices are slightly different, people will prefer the cheaper one. A one-tenth of a cent difference in the price of two nonbrand cigarettes will lead the marginal user to substitute the cheaper one for the more expensive. Once brands are placed on the cigarettes, however, the products are not identical, and a small price difference will not induce the same amount of substitution. The degree of substitution is a function of the difference in price in relation to the desire for the two goods.

Thinking about criminal law, it is not difficult to understand how crimes committed for profit are ripe candidates for substitution analysis, but it is hard to imagine how other crimes can be analyzed in such terms. Yet even crimes of passion may, perhaps, be examined in terms of substitution.<sup>13</sup> Passion, after all, comes in different forms, and a penalty structure may induce people to act in particular ways by assigning costs to particular passionate activities. As Richard Herrnstein puts it, when husbands and wives start throwing dishes at each other, they do not usually throw the fine china.<sup>14</sup>

Take what seems like the quintessential example in which substitution would not occur: rape. Insofar as these categories are separable — and the argument does not depend on their separability — is rape a crime of sex, violence, or domination? If rapists seek sex, it might follow that lowering the penalty for prostitution will reduce the frequency of rape. If they seek to dominate and humiliate, legalized prostitution may provide a substitute as well. To the extent that rapists seek violence, lowering the penalties for other violence, say, assaults, may reduce the commission of rapes. Conversely, a high penalty for rape may mean that there are more instances of spousal abuse and other violence. These ideas are not policy suggestions, only possible illustrations of substitution at work. There may be many reasons why legalized prostitution is

---

13. See HERBERT L. PACKER, *THE LIMITS OF THE CRIMINAL SANCTION* 41-42 (1968) (arguing that Bentham's rational-actor deterrence model helps analyze even "irrational" and "impuls[ive]" crimes).

14. See JAMES Q. WILSON, *THINKING ABOUT CRIME* 127 (rev. ed. Vintage Books 1985) (1975).

problematic.<sup>15</sup> But the complementarity between prostitution and rape itself suggests that interrelationships between behavior cannot be ignored.<sup>16</sup>

Even when the penalties for a crime are so high that it appears that all the crime that can be deterred will be deterred, substitution presents possible problems. After all, one result of a higher penalty may be an *increase* in criminal activity — both of the particular crime and of other crimes. As explained above, at high prices, many crimes may substitute for one crime. The rapist who is deterred to rape a particular woman and is not deterred by a high penalty for rape may go out and commit other crimes. He may first rape the woman, then kill her, and finally assault unrelated others, because the cost of future criminal activity is negligible. If, on the other hand, the penalty for rape is not high, the marginal cost of additional criminal activity may be much higher.

To take another example, imagine the potential consequences of the “three strikes you’re out” rule.<sup>17</sup> If offenders know that, on their third offense, they will be jailed for life, they may be less likely to commit that third offense; but if they do, they might make it a drastic one. Indeed, they may even decide to kill the witnesses to their crimes because — at least in states without a death penalty — there is nothing more that the government can do to them.<sup>18</sup>

Viewed in these terms, the death penalty could provide an incentive for additional crime. The person who has already killed a child in a state where such action qualifies for the death penalty will not have a legal incentive, or at least not a very strong one, to re-

15. Consider, for example, the potential complementarity between legalized prostitution and rape and the harm legalized prostitution might do to the status of women. See generally Neal Kumar Katyal, Note, *Men Who Own Women: A Thirteenth Amendment Critique of Forced Prostitution*, 103 YALE L.J. 791 (1993).

16. Although I have expressed, through the train and salt examples, the belief that at a high enough price, virtually anything can be a substitute, I specifically do not make the claim that all crimes are in fact substitutes for each other. I simply do not have the necessary statistical information as to criminals’ preferences for committing different crimes in order to determine at what point substitution may or may not occur.

17. See, e.g., 18 U.S.C. § 3559 (1994); CAL. PENAL CODE § 1170.12 (West Supp. 1997).

18. See Nkechi Taifa, “Three-Strikes-and-You’re-Out” — Mandatory Life Imprisonment for Third Time Felons, 20 U. DAYTON L. REV. 717, 724-25 (1995); Victor S. Sze, Comment, *A Tale of Three Strikes: Slogan Triumphs Over Substance as Our Bumper-Sticker Mentality Comes Home to Roost*, 28 LOY. L.A. L. REV. 1047, 1070-71 (1995); Timothy Egan, *A 3-Strike Law Shows It’s Not as Simple as It Seems*, N.Y. TIMES, Feb. 15, 1994, at A1; Lisa Leff, *Chief’s 3 Strikes Theory Generates Mixed Reaction*, L.A. TIMES, Mar. 14, 1996, at 1; Edwin Meese III, *Three Strikes Laws Punish and Protect*, INSIGHT MAG., May 16, 1994, at 18, available in 1994 WL 11601844; William Tucker, *Three Strikes and You’re Dead*, AM. SPECTATOR, Mar. 1994, at 22, 26 (“Three-strikes-you’re-out will only turn more victims of violent crime into murder victims. Dead men tell no tales.”); Thaa Walker, *Police Concerned About ‘3 Strikes’ Law*, S.F. CHRON., Mar. 14, 1994, at A15.

frain from killing again. If a legal incentive exists, it is simply to avoid getting caught. But because deterrence is a function of both the sanction level and the probability that it will be imposed, the disincentive is lower for the repeat murderer than it is for the first-time one.<sup>19</sup> Since the penalty for one, two, or even three more murders is the same, the penalty itself does not work to provide additional deterrence.

On the other hand, Chief Judge Posner has used similar ideas to make an interesting argument for capital punishment. Posner suggests that the threat of death creates an additional rung on the ladder to increase the range of punishment:

Capital punishment is also supported by considerations of marginal deterrence, which require as big a spread as possible between the punishments for the least and most serious crimes. If the maximum punishment for murder is life imprisonment, we may not want to make armed robbery also punishable by life imprisonment, for then armed robbers would have no additional incentive not to murder their victims.<sup>20</sup>

Yet the problem with Posner's argument is that the death penalty eliminates the "spread" for all death-eligible crimes. As we have seen, imposing a death penalty for a particular crime gives an incentive to those who commit it to commit further crimes. Posner's inattention toward this problem is, I imagine, related to his reliance on the principle of marginal deterrence — itself a nice demonstration of the narrow focus of marginal deterrence and the broader perspective substitution engenders. Once it is understood that the death penalty creates equality between a variety of crimes, then the death penalty might not be as helpful as Posner suggests. If the question is between deterring  $X$  and  $X + 1$  — as Posner puts it, between armed robbery and murder + armed robbery — and those who commit  $X$  receive a life sentence, then perhaps those who commit  $X + 1$  should get the higher sentence of death.<sup>21</sup> But once substitution is considered, the question is not about simply deterring  $X + 1$ , but deterring crimes  $Y$  and  $Z$ , and therefore, imposing the

---

19. The possibility of being caught may not be constant, as it might increase if the murders take a particular pattern, or if the murders are committed in one place, because the police may devote more resources to such a crime. But then, the logical response might be to simply substitute other crimes that have nothing to do with the initial one.

20. Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 COLUM. L. REV. 1193, 1210-11 (1985).

21. This analysis, of course, does not consider other arguments against the death penalty, such as those based on morality. I should also note that a jury might be more likely to convict someone for one murder if evidence is introduced that he also committed other murders. This issue arises in, *inter alia*, debates about prior bad acts and prejudicial indictments that allege a laundry list of criminal activities to smear a defendant.

death penalty for the activity of  $X + I$  could cause  $Y$  and  $Z$  to be committed instead. Jail, unlike death, has many different rungs — ranging from duration of sentence to the size and characteristics of one's jail cell.<sup>22</sup>

The limitations of models that do not incorporate substitution also explain some of the rather odd conclusions of academics writing in criminal law. For example, Robert Cooter has argued that the government could afford to get sanctions wrong, but that it could not afford to make mistakes when it imposed prices on activities.<sup>23</sup> More recently, Dan Kahan has argued that the lenity doctrine should be abolished because people who break the law know that what they are doing is wrong.<sup>24</sup> The problem that both authors face is that they assume that people will either commit crime  $X$  or no crime at all. If people are picking between a range of crimes, however, then the government, *contra* Cooter, cannot afford to make mistakes about a sanction. Such mistakes create perverse incentives and may lead people to engage in conduct that is more harmful than what they would have done otherwise.<sup>25</sup> Similarly, the lenity doctrine may provide legislators with incentives to draft laws that specify precise punishments so that actors know the relative severity of different activities. And specification may permit the government to channel law-breaking behavior into the least harmful forms by increasing the expected sanctions on harmful activities. As Brissot once wrote, the laws must be clear “so that each member of society may distinguish criminal actions from virtuous

---

22. There are ways to create rungs with the death penalty. As one eighteenth-century French tract described it:

some prisoners may be condemned to be hanged, others to having their hands cut off or their tongues cut out or pierced and then to be hanged; others, for more serious crimes, to be broken alive and to die on the wheel, after having their limbs broken; others to be broken until they die a natural death, others to be strangled and then broken, others to be burnt alive, others to be burnt after first being strangled; others to be drawn by four horses, others to have their heads cut off, and others to have their heads broken.

FOUCAULT, *supra* note 4, at 32 (internal quotation marks omitted) (quoting JEAN ANTOINE SOULATGES, *TRAITÉ DES CRIMES* (1762)). These options are, of course, constrained by our Eighth Amendment — which functions as an explicit morality limit on the outer bounds of deterrence theory. See U.S. CONST. amend. VIII.

23. See Robert Cooter, *Prices and Sanctions*, 84 COLUM. L. REV. 1523, 1550 (1984) (“Fortunately, if the officials make a mistake and attach the wrong sanction to a crime — either too high or too low — the behavior of most people will not be affected by the error, because the cost of crime far exceeds the benefit.”).

24. Dan M. Kahan, *Lenity and Federal Common Law Crimes*, 1994 SUP. CT. REV. 345, 396-97, 400-03.

25. In addition, to the extent punishments are meted out based only on the harmfulness of an activity, without a discount for substitution effects, criminal enforcement resources are likely to be misdeployed. See *infra* section II.B.

actions.”<sup>26</sup> Yet the laws must also permit distinctions among criminal actions to provide proper incentives.

When considering proper incentives, the law must also take account of potential complementarities between activities. Certain acts lend themselves naturally to the commission of other acts, and even if the initial act does not itself harm society, the encouragement the initial act gives to these others may provide a basis for criminalization. The possession of a set of burglar's tools does not, in itself, create much social harm. Nevertheless, the law may seek to criminalize possession because of the complementarity between possession and burglary. Indeed, many acts are punished, or punished severely, not because of their inherent harm, but because of the likely complementarity with other crimes (such as prohibitions on drunk driving).<sup>27</sup> But, as previously noted in the three-strikes and death penalty examples, increasing the expected sanction might, in limited circumstances, increase complementarity between crimes as well.<sup>28</sup>

Increasing the expected sanction can be understood in economic terms as a decrease in the “wage” for a criminal activity, an increase in “tax” for that activity, or, more simply, as an increase in the “cost” of that activity. Naturally, the expected sanction is only one of many costs for criminals. Take the example of a drug dealer. Other costs may be the price at which the dealer buys the drugs and the cost of other materials such as crack vials or plastic glassine bags — all of which affect substitution. These costs, analogous to raw materials, are “variable” in the sense that a dealer incurs additional costs when she sells an additional unit. The dealer may also have “fixed” costs — for example, the price of a scale or other weighing device, rent on an apartment or other place to sell drugs, and so on.<sup>29</sup>

In addition to raw materials and fixed costs, two other costs deserve consideration. First, the aforementioned expected sanction can be understood as a cost. Because in most cases the likelihood of getting caught increases with each additional drug deal, the cost can vary. Because the expected sanction increases due to the quantity of a given sale — a larger sale increases jail time and may be

---

26. 1 J.P. BRISSOT, *THÉORIE DES LOIS CRIMINELLES* 24 (1781). I am grateful to Aaron Panner for the translation.

27. See PACKER, *supra* note 13, at 270.

28. See *supra* text accompanying notes 17-22.

29. Some of these raw materials may also have legal costs. To the extent items such as scales are considered contraband, or reliable indicia of criminal behavior, they can possess costs beyond their purchase price.

the subject of additional law enforcement — the cost varies not only by the raw number of sales, but by the number of units sold as well.<sup>30</sup>

Second, criminal behavior entails opportunity costs. The most obvious such cost is that the amount of time one spends on illegal activity detracts from the time one could spend on lawful employment.<sup>31</sup> (Note that as the returns from legal activity increase, the opportunity cost of criminal activity increases as well — a fact that may partially explain why some criminal activities tend to be performed by poorer, unemployed individuals.)

Criminal behavior also foists indirect opportunity costs, of both an economic and social nature, on individuals. One possible cost is that those who undertake criminal behavior may internalize the legal system's view of them as "outlaws" and believe that they are not suited to lawful employment. The law-abiding world, for its part, may not want to hire people who are or were engaged — or suspected to be engaged — in illegal activity.<sup>32</sup> The fact that persons only suspected of criminal activity may be affected is important, for it suggests that even when someone has not been arrested for a particular activity, the unlawfulness of that activity may still impose costs on that person.<sup>33</sup>

Such stigmatization costs can be said to be opportunity costs since the cost of criminal activity is the forfeiture of some lawful employment. This type of opportunity cost, however, is not a function of the attributes of the particular activity (such as the time spent in performing it), but of the fact of criminalization, which creates opportunity costs that would not exist if the activity had been legal.<sup>34</sup>

---

30. Therefore, to the extent that a given dealer prefers a few large risks — as opposed to a smaller number of minor risks — she will tend to make larger sales.

31. For those without lawful employment, the opportunity cost could be reduced leisure.

32. Professor Opp, for example, notes that punishment may label particular people as criminals, stigmatizing those individuals and channeling them into criminal careers. See Karl-Dieter Opp, *The Economics of Crime and the Sociology of Deviant Behaviour: A Theoretical Confrontation of Basic Propositions*, 42 KYKLOS 405, 420 (1989). And we must add to this cost of punishment the fact that imprisonment often breeds crime because of the unsavory contacts one meets while imprisoned, contacts that may reduce the cost of further criminal activity. This suggests that a penalty on first-time offenders should be low to avoid labeling effects. See Stephan M. Panther, *The Economics of Crime and Criminal Law: An Antithesis to Sociological Theories?*, 2 EUR. J.L. & ECON. 365, 375 (1995).

33. The concept of "stigma" is developed in detail in section III.B.3.

34. There are other forms of opportunity cost that may exist regardless of a legal sanction. If hard drugs were legalized tomorrow, for example, employers still might not want to hire those people suspected of part-time drug dealing. The opportunity costs arising from stigmatization, therefore, are not perfectly symmetrical to the legal sanction of that particular

Another related cost may follow in some cases. For a person who has two "jobs," one of which then becomes illegal, criminalization could either lead him to decrease his now-unlawful employment (the deterrence hypothesis) or, ironically, to decrease the time he spends in his lawful occupation. The latter option may be a reasonable reaction because his lawful employment may make it more difficult to carry out his illegal one (for example, his law-abiding coworkers could inform the police). Put in terms that are now familiar, the social and economic impact from the criminalization of an activity may engender complementary relationships with other illegal acts.<sup>35</sup>

The stigmatization opportunity cost can be said to be variable in that the more criminal the activity, the greater the stigma. Burger King is less likely to hire big-time, rather than small-fry, drug dealers. But there may also be a minimum fixed stigmatization cost of committing a crime. When neighbors derogatorily say that someone is a "drug dealer," their disapproval generally is not adjusted to take account of the fact that he sold only one vial of crack last week. In other words, stigma itself has cliffs. The fact that a person is a drug dealer may impose fixed opportunity costs on that individual. Such costs may vary from individual to individual — for example, by how susceptible the individual is to internalization, or by how many people know about the person's status as a dealer. But for some people, at least, such costs may be fixed at a minimum floor amount. Increasing the amount of criminal activity one undertakes may raise these costs, but they won't drop below that floor. This suggests that once a lawbreaker faces stigma — either from the community, individuals, or the law — she may capitalize on her sunk cost and increase her criminal activity. A proper analysis of stigma requires knowledge about the interrelationships between law and social norms, raising issues that are well beyond the confines of traditional economics and that will be taken up later in this article.<sup>36</sup>

An actor can be expected to conduct that amount of criminal activity for which the marginal gains obtained from an additional

---

activity. They are intermediated by a number of social norms — norms that have a complicated relationship with the law. See *infra* section III.B.

35. Complementary relationships may also exist because of something inherent about a particular type of crime. Auto thefts, for example, may be a complement to burglaries because stolen cars are often used in committing other types of crime. See Simon Hakim et al., *Substitution, Size Effects and the Composition of Property Crime*, 65 Soc. Sci. Q. 719, 731 (1984).

36. See *infra* section III.B.



crime are exactly equal to the marginal cost of undertaking that additional crime. The expected sanction, however, differs from a “cost” in at least one crucial sense — it is only probabilistic. If the criminal doesn’t get caught, she doesn’t incur the expected sanction. She may suffer other costs as a result of the expected sanction, even if she isn’t caught — for example, she may pay a higher price for the drugs or she may suffer stigmatization opportunity costs. But the expected sanction itself may not be imposed.

These terms — elasticities of substitution, variable costs, opportunity costs, and so on — are relatively unfamiliar concepts in contemporary criminal law. Contrast, for example, two other legal fields, torts and taxation. Ever since Calabresi and Posner’s writings,<sup>37</sup> the question of whether to impose liability on a tortfeasor has been governed in part by an analysis of whether the imposition of liability will create substitution incentives that make society worse off than it would be by letting the loss lie where it falls. If imposing strict liability on doctors for malpractice will lead to fewer working doctors — because doctors will substitute leisure or other income-producing activity — strict liability may be counterproductive.<sup>38</sup>

The notion that increased liability will have counterproductive results also brings to mind Laffer’s claim that an increase in a tax rate actually results in less tax revenue.<sup>39</sup> The intellectual pedigree for his claim depends, in part, on substitution as well. At about the same time that Calabresi and Posner were incorporating substitution into torts, tax policy received a large boost with the publication of James Mirrlees’s 1971 article.<sup>40</sup> Mirrlees, extending the landmark work of Frank Ramsey,<sup>41</sup> developed a theory of optimal taxation that explicitly incorporated substitution. In examining what income tax rate structure would maximize social welfare, Mirrlees argued, one must take account of the substitution between income-producing activity and leisure.<sup>42</sup> Optimal tax theory there-

---

37. GUIDO CALABRESI, *THE COSTS OF ACCIDENTS* (1970); RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 78, 84-88 (1st ed. 1972).

38. As I contend below, however, there are important differences between tort and criminal law. See *infra* text accompanying notes 183-200. (arguing that taste-shaping helps explain the tort/crime distinction and why substitution may be even more pronounced in criminal law).

39. See JOHN B. TAYLOR, *PRINCIPLES OF MICROECONOMICS* 483-84 fig.15.7 (1995).

40. J.A. Mirrlees, *An Exploration in the Theory of Optimum Income Taxation*, 38 *REV. ECON. STUD.* 175 (1971).

41. F.P. Ramsey, *A Contribution to the Theory of Taxation*, 37 *ECON. J.* 47 (1927).

42. For one discussion of the Mirrlees model and the substitution between income-producing activity and leisure, see Joseph Bankman & Thomas Griffith, *Social Welfare and*

fore broke away from what Edward McCaffery calls "[t]he traditional ideal" that "certain presumptively equal things . . . be taxed on nominally equal footing."<sup>43</sup>

Optimal tax theory began by examining whether taxation would impose deadweight losses. The deadweight loss, or excess burden, is the difference between the revenue raised by the government and the burden imposed on taxpayers.<sup>44</sup> The revenue raised is a product of the income effect, so any additional burden on the taxpayers is the substitution effect. If the substitution effect outweighs the income effect, then the tax imposes a deadweight loss and should not be imposed. For this reason, it is said that optimal taxation turns on the compensated, or Hicksian, elasticity — the substitution effect.<sup>45</sup>

Optimal tax theory has moved on to consider the impact of taxation on different income-producing activities. As one of the leading analyses of optimal tax describes it, the idea is that "[a]ctivities or commodities for which substitution effects are the smallest ought to be taxed more heavily."<sup>46</sup> Of course, all of this is tough to do<sup>47</sup> — but the difficulty does not mean that tax analysts should ignore optimal taxation altogether.

Both torts and taxes are areas where the government attempts to price conduct in ways that minimize certain distortions in behavior. The insights of substitution from these areas of law suggest a rather different way of thinking about deterrence in criminal law; instead of examining whether a penalty deters a particular activity, it is also important to inquire about the cost of that deterrence. Just as torts analysts ask what the price of strict liability is, and just as

---

*the Rate Structure: A New Look at Progressive Taxation*, 75 CAL. L. REV. 1905, 1962-65 (1987).

43. Edward J. McCaffery, *Taxation and the Family: A Fresh Look at Behavioral Gender Biases in the Code*, 40 UCLA L. REV. 983, 1035 (1993).

44. Alan J. Auerbach, *Excess Burden and Optimal Taxation*, in HANDBOOK OF PUBLIC ECONOMICS 61, 67 (Alan J. Auerbach & Martin Feldstein eds., 1985).

45. See Edward J. McCaffery, *Slouching Towards Equality: Gender Discrimination, Market Efficiency, and Social Change*, 103 YALE L.J. 595, 659-60 (1993).

46. Walter Hettich & Stanley Winer, *Blueprints and Pathways: The Shifting Foundations of Tax Reform*, 38 NATL. TAX J. 423, 428 (1985); see also David F. Bradford & Harvey S. Rosen, *The Optimal Taxation of Commodities and Income*, AM. ECON. ASSN. PAPERS & PROC., May 1976, at 94; McCaffery, *supra* note 43, at 1047 ("[O]ptimal taxation can be understood as an elaborate device for minimizing the impact of taxation on free market allocations."); Agnar Sandmo, *Optimal Taxation: An Introduction to the Literature*, 6 J. PUB. ECON. 37 (1976) (considering optimal tax rates by considering income distribution, work incentives, and capital formation).

47. See McCaffery, *supra* note 43, at 1038 ("Elasticities are a tricky and treacherous business; they vary among the short and long terms, are difficult to measure, incorporate numerous expectancies regarding the future, and are highly particularistic.").

the tax wonks examine whether particular changes to the tax code will change income-producing behavior, scholars and policymakers might gain a fuller appreciation of the impact of deterrence through the vehicle of substitution.<sup>48</sup>

### C. *The Example of Crack Cocaine*

I want to put this abstract discussion in concrete terms. To this end, this section brings forth some background material regarding the new mandatory-minimum sentences for crack cocaine so that readers may use such penalties as a concrete example to think about substitution theory and its variants. My goal is not empirical, as there is no adequate study of drug consumption among the general population. It is only theoretical — I want to raise the possibil-

---

48. This idea, though, does not by itself eliminate other reasons for punishment besides deterrence. At least one of these, incapacitation, is subsumed within this model. See John C. Coffee, Jr., *Does "Unlawful" Mean "Criminal"?: Reflections on the Disappearing Tort/Crime Distinction in American Law*, 71 B.U. L. REV. 193, 223-27 (1991); Steven Shavell, *Criminal Law and the Optimal Use of Nonmonetary Sanctions as a Deterrent*, 85 COLUM. L. REV. 1232, 1232 & n.1 (1985). If incapacitation strategies lead the government to put people who commit a minor offense in jail for a longer time than others who commit more harmful crimes, it will encourage commission of those more harmful crimes — and will mean that those who should be incapacitated for longer periods will not be. The story with rehabilitation and just deserts is a bit more complicated, but both could fit within the substitution model. One cannot consider adopting a penalty for rehabilitative purposes without understanding the effect that the penalty will have on a potential lawbreaker's behavior. If, as I have suggested, the penalty scheme channels the activity into an area that is even worse (for now, this may be defined in terms of harm to society, where such harm includes the probability of similar future acts because the defendant has not been rehabilitated), then the substitution effect works in much the same way. To put the point slightly differently, if rehabilitation is concerned with the use of rewards or education to change behavior, see GEORGE P. FLETCHER, *RETHINKING CRIMINAL LAW* 414 (1978), then it becomes highly important to understand whether such a system may distort behavior by shifting activity to comparatively more dangerous activity.

In addition, the methods of rehabilitation may create substitution-like results. For example, if part of rehabilitation is to train those in jail for new jobs, and one part of that is to teach people how to become good actors, the number of crimes where good actors are needed (e.g., fraud schemes) may increase. If the government tells those in jail that violence is wrong, the consequence may be an increase in theft. And if the government tells thieves that theft is wrong, it may increase violence. If sodomy is against the law and the law seeks to reduce the instances of sodomy, putting people in jail may shape tastes and lead to an increase in sodomy. The jail also permits criminals to meet others like them and may encourage future crime. See FOUCAULT, *supra* note 4, at 267 ("The prison makes possible, even encourages, the organization of a milieu of delinquents, loyal to one another, hierarchized, ready to aid and abet any future criminal act . . .").

Similarly, if the law is concerned with punishing those who are morally blameworthy, see, e.g., LEO KATZ, *BAD ACTS AND GUILTY MINDS* 27-28 (1987), then the penalty scheme might want to ensure that a set of penalties is not likely to induce a potential lawbreaker to commit a more blameworthy act. Again, the opportunity-shaping effect of criminal law may run counter to the very goals of criminal law. The one exception to this is a Kantian conception of criminal punishment that focuses only on the initial blame of an act and not on any consequences that follow from the act or its punishment. To the extent that this reasoning underlies our criminal law (and it may in cases such as murder), substitution does not change the reasoning, it only explains what the result of such thinking may be.

ity that substitution effects from the crack penalties might have led to an increase in heroin consumption.

Eleven years ago, Congress enacted mandatory-minimum sentences for the possession and dealing of crack cocaine.<sup>49</sup> Prominent lawmakers, believing that crack "make[s] people into slaves,"<sup>50</sup> made possession of five grams of crack a felony offense that carries a mandatory five-year jail term. Congress acted quickly — only a few months after they learned of crack's existence<sup>51</sup> — and rushed to pass the legislation before Election Day. The mandatory-minimum scheme Congress enacted provides that a minor crack dealer caught with five grams of crack will be in jail for at least sixty months, even on a first offense.<sup>52</sup> Additionally, the Anti-Drug Abuse Act of 1988 criminalized simple possession of a five gram amount with a mandatory-minimum sentence of sixty months in prison.<sup>53</sup> In contrast, simple possession of any quantity of any other substance by a first-time offender is a misdemeanor punished by a maximum of twelve months in prison.<sup>54</sup>

Simply by weight, the ratio of crack to heroin penalties is approximately 20:1. The Guidelines double the sentence, roughly, for each six offense levels.<sup>55</sup> For a first-time offender, the base levels are the following: 80–100 grams of heroin is punished by 51–63 months (whereas 4–5 grams of crack receives that penalty); 100–400 grams of heroin is punished by 63–78 months (whereas 5–20 grams of crack receives that penalty); 400–700 grams of heroin is punished by 78–97 months (whereas 20–35 grams of crack receives that penalty); 700–1000 grams of heroin is punished by 97–121 months (whereas 35–50 grams of crack receives that penalty), and so on.<sup>56</sup>

---

49. See 21 U.S.C. § 841(b)(1) (1994).

50. 132 CONG. REC. 26447 (1986) (statement of Sen. Chiles).

51. Crack was first mentioned on the floor of Congress on March 12, 1986. Senator Hawkins, borrowing from the latest issue of *Newsweek*, described crack as resulting in "almost instantaneous addiction, whereas if you snort coke it can take two to five years before addiction sets in." 132 CONG. REC. 4412, 4418 (1986) (internal quotation marks omitted) (quoting *Kids and Cocaine*, NEWSWEEK, Mar. 17, 1986, at 58, 58 (quoting Arnold Washton)).

52. The base level under the U.S. Sentencing Guidelines for five grams of crack is level 26. U.S. SENTENCING GUIDELINES MANUAL, § 2D1.1(c)(7) (1995) [hereinafter USSG]. A level 26 offense earns between 63 and 78 months for the first offense. Five grams of crack is equivalent to ten to fifty doses, and costs between \$225 and \$750. See U.S. SENTENCING COMM., COCAINE AND FEDERAL SENTENCING POLICY at viii (1995) [hereinafter SENTENCING COMM. REPORT]. This 200-plus page report, which is dedicated to eliminating the disparity between crack and powder cocaine, does not breathe a word about the impact of high crack sentences on heroin use (or even powder cocaine use for that matter).

53. See 21 U.S.C. § 844(a) (1994).

54. See 21 U.S.C. § 844(a) (1994).

55. See U.S. SENTENCING GUIDELINES MANUAL ch.1, pt.A, intro.

56. See U.S. SENTENCING GUIDELINES MANUAL § 2D1.1(c).

A dealer can carry 375 grams of heroin and be punished at the same level, 5 to 6½ years, as the 5-gram crack dealer.

When drug dosage is factored into the equation, the crack to heroin punishment ratio is somewhere between 80:1 to 400:1. This is so because a gram of heroin produces four to twenty times more doses than does a gram of crack.<sup>57</sup> The incentives created by this penalty structure are fairly clear. Given this sentencing scheme, it should come as no surprise that crack defendants receive the longest sentences.<sup>58</sup>

It is, however, very difficult to test whether heroin is being substituted for crack.<sup>59</sup> In part, this is because the data regarding drug

---

57. Compare *United States v. Kinder*, 64 F.3d 757, 764 n.7 (2d Cir. 1995) (Leval, J., dissenting) (observing that single dose bags of heroin are generally 20 to 50 milligrams), *cert. denied*, 116 S. Ct. 931 (1996) and OFFICE OF NATL. DRUG CONTROL POLY., HEROIN USERS IN NEW YORK, CHICAGO, AND SAN DIEGO 27 (1994) ("A bag of heroin in New York typically contains around 25 milligrams of pure heroin.") with *The U.S. Sentencing Commission and Cocaine Sentencing Policy: Hearing Before the Senate Comm. on the Judiciary*, 104th Cong. 26 (1995) (statement of Michael Goldsmith, Commissioner, U.S. Sentencing Commission) (stating that crack dosages range from 100 to 500 milligrams) and *The U.S. Sentencing Commission and Cocaine Sentencing Policy: Hearing Before the Senate Comm. on the Judiciary*, 104th Cong. 20 (1995) (statement of Judge Deanell Reece Tacha) (same) and SENTENCING COMM. REPORT, *supra* note 52, at viii (same) and Drug Enforcement Admin., U.S. Dept. of Justice, "Crack" Price Data, in DRUGS & CRIME DATA; CRACK FACTS AND FIGURES 50 (Office of Natl. Drug Control Policy Drugs & Crime Clearinghouse Packet No. PK25, 1996) (same). It is not clear to what extent these fractions control for purity.

58. The mean length of imprisonment for crack offenses during one year studied was 133.4 months, and the median was 102 months. See U.S. SENTENCING COMM., 1994 ANNUAL REPORT UNITED STATES: JANUARY 1992-MARCH 1995, at 115. Heroin offenders, on the other hand, received a mean sentence of 76.2 months and a median of 51 (both approximately half that of the crack sentences). See *id.*

59. Some softer data and reports suggest an increase in heroin consumption. See generally *Enforcement of Federal Drug Laws: Strategies and Policies of the FBI and DEA: Hearing Before the Subcomm. on Crime of the House Comm. on the Judiciary*, 104th Cong. 25 (1996) (statement of Thomas A. Constantine, Administrator, DEA) ("We are also seeing a resurgence of heroin. It's now available in more cities, and at lower prices and higher purities, than ever before in our history."); *Heroin Production and Trafficking Trends: Hearing Before the Subcomm. on Crime and Criminal Justice of the House Comm. on the Judiciary*, 103d Cong. 30 (1995) (statement of Thomas A. Constantine, Administrator, DEA); OFFICE OF NATIONAL DRUG CONTROL POLY., PULSE CHECK: NATIONAL TRENDS IN DRUG ABUSE, Spring 1995, at 7 [hereinafter PULSE CHECK 1995] (documenting increase in heroin use); *id.* at 8 ("Sources in New York, New Jersey, Delaware and Connecticut mention that some crack users are switching to snorting heroin because it is cheaper, more plentiful and less stigmatized than crack."); OFFICE OF NATL. DRUG CONTROL POLY., PULSE CHECK: NATIONAL TRENDS IN DRUG ABUSE, Dec. 1994, at 5 [hereinafter PULSE CHECK 1994] ("More teenagers and young adults nationwide are using heroin, and some are also shifting to injecting as a primary route of administration. More middle and upper-middle class people are using heroin."); Trip Gabriel, *Heroin Finds a New Market Along Cutting Edge of Style*, N.Y. TIMES, May 8, 1994, at 1; David Lipsky, *The Hard-Core Curriculum*, ROLLING STONE, Oct. 19, 1995, at 99; Alan Lupo, *Heroin Makes a Comeback in Hub*, BOSTON GLOBE, June 18, 1995, at 1; Sam Vincent Meddis, *Smack's Back*, USA TODAY, May 25, 1994, at 3A.

The common justification for the increase in heroin consumption — the rise in snortable heroin — probably does not fully explain the increase. Snorting has been around as a way to hook new users for at least two decades. See MARK LIEBERMAN, *THE DOPE BOOK* 117 (1971) (describing how pushers and "friends" turn new users on by encouraging them to

consumption is weak — for somewhat understandable reasons — and because crack was only added to the surveys in 1987. The two main annual studies are the University of Michigan's *Monitoring the Future* study and the U.S. Department of Health and Human Services's *National Household Survey on Drug Abuse*. The former primarily measures consumption of drugs among eighth, tenth, and twelfth graders, which will not yield an accurate portrait of drug consumption throughout the general population.<sup>60</sup> The latter does not concentrate on any particular age group, but its results are almost certainly skewed because its surveyors conduct their interviews while the subjects are in their homes, which creates self-reporting problems and excludes a large percentage of the relevant drug-using population (for example, the homeless).<sup>61</sup> Moreover, because the study changed its methodology after 1993, the data from earlier surveys is not generally comparable to the data from the more recent studies.

The *Monitoring the Future* survey reveals that, in general, heroin consumption has been increasing while crack consumption has been declining. The survey shows that the percentage of twelfth graders admitting to crack use in their lifetime was 5.4% in 1987 and 3.3% in 1996, whereas the percentage of twelfth graders admitting to heroin use was 1.2% in 1987 and 1.8% in 1996 (the last time heroin prevalence reached such high levels was in 1977).<sup>62</sup> Results were similar when the twelfth graders were asked about whether they consumed crack or heroin in the past year or past month.<sup>63</sup> What's

---

snort); OFFICE OF NATL. DRUG CONTROL POLY., *supra* note 57, at 17-18 (reporting that one-half of first-time users snorted heroin and that most users studied had first exposure before 1985).

60. Indeed, it might even be the case that harsh penalties create another substitution of sorts, the targeting of relatively innocent children and teenagers as customers. Because children are less likely to turn in a dealer — due to their naïveté, susceptibility to fear and coercion, or belief that it would upset their community norms — and because it is harder to use undercover buyers in such markets, harsh general penalties that do not escalate punishment for targeting youth might encourage substitution towards that market. For a similar point, see *infra* text accompanying notes 107-09 (discussing distributional effects of penalization).

61. See U.S. DEPT. OF HEALTH AND HUMAN SERVS., NATIONAL HOUSEHOLD SURVEY ON DRUG ABUSE: MAIN FINDINGS 1993, at 9 (1995) (describing these problems).

62. See OFFICE OF NATL. DRUG CONTROL POLY., HEROIN: FACTS AND FIGURES 7 tbl.2 (1997) (citing Inst. for Social Research, Univ. of Michigan, *Monitoring the Future* Study).

The study changed the heroin question in 1995 on half of the forms so that separate questions were asked for use with injection and without injection. See *id.* at 11 n.j. This change might have biased some of the results.

63. The data shows that 4.1% of twelfth graders admitted to crack use in the past year in 1987 (compared with 2.1% in 1996), and that 0.5% admitted to heroin use in 1987 (compared with 1.0% in 1996). See *id.* at 8 tbl.3. Again, before the crack penalties came into effect, the last time heroin prevalence was so high was in 1975, when 1.0% admitted to heroin use in the past year. See *id.*; see also *id.* at 9 tbl.4 (showing similar trends in thirty-day prevalence).

more, the study reveals that a significantly higher percentage of twelfth graders think heroin is "fairly easy" or "very easy" to get than at any time since the study's inception in 1975.<sup>64</sup> The *National Household Survey* shows a similar drop in crack consumption, but its heroin consumption results demonstrate a slight drop as well.<sup>65</sup>

The correlation between the increase in heroin consumption and the decrease in crack consumption may not be a coincidence. The truth is, no one has studied the connection between the two drugs, and a regression study that attempts to analyze the consumption patterns of the two drugs is needed. From a theoretical standpoint, at the very least, crack and heroin provide us with one situation in which to apply some of the refinements to our understanding of deterrence. If the new penalties induce a substitution toward heroin, some might even argue that the law is creating harm because heroin could be more dangerous than crack.<sup>66</sup>

One of the most important lessons of substitution theory is that criminalization can cause substitution even when the goods may be complements in an unregulated market. Again, what a substitute or complement is depends on one's taste and on the price difference of the goods. Cheese and pasta, for instance, may be complements when the price is low, but may be substitutes if the price for each is relatively high. If there were no penalties on heroin and crack, they could be complements or even independent goods.<sup>67</sup> They have

---

64. In 1996, 32.2% said that heroin was "easy" or "fairly easy" to get. This compares with 23.7% in 1987. The data reveals a sharp rise in perceived availability after 1987. Before 1987, the percentage generally hovered in the high teens or low twenties (1976: 18.4%; 1980: 21.2%; 1985: 21.0%). After 1987, the percentages rose dramatically (1988: 28.0%; 1989: 31.4%; 1992: 34.9%; 1995: 35.1%). See *id.* at 15 tbl.11.

65. See Jennifer A. Neisner, CRS REPORT FOR CONGRESS, HEALTH CARE FACT SHEET: ILLICIT DRUG USE IN THE U.S., Sept. 17, 1996, at 59 (reprinting preliminary data from 1996 survey) (showing that crack consumption dropped from .7% of the population in 1988 to .5% in 1995 and that heroin consumption dropped from .3% in 1988 to .2% in 1995). For a discussion of the problems with the survey's heroin data, see U.S. DEPT. OF HEALTH AND HUMAN SERVS., *supra* note 61, at 9.

66. Unlike crack, heroin is physiologically addictive. See SENTENCING COMM. REPORT, *supra* note 52, at 24. For discussions about the drugs' other effects, see Nancy P. Fieldman et al., *Dimensions of Self-Concept: A Comparison of Heroin and Cocaine Addicts*, 21 AM. J. DRUG & ALCOHOL ABUSE 315 (1995); Dorothy K. Hatsukami & Marian W. Fischman, *Crack Cocaine and Cocaine Hydrochloride: Are the Differences Myth or Reality?*, 276 JAMA 1580 (1996); James A. Inciardi & Anne E. Pottieger, *Crack-Cocaine Use and Street Crime*, 24 J. DRUG ISSUES 273, 274, 288-89 (1994); Sung-Yeon Kang et al., *Correlates of Cocaine/Crack Use Among Inner-City Incarcerated Adolescents*, 20 AM. J. DRUG & ALCOHOL ABUSE 413 (1994); Reginald G. Smart, *Crack Cocaine Use: A Review of Prevalence and Adverse Effects*, 17 AM. J. DRUG & ALCOHOL ABUSE 13 (1991).

67. Crack users often like to use heroin to moderate the intensity of the drug. See *Heroin Comes Back*, TIME, Feb. 19, 1990, at 63. One study shows that most crack users used crack before they used heroin. See Andrew Golub & Bruce D. Johnson, *Cohort Differences in Drug-Use Pathways to Crack Among Current Crack Abusers in New York City*, 21 CRIM.

vastly different effects — one intensifies experience and the other numbs it. But because both are illegal, consumers and dealers tend to group them together more than they would in a free market. Illegality defines the product market, and therefore encourages substitution. Indeed, Colombian heroin traffickers are known to “persuad[e] their established cocaine distributors to purchase and sell heroin as a condition of doing business.”<sup>68</sup> And the point is generalizable — illegality often makes activities substitutes when they wouldn't ordinarily be so.<sup>69</sup>

Even within the drug context, substitution may be at work in ways other than crack and heroin shifts. For example, the rise in so-called “designer drugs” might be explained by the criminalization of marijuana and other soft drugs. The dangers of designer drugs — many of which are made by amateur teenage chemists and are deadly — arguably dwarf the health dangers of marijuana use.<sup>70</sup> A more obvious substitution may be excessive teenage cigarette smoking and drinking, perhaps in part the result of the high price of

---

JUST. & BEHAV. 403 (1994) (finding that only 10% of crack users born after 1967 had previously injected heroin).

68. Intelligence Div., U.S. Dept. of Justice, *South American Heroin Trafficking in the United States, in DRUGS & CRIME DATA: HEROIN INFORMATION PACKET 5, 6* (Office of Natl. Drug Control Policy Drugs & Crime Clearinghouse Packet No. BC0005670H, 1995); see also PULSE CHECK 1995, *supra* note 59, at 5 (“[I]n Atlanta and other areas where heroin is becoming increasingly popular . . . crack and cocaine dealers are changing their product lines completely; that is, getting out of cocaine and into heroin.”). Of course, substitution will be even more pronounced for dealers than for consumers, since their involvement with drugs is motivated by money rather than by something inherent about the activity. Cf. *Mixed Signals on Possible Heroin Upsurge, ALCOHOLISM & DRUG ABUSE Wk.*, June 26, 1991, at 4 (“[T]he recently increased marketing of heroin is a result of former cocaine dealers moving away from the ‘violence and rip-offs’ associated with that drug to the more ‘mature and stable heroin market.’”).

69. For example, prostitutes often refuse to take jobs in legal industries after a while in the “life.” If they quit prostitution, they often take a job in another illegal industry such as drugs. See generally KATHLEEN BARRY, *THE PROSTITUTION OF SEXUALITY* 199-201, 215, 268, 305-116 (1995). Part of this substitution for one activity, prostitution, of another, such as drug dealing, may occur because the legal system labels them “criminals.” Prostitutes who want to leave internalize their criminal status, and the law-abiding outside world reinforces it by not wanting to hire them. See generally Katyal, *supra* note 15. A system of decriminalization, therefore, may prevent women from becoming drug dealers. Indeed, there may be a dual economy composed of legal jobs and illegal ones. Those in the illegal economy, such as prostitutes, fear that they cannot enter the legal one. They see themselves as “specialists” in illegal activity. The law's ability to attach the label “criminal” to an activity, therefore, may actually be *creating* crime in other areas. See *infra* section III.B.3.

70. Similarly, young kids may tend to use inhalants because they are more accessible than penalized drugs like alcohol and marijuana. See Richard L. Peck, *Are Drugs Making a Comeback with Kids?*, BEHAV. HEALTH MGMT., May-June 1994, at 12, 14 (interview with Dr. Zili Sloboda, Director, Division of Epidemiology and Prevention Research, National Institute on Drug Abuse) (finding that one in five eighth graders uses inhalants and that they “are more readily available to these teens”). This leads me to speculate that the prohibition policy on drugs could have been the greatest single boon for the cigarette and alcohol industries in history.



other types of drugs. Another somewhat less obvious form of substitution may be the increase in drug purity. Because the penalty structure uses the weight of a drug as the relevant factor in sentencing, drug dealers have compensated for the increased risk of sentences by increasing the purity of the drug.<sup>71</sup>

The substitution argument has general applicability in areas of the law besides drugs. Take, for example, penalties for white-collar crime. Substitution explains why a manager bent on making some extra money may be more likely to embezzle \$100 than she is to take a \$100 kickback because the kickback is punished at twice the sentencing level.<sup>72</sup> It might be the case that reducing kickbacks is “worth” the cost — more embezzlements — but the validity of this argument is not immediately obvious. This type of analysis can be done throughout the penalty structure. For example, a recent *New York Times* article details how the New York City mafia has shifted away from extortion and bid-rigging and toward white-collar crime because of increased enforcement.<sup>73</sup> In this article I concentrate on the simple example of crack and heroin, but readers should treat this as only one illustration of substitution.

#### D. Further Refinements

##### 1. A Model of Economic Substitution

At this point, a formal model may simplify matters. The economic substitution argument states that the expected punishment

71. For these reasons, the Supreme Court’s decision last year in *Neal v. United States*, 116 S. Ct. 763 (1996), which held that the carrier medium for LSD would be weighed in determining a mandatory-minimum sentence, might exacerbate the drive to increase purity to reduce the expected punishment. The Court eschewed such policy arguments and decided the case on statutory grounds. See 116 S. Ct. at 769.

72. Section 2B1.1 of the Sentencing Guidelines penalizes the embezzlement at a base level of 4, and section 2B4.1 penalizes commercial bribery at a base level of 8. For first-time offenders, both offense levels are punished between 0-6 months, and the guidelines do not draw a difference between the two. If the defendant has already committed two offenses, then the punishment diverges, with base level 4 punished between 0-6 months and base level 8 punished between 4-10 months. See U.S. SENTENCING GUIDELINES MANUAL ch. 5, pt. A (1995). This means that the sentencing structure likely will influence the behavior of repeat offenders. Since repeat offenders, as repeat players, tend to be some of the most sophisticated criminals in terms of knowledge of the law, the Guidelines, ironically, encourage substitution for the class they are most likely to influence. While this is admirable when the sentencing structure accurately reflects the goals of punishment, it is pernicious when the structure does not.

73. See Selwyn Raab, *Officials Say Mob is Shifting Crimes to New Industries*, N.Y. TIMES, Feb. 10, 1997, at A1 (“With their longtime control of extortion and bid-rigging rackets in New York City and New Jersey weakened by years of relentless prosecutions and by regulatory crackdowns, the region’s Mafia crime families are switching increasingly to white-collar crimes as ripe sources of plunder . . . . With a tinge of irony, law enforcement officials said that their successes in eliminating long-established rackets had caused the crime families to mine new fields for booty.”).

— determined by considering the probability of detection multiplied by sentence and fine imposed by law — for one crime may prompt actors to shift to other crimes. Assume the following:

$b$  = benefit to individual from committing the act, where  $b \geq 0$ ;

$p$  = probability of detection; and

$s$  = sentence and fine imposed by law.

Someone will commit an illegal act if the benefit to him exceeds the expected punishment ( $b > p * s$ ). If it does not, then the expected punishment of other crimes is irrelevant; so long as  $p * s > b$ , a person will not commit the act. In other words, if the expected penalty for crack exceeded its benefit to the individual ( $b_c < p_c * s_c$ ), it does not matter that the expected penalty for heroin exceeded its benefit as well ( $b_h < p_h * s_h$ ). But when the expected penalty for one activity is significantly different from its substitute, some consumers may prefer the substitute.<sup>74</sup> When  $p_c * s_c > p_h * s_h$ , substitution of  $h$  for  $c$  may occur. Under these conditions, substitution is not limited to the condition that  $b_h \geq b_c$ . The proper predictor of substitution is when  $b_h - (p_h * s_h) > b_c - (p_c * s_c)$ . When the benefits of heroin less the expected punishment of heroin exceed the benefits of crack minus its expected punishment, the marginal user will substitute heroin.

Substitution may not necessarily occur when the benefits of both activities exceed their expected punishments. If  $p_h * s_h$  and  $p_c * s_c$  are negligible — say, under a decriminalization regime — then little or no substitution may occur, depending on the degree of the substitutability. Again, the equation above holds, and substitution would occur when the difference in punishments outweighs the difference in benefits.

In addition to  $p$  and  $s$ , another variable,  $l$  — representing the expected loss to the individual from the act — must be introduced. Some crimes, particularly drug crimes, have a paradoxical quality in that they have both benefits and harms for individuals. Variables  $l$  and  $b$  can be framed in several ways. For drugs,  $l$  may be the perceived risk of addiction and the health risks involved, while  $b$  may be the attainment of euphoria. In other words, one element of “price” is the harm to the user. If crack has the potential to send users on paranoid hallucinogenic trips, that is a cost to the user. While  $b$  can incorporate  $l$ , it is useful to separate them out because they suggest different concepts. One goal of law is to shape atti-

---

74. Implicit in this model is the assumption that the benefit from the two activities is commensurate. To the extent that the activities produce such radically different forms of benefit that they cannot be evaluated on a single  $b$  axis, substitution theory does not apply.

tudes towards  $l$ .<sup>75</sup> Moreover,  $l$  may not correlate with actual harm, since it only measures perceived harm. Given this refinement, we can say that the marginal user will substitute when  $b_h - l_h - (p_h * s_h) > b_c - l_c - (p_c * s_c)$ .

Matters are complicated further because  $p$  may increase over time. A criminal known for dealing crack may have a higher  $p$  on his tenth offense than on his first. While this does not change the basic equation, it acknowledges that substitution may occur over time even when expected punishments do not initially exceed the difference in benefits. Criminals benefit from diversification just like everyone else.<sup>76</sup> The known crack dealer may shift to heroin to walk a different beat.

Another complication must be introduced:  $p * s$  may positively correlate with  $b$ . That is, when the risk and severity of punishment increases, the benefits from undertaking the crime — such as profit — may increase.<sup>77</sup> For this reason, rather stiff penalties may not deter commission of the offense, because the incentive to commit the crime increases as well. For most crimes, this is not a 1:1 replacement, so deterrence at some level will outstrip benefit.

The preceding analysis assumes that  $b_h < b_c$ , an assumption that may not play out in real life. If, however,  $b_h > b_c$ , then substitution will not occur because people will already be using heroin, assuming that the  $(p * s) + l$  difference is favorable to heroin. In reality, both happen — some prefer crack and others heroin. For those users who believe  $b_h < b_c$ , substitution is predicted. All that matters for my argument is what the marginal user will do.

## 2. Probability Versus Sentence

A strong economic objection could now be made that the model should reflect enforcement costs and that an efficient policy would

75. As I will show later, this shaping may occur regardless of whether individuals actually know that the law,  $s$ , is motivating their views of  $l$ . See *infra* section III.B.1.

76. See FRANK H. EASTERBROOK & DANIEL R. FISCHEL, *THE ECONOMIC STRUCTURE OF CORPORATE LAW* 29–30, 122–24 (1991) (discussing the benefits of diversification in financial investments).

77. Some crimes are market-based, where an increase in  $p * s$  will increase  $b$ . Others are not market-based, and an increase in  $b$  does not occur. See *infra* text accompanying notes 145–48.

impose a very high  $s$  and low  $p$  to save such costs.<sup>78</sup> In other words,  $s$  and  $p$  are *themselves* substitutes and  $p$  costs more than does  $s$ .<sup>79</sup>

We will encounter one significant rejoinder to this argument later in this article, the inverse sentencing effect, which predicts that an increase in  $s$  may decrease  $p$ .<sup>80</sup> A second response to the economic objection attacks its premise, that  $p$  and  $s$  can be independently controlled. The models Becker and his followers use do not recognize that higher sentences may create higher, not lower, enforcement costs. This complementary relationship may be the result of several different forces. For one, a high sentence on a particular activity may send a signal to the police that they should concentrate on that activity.<sup>81</sup> For another, police may, perhaps unconsciously, be influenced by a penalty scheme and come to believe that the acts with the highest penalties are the worst ones and thus the top enforcement priorities. What's more, to the extent the police endeavor to keep as many lawbreakers off the streets as possible, pursuing those criminals who commit high penalty crimes maximizes that goal since each arrest keeps a criminal in jail for a longer period of time. A somewhat separate, but important point, is that individuals may not be able to disaggregate relatively similar probabilities of detection and may perceive them to be equal. As Harsanyi has argued, people reduce complexity in decisionmaking by treating some outcomes as certain and will not distinguish between 80% and 100% probabilities of an event happening.<sup>82</sup>

---

78. As Posner observed, eighteenth-century England punished nonserious crimes with capital punishment but lowered the probability of being caught (there was no organized police force) so that the expectation of punishment was low. See POSNER, *supra* note 37, at 230. The English solution might have been appropriate for the time and place. But today, when the government has many resources at its disposal and when those resources can be selectively targeted to attack particular crimes, the English strategy seems counterproductive. See *supra* text accompanying notes 21-22 (discussing the potential ineffectiveness of the death penalty when full substitution options are considered) and *infra* text accompanying notes 90-94 (explaining the comparative effectiveness of targeted and wide-band approaches).

79. In some ways, the economic proposal here is similar to a debate in tax policy — should the government tax the small number of rich people a lot or should it tax a larger number of people less?

80. See *infra* text accompanying notes 212-22.

81. Indeed, the stiff crack penalties were created for precisely this reason. The House Subcommittee on Crime determined that mandatory minimums create the proper incentives for the Department of Justice to use its "most intense focus" on "major traffickers" and that "[o]ne of the major goals of this bill is to give greater direction to the DEA and the U.S. Attorneys on how to focus scarce law enforcement resources." H.R. REP. No. 99-845, pt. 1, at 11 (1986).

82. John C. Harsanyi, *Practical Certainty and the Acceptance of Empirical Statements*, in RECENT DEVELOPMENTS IN THE FOUNDATIONS OF UTILITY AND RISK THEORY 27, 30-34 (L. Daboni et al. eds., 1986). While this idea does not necessarily track the rational-actor assumption, it may be consistent with it because differentiating between 80% and 100% probabilities is too costly.

(Though it might be that people also reduce complexity by not distinguishing between eight- and ten-year sentences as well.) Nevertheless, the standard law and economics view that reducing the probability of detection can be compensated by increasing  $s$  may not be realistic.<sup>83</sup>

Lucian Bebchuk and Louis Kaplow have provided a third rejoinder to the economic objection.<sup>84</sup> They contend that individuals vary in their assessments of the probability of detection and that even small variations in this assessment mean that it is not optimal to set a sanction at the highest possible level. To illustrate their claim, they posit an act with a harm of 10 and a maximum sanction of 500. With a sanction level of 500, a two percent probability of detection would deter the act. And with a sanction level of 100, a ten percent probability of detection deters comparably. Imagine that half of the people overestimate the probability of detection by one percent, and half underestimate it by one percent. In the two percent regime, half will face an expected sanction of 15, and half will face one of 5. With the ten percent solution, however, half face an expected sanction of 11 and half face 9 — producing much less over and underdeterrence. If the cost of increasing the probability of detection by eight percent is less than the harm caused by under and overdeterrence, the probability of detection should be set at ten percent.

A final set of responses contends that, once substitution is considered, variations in the probability of detection may not matter as much as those in the expected sanction. This is because the actual and perceived differences in the probabilities of detection for heroin and crack may not be large enough to change behavior.<sup>85</sup> Catching a dealer is analogous to an “experience good”<sup>86</sup> — the

83. Another reason that probability of detection and sentence may positively correlate is because preventive deterrence may create a feedback loop. As citizens internalize the lore created by the punishment scheme, they will push for increased enforcement to combat those activities with the highest penalties. This idea requires incorporation of sociological theories of deterrence and is discussed *infra* Part III.

84. Lucian Arye Bebchuk & Louis Kaplow, *Optimal Sanctions When Individuals Are Imperfectly Informed About the Probability of Apprehension*, 21 J. LEGAL STUD. 365 (1992).

85. For an analysis of how actual and perceived probabilities of punishment differ, see Raaj K. Sah, *Social Osmosis and Patterns of Crime*, 99 J. POL. ECON. 1272, 1273 (1991). Sah notes that perceptions vary locally and that there is a large variance in perceptions among social groups even when the crime is narrowly defined. See also Claude Montmarquette & Marc Nerlove, *Deterrence and Delinquency: An Analysis of Individual Data*, 1 J. QUANTITATIVE CRIMINOLOGY 37 (1985) (finding that perceptions of expected punishment have an impact on crime rates).

86. An “experience good” is one in which its characteristics are known only after the good is purchased. See JEAN TIROLE, *THE THEORY OF INDUSTRIAL ORGANIZATION* 95, 106 (1988); Phillip Nelson, *Information and Consumer Behavior*, 78 J. POL. ECON. 311 (1970).

police must invest capital in busting people at a time when what the dealer carries is an unknown. The police only discover what the dealer is dealing when they catch him. As such, the differential enforcement between crack and heroin is not likely to matter much. Contrast this with larger dealers known for a particular drug — say, a Colombian drug lord — where the police know in advance that the drug lord exports cocaine. In this case, analogous to “search goods,” the differences in enforcement may matter a great deal.<sup>87</sup>

The experience goods point suggests that the probability of detection for heroin may increase when the probability of detection for crack increases. In other words, because targeting crack may necessarily result in additional heroin arrests, the two variables may be positively correlated. What's more, the *perceived* difference in enforcement may be negligible. Heroin dealers, for example, may self-enforce the laws in response to an increase in crack enforcement and refuse to deal.<sup>88</sup>

This argument, however, presupposes a lack of dealer sophistication that is probably untenable for repeat players. If, because heroin and crack enforcement are not independently controlled, an effort to arrest crack dealers results in an increase in enforcement against heroin dealers, the substitution effect may be minimized. If not, then the result of targeting particular crimes will be an increase in substitution. There are examples, however, of situations in which untargeted police crackdowns have prevented the commission of more serious crime. James Q. Wilson and George L. Kelling call this the “broken window” effect and argue that the punishment of all minor crimes will demonstrate the government's “get tough” strategy to lawbreakers.<sup>89</sup> By using an untargeted, or wide-band, approach, the government may reduce the complementarities between one crime and another. Recent empirical support for this proposition can be found in New York City, where, since 1993, the murder rate has fallen by almost forty percent and the robbery rate

---

This contrasts with a “search good,” in which the characteristics of the product are known before purchase. See TIROLE, *supra*, at 106. A dress is an example of the latter; canned food is an example of the former. I thank Ian Ayres for introducing me to this terminology.

87. To put the point in economic terms, the enforcement effort is independently controlled for small-time dealers, but not for larger dealers. Of course, because small-time dealers operate as retail merchants, their choice of wares will have an impact on the importation choices of bigger “wholesalers.”

88. This argument parallels one made in the First Amendment context, that restrictions on speech will “chill” conduct outside the technical scope of such restrictions.

89. James Q. Wilson & George L. Kelling, *Broken Windows*, ATLANTIC MONTHLY, Mar. 1982, at 29.

has fallen by over thirty percent.<sup>90</sup> These statistics might be the result of the intense campaign a new police commissioner waged to punish vigorously all “public order” offenses.<sup>91</sup> By adopting a non-selective enforcement regime, New York City sent a signal that all crimes — big or small — would be prosecuted.

The New York City approach also has its problems. To the extent that substitution suggests that expected punishments should be tailored to particular crimes, equalizing the probability of detection may not be desirable. The overall crime rate may drop, for example, but those crimes that do occur may be particularly heinous. One way to minimize such distortion is to equalize the probability of detection but to permit the expected sanction to vary. This strategy provides criminals with an incentive to refrain from more harmful activity.<sup>92</sup> But the range of sanction levels may be subject to a maximum sanction constraint — either because there is no room for increased penalty (beyond death) or because such equality in punishment would contravene other, moral, theories of punish-

---

90. See Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 VA. L. REV. 349, 367-68 n.68 (1997).

91. See William J. Bratton, *The New York City Police Department's Civil Enforcement of Quality-of-Life Crimes*, 3 J.L. & POLY. 447, 448 (1995).

92. Other economic concerns, however, may suggest that this wide-band strategy should be slightly altered to take account of victim precautions. See Gary S. Becker & George J. Stigler, *Law Enforcement, Malfeasance, and Compensation of Enforcers*, 3 J. LEGAL STUD. 1, 14 (1974) (noting that “[t]he amount of victim enforcement would be optimal if successful enforcers were paid the amount that they had suffered in damages, excluding their enforcement costs, divided by the probability that they are successful”).

If the police were to enforce all crimes equally, it would reduce the incentive for potential crime victims to take precautions. For example, if the police devoted as many resources to catching house robbers and car thieves as they did to the war on drugs, people would not have as much of an incentive to install deadbolts on their houses and to buy The Club for their vehicles — even though these devices are a much more efficient way to reduce such crimes. Alternatively, if the police devote large sums to prosecuting car thefts where the victims leave their keys in their cars, it may increase the number of those who leave their keys in the ignition. In such circumstances, it may be efficient to reduce public enforcement so that people do not act carelessly. An optimal enforcement strategy, therefore, would create enforcement differentials to take account of the possibility of self-enforcement. If done correctly, this strategy would create just as much of a wide-band effect as unselective enforcement. The only difference would be that it would be more efficient because it would recognize that private enforcement plays a role in the prevention of crime. See generally Omri Ben-Shahar & Alon Harel, *Blaming the Victim: Optimal Incentives for Private Precautions Against Crime*, 11 J.L. ECON. & ORG. 434 (1995).

An important caveat must be added: Self-enforcement may at times be counterproductive. If the government, for example, reduced street patrols in an area believing that the victims were the cheapest cost avoiders, the strategy might increase the incentive for crime because fewer people would walk the streets, thereby reducing the number of people who could (1) intervene to prevent the crime, (2) call the police, and (3) testify as an eyewitness at trial. Assessments of who is the cheapest cost avoider, therefore, should consider the positive externalities of what may initially appear to be inefficient behavior on the part of victims.

ment.<sup>93</sup> For these reasons, sometimes the only way to increase an expected penalty is to use a strategy that increases the probability of detection by increasing enforcement as the severity of the crime increases.<sup>94</sup>

The relationship between the probability of detection and sanction level has other important lessons for criminal law. As we have seen, criminalization raises the cost of an activity. A decision to make heroin illegal raises the cost to the seller and to the user. For the user, it increases cost because dealers will charge a premium for incurring legal risk. The user also bears her own costs, such as the threat of jail. But the threat of jail is probabilistic — she may or may not get caught. The price, on the other hand, is certain. An increase in monetary price may therefore create substitution even when the law does not deter consumption through its other costs — such as the risk of jail.

To the dealer, on the other hand, the threat of a jail sentence will push many potential competitors out of the market. For those criminals who are not as risk averse,<sup>95</sup> a penalty on drug dealing confers monopoly power.<sup>96</sup> Criminalization thus creates the possibility for entrepreneurship and for *true profit*.<sup>97</sup> Frank Knight contended that all profit is created by unleveraged risks, that if one could insure against an outcome then it would not beget profit.<sup>98</sup> When possession or sale of a good is criminalized, the inherent un-

---

93. Another maximum sanction constraint is the inverse sentencing effect discussed *infra* section III.B.

94. Such a policy may raise intricate ethical concerns because those who are punished are punished not only for their acts, but for those of uncaught others. As Kant wrote, “[p]unishment can never be administered merely as a means for promoting another good. . . . For one man ought never be dealt with merely as a means subservient to the purpose of another.” IMMANUEL KANT, *PHILOSOPHY OF LAW* 195 (W. Hastie trans., Augustus M. Kelley Publishers 1974) (1796); see also Johannes Andenaes, *The Morality of Deterrence*, 37 U. CHI. L. REV. 649 (1970).

95. Gary Becker's examination of the relative elasticities of crime rates to changes in the expected punishment led him to conclude that criminals must prefer risk. See Becker, *supra* note 2, at 178. But, as Michael Block and Robert Lind have argued, this violates Becker's assumption that criminals behave like noncriminals do. See Michael K. Block & Robert C. Lind, *An Economic Analysis of Crimes Punishable by Imprisonment*, 4 J. LEGAL STUD. 479, 480 (1975). Instead, it may be better to differentiate on the basis of degree of aversion to risk. For other treatments of this issue, see A. Mitchell Polinsky & Steven Shavell, *A Note on Optimal Fines When Wealth Varies Among Individuals*, 81 AM. ECON. REV. 618, 618 (1991); A. Mitchell Polinsky & Steven Shavell, *The Optimal Tradeoff Between the Probability and Magnitude of Fines*, 69 AM. ECON. REV. 880, 884-85 (1979); Stephen J. Schulhofer, *Criminal Justice Discretion as a Regulatory System*, 17 J. LEGAL STUD. 43, 48 (1988).

96. See PACKER, *supra* note 13, at 279-81.

97. By true profit I mean what Frank Knight did — the profit one gets for taking an uninsured risk. True profit is not payment for labor, return from capital, or monopoly rents. See FRANK KNIGHT, *RISK, UNCERTAINTY AND PROFIT* 19-21, 35-48 (1964).

98. See *id.* at 46-47.



certainty of being caught permits risk-takers to charge a high price for that good. Increasing the expected sentence thus creates opportunities for true profit. Dealers are high-risk takers whose markets and profits are *expanded* by criminalization.

### 3. *Two Types of Deterrence*

All of this talk about probability of detection and sanction level matters a great deal for criminal law, and I shall sketch out three implications here. Note that among drug users, the threat of jail has different meanings for different people. For lawyers, going to jail may mean a huge loss in social reputation and the loss of a legal future. For others, particularly in communities where a high percentage of the population is behind bars, the threat of jail may not matter nearly as much and may even be seen as a positive benefit by some.<sup>99</sup> In other words, the social "price" of the expected punishment may vary tremendously. Nevertheless, both might pay a dealer the same monetary "price" for a gram of heroin.<sup>100</sup>

The first implication that follows is that the law creates two different types of deterrence. For users of moderate wealth, the threat of jail may provide more of a deterrent than the monetary cost. For low-income users, the threat of jail could matter less than the monetary cost. In other words, the price of a drug may be the same to two users, but the makeup of that price is different. One user worries about the legal and social price (the sentence), the other worries about the monetary price (the dollars).<sup>101</sup> This suggests that proposals such as Becker's<sup>102</sup> that would replace incarceration of the wealthy with high fines may not provide adequate deterrence if a high fine does not carry the same stigma as does jail.

This point is not only about economic class. Rather, people who "invest" a great deal in their reputations are likely to forgo utility-producing acts that tarnish their social standing. If they do risk stigma, they will often pay a monetary premium to avoid legal risk. For others, the utility of reputation is lower. These individuals are willing to face a further loss in their reputation in exchange for monetary gain. People invest in different areas, and differences in

---

99. I defer until section III.B the situation where reputations are enhanced, rather than hurt, by crime.

100. This is not always the case. As I later contend, differentials in monetary price can be used to enhance deterrence. See *infra* section II.D.2.

101. See *supra* text accompanying notes 31-34 (noting that the opportunity cost of unlawful behavior varies by return from legal activity).

102. Becker, *supra* note 2, at 193-98.

investment strategies have distributional consequences. Some will be more likely to gamble their reputation for high profit. Once their reputation is tarnished, they may then increase their criminal activity because they can suffer no further significant damage to their status in society.<sup>103</sup>

The difference between social and monetary price may therefore explain why certain crimes are associated with certain types of people.<sup>104</sup> For those who do not invest in their reputations, more high-risk activity is predicted. And, as will be shown later, because the stigma associated with particular acts varies in different communities, the price of an activity varies by community as well. This will also produce distributional variations.

The second implication that follows from the two types of deterrence concerns sellers. Because only some sellers will take uninsured risks, criminalization confers monopoly or quasi-monopoly power on those that do.<sup>105</sup> Such individuals tend to be relatively less risk-averse than the general public, a trait that may correlate, for example, with a tendency towards violence.<sup>106</sup>

Monopoly power also has other dangerous ramifications. It may, for instance, lead to greater economies of scale in criminal operations. Such economies of scale may in turn increase distribution and marketing techniques, driving up consumption of illegal drugs and inducing more people to break the law. It may also cre-

---

103. Other forms of distributional cost are associated with penalties. If criminalization of an activity labels many people in a particular community lawbreakers, members of that community could internalize this status. See *supra* note 69 and *infra* section III.B.3. This idea links up to a central concern in Foucault's work, that punishment systems are not only negative mechanisms to repress crime but also positive procedures designed to maintain the social order. See FOUCAULT, *supra* note 4, at 24. Punishments are more than social phenomena reflecting the fundamental ethical choices of society or its juridical structure; they are part of a system of inculcating behavior. Since the ways in which penalties affect individuals vary, and these variations are in part a product of income, penalties have distributional consequences that may disadvantage certain subgroups. See *infra* text accompanying notes 109-12 and section II.D.

104. In addition, *who* commits what *type* of crime may vary due to offenders' economic situations. Furlong and Mehay found, for example, that elasticity of crime with respect to the male unemployment rate was three times the elasticity of crime with respect to the police clearance rate. William J. Furlong & Stephen L. Mehay, *Urban Law Enforcement in Canada: An Empirical Analysis*, 14 CANADIAN J. ECON. 44, 52 (1981). This finding demonstrates that deterrence of property crimes may be more a function of economics than of expected sanctions. And the type of crime may be a function of general economic effects. For example, sharp increases in gold and silver prices and the frequency of urban house burglaries are apparently closely correlated. See Isaac Ehrlich, *The Market for Offences and the Public Enforcement of Laws: An Equilibrium Analysis*, 21 BRIT. J. SOC. PSYCHOL. 107, 110 (1982).

105. See *supra* text accompanying note 97.

106. The fact that dealers have to break the law to sell may itself have the complementary effect that those who sell are less hostile to the idea of breaking laws and may indeed be more likely to break other laws, such as those prohibiting violence.

ate the potential for corruption and bribery, because as resources grow, the ability of a criminal enterprise to use nefarious means to thwart law enforcement grows as well. Economies of scale may also induce such organizations to diversify and commit other nondrug crimes — another complementary relationship. This is evident in the case of juvenile gangs, where opportunities for true profit lead individuals to band together and reinforce their tendency towards all sorts of crime.<sup>107</sup> And economies of scale may create turf wars and violence as rival dealers vie for full monopoly power.<sup>108</sup>

A third possible implication flows from the conceptualization of two levels of deterrence. Some proponents of harsh drug penalties, when faced with evidence that drug prices have dropped under these harsh penalties, might argue that the price decline evidences success of the drug war. That is, the price has dropped because *demand* has dropped. If this argument were correct, however, it would not necessarily follow that criminalization will increase social utility. After all, those who don't use drugs because of their high monetary cost are going to be more likely to use them if the price drops.<sup>109</sup> What's more, demand from the richer segment may drop even further because some rich people use the drug *because* of its high price.<sup>110</sup> Thus, to the extent that drug policies, from education to penalties, reduce the demand of some individuals for drugs by increasing the social stigma attached to drug use, they may increase demand from others by decreasing the monetary price of drugs. It might be the case that penalties only shift demand from one set of consumers to another and do not reduce it overall.

Wealth effects make a pernicious result even worse. If poorer citizens tend to be deterred by price, and richer communities by legal punishment, a price decrease would tend to increase drug consumption in poor communities. A "just say no" media campaign could, for example, have a greater educational effect in richer communities because stigma matters more to members of such communities. The resulting drop in demand in those communities would mean that the price would drop, thus increasing consumption of the

---

107. See Martin Sanchez Jankowski, *Getting into Gangs*, in *DEVIANCE: THE INTER-ACTIONIST PERSPECTIVE* 279, 281-83 (Earl Rubington & Martin S. Weinberg eds., 6th ed. 1996) (arguing that joining a gang is "a rational decision to maximize self-interest" due to opportunities for profit and that gangs reinforce criminal tendencies).

108. Removing one dealer when many others are around will not cement a monopoly, but when only a few dealers exist in a market, the return from violence is much higher.

109. As more people use the drug, the price may rise again, which will reduce the rate at which price-sensitive users will consume the drug, and so the cycle repeats itself.

110. See generally Gary S. Becker, *A Note on Restaurant Pricing and Other Examples of Social Influences on Price*, 99 J. POL. ECON. 1109, 1109-11 (1991).

drug in poorer communities where price is more of a deterrent. The distributional effects of the drug war, therefore, could be traumatic.<sup>111</sup>

It is even possible that criminalization may increase the number of overall users by creating a more than 1:1 replacement from rich to poor communities. As we have already seen, some drug policies may unintentionally shift consumption from wealthy communities to poor communities due to disparities in the cost of stigma. Even if stigma has the same cost to rich and poor users, amplified distributional inequalities could still occur because consumption by the richer demand segment has fallen. That richer segment dictates the market price, and as demand falls in that segment, prices drop accordingly. Sellers then can make up for the revenue loss through sales to another, poorer demand segment.

The effect of this shift in demand segments is not only that the poor consume what the rich previously did; it is also that the poor may consume even more of the drug than what the rich previously consumed. Because users in poorer communities have less income, each user may not be able to consume at the same high levels of their richer counterparts. As a result, dealers may be forced to deal to greater numbers of people in an attempt to make up for the revenue shortfall.<sup>112</sup> Apart from the obvious distributional consequences, the reduction in price may therefore lead to a greater number of users than existed before the policy, because the price drops more than compensated for the loss in demand. The overall units consumed may not increase under harsh penalties or education, but the composition of the consumption will change; a greater number of people will now consume the drug (though each user

---

111. This is particularly so if poorer individuals tend to lack representation in the political process, so that richer individuals are simply externalizing their problem onto those who do not have a voice.

112. The idea here is that there are two discrete demand segments, one composed of richer consumers and the other of poorer consumers. When the richer segment has high demand, dealers may prefer a Nieman-Marcus-style strategy (low volume; high price). But when demand among that segment drops, for whatever reason, dealers may have to shift to a K-Mart-style strategy (high volume; low price). Many more men, for example, seem to be wearing Calvin Klein suits today as compared to ten years ago. Despite this increase in consumption, prices of the suits have fallen dramatically. While there may be many reasons for this increase in consumption of Calvin Klein suits, it is not unlikely that many people are wearing them only because the price has fallen. And we can speculate that one reason the price fell is because the richer demand segment dropped a great deal due to newcomers like Giorgio Armani. This meant that Klein probably had to shift to a K-Mart-style strategy to maintain profits. I do not know if any of this actually happened, but the fact it could have happened is enough to illustrate the general point. On distinctions among forms of consumption, see *infra* text accompanying notes 166-67 (outlining two different types of consumption, horizontal consumption — the overall number of people consuming a drug — and vertical consumption — the amount of a drug each user consumes).

may consume less of it) because the monetary price deterrent has weakened.

## II. SOME NORMATIVE RESULTS

### A. *The New Benefit of the Sentencing Guidelines*

It is now time to begin thinking about some of the normative conclusions that might follow from substitution effects.<sup>113</sup> I begin with one rather simple insight, that a uniform system of sentencing will be superior to an ad hoc one insofar as such a uniform system can better respond to substitution effects. By showing how legislators can channel crimes into activities that are more desirable, a deterrence perspective grounded in substitution theory may provide an important justification for the Guidelines. A uniform system can shape behavior in ways that ad hoc sentencing legislation often cannot. In particular, piecemeal approaches typically will not address substitution and complementarity correctly. This idea supports thinking about criminal law somewhat like the way in which we analyze questions in taxation.<sup>114</sup>

As the 1986 Congressional debate on crack illustrates, Congress creates penalty schemes with tunnel-vision.<sup>115</sup> No member of Congress wants to be cast in the position of stalling a vote to penalize what the media has dubbed the latest threat to Americans. It is very difficult for legislators in such situations to think about the impact of a penalty on other behavior. A sentencing commission that proposes sentences on a whole host of crimes at once, on the other hand, may be in a better position to design penalties that provide the correct incentives. The only treatment of this issue, to my knowledge, is Jeffrey Standen's.<sup>116</sup> Standen advocates the creation of specific charging guidelines so that variance in prosecutorial discretion will not undermine marginal deterrence. Standen's point is well taken, but is only one instance of a much larger notion. For the substitution theorist, sentencing guidelines themselves, insofar as they provide concrete knowledge about expected penalties, lead

---

113. This Part retains the assumption that preferences are fixed and that people act rationally on those preferences. See *supra* text accompanying note 2. Once complications regarding taste shaping are introduced, however, it might be impossible to determine the deterrence and substitution effects of a penalty. Therefore, the policy suggestions in this part are offered only under the economic assumptions that are specified — assumptions that may prove to be unrealistic. See *infra* Part III.

114. See *supra* text accompanying notes 39-48.

115. See *supra* note 51 and accompanying text.

116. Jeffrey Standen, *Plea Bargaining in the Shadow of the Guidelines*, 81 CAL. L. REV. 1471, 1523-25 (1993).

those who commit crimes to commit less serious ones from the entire panoply of possible crimes. The substitution theorist seeks to do to criminal sentencing what some optimal tax advocates did to the field of taxation. The critics of the Guidelines who contend that sentencing is so difficult that it should be done by judges<sup>117</sup> prove the point. It is precisely because such determinations are so difficult, and the effects of penalization so varied, that one body should set uniform guidelines.

A uniform increase in a sentence will also minimize so-called "spillover" effects. Without uniformity, increasing the expected sanction of a crime in one area may simply lead criminals to commit the same crime in other areas — a kind of geographic substitution.<sup>118</sup> Interjurisdictional variations in sentences and the probability of detection therefore mitigate the effectiveness of increasing the expected sanction in one locale — unless the goal of punishment is to externalize the crime onto neighboring jurisdictions.<sup>119</sup> By adopting a set of uniform penalties, geographic substitution is constrained.

I do not want to defend the entire project of the Sentencing Guidelines, but only to suggest that substitution provides an argument in favor of them. Substitution sometimes may shed light on

117. See, e.g., Jose A. Cabranes, *Incoherent Sentencing Guidelines*, WALL ST. J., Aug. 28, 1992, at A11; Jose A. Cabranes, *Sentencing Guidelines: A Dismal Failure*, N.Y. L.J., Feb. 11, 1992, at 2.

118. See Samuel Cameron, *The Economics of Crime Deterrence: A Survey of Theory and Evidence*, 41 KYKLOS 301, 303 (1988); Furlong & Mehay, *supra* note 102; Simon Hakim et al., *Interjurisdictional Spillover of Crime and Police Expenditure*, 55 LAND ECON. 200 (1979). The latter study found interjurisdictional spillover for property crimes such as auto theft and breaking and entering, but not for violent crimes. See *id.* at 211. This also provides further evidence of the nonmarket crime hypothesis.

For one example of geographic substitution, see THE NNICC REPORT 1994: THE SUPPLY OF ILLEGAL DRUGS TO THE UNITED STATES:

A combination of factors — saturated markets, low prices, violent competition, and/or effective police pressure in major urban areas — has forced some crack distribution groups, in conjunction with local gangs, to develop new markets in smaller towns and rural areas. . . . The more established distribution groups are crisscrossing the nation to find new markets . . . .

U.S. DEPT. OF JUSTICE, DRUG ENFORCEMENT AGENCY, THE NNICC REPORT 1994: THE SUPPLY OF ILLEGAL DRUGS TO THE U.S. at 7 (1995). Geographic substitution can also be driven not only by high enforcement, but also by high sentences. If one jurisdiction decides to punish an activity at a higher level than another, people may commit the crime in the latter jurisdiction.

119. Hakim et al. also show how a one dollar per capita increase in police expenditures in neighboring communities will generate a five cent per capita increase in police expenditures in a given community. See Hakim et al., *supra* note 118, at 211. The authors assume that this is evidence of geographic substitution, but this phenomenon may also reflect the influence of norms. When a community decides to devote resources to prevent a type of crime, other communities may follow suit, not only because of the possible externalization of crime, but because they follow the example set by their neighbors.

the problems with the Guidelines. For example, the Sentencing Guidelines contain a multiplier effect that has heretofore gone unnoticed. As previously mentioned, the Guidelines provide that a sentence will approximately double for each six levels.<sup>120</sup> The Guidelines' enhancements and reductions, however, work by increasing and decreasing the level of an offense. This can create enormous disparities in sentencing that enhance substitution. Imagine an individual considering becoming a leader in a drug gang. The Guidelines provide for a four-level increase from the base offense level if the convicted person is a leader.<sup>121</sup> The way the Guidelines work, four levels mean something drastically different when the offense is a low-level one than it does when the offense is a high-level one. Leaving aside leadership enhancements, one kilo of crack yields a 188–235 month sentence and one kilo of heroin yields 121–151 months.<sup>122</sup> The four level enhancement increases a crack sentence to 292–365 months — an average *increase* of about ten years. The enhancement increases a heroin sentence, however, to 188–235 months, a much smaller increase of about six years. Under these circumstances — and assuming that profit did not increase as a result of the criminalization<sup>123</sup> — the individual would have to be a fool to lead a crack distribution ring. By pegging the enhancement to the underlying crime, the Guidelines create grave sentencing disparities that both undermine their purpose and magnify the substitution problem.

Yet why should the government not prevent this substitution problem by penalizing every crime at the highest possible level? The “Singapore strategy” has something to commend it — if the government punished everything with death, then people would commit fewer crimes. As we have seen, however, substitution explains why those who then committed crimes would commit them on a greater scale than they would under a proportionality regime.<sup>124</sup> This strategy also has another flaw. Leaving aside both

---

120. See *supra* note 55 and accompanying text.

121. See U.S. SENTENCING GUIDELINES MANUAL § 3B1.1 (1995).

122. See U.S. SENTENCING GUIDELINES MANUAL § 2D1.1 (1995).

123. See *supra* text accompanying note 97.

124. This is true, actually, in Singapore. As substitution theory would predict, some reports show that serious crimes, such as murder and rape, are disproportionately high in Singapore. See Philip Bowring, *In Singapore, Unusual Law Doesn't Bring About Unusual Order*, INTL. HERALD TRIB., Apr. 20, 1994, at 6, available in LEXIS, News Library, Non-US File. (“Despite all these stringent laws, Singapore is not extraordinarily crime-free. . . . [I]t has far more murders per head than Australia or South Korea, both highly urbanized societies.”); Philip Sherwell, *Island Makes No Apology for Draconian Penal Code*, DAILY TELEGRAPH (LONDON), Mar. 18, 1995, at 21; see also *Singapore Crime Drops, But Some Sexual Offences Up*, DEUTSCHE PRESSE-AGENTUR, Mar. 21, 1997, available in LEXIS, News Li-

the obvious moral problems and the excessive cost involved — which was Becker's reason for rejecting such an approach<sup>125</sup> — the strategy may create over-deterrence.<sup>126</sup> A death penalty for those who cause fatal car accidents will mean that fewer people will drive. High penalties on criminal activity have chilling effects that radiate beyond the core of the conduct for which the penalty is sought and thus may inhibit lawful and utility-producing behavior. It will not generally be feasible to set penalties at extreme levels without entailing some significant costs.<sup>127</sup>

A difficult question now presents itself: How does one determine what activities are harmful enough to merit a particular sanction? That is, how are sentencing commissioners to determine appropriate punishment?

### B. *Redefining Optimal Penalties*

Substitution yields three important normative conclusions about optimality. First, no matter what utilitarian criteria a penalty scheme uses to determine what crimes merit what penalties, the punishment should fit the crime vis-à-vis other crimes to avoid perverse consequences.<sup>128</sup> Second, penalties cannot be set only in light of the harm an undesirable act causes, but also must take account of substitution effects. Third, a different deterrence strategy should be used for those crimes where ease of substitution and elasticity of offenders are high — I call such crimes *market-based*.

A government that punishes very harmful activity lightly and less harmful activity strongly will encourage the commission of the harmful activity, thus imposing a net harm on society. What is "very harmful" should depend not only upon the particular attributes of the act, but also on whether the crime will induce complementary relationships with other crimes, either directly (for

---

brary, DPA File; *Singapore's Crime Rate Up First Time in Seven Years*, AGENCE FRANCE-PRESSE, Sept. 11, 1996, available in 1996 WL 12136162.

125. See Becker, *supra* note 2, at 180-81.

126. See *supra* text accompanying notes 37-38 (explaining how high tort liability will have chilling effects).

127. These costs do not, in general, mean that crimes should go unpunished on deterrence grounds. Just as Foucault realized that the prison, though it produces delinquency, may channel that delinquency into "politically or economically less dangerous" forms, see FOUCAULT, *supra* note 4, at 277, sentencing structures can reduce the most harmful crimes by encouraging the commission of less harmful acts.

128. Specifically with respect to rehabilitation, policies aimed at reconditioning offenders may have counter-deterrent effects. If an individual knows that he is going to get job training in jail, the price of committing a crime is reduced to him. And if the penal system aims to remove the stigma on those who serve time — for example, through measures to integrate ex-convicts into society — the ex ante price of the crime is reduced even further.



example, drug dealing leading to money laundering) or indirectly (for example, by destabilizing the rule of law)<sup>129</sup> — and whether it will set an example and lead others to commit the crime through complementarity.<sup>130</sup>

Imagine that the purpose of the criminal law is to deter physical harm to others. Behavior *A* imposes one unit of physical harm on a victim, and Behavior *B* imposes two units of such harm. Common sense would dictate that Behavior *B* should be punished at a level somewhat above Behavior *A*.<sup>131</sup> If not, and a person's taste for *B* is higher than for *A*, the government will induce substitution effects.<sup>132</sup>

But now we need a fuller account of what substitution — at least the narrow economic version of the substitution argument — does to punishment calculations. Here, substitution should be viewed as a *cost* of punishment. Every time an activity is sanctioned, the sanction has the potential to create substitution effects. Some of those effects will ripple downwards — inducing people to commit less harmful crimes or no crime at all if lawful behavior is substituted. Some of these effects, as we have seen, can work to increase harm to society.<sup>133</sup>

129. As the Italian theorist Gaetano Filangieri put it: “The proportion between the penalty and the quality of the offence is determined by the influence that the violation of the pact has on the social order.” 4 GAETANO FILANGIERI, *LA SCIENCE DE LA LÉGISLATION* 214 (J.A. Gauvin Gallois trans., Chez Cuchet 1786) (1784), quoted in FOUCAULT, *supra* note 4, at 92-93.

130. Crimes that shock the conscience may actually not be as harmful as those that everyone tolerates. Cf. FOUCAULT, *supra* note 4, at 93 (“There is a scarcity of great crimes, on the other hand, there is the danger that everyday offences may multiply.”).

131. Determining what is “twice” as bad as something else can often be quite difficult. For example, George Stigler, in his classic article, argued that the theft of \$1000 is more than twice as harmful as the theft of \$500. Stigler, *supra* note 3, at 529. This is far from clear, for two \$500 thefts may in some circumstances impose more societal harm than one theft of \$1000. For example, those who have \$1000 to steal may be better able to bear a loss than those who have only \$500. Not only may the wealthy be better able to bear a loss, they also may have opportunities for self-enforcement that the poor do not — for example, locks, guards, and alarm systems. Moreover, the occurrence of two thefts means two investigations and two prosecutions — all of which will soak up additional government money. Larger thefts also may be easier to detect and may therefore merit a slightly lower penalty because the probability of getting caught is higher.

132. This type of analysis can be replicated under other theories of punishment as well (e.g., if behavior *B* was twice as immoral as behavior *A*, then the government should punish *B* at a level higher than *A*).

133. This point gives rise to another question: Should the government base the sanction on the harm a particular activity causes to society or on the gain an individual reaps from that activity? In many cases, there may not be much of a difference since the gain to the criminal may be directly influenced, or precisely equal to, the amount of harm caused to society. But there are occasions in which this equivalence may not hold. Selling one vial of crack, for example, may result in a five dollar profit to an individual, but a much larger harm to society. Criminal insider-trading, on the other hand, may provide an individual with a one million

Of course, this all turns on how actors weigh crimes. If someone believes in ignoring all expected sanctions, then it will not matter at what level a penalty is set. Put differently, even with a maximal sanction, some crime will still occur because some will have a taste for it or because it maximizes their self-interest. The amount of this residual crime will be, in part, a function of whether the particular activity is a market-based one or whether it is not.<sup>134</sup> A host of other factors, from biological and sociological influences to "over-estimation" of the crime's particular benefits, may also influence its prevalence.<sup>135</sup>

None of this undermines criminal deterrence once its function is properly understood. Torts occur despite a system of tort liability, yet our faith in deterrence is not completely shaken.<sup>136</sup> Similarly, the fact that crimes still occur is not a per se argument against the validity of deterrence. It is just plain silly to think that criminal law can deter all crime, but it is just as absurd to think that the mere fact that some crime occurs under any system of sanctions means that deterrence is bankrupt.<sup>137</sup>

In fact, even if you think that deterrence is generally bankrupt, you might still be concerned about substitution. After all, under

---

dollar gain, but arguably may not impose any cost to society. See EASTERBROOK & FISCHEL, *supra* note 76, at 345-48.

The answer in tort and contract law primarily has been that liability should be based on the harm the victim suffers. The Sentencing Guidelines, however, provide that the size of a fine should "reflect the seriousness of the offense (including the harm or loss to the victim and the gain to the defendant)." U.S. SENTENCING GUIDELINES MANUAL § 5.E1.2.(d)(1) (1995). As Polinsky & Shavell have shown in the tort context, gain-based liability is problematic when legal error is present because such liability will fail to deter many socially undesirable acts. A. Mitchell Polinsky & Steven Shavell, *Should Liability Be Based on the Harm to the Victim or the Gain to the Injurer?*, 10 J.L. ECON. & ORG. 427, 428-29 (1994). This is because an underestimation of gain will lead individuals to commit an undesirable act. But because harm generally exceeds gain by a large amount, errors in assessing harm are not as likely to create the same incentives, and so harm-based liability is generally preferable. See *id.*

An analysis of the proper type of liability based on the government's errors in determining gains would seem to apply in the criminal context as well, although it may be harder to determine social harm and individual gain. The Polinsky & Shavell theory, in addition, does not incorporate perception errors in processing information about gain and harm from expected punishment. The extent to which these errors affect the analysis is beyond the scope of this article.

134. See *infra* text accompanying notes 145-49.

135. On possible biological influences on crime, see ADRIAN RAINE, *THE PSYCHOPATHOLOGY OF CRIME: CRIMINAL BEHAVIOR AS A CLINICAL DISORDER* (1993); Roger D. Masters, *Environmental Pollution, Neurotoxicity, and Violent Crime*, in *ASPECTS OF ENVIRONMENTAL TOXICOLOGY* (J. Rose ed., forthcoming 1997).

136. See WILLIAM M. LANDES & RICHARD A. POSNER, *ECONOMIC STRUCTURE OF TORT LAW* 72-73 (1987); STEVEN SHAVELL, *THE ECONOMIC ANALYSIS OF ACCIDENT LAW* 83-84 (1987).

137. For a similar argument in tort law, see Gary T. Schwartz, *Reality in the Economic Analysis of Tort Law: Does Tort Law Really Deter?*, 42 UCLA L. REV. 377 (1994).

the conventional inquiry, a penalty is said to deter only when it induces law-abiding behavior. But this is the ultimate form of substitution, and thus, the most difficult. To grasp this point, it might be helpful to think once again about the salt example.<sup>138</sup> A high salt price could mean that people forgo salt altogether. But other spices might be substituted as a response to an increase in the price of salt. Indeed, other spices might be a closer substitute than nothing at all. The same idea might hold in criminal deterrence. A penalty system might not cause a criminal to forgo all illegal activity, but it might influence which illegal activity a lawbreaker picks, just as the pricing system influences which spice to use.

So even if criminals act “irrationally” by taking a risk where the expected sanction exceeds the expected benefit, they may still take the course of action with the lower cost-to-benefit ratio. A person bent on defrauding clients may not calculate the costs and benefits of his activity properly (or even at all), but may still decide that defrauding the government is an additional risk that he does not want to take. Deterrence still matters — even for those who will not be deterred from committing a crime. To modify Herrnstein’s suggestion, you may not be able to stop the husband from throwing the plates at the wall, but you may be able to stop him from throwing the china.<sup>139</sup>

The above point suggests, in part, a modification of the category economists call “nondeterrable crime.” Nondeterrable crime refers to situations where people so enjoy the crime they commit that no criminal sanction will induce them to change their behavior.<sup>140</sup> Instead of looking to one particular crime, substitution focuses the question on situations where sanctions will not be able to deter crime without creating a shift to another crime.

This raises the question of whether the existence of substitution effects inexorably leads to the conclusion that penalty should correlate with harm. If, hypothetically, penalty structures perfectly correlate with harmfulness of activities, does substitution have any importance? The answer is “yes.” Even under the unrealistic assumption that the government could devise perfect penalties, substitution suggests that punishment calculations cannot be premised only on the social harm that an activity creates. Even if the harmfulness of particular crimes can be calculated correctly, re-

---

138. See *supra* text accompanying note 10.

139. See *supra* text accompanying note 14.

140. See Kenneth G. Dau-Schmidt, *An Economic Analysis of the Criminal Law as a Preference-Shaping Policy*, 1990 DUKE L.J. 1, 4 n.21.

sources may be misdeployed due to the narrow focus on social harm. Instead, sentencing commissioners must examine the likely reaction of would-be criminals to a particular sentence.<sup>141</sup> If punishing crack means that offenders will substitute heroin, the law is not doing much good. The converse is also true: substitution creates the possibility of the government creating less harmful substitutes to dangerous activity and, thus, increasing social welfare through deterrence.

This curse blessing follows from the realistic and uncontroversial proposition that no system of penalties, however sophisticated, will deter all crime. As Beccaria wrote, “[i]t is impossible to foresee all the mischiefs which arise from the universal struggle of the human emotions.”<sup>142</sup> There are at least three reasons for this: individuals vary in their responsiveness to sanctions; some crimes are more amenable to deterrence than others; and some crimes become more attractive as the punishment increases. As to the first, Alexander and Staub have written of the distinction between the “neurotic criminals” who “cannot help” committing crimes and the “normal criminals” who are deterred by the “fear of painful consequences.”<sup>143</sup> While this binary division is artificial, the distinction helps clarify the behavioral extremes.<sup>144</sup>

---

141. The commissioners appear not to have considered such questions. See, e.g., Symposium, *Equality Versus Discretion in Sentencing*, 26 AM. CRIM. L. REV. 1813, 1822 (1989) (comments of Stephen Breyer) (“A theft of \$10,000 typically led to a prison term, but a fraud of \$10,000 led more often to probation and no prison, or to lighter prison sentences. What is the difference between fraud and theft? Fraud is larceny by trick, which is a form of theft. Why should there be this discrepancy? The Commission could think of no reason. Therefore, it raised sentences for white collar crimes . . .”).

142. BECCARIA, *supra* note 4, at 19.

143. FRANZ ALEXANDER & HUGO STAUB, *THE CRIMINAL, THE JUDGE, AND THE PUBLIC: A PSYCHOLOGICAL ANALYSIS* 209-11 (Gregory Zilboorg & Franz Alexander trans., The Free Press 2d ed. 1956) (1931). Wechsler & Michael do much the same when they write of “the class of non-deterable persons.” Herbert Wechsler & Jerome Michael, *A Rationale of the Law of Homicide*, pt. 1, 37 COLUM. L. REV. 701, 759 (1937).

144. Differences in behavior can occur for a variety of reasons. See FRANKLIN E. ZIMRING & GORDON J. HAWKINS, *DETERRENCE: THE LEGAL THREAT IN CRIME CONTROL* 96-141 (1973). First, as previously discussed, people vary in their desire for risk. Those who are risk-preferrers are more likely to commit crimes than those who are not. Second, those who are “future-dwellers” may be more worried about future punishment than are “present-dwellers,” who find the immediate gains from crime more attractive. See MARGERIE FRY, *ARMS OF THE LAW* 82-84 (1951). People do not all discount the future with the same yardstick, and differences in discount rates may explain some criminal behavior. Third, people may vary in their degrees of optimism and pessimism about the likelihood of punishment. As Daniel Claster found, delinquent boys believed that they had a “magical immunity” mechanism that protected them from punishment. Daniel S. Claster, *Comparison of Risk Perception Between Delinquents and Non-Delinquents*, 58 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 80, 84 (1967). Fourth, those who are more impulsive may be less amenable to deterrence than are those who are more deliberative. See Johannes Andenaes, *Deterrence and Specific Offenses*, 38 U. CHI. L. REV. 537, 539 (1971) (stating the “old proposition” that “carefully

In addition, one might expect particular crimes to be more easily deterred than others. This variation is not only the result of substantial differences between particular crimes, but also of the type of people drawn to particular forms of lawbreaking. If the highest risk-takers are drawn to dealing drugs, while the lowest risk-takers are drawn to petty shoplifting, this stratification will have an impact on the law's ability to deter each type of crime. The particular attributes of crimes also will affect deterrence. A crime that is highly impulse-oriented — say, murder in the heat of the moment — may be more difficult to deter than one that is less impulsive — say, bank fraud.

On the other hand, crimes that are impulsive are paradoxically ones where deterrence sometimes can function well in a systemic sense, because they are *nonmarket crimes*. The law's ability to deter depends not only on the characteristics of a particular crime, but also on whether a market for that crime exists.<sup>145</sup> Criminalization of some activities, such as drug dealing, results in an increased benefit to drug dealers — a higher price — that creates an opportunity for true profit.<sup>146</sup> And if the elasticity of the supply of offenders is high, punishing one criminal will not accomplish much because another one will simply take his place.<sup>147</sup>

It is theoretically possible that such a replacement may be at a level higher than 1:1. If some of the resources expended by a drug ring, for example, are devoted to fighting rival rings, busting a drug ring may actually increase the potential for crime by permitting a surviving ring to earmark more of its resources for selling and less for physical protection. Such a result is by no means inevitable. Just as modern warfare between nation-states has partially moved away from emphasis on the physical to the economic, drug rings in some markets may fight each other through lower prices instead of

---

planned acts are more easily deterred than those that result from a sudden, emotional impulse”).

145. This distinction is, in some ways, similar to one used by Professor Packer. See PACKER, *supra* note 13, at 281 (separating penalties on some crimes, such as those regarding performance of then-illegal abortions, from others, such as bank robbery, because “[t]he harder we work to make the sale of abortions risky, the higher we drive the price that makes the risk worthwhile” yet “[w]e do not make bank robbery more attractive by punishing the bank robbers whom we manage to catch. The potential gain is unaffected by the offsetting risk of punishment”). Packer's distinction is valid, but his example is not. A penalty on bank robbery would, I imagine, create a market for skilled bank robbers who could avoid detection, for strong-armed criminals who could prevail in a confrontation with bank security, and so on. But the basic point, that the gain from some — not all — crimes increases with expected punishment, remains valid.

146. See *supra* note 97.

147. See Posner, *supra* note 20, at 1216-17.

violence. If so, taking these dealers out of circulation will have the price effect sought by the government — higher prices via reduced competition among dealers. On the other hand, such a policy could result in greater economies of scale and monopolization, with their own attendant problems.<sup>148</sup> Such problems could include a further increase in crime — because removing some dealers increases the price of the drug and hence the returns, particularly given possible monopolization.

Another way that punishment may increase crime is if those caught tend to be less competent and are replaced by more adept criminals.<sup>149</sup> If members of a drug ring are simply replaced when they are caught, then the deterrent effect is virtually nil. And if a person is caught because she happened to be the worst drug dealer, the addition of a more cunning replacement to the ring would increase, rather than reduce, crime. Law enforcement would thus serve as a sorting mechanism to aid criminals in their search for competent associates.

For “nonmarket” crimes, however, a market does not develop. Target-specific violence, such as most rape and murder, are not crimes where a market can be expected to develop. Other examples include assault, battery, and child abuse. An increase in the penalty for these crimes will not generally result in an increased benefit to the criminal.<sup>150</sup> No market is created, and a new actor does not replace a punished criminal — the elasticity of the supply of offenders is low. Deterrence, therefore, may be a more effective means of preventing some nonmarket, rather than market, offenses.

Because people and crimes vary in their amenability to deterrence, it is too large an abstraction to speak of a system of perfect penalties that will deter all crime. It is helpful, instead, to think about whether a penalty structure will provide the most deterrence at the cheapest cost. To do this, the penalty structure cannot be calibrated to those who are truly nondeterrable persons. If the law has no impact on someone, then it makes little sense to elevate a penalty to the highest possible level in a futile attempt to deter him.

---

148. See *supra* text accompanying notes 105-08.

149. See Stephen S. Brier & Stephen E. Fienberg, *Recent Econometric Modelling of Crime and Punishment: Support for the Deterrence Hypothesis?*, in *INDICATORS OF CRIME AND CRIMINAL JUSTICE: QUANTITATIVE STUDIES* 82, 83 (Stephen E. Fienberg & Albert J. Reiss, Jr. eds., Bureau of Justice Statistics, U.S. Dept. of Justice, NCJ-62349, 1980); Phillip J. Cook, *Punishment and Crime: A Critique of Current Findings Concerning the Preventive Effects of Punishment*, *LAW & CONTEMP. PROBS.*, Winter 1977, at 164.

150. Sometimes it can, though. Imagine contract murder or a gang that “rewards” the member who risks killing the leader of a rival gang.

That will not deter him and will simply create substitution problems for everyone else.<sup>151</sup> This situation, naturally, is rare. More common is the case of a nondeterrable person who will commit some crime, but has not decided which particular crime to commit. Substitution will provide an answer for these cases.

Imagine that Abe, who lost his shirt in the stock market, will either commit embezzlement or take kickbacks to make up for his loss. He is a nondeterrable person in the sense that he will commit one of those two crimes, but this fact should not blind us to the possibility that the legal system may affect which of the two he eventually commits. Imposing a higher penalty on one may induce him to do the other. On the other hand, if Ben is committed to assaulting an ex-girlfriend, the possibility for deterrence is not as great because he is a nondeterrable person committed to a nondeterrable crime.<sup>152</sup>

People do not, however, generally face a choice between two crimes like embezzlement or bank fraud. Instead, they have to contend with a smorgasbord of options. In view of the range of these choices, policymakers must gauge the likely response of individuals to the enactment of particular punishments. If the only benefit of a penalty is to lead criminals to substitute a slightly less harmful offense, then the penalty is not accomplishing a great deal. In other words, calculations about optimality must be refined to include substitution, and discussions about punishment cannot revolve simply around the harm of the activity being punished. Crack may be the worst, or second worst, thing around, but spending enforcement dollars on it may not confer much benefit because of substitution effects. Even if the penalty for heroin dealing or use is raised to compensate for the substitution, other illegal substitutes are available. Even after penalties are raised on all illegal substitutes, there are legal drugs that individuals abuse — model glue, nitrous oxide, and so on. On the other hand, a high penalty on a nonmarket crime such as murder may not create as much substitution. If murder is an impulse-driven and target-specific crime, then allocating enforcement and punishment to prevent it is optimal.

---

151. Incapacitation, however, may be a powerful solution for the undeterrable. Lifetime imprisonment is, of course, the ultimate specific deterrent. But again, the *ex ante* perspective requires that such imprisonment be used only in circumstances where the benefit of incarcerating the undeterrable for a given crime outweighs the substitution, framing, and other effects that would ensue.

152. See *supra* text accompanying notes 136-40 (arguing that there is a range of crimes that can be deterred even for the nondeterrable) and *infra* text accompanying notes 155-60 (discussing the concept of *Y*-optimality, that high prices on activities can have income effects that deter other, unrelated acts).

More concretely, if my earlier conjecture proves wrong and the conventional wisdom — that crimes like rape and assault have no substitutes — is true, a clear way to enhance optimality will exist. The unstitutability of nonmarket crimes would mean that enforcement resources should be allocated to prevent them. This idea is best understood through a numerical example. Suppose that an instance of crack use imposes five disutility units on society. Imagine that a penalty is devised that prevents most crack consumption. If the effect of the penalty is to encourage people to use, say, morphine, which imposes three disutility units, society profits by a reduction of two disutility units.<sup>153</sup> Now take a crime where substitution is unlikely. Imagine, for example, that the crime is simple assault and that simple assault imposes three disutility units on society. Enforcing laws against simple assault, then, may yield more of a benefit to society (three units) than enforcing laws against crack (two units). This is so even if we assume that crack dealing is *worse* for society than simple assault.

Substitutes can be a benefit, however, and not just a burden. The possibility for substitution to facilitate compliance with the law should not be overlooked. If ready substitutes exist, the substitution effect may *enhance* deterrence. Which crime an individual picks depends on her desire for the crimes relative to their expected punishments.

The simple assault penalty discussed above may not yield much of a benefit if it will not deter many people. Deterrence may be unlikely if the actor has no alternative to — that is, no substitute for — the simple assault. To examine the effectiveness of a penalty, it is not enough to look only at substitution effects. A discount for those whom the penalty does not deter is also needed. This suggests a new role for the criminal law — creating substitutes. If the law manufactured less harmful alternatives to activity that it wants to discourage, deterrence might be improved. This is, after all, a similar idea to what is at work in the concept of rehabilitation — the teaching of useful behavior to replace criminal proclivities. But the substitution-creating role extends well beyond the traditional goals of rehabilitation and suggests that the law should, at times, permit harmful activity in order to reduce activity that is even more harmful.

---

153. Of course, if the effect of the penalty is to encourage people to do something that imposes more than five units of harm, then the penalty is counterproductive.



To illustrate, let us return to our crack-morphine hypothetical. Suppose that crack and morphine were punished in accordance with the discussion above and that marijuana consumption causes one unit of harm. If marijuana use were legalized, it might encourage those who would have used morphine to use marijuana instead, because the legal price of marijuana is much lower than that of morphine, even though the benefits may not be as great. In so doing, it would yield a net benefit of two units to society. In addition, this marijuana scheme may also help draw people away from crack. In its absence, it was assumed that everyone would substitute morphine for crack. This assumption was, of course, unrealistic. Because some users may find that the benefits of crack minus its expected penalty outweigh the benefits of morphine minus its expected penalty, some will stick with crack. But when the expected penalty of marijuana is dropped to zero — or, factoring in the monetary cost, an amount close to zero — marijuana may be more of a substitute than morphine.<sup>154</sup>

The possibility for law-enhancing substitution shows that these concepts do not always portend a bleak future for criminal law. By manipulating penalties to prevent those crimes that are both harmful and preventable at a low cost, substitution opens new doors for the government.

### C. *The Income Effect and Y-Optimality*

Until now, we have been thinking about one half of the equation, the substitution effect. But the discussion would be incomplete without considering its other half, the income effect. Specifically, the income effect, when applied to criminal law, suggests that a high price — whether monetary or legal — for one crime may *decrease* the commission of other crimes. We can see this phenomenon most easily in consumption crimes, like drug use. If the monetary price of heroin increases as a result of greater penalties, those who continue to consume heroin likely will reduce their consumption of other illegal drugs. This is because heroin users' real income drops when the heroin price rises, so they do not have the purchasing power they did before the price increase.

---

154. In addition, as section III.C argues, research in cognitive psychology shows that the addition of a salient alternative, such as marijuana, may increase the perceived benefits of morphine vis-à-vis crack. By adding a viable third option — marijuana — extremeness aversion may lead people to pick morphine instead of the more extreme crack option. See *infra* text accompanying notes 264-72.

Although the income effect of a heroin price increase — a loss of purchasing power — will tend to reduce a heroin user's consumption of other narcotics, the substitution effect may work to encourage heroin users to switch to other drugs. That is, the high heroin price decreases the relative price of heroin substitutes and heroin users may try to maintain their old purchasing power by substituting other drugs. Therefore, we can see both substitution and income effects at play. If constraints on substitution exist due to, for example, an intense preference for heroin or an unavailability of substitutes, a price increase may have an income effect that reduces consumption of other drugs. This yields the interesting result that a heroin penalty actually might not deter heroin consumption but might deter other crimes.

Other income effects are possible as well. Consider a few quick examples. First, if criminals have a target income, deterrence measures that reduce the return from each instance of criminal activity may increase the amount of crime required to maintain the target. Second, if criminals were punished by fines, it could increase crime because some individuals will resort to crime to pay their fines or to insure against getting caught and imprisoned in the future.<sup>155</sup> Third, the income effect may mean that penalties for consumption crimes may have exactly the opposite effect from that intended — criminalization may expand consumption. In particular, criminalizing drugs may increase their price, which in turn may encourage users to hook new consumers so that they can deal to them to generate additional income with which to buy drugs. Criminalization could, similarly, lead to theft and other crimes.<sup>156</sup>

The most exciting application of the income effect in criminal law regards what I term *Y-Optimality*.<sup>157</sup> The previous section considered what should be done to a person with a taste for criminal activity that is so high that no penalty will deter the person. Ordinarily, criminal lawyers throw up their hands and believe that deterrence has no role to play against such “undeterrables.” But, as the tea and foot powder example shows,<sup>158</sup> when someone has a strong taste for a good — that is, when there is no viable substitute — a price increase in that good will decrease consumption of other goods. A *Y-Strategy* uses this insight to enhance deterrence of several crimes through a penalty for one crime. Like an increase in the

---

155. See Cameron, *supra* note 118, at 303-04.

156. See *infra* text accompanying note 167.

157. The *Y* effect is economic shorthand for the income effect.

158. See *supra* text accompanying note 1.

price of tea that reduces consumption of foot powder, an increase in a penalty for one crime may have an income effect that reduces the commission of other crimes — *even when the penalty increase does not deter commission of that particular crime.*

Y-optimality occurs when the penalty for an activity is set at the highest level where the activity will still occur.<sup>159</sup> Under these circumstances, a penalty — even though it does nothing to reduce the occurrence of the targeted activity — may reduce the commission of other crimes.

For example, imagine that Jim commits one arson and one robbery per week and that the penalty is five years imprisonment for each of these crimes. Suppose that the government decides to increase the arson penalty to seven years. The additional punishment may not reduce Jim's proclivity to commit arson, but may reduce his commission of robbery. For this to be true, the benefits from arson must be such that Jim would not simply substitute additional robberies to replace his previous arsons.<sup>160</sup> On the other hand, Jim may no longer commit robbery, but may decide instead to commit an additional smaller crime — for example, a mugging. A high *arson* penalty is thus *detering robbery* and *increasing mugging* — two crimes traditionally believed to have nothing to do with arson.

Put slightly differently, if the penalty is set at a level where substitution will not occur and the marginal user will still commit the crime, the income effect suggests that commission of other crimes may drop. Because the dominant paradigm in criminal law ignores the interrelationships between crimes and their punishments, Y-Optimality has been unrecognized and ignored. This is one more example of how the key break from marginal deterrence towards the economics of substitution takes criminal law far beyond its current assumptions.

## D. *Income Effects and Drugs*

### 1. *Giffen Goods*

Income effects are also important for other reasons. For example, one might think that if heroin is becoming a substitute for

---

159. Readers will note that this necessarily means that the taste for the activity must be larger than the utility derived from any substitutes.

160. For example, this may occur if Jim earned twice as much committing arson as robbery. Of course, Jim's opportunities to commit arson must be limited by the risk of nondiversified criminal activity or by some other external constraint, such as the number of willing "clients" who pay him to torch buildings, for example. In the absence of such constraints, Jim would be committing only arson in the first place.

crack, heroin penalties should be increased to compensate for the disparity. But yet again, simple and elegant conclusions do not always follow from complicated phenomena. If government policies increase the price of heroin to the point where substitution would not occur, it may wind up increasing drug consumption in a certain segment of the population as a result of income effects. To understand this ironic — albeit perhaps only hypothetical — effect, it is necessary to introduce the concept of Giffen goods.

In general, when a consumer's real income rises, the consumer will increase the quantity of each good that she purchases. Conversely, when her real income declines, the quantity of each good bought decreases. When the price of a good increases, however, both income and substitution effects are at play. The substitution component will lead a consumer to substitute the good that has become relatively cheaper for the one that has become relatively more expensive.<sup>161</sup> The income component means that a price change will change purchasing power, decreasing the individual's overall consumption accordingly.

Put a slightly different way, when the price of a good rises, substitution effects work together with income effects to produce a decline in demand for that good. Income effects, however, are sometimes hard to predict. An impoverished immigrant from France who is used to eating pork and who prospers in America will either eat more pork or substitute steak — depending on his taste. Pork suggests an exception to the general rule about the impact price changes have on demand. The consumption of some goods, inferior goods, relates inversely to income. Examples might include rot-gut whiskey and second-hand clothing. As an individual's income declines he consumes more of these goods.<sup>162</sup>

For some inferior goods, a price increase could have the counterintuitive effect of *increasing* consumption. These goods are called Giffen goods. For such a good to exist, the income effect must be large enough to outweigh the substitution effect and the good must be inferior. The classic example, used by Victorian economist Robert Giffen, concerned the Irish potato blight.<sup>163</sup> Before

---

161. In other words, the individual will reallocate her consumption in order to equate the marginal rate of substitution between the two goods to the new price ratio. The marginal rate of substitution reflects how much of one good an individual is willing to give up in return for one more unit of another good. Because a consumer's relative preference for each good may depend on how many units of each good the consumer purchases, the marginal rate of substitution varies depending on the quantity of the good purchased.

162. See NICHOLSON, *supra* note 1, at 85.

163. *Id.* at 95-96.

the blight, the typical Irish family ate a diet consisting mostly of cheap potatoes and a little bit of meat, which was considerably more expensive than potatoes. When the blight hit, potato prices rose and the real income of the Irish plummeted. Had potatoes been superior goods, one would expect that the consumption of potatoes would have decreased because their price increased. But Giffen observed that potato consumption increased; the Irish ate *more* potatoes than they did before the blight, because the high potato price reduced income to the point where meat had become prohibitively expensive. Because there were no available substitutes for meat besides potatoes, the price increase led the Irish to become more dependent on potatoes than they were previously. The positive income effect of the potato price increase had dwarfed the negative substitution effect.<sup>164</sup> There are, therefore, three types of goods: superior goods, where a price increase in the good will reduce consumption of the good; inferior goods, where a decrease in income will increase consumption of the good; and Giffen goods, where an increase in the price of a good will increase consumption of the good.

I want to raise tentatively the possibility that heroin might be a Giffen good and that a large penalty on heroin may work havoc similar to the Irish potato famine. Some evidence suggests that consumption of heroin may be inversely proportional to income.<sup>165</sup> By increasing its price, a high heroin penalty reduces real income. Such reductions in real income may *create* drug use by making users even poorer than they were before the penalty. This poverty, in turn, further contributes to the need for additional heroin.

To my knowledge, no one has described heroin as a Giffen good. But the description may be appropriate for those users who are addicted. We all know about cycles of drug addiction where a user becomes so impoverished that she loses hope for her future. Drugs become the user's only escape. If drug prices were low enough that

---

164. Whether this actually happened in Ireland is a matter of some historical debate, although, as Stigler points out, the fact that it could have happened this way suffices to illustrate the theoretical point. See GEORGE J. STIGLER, *THE THEORY OF PRICE* 24, 62-63 (3d ed. 1966).

165. For example, one study measured the income of 201 heroin users in Harlem during 1978-1982. It found that "daily" users had an average noncriminal income of \$5607 and an average criminal income of \$11,974. Less frequent "regular" users had an average noncriminal income of \$5,897 and an average criminal income of \$11,203. "Irregular" users had an average noncriminal income of \$5,952 and the lowest average criminal income, \$4,451. See BRUCE D. JOHNSON ET AL., *TAKING CARE OF BUSINESS: THE ECONOMICS OF CRIME BY HEROIN ABUSERS* 81, 89 (1985). Unfortunately, the study, even if it were representative of heroin users in general, cannot be determinative since it is unclear whether the loss in income is caused by increasing heroin use, or vice-versa.

users would not be impoverished, consumption might actually decline.<sup>166</sup> Addicted users — those who are least likely to opt for a substitute — may increase their consumption of heroin as their misery, exacerbated by their decline in real income, increases.

A related point concerns types of drug consumption. The main goal of current drug penalties has been to prevent what I suggest we label *horizontal consumption* — the number of people using drugs — and not to prevent *vertical consumption* — the amount of a drug that a user consumes. But drug penalties, even if they decrease horizontal consumption, may increase vertical consumption. If heroin is a Giffen good, high heroin penalties increase vertical consumption through income effects.

It is not immediately obvious that trading horizontal consumption for vertical consumption promotes social welfare. High drug dosages contribute to fatal overdoses, antisocial behavior, loss of productivity, broken homes, and all of the other things that are commonly used as justifications for punishing drug use. It is not clear, however, whether consumption at lower dosages creates these problems. Drug war proponents thus far have succeeded in identifying situations in which high vertical consumption creates problems — but these vertical consumption problems might be, if the Giffen goods argument is correct, an *effect* of the drug war itself. Of course, horizontal consumption is worrisome because of its potential to lead to increased vertical consumption. But the strength of the complementary relationship between the two cannot be assumed and requires empirical support, particularly when costs are high for unaddicted users.

A focus on income effects also suggests that one of the perennial justifications for the drug war, that drugs increase crime, may actually draw the causality arrows in the wrong direction. Because of increases in the price of drugs, addicts may turn to mugging and other crimes so that they can continue to afford the drugs.<sup>167</sup> High

---

166. An interesting question, beyond the scope of this article, concerns whether and to what extent substitution varies by income. For example, rich persons may use more cocaine than poor ones because there are more activities that richer folks can do while on the drug, such as trade stocks, that are not available to less wealthy individuals. In other words, wealth effects might make certain types of crime more or less common for certain socioeconomic groups. Again, my point is not to suggest policy implications but simply to show how income combines with substitution in interesting ways. For other examples of potential interactions with wealth effects, see *supra* text accompanying notes 31-34 (regarding interactions between criminal behavior and the workplace) and text accompanying notes 109-11 (discussing how penalization might create demographic shifts in the drug-taking population because of the two types of deterrence).

167. See JOHNSON ET AL., *supra* note 165, at 4-5 (citing studies showing that heroin addicts resort to theft, robbery, and other crime to support their habits); *id.* at 186 (presenting

drug prices may also lead to other negative externalities. For example, if drug prices increase, drug consumers may be forced, because of the decline in real income, to begin dealing drugs to support their habit.<sup>168</sup> Some of this increase in dealing will take the form of exposing neophytes to the drug, thereby increasing horizontal consumption. Again, it is possible that high drug prices increase, rather than reduce, the drug use.<sup>169</sup>

This observation, even if true, does not require America to call off the drug war. After all, drugs can only be Giffen goods for those who use them. Unlike potatoes, drugs are not a necessity to those who are not addicted.<sup>170</sup> Proponents of the drug war are correct to argue that high monetary prices could deter new users from taking heroin, but they err in extrapolating their analysis to addicts. This suggests that an optimal deterrence strategy would prevent new users from being exposed to the drug and simultaneously ensure that prices are low for those who are already addicted. In economic terms, the goal is to create price discrimination.

---

data showing that as users decrease heroin use they decrease the number of robberies they commit); Inciardi & Pottieger, *supra* note 66, at 289 ("Just as use of more heroin is associated with commission of more crimes, increased levels of crack use are also clearly correlated with a greater level of crime involvement."); see also PACKER, *supra* note 13, at 332.

168. One important study of Miami crack users found that the 114 men interviewed engaged in 683,595 drug sales in a period of ninety days. The 84 women interviewed had engaged in 258,849 drug sales, with 94% reporting that they had sold drugs during the time period. See Inciardi & Pottieger, *supra* note 66, at 280; see also JOHNSON ET AL., *supra* note 165, at 183 ("In addition to drug sales, the heroin business encompasses a repertoire of roles by which heroin abusers distribute drugs to others like themselves. Steering, touting, coping, and other roles . . . are central to the lives of heroin abusers. Performing these low-level roles is the bread-and-butter of their drug sustenance."); Edward Preble & John J. Casey, Jr., *Taking Care of Business — The Heroin User's Life on the Street*, 4 INTL. J. ADDICTIONS 1, 3, 21 (1969) ("The cost of heroin today is so high and the quality so poor that the street user must become totally involved in an economic career. . . . [T]he street heroin user is an active, busy person, preoccupied primarily with the economic necessities of maintaining his real income — heroin."). Many studies conclude that addicts are full-time or part-time dealers. See, e.g., John C. Ball et al., *The Criminality of Heroin Addicts When Addicted and When Off Opiates*, in THE DRUGS-CRIME CONNECTION 39 (James A. Inciardi ed., 1981); John C. Ball et al., *The Day-to-Day Criminality of Heroin Addicts in Baltimore — A Study in the Continuity of Offence Rates*, 12 DRUG & ALCOHOL DEPENDENCE 119 (1983).

169. An increase in crime and drug dealing may be able to offset the Giffen good effect. If addicts are able to make up for their lost income through crime and drug dealing, their vertical consumption may not rise because their income remains steady. If they do not have opportunities to make up their lost income, however, they may increase their consumption in response to a drop in their real incomes. It may be that some users will maintain their real income levels by turning to theft or other crimes, while other users will be subject to a Giffen goods phenomenon.

170. As William Chambliss has written, "evidence, then, suggests that drug addiction, like murder, is relatively unaffected by the threat or the imposition of punishment." William J. Chambliss, *Types of Deviance and the Effectiveness of Legal Sanctions*, 1967 WIS. L. REV. 703, 708.

## 2. *Price Discrimination*

Price discrimination occurs when a firm with a monopoly in two markets charges different prices for its output in those markets.<sup>171</sup> Generally speaking, price discrimination seems fanciful. We all know that the existence of price differences leads to arbitrage — wholesalers will buy on the cheap and sell high — which in turn reduces the price differential. But for an illegal narcotics market, price discrimination could be created through redesigning law enforcement.

The government can manufacture price discrimination by creating a strategy that makes it very costly for dealers to sell to new customers. Recall that new customers who approach a dealer are met with a great deal of scrutiny, as the dealer tries to determine whether the customer is an informant or a police officer.<sup>172</sup> One way to raise dealers' information costs is to have undercover police pose as new users. By increasing the likelihood that a street sale to an unknown customer will lead to jail time, dealers may be forced to concentrate on known customers.<sup>173</sup> Street sales to unknown customers will be at higher prices to compensate for the increased risk borne by dealers.<sup>174</sup>

Similarly, there are ways to raise the search costs for buyers. First, undercover police could pose as dealers and arrest people attempting to buy drugs from them. If the scheme deploys these "dealers" in areas known for drug dealing, it will increase the search costs a new consumer would incur. A well-designed strategy could easily avoid entrapment challenges — for example, ensure that would-be buyers initiate the transactions. Moreover, because new users risk the possibility that their purchase will be low-quality or even fake heroin, impugning the reputation of a particular area as a market for good heroin may be an effective strategy to increase search costs. Second, the police could target open-air drug markets

---

171. See NICHOLSON, *supra* note 1, at 339.

172. See Mark H. Moore, *Policies to Achieve Discrimination on the Effective Price of Heroin*, AM. ECON. REV., May 1973, at 270.

173. This also suggests that using addicts as informants may be counterproductive because it may undermine the price discrimination regime.

174. Heavy heroin users account for the bulk of money spent on heroin. See OFFICE OF NATL. DRUG CONTROL POLY., WHAT AMERICA'S USERS SPEND ON ILLEGAL DRUGS, 1988-1993, at 15 n.21 (1995) (finding that 87% of heroin expenditures are made by heavy users). This provides some incentive to dealers to concentrate on heavy users. When law enforcement targets sales to new users, dealers deciding whether to expose new users may confront collective action problems — turning a new user on is a large risk to the particular dealer, while the "benefit" of the additional use (a larger market) is susceptible to competition from other dealers.



to drive crime indoors. Again, this will raise search costs because it will be more difficult for new buyers to locate dealers.<sup>175</sup>

A third strategy may be to create an inverse scale of penalties so that a first offense carries a greater penalty than does a second or third. This would help create price discrimination by making the monetary price a neophyte pays reflect a premium for security that the dealer is not a cop. It would also reduce complementary effects by preventing people from turning into new users and going down a path that may lead to additional crime. This “three strikes you’re safe” rule inverts standard thinking about criminal penalties — a first offense is, in every area of the law, the offense with the lowest penalty. It is, yet again, a new way of thinking about criminal penalties that flows from the application of substitution economics to criminal law.<sup>176</sup>

These policies would concentrate law enforcement efforts on deterring people from becoming users. In so doing, the price for neophytes would rise, but the price for known customers would drop. A strategy that created dual pricing for heroin may be a powerful way to address the Giffen goods problem. If the price for addicts and repeat users were low, the factors that drive people to increase their consumption of the drug may evaporate. A low price for addicts could also reduce the other harmful effects of high drug prices, such as the increase in crime rates. Because the elasticity of demand for new users is relatively high — because they are not yet addicted — a price increase for new users may reduce heroin consumption.<sup>177</sup>

Finally, note that there is one other potential solution to the Giffen goods problem that does not rely on price discrimination. We have already seen how the law’s substitution-creating ability might be able to reduce substitution effects.<sup>178</sup> If the law created viable substitutes for those who are addicted — for example, methadone treatment — this could prevent the Giffen goods problem as well. Because of the availability of a substitute, the income effect

---

175. See Moore, *supra* note 172, at 271-73.

176. As I shall suggest later, however, when sociological insights are superimposed onto this economic analysis, it is not so clear that “three strikes you’re safe” makes sense. See *infra* section III.B.3 (arguing that stigmatization from penalization may increase crime). Moreover, “three strikes you’re safe” might encourage those who do engage in some criminal activity to engage in a lot to reach “safety.”

177. Some price discrimination probably already exists because of the monopolization effect identified earlier. See *supra* text accompanying note 97 (explaining true profit incentive). My point is only that price discrimination can be increased even further through manipulation of existing legal machinery.

178. See *supra* text accompanying notes 153-54.

from a price rise in heroin would be tempered, and the increase in consumption would not occur.

I am not advocating price discrimination in the heroin markets; we do not yet know whether heroin is a Giffen good. As I have suggested, this is a plausible hypothesis. But policymakers and scholars, too steeped in the assumption that a higher penalty reduces all evil in the world, have not yet asked the question. The reconceptualization of criminal deterrence I advocate allows policymakers to focus on these questions by pointing out that substitution is a consequence of higher penalties, and that higher penalties could have pernicious income effects.<sup>179</sup>

### III. SHAPING SUBSTITUTION

Thus far, my analysis has assumed that preferences are fixed — that is, that one's taste for a particular activity is not influenced by its cost. This assumption corresponds to a cornerstone concept in traditional economics, that preferences are given and that opportunity is constrained by cost. This has yielded a tidy understanding of deterrence — increasing the cost of an activity will prompt less people to do it and more people to do something else. But this model seems beset with problems. First, isn't the degree of substitution much narrower? Why will someone who has learned how to break safes switch to arson? Second, aren't social norms much more important than criminal law in reducing crime? Third, substitution theory itself specifies no mechanism by which people choose between options, except by what is "rational" — a theory that is at odds with much of what we know about human behavior. So how do people really select crimes? Fourth, where does the impulse to commit crime originate, at least for those crimes that are not obviously rational?

Each of the above problems suggests that substitution must incorporate knowledge about preferences. Instead of taking them as given, a theory of how preferences adapt to the legal topography may yield a fuller account of the likely reaction to particular penalties. This approach, marvelously put forth by Jon Elster, stresses the adjustment of wants to possibilities.<sup>180</sup> In Elster's view, what

---

179. Substitution may show that drug penalties have other perverse effects. In addition to the standard libertarian line that criminalization leads to underground markets and violence, one could plausibly argue that prohibition has driven a shift to newer, more dangerous drugs. Indeed, one may speculate as to whether the rise of crack in the first place may have been caused by efforts to clamp down on cocaine.

180. JON ELSTER, *SOUR GRAPES: STUDIES IN THE SUBVERSION OF RATIONALITY* 25 (1983).

people want may be a product of what they can get. The fox does not want sour grapes because it cannot get them.<sup>181</sup> This idea is central to the work of Cass Sunstein and those who read him.<sup>182</sup> In the sections that follow, I go beyond this familiar legal exegesis to explain *how* preferences adapt to criminal punishment. My description of the taste-shaping mechanism of punishment will take readers through sociological theories of control as well as psychological analyses of preferences to analyze the societal reaction to a set of penalties. The upshot is that the simple and elegant economic model of deterrence needs reexamination.

### A. *The Preference-Shaping View of Criminal Law*

Consider a limitation of the model in Part I, namely, that many crimes do not seem to have substitutes. Will an arsonist really become a white-collar criminal simply because of an increase in the penalty for arson? Sometimes criminal activity has sunk costs — the equivalent of learning a new profession — and criminals may not be able to transfer their skills to other areas. By pointing out that the incentive to substitute different crimes to compensate for a penalty increase will not always be strong, the terms of the objection beg a more fundamental question. After all, is the premise that we have been laboring under, that criminal law focuses on reducing *incentives* to commit crime, really a precise statement of what criminal law is? If so, why is it any different than tort law? Why are governments everywhere failing to follow Becker's classic solution to crime — take away all the wealth of those who commit crime but do not imprison them?<sup>183</sup>

The answer to such questions lies in the nature of criminal law. Classical economics understands punishment as designed to constrain the opportunity for crime.<sup>184</sup> But some laws serve a more complicated function in that they act as exemplars of good behavior. They seek, in effect, to influence *tastes* or *preferences*<sup>185</sup> rather than to constrain opportunity. Outside criminal law, for example,

---

181. *See id.* at 109.

182. Cass R. Sunstein, *Legal Interference with Private Preferences*, 53 U. CHI. L. REV. 1129, 1147 (1986).

183. Becker, *supra* note 2, at 193-98. In the recent O.J. Simpson trials, however, two California juries adopted Becker's solution de facto. The extent to which civil sanctions may deter murder and other crimes is a fascinating question that I leave for another day.

184. Cf. T.A. Marschak, *On the Study of Taste Changing Policies*, AM. ECON. REV., May 1977, at 386 (describing classical economics' hostility to theories based on endogenous preferences).

185. I use "taste" and "preference" interchangeably in this article.

the minimum wage, an opportunity-shaping law, may induce employers to extract additional concessions from employees while sexual harassment laws, which are taste-shaping, may not.<sup>186</sup>

Criminal law may be said to set itself apart from many other areas of the law because it concentrates less on constraining opportunities and more on shaping tastes. As Professor Dau-Schmidt has argued so well, the criminal law seeks to reduce crime by minimizing the taste for it through imprisonment rather than fines.<sup>187</sup> Criminal law, unlike torts, aims not to price conduct, but to change people's preferences by requiring imprisonment and accompanying social ostracism, regardless of the willingness of an individual to pay a fine.<sup>188</sup> No matter what amount of money a person will spend, the law says, the individual cannot commit the act. Criminal law is a powerful illustration of Bob Cover's depiction of the state as an entity that uses its coercive power to suppress alternative conceptions of law and rules of behavior.<sup>189</sup>

Both tort and criminal law impact upon people's preferences, but only the strong-arm of the criminal law focuses on *intent*, which can be understood as a proxy for taste. Two people who commit the same act, killing someone, are treated in different ways under the law because of their internal preference for the act. Criminal law, then, can shape taste by punishing undesirable preferences, not simply undesirable acts. The intent requirement ensures that the commission of an act alone is not enough to impose criminal punishment.<sup>190</sup> As the next section will argue, criminal law not only shapes the taste of those it punishes, its mechanisms of punishment aim to shape *societal* tastes through stigma, lore, and a host of other methods. My work therefore takes Dau-Schmidt's suggestion further in two important respects. First, it attempts to provide an account of the devices criminal law uses to shape preferences, not only for those who are punished, but for society at large. Second,

---

186. See Sunstein, *supra* note 182, at 1137. Sunstein also stresses the "endowment effect" — the tendency for people to value the things they own more than they do when the same things are owned by others. See *id.* at 1150. As he recognizes, however, this is really a "generalization" of the adaptive preference phenomenon. See *id.* at 1151.

187. See Dau-Schmidt, *supra* note 140. See generally Sunstein, *supra* note 182, at 1146 ("It is hard to imagine a preference not shaped in part by legal arrangements.")

188. One can now begin to understand why Becker's proposal to replace imprisonment with fines to the extent possible seemed so implausible when it was stated. Becker assumed that the point of criminal law was to reduce opportunity, not to shift preference. Then, however, criminal law looks no different than tort law. Becker's assumption that preferences are static harmonized with modern economics, but not with criminal law.

189. Robert M. Cover, *The Supreme Court, 1982 Term — Foreword: Nomos and Narrative*, 97 HARV. L. REV. 4, 14-17, 46-47, 53-54 (1983).

190. See Dau-Schmidt, *supra* note 140, at 25-27.

my argument will center on the reverse of the Dau-Schmidt hypothesis by showing how penalties shape tastes *towards* crime under certain conditions.

Taste shaping explains why potential substitutes for a particular crime may radiate well beyond crimes with similar characteristics to the original one. A penalty structure has importance not only for current criminals, but for future ones. By shaping preferences, a penalty structure therefore may encourage people to choose certain lines of "work" — much the way that opportunities for profit guide many college students and channel them into certain jobs. A drastic change in the profitability of a career, say law, may not induce those who are already lawyers to switch to another career, but it may prevent many students from becoming lawyers in future years — not only because of profit, but because people internalize the belief that they do not "want" to become lawyers. In a similar way, the point about substitution must be taken not only in terms of what current criminals will do, but what future criminals will do. Substitution theory expands on this insight by demonstrating that the law may shape tastes in perverse ways. If the penalty for consumption of one drug induces people to use other drugs, for example, these penalties are altering those people's desires.<sup>191</sup>

Substitution theory also explains why the law cannot easily move back to its starting point once punishment has shaped tastes. Once people are addicted to heroin, for example, it may be difficult to undo the damage that the criminal law has wrought. Or, to move this discussion out of the drug context to demonstrate the broader applicability of substitution, envision the cost when the law calls someone a felon. This action reduces the cost of future criminal activity to that person — it costs much less to be called a felon again once you have already been called one — an instance of complementarity at work.<sup>192</sup> And, even more problematical, it may cre-

---

191. For example, to the extent that the price-substitution effect induces marginal users to shift to heroin, the taste for heroin is increased — the user is addicted.

192. Complementarity also exists, however, in letting crime go unpunished. Those who get away with a crime may be more likely to commit that crime, or others, again. Consider, in this regard, the words of Marion Le Goff, a famous bandit leader in Brittany in the mid-eighteenth century, who, the perhaps apocryphal story goes, cried out from the scaffold:

Fathers and mothers who hear me now, watch over your children and teach them well; in my childhood I was a liar and good-for-nothing; I began by stealing a small six-liard knife. . . [.] Then I robbed pedlars and cattle dealers; finally, I led a robber band and that is why I am here. Tell all this to your children and let it be an example to them.

FOUCAULT, *supra* note 4, at 66.

ate a taste for the criminal behavior as preferences adapt to the social milieu of an underworld where lawbreaking is common.<sup>193</sup>

Punishment, therefore, can breed crime by increasing the taste for it *and* by reducing the “price” of future criminal activity.<sup>194</sup> The former effect is particularly pronounced when the law punishes behavior that is commonplace in particular communities. When the law is out of step with the norms in a given community, and it labels “ordinary” citizens lawbreakers, the ability of the law to shape the behavior of that community is compromised. The individual lawbreaker — whose reputation may even have been enhanced by the skirmish with the police — is not as likely to heed a law-following message as a resident of a community where the law tracks its norms.

The natural demographic areas to apply such thoughts are areas of inner cities where drug laws appear to have created a world of lawbreakers with little respect for the legal system.<sup>195</sup> Such laws may induce complementary relationships and increase crime.<sup>196</sup> This idea may yield important results in other areas as well. It may show, for example, that the “gateway studies” — studies that show that marijuana and alcohol use lead to the use of hard drugs — are useless because they get the question backwards. It may be that the legal system, by branding kids who take the relatively common steps of drinking and smoking pot with the criminal label, shapes preferences by reducing the cost of being called a criminal a second time.<sup>197</sup>

---

193. See *infra* section III.B.3. This idea suggests one perverse consequence of workplace drug testing. While such testing may provide some deterrent effect, for those who fail the tests the result may be to increase drug use. By stigmatizing (firing) those people who fail, the system may encourage them to increase their drug use. Drug use has a lower cost for the unemployed — both because use does not interfere with the user’s job and because employment may “incapacitate” a worker from criminal activity because of the time the actor spends at work. On the latter claim, see Panther, *supra* note 32, at 372.

194. One could adjust punishment to compensate for this complementarity. For example, the law could subject repeat offenders to higher lockstep penalties and could conduct increased police surveillance of released felons. This strategy of deterring crime through high penalties for recidivists has been validated by some empirical evidence. See Maurice Cusson & Pierre Pinsonneault, *The Decision to Give Up Crime*, in *THE REASONING CRIMINAL* 72, 77 (Derek B. Cornish & Ronald V. Clarke eds., 1986) (“We know that the more active a defendant’s criminal file, the greater the risk of a long sentence. Criminals know this as well. The majority of our subjects told us that the fear of incurring a long prison term the next time had an influence on their decision to stop.”).

195. Even if people have little respect for the legal system, this does not mean that the law does not deter, because penalties still exert an opportunity-shaping presence. But the more significant power of the law to shape tastes is weakened in such instances.

196. See generally W.A. GOVE, *THE LABELLING OF DEVIANCE* (1975); EARL RUBINGTON & MARTIN S. WEINBERG, *DEVIANCE* (1981).

197. See RICHARD R. CLAYTON & HARWIN L. VOSS, *YOUNG MEN AND DRUGS IN MANHATTAN: A CAUSAL ANALYSIS* 62-64, 99-113, 138, 161 (National Institute on Drug Abuse

What follows from this discussion of preferences, apart from broadening the reach of substitution, is an interesting result about *education*. Both punishment and education create substitution effects. If teens are educated about the dangers of crack all the time, for example, they may read this as an indication that heroin is less harmful than they previously thought. Crack is stigmatized, and people will try something that is not stained by the stigma. Or, to take another example that invites exploration, the intense educational campaign against teenage drunk driving may have led teens to substitute driving while under the influence of marijuana. Substitution theory, as modified by taste shaping, does not yield conclusions only about *punishment*.<sup>198</sup>

Finally, taste-shaping helps clarify the complicated relationship between a penalty and attitudes towards loss and provides another answer to Becker's claim that sentences should be increased and the probability of detection should be lowered.<sup>199</sup> The criminal law sends a message that those activities with a high expected penalty may have a high loss. For these individuals, taste-shaping provides deterrence above and beyond the legal price imposed by a penalty. The government is saying, for example, that consuming this drug imposes huge losses on your health. Even if circumstances thus present themselves that eliminate the expected sanction, such as being in the company of a very close friend, many people will not use the drug because of possible health effects.

Again, though, penalties have paradoxical effects at times because a strong punishment could suggest low loss levels. "The only reason the government is regulating this stuff is because it is not harmful, because were it harmful, people wouldn't be doing it." Indeed, one result of high sentences is that some individuals may begin to associate the "price" of an activity, such as taking a drug, with only its legal cost. When the opportunity presents itself and a

---

Research Monograph No. 39, 1981) (providing data suggesting that, while marijuana is a gateway drug, the complementarity may arise from its illegality, which pushes users into dealing drugs and in turn leads to consumption of other drugs).

If certain activities cause complementarity, the law could take advantage of this escalator effect to increase the penalty for undertaking such activities. For example, if the use of marijuana in an unregulated market often leads to consumption of harder and more dangerous drugs, the government may want to increase the penalty for marijuana to deter people from using marijuana in the first place — not because of the harm marijuana use, by itself, causes, but because of the potential for marijuana use to lead to consumption of more harmful drugs.

198. A somewhat similar story might be told under the opportunity-shaping view. One could say that educational campaigns simply increase awareness of the costs of an activity, and that the opportunity-shaping conception is thus viable because this education reduces the transaction costs to obtaining information about penalties.

199. On attitudes toward loss, see *supra* text accompanying note 75.

tempted person is unlikely to be caught consuming, a person may use the drug because the cost appears low. In other words, one casualty of using a high-sanction strategy may be a decrease in perceived loss — at least for those users whose preferences are not shaped by a penalty. When the opportunity to “get away” with an act presents itself, the harm of the act may not appear to be as dangerous as it once was. The crime is essentially “on sale,” and we all know — too well — that discounts are powerful inducements to act.<sup>200</sup> The deterrent effect from penalization varies, then, by people’s internalization of a penalty’s taste-shaping component.

We have now seen that substitution is not simply about reducing the opportunity to commit harmful acts, but also about shaping tastes to reward less harmful activity. This theory predicts that the law will encourage people to refrain from crime even when there is no chance that they will get caught or suffer any adverse consequences. But legal cost is not everything in determining behavior. Internal preferences matter too, in a number of ways.

## B. *Sociological Norms and the Coercion of Preferences*

### 1. *Know Your Law*

Perhaps the above discussion smacks of law-fetishism. Preferences are not, after all, simply a product of the legal system. Criminal law must therefore try to understand the interrelationship between preferences, norms, behavior, and the law. Implicit in the discussion up to this point was the assumption that people actually *know* the cost of an activity despite the costs of obtaining such information.<sup>201</sup> The substitution skeptic is rightfully concerned — how can policymakers expect would-be lawbreakers to know such details? This is a recurring theme in any legal-economic analysis,<sup>202</sup> and an answer here is no more difficult than it is anywhere else in law. Yet a full answer to this critique deserves a separate article, and so it is best here to only give readers a taste of my response.

Before introducing the response, however, we must first situate the criticism in light of deterrence generally. If people do not know

---

200. This conclusion follows from work in cognitive psychology, which I discuss *infra* text accompanying notes 260-72.

201. See, e.g., Assembly Comm. on Criminal Procedure (Cal.), *Public Knowledge of Criminal Penalties*, in PERCEPTION IN CRIMINOLOGY, 74, 78 (Richard L. Henshel & Robert A. Silverman eds., 1975) (“[P]enalties cannot act as deterrents since these are unknown until after a person has committed a crime or become a prisoner.”).

202. For one example, a powerful and rich analysis of the issue in tort law, see Alan Schwartz, *Proposals for Products Liability Reform: A Theoretical Synthesis*, 97 YALE L.J. 353, 371-84 (1988).



the law and do not understand the penalties, then it is tough to see how increasing the penalties will ever make a difference. Perhaps the substitution skeptic can respond by modifying her position to argue that people know the law in general, but that they do not compare differences in penalties. The skeptic's gambit seems tenuous here. We all know that some crimes are punished more than others — littering versus robbery, for example. And it may be exactly the reverse, that people may not know precise penalties, but know that one activity carries a relatively higher penalty than another.<sup>203</sup> The skeptic's better argument, therefore, is that for one-shot crimes, where only a low chance exists that the perpetrator will repeat the activity, substitution may not occur. Because she does not plan on committing the crime again, the perpetrator has little incentive to “invest” in gathering information about penalty differences.

Yet the substitution skeptic is asking the wrong question by assuming that an actor's knowledge of the law is determinative. My contention in this section is that high criminal penalties may deter crimes even when people do not know what those penalties are. Traditional economists, too focused on the *price* of criminal conduct, have not understood that preferences may be shaped even when the price is unknown to actors. People who have never eaten caviar, for example, do not need to know its cost for their preferences to be affected by the price. This would be particularly so if the high price of caviar put a stigma on caviar-eaters — that they were greedy and selfish, for example. In such circumstances, even if the monetary price of the good is unknown, the social price (which is in part a function of the monetary one) will deter consumption.

My idea is an extension of the pathbreaking work of Johannes Andenaes, who argued that criminal law deters crime through its educative impact.<sup>204</sup> He believed that penalties send out “messages” to members of society and that these messages exert a moral influence to inculcate social norms. This theory of messages thus gives meaning to James Fitzjames Stephen's statement that [s]ome men, probably, abstain from murder because they fear that, if they committed murder, they would be hung. Hundreds of thousands abstain from it because they regard it with horror. One great reason why they regard murder with horror is, that murderers are hung with the hearty approbation of all reasonable men.<sup>205</sup>

---

203. See *supra* text accompanying notes 138-39.

204. JOHANNES ANDENAES, PUNISHMENT AND DETERRENCE (1974).

205. JAMES FITZJAMES STEPHEN, A GENERAL VIEW OF THE CRIMINAL LAW OF ENGLAND 99 (1863).

Stephen wisely realized that a penalty can have an unconscious deterrent effect — through a subtle changing of people's mores.

Stephen's words, therefore, mark him for more than the general deterrence theory for which he is cited today. Stephen believed that a penalty can affect the behavior of more than the individual punished — the general deterrence point. But he also argued that the criminal law has an educational effect and that this effect may dwarf general deterrence.<sup>206</sup> Note also Stephen's important assumption about taste-shaping, that murder is "regard[ed] with horror" because of the penalty structure.<sup>207</sup>

It is now not difficult to understand how penalties may deter behavior even when such punishments are unknown to would-be lawbreakers. The educational impact of the criminal law is not a narrow Skinnerian stimulus and response, but rather one that works through a complex process of social interaction. A small group of people may look at the sentencing structure and be influenced by its relative treatment of crimes. As time passes, the information this group possesses will trickle down, but now in a way no longer tied to sentencing. Instead, it may simply be said that activity *X* is worse than activity *Y*.

This trickle-down theory leads me to posit the existence of *information vanguards* — people who "get the message" first and then transmit it to others. These information vanguards take in information, digest it, and pass it along to the rest of the world. They may relay the message as they first heard it — committing murder has a 20 year jail sentence — or they may pass it along in a processed form as *lore* — committing murder is simply bad.<sup>208</sup> It does not matter for deterrence purposes which one of these actually happens. My word choices may be shaped by the influence of today's great writers, but I do not have to be aware of the writers to be influenced by them. Consumers do not need to know that an

---

206. And for this reason, Stephen claimed that "the sentence of the law is to the moral sentiment of the public in relation to any offence what a seal is to hot wax." 2 JAMES FITZJAMES STEPHEN, *A HISTORY OF THE CRIMINAL LAW OF ENGLAND* 81 (1883).

207. See Dau-Schmidt, *supra* note 140, at 18 n.88. As Paul Robinson and John Darley have recently put Stephen's point: "Most people obey the law not because they fear punishment but because they see themselves as persons who want to do the right thing." PAUL H. ROBINSON & JOHN M. DARLEY, *JUSTICE, LIABILITY, & BLAME* 201 (1995).

208. Some of the vanguard will be composed of people who have already been arrested and subjected to the penalty. Such people, and their loved ones, may communicate the penalty schemes directly to others or may pass their knowledge on as lore. Others who may perform an information vanguard role may include community leaders, the media, and drug dealers.

activity is illegal to refrain from doing it. They need only know that it is “bad” or “dangerous.”<sup>209</sup>

To put the idea slightly differently, this lore — arising in the shadow of the law — influences behavior regardless of whether its origin is known. The cross-pollination of norms explains why lore need not be heard by everyone for it to affect everyone.<sup>210</sup> This may explain why death penalty studies can never adequately measure the deterrent effect — lore created in states that have the death penalty may even influence people in states without it.

The generalized educational effect, when combined with substitution, yields a powerful explanation of deterrence. In particular, we can now understand how relative differences in punishment can unconsciously influence criminal behavior through the vehicle of lore. No one, not even Andenaes, has yet examined whether penalty schemes can invert the educational mission by overdetering some crimes and underdetering others. Economists have shied away from such thoughts because of a belief that sociology and discussion of lore and norms is “too fuzzy.”<sup>211</sup> Sociologists, for their part, have not incorporated the rich insights of substitution and economics into their approaches.

A theory of lore and norms will also illustrate other defects in the opportunity-shaping view of behavior. Ever since Becker, a standard law and economics assumption has been that reducing enforcement costs and increasing penalties creates optimal deterrence. But this approach ignores the way in which people react to high penalties. Such penalties create what may be termed an *inverse sentencing effect*. High penalties, instead of increasing conviction rates, may decrease them. As penalties increase, people may

---

209. The question remains whether the relative differences in penalties can be transmitted through the vehicle of lore. The substitution skeptic can concede that penalties can have an educational effect that deters crime in general but can quibble with the claim that actors distinguish between crimes on the basis of an indirect educational effect. But we can dismiss this argument for the same reasons we rejected the skeptic's gambit earlier. One may know that dealing marijuana and crack may both be “bad” — and still have a clear idea that one is “worse” than the other — even without direct knowledge of the penalty scheme for each. See *supra* text accompanying note 203.

210. One interesting avenue for exploration is whether the information vanguard for various crimes is composed of the optimal characters. If drug dealers compose the brunt of the information vanguard, the message that they transmit could be something like: “Heroin isn't as bad as crack.” If, instead, community leaders play this role, the message could be: “Heroin and crack are both equally evil — it's just harder to catch crack dealers — and you should not use either poison.” This suggests that the goals of criminal law may be enhanced if social leaders explained the rationale behind policy decisions and attempted to counteract the “spin” that savvy and unsavory individuals put on criminal sanctions.

211. See Opp, *supra* note 32, at 426.

not be as willing to enforce them because of the disproportionate impact on those caught.

Several different mechanisms are responsible for the inverse sentencing effect. When the penalties are high, for example, the public may not be willing to turn lawbreakers in, police and prosecutors may not want to prosecute, and jurors may not vote to convict.<sup>212</sup> Beutel observed this phenomenon in his study of bad check laws.<sup>213</sup> He found that Nebraska's severe punishment for bad checks hampered enforcement and conviction. In Colorado, by contrast, he found that fewer bad checks were written because the punishment was weaker but enforcement was more consistent.<sup>214</sup>

Across the Atlantic, when England tried to increase the penalty on prostitutes for solicitation in the 1959 Street Offenses Act, "many police forces adopted a practice of cautioning women for soliciting on the first and even the second occasion; and . . . a prostitute could usually escape prosecution by moving to another police district when she knew she could not expect another caution."<sup>215</sup> In the eighteenth and nineteenth centuries, even though the number and severity of English penal laws had increased, English jurors regarded the penalties as excessive and were lenient in applying them.<sup>216</sup>

Andenaes noted this phenomenon in his discussion of strong sentences,<sup>217</sup> but he did not explain how this observation could be

212. For an analysis of a somewhat similar issue, jury nullification, in the context of race, see Paul Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 *YALE L.J.* 677 (1995).

213. See FREDERICK K. BEUTEL, *SOME POTENTIALITIES OF EXPERIMENTAL JURISPRUDENCE AS A NEW BRANCH OF SOCIAL SCIENCE* 366 (1957). Recent studies show, moreover, that one effect of three-strikes laws is that prosecutors tend not to use them. See Henry J. Reske, *Hardly Hardball: Prosecutors in Most of 22 States Studied Are Not Using Three-Strikes Laws Against Repeat Offenders*, *A.B.A. J.*, Dec. 1996, at 26. Other evidence shows that those who are charged under three-strikes laws refuse to plea bargain and clog the courts, which in turn prevents the administration of swift sentencing. See Cyndee Fontana, *'Three Strikes' Law is Bearing Down on Fresno Courts*, *FRESNO BEE*, Jan. 21, 1996, at A1.

214. See BEUTEL, *supra* note 209, at 366-67. The inverse sentencing effect could also vary by community. Some communities may be more receptive to the norms a criminal penalty engenders than others. It may also vary by the particular targets of the crime. For example, in eighteenth-century Europe, peasants who fled from their masters received much sympathy, while a man who "committed crimes at the expense of this population, the vagrant beggar . . . who robbed and murdered, easily became the object of a special hate." FOUCAULT, *supra* note 4, at 83.

215. NIGEL WALKER, *CRIME AND PUNISHMENT IN BRITAIN* 241 n.3 (2d rev. ed. 1968).

216. See FOUCAULT, *supra* note 4, at 14. Within the economic tradition, James Q. Wilson has pointed out that high sentences also mean that defendants are unlikely to plead guilty, which may reduce the speed with which sentences are imposed, itself an important factor in enhancing deterrence. See WILSON, *supra* note 14, at 134-35.

217. ANDENAES, *supra* note 204, at 61-62.

consistent with his overarching theory of criminal law. If the criminal law sends out messages that shape the moral behavior of society, then how can it be that high penalties can lead to a lower expected punishment? Shouldn't those high penalties shape behavior and prevent inverse sentencing? Andenaes's failure to answer these questions threaten his whole theory, for it suggests that morality may be exogenous to law.

But there are ways to rehabilitate the theory. My answer emphasizes the *lag of lore* — the tendency for old messages from criminal law to stay entrenched when new messages take their place.<sup>218</sup> The lag of lore predicts that the effect will be most pronounced when an old message is entrenched into public consciousness. But for crimes where the old message is weak — for example, anti-sodomy laws — or perhaps nonexistent — such as the widespread ignorance of crack in 1986 — the lag is weaker.<sup>219</sup> This lag explains why the current penalties for crack may not have had an inverse sentencing effect and why an increase in the penalty at this time may.

If crack penalties were increased further today — say, dealing a one-gram amount means life imprisonment — an inverse sentencing effect is possible. This is because the message created by the 1986 laws has become entrenched, and the immorality of using crack, in large part, has been determined. An increase in the punishment now may therefore create a much greater inverse sentencing effect than would the same increase had it been implemented in 1986. Indeed, the lag of lore may yield a powerful explanation of why Prohibition was a failure — Americans had a preexisting message that alcohol use was tolerable.<sup>220</sup> Scholars such as Byse may, then, be wrong to say that the Prohibition experience is evidence that “goodness cannot be legislated into men.”<sup>221</sup> Had

---

218. The concept is similar to the inherent obstacles caused by taste-shaping when the law attempts to return to a starting point. See *supra* text accompanying note 188.

219. One could even take this point to suggest that an increase in legal penalties may reduce aggregate deterrence. As the O.J. Simpson trial suggests, even unsuccessful criminal enforcement can deter crime through the creation of stigma. In other words, detection results in costs exogenous to the legal system — such as social stigma — that the level of penalty does not always influence. Such areas tend to be ones where lore has developed and is entrenched. In these cases, the probability of detection may be more important than the length of sentence, and reducing the probability of detection to compensate for an increase in the sentence may be counterproductive.

220. The inverse sentencing effect is thus related to anomie theory. By providing a mechanism for understanding how norms develop when penalties do not deter actors, it explains why social norms may not fully correlate with penalties. See *infra* text accompanying notes 224-28.

221. Clark Byse, *Alcoholic Beverage Control Before Repeal*, 7 LAW & CONTEMP. PROBS. 544, 569 (1940).

America waited out the lag period, Prohibition may even have turned into a success.

The influence of lore means that policymakers should be particularly attuned to first determinations of penalties because lag may constrain the potential to correct penalties later. This theory provides an explanation of why the 1986 crack law probably has not led to an inverse sentencing effect — the absence of a preexisting message.<sup>222</sup> All of this suggests that the law should invest more in determining optimal penalties at the outset, rather than trying to adjust them later. It also predicts the effectiveness of campaigns that work to influence the preferences of young children before they develop strong tastes of their own — such as President Clinton's antismoking policies — and the weakness of ex post penalization schemes.

The divide between sociology and economics creates other casualties as well. We have already seen that when the cost of a particular crime is lessened, more people may undertake that crime. Yet the economist's definition of "cost" — focusing on the expected legal sanction — is too law-centric. Instead, one must also factor in the social "cost" of crime — the ostracism a community levies upon a lawbreaker.<sup>223</sup> This is, of course, related to a sociological perspective on crime, anomie theory.<sup>224</sup> Anomie theory posits that crime will occur more frequently when members of a group desire a certain goal and it is difficult to use lawful means to obtain the goal.<sup>225</sup>

---

222. Of course, first determinations may matter more for other reasons, too, such as the fact that criminals incur sunk costs so that, as time passes, the law's ability to encourage substitution to less harmful activities will be weaker.

223. Community ostracism may not be as much of a deterrent as the disapproval of those close to the person contemplating crime. As Marcus Felson has put it, most actors tend to have a relationship with another individual, the "intimate handler," and the disapproval of this handler strongly influences the actor's legal and illegal activity. Marcus Felson, *Linking Criminal Choices, Routine Activities, Informal Control, and Criminal Outcomes*, in *THE REASONING CRIMINAL*, *supra* note 194, at 119, 121-23.

224. See ROBERT MERTON, *SOCIAL THEORY AND SOCIAL STRUCTURE* 185-248 (rev. ed. 1968).

225. I note that this is only one understanding of anomie, and it represents a peculiarly American interpretation of the theory. For a comparison of this interpretation with others, see MARCO ORRU, *ANOMIE: HISTORY AND MEANINGS* 2, 123-24, 129-47 (1987). Durkheim himself did not appear to consider the substitution point:

Thieves are as strongly disposed to theft as murderers to homicide. The resistance shown by the former category is in no way weaker than that of the latter. Thus, to overcome it, we should have recourse to the same means. . . . The punishment should vary only according to whether the subject is more or less hardened a criminal, and not according to the nature of the criminal act. An incorrigible thief should be treated like an incorrigible murderer.

EMILE DURKHEIM, *DE LA DIVISION DU TRAVAIL SOCIAL* 52-64 (10th ed. 1978), *excerpted and reprinted in* DURKHEIM AND THE LAW 59, 62 (Steven Lukes & Andrew Scull eds. & W.D. Hall trans., 1983).

Legal means may be difficult because social norms do not reward such activity — and may actively discourage it — or because such means are not available to members of the group — for example, a weak educational system makes it difficult for children in certain communities to obtain their goal — money — through lawful employment.

Thus, anomie theory may be stated, at least partially, in economic terms: as the costs of achieving one's goal through lawful activity grow — costs that may be social or systemic entry barriers — unlawful activity to fulfill that goal is more likely to occur.<sup>226</sup> But anomie theory also contains the further insight that social norms are not static and that the gap between preexisting goals and lawful means may influence them.<sup>227</sup> If children growing up on the South Side of Chicago do not believe that they will be able to earn money lawfully, some may turn to illegal means, and this behavior may further erode social taboos against such means. This nicely tracks Elster's conception of adaptive preferences, which holds that preferences will adapt to a situation where lawful employment does not permit fulfillment of one's monetary goals, not by reducing the taste for money, but by *increasing* the taste for crime. For these segments of society, a subnorm towards crime develops.<sup>228</sup>

In some respects, this insight may be seen as the converse of the taste-shaping economic approach to criminal law. That economic approach focuses on a penalty's ability to influence the tastes of actors. Anomie concentrates on what happens when a penalty does not deter unlawful behavior and the taste-shaping effect that occurs when, and because, the penalty fails. Anomie thus shows how the economic approach, even with its taste-shaping modifications, cannot fully explain deterrence. Instead of only concentrating on those criminals that a penalty deters, as the economic approach does, an-

---

226. This may provide one explanation for why people take drugs. As Cloward and Ohlin have emphasized, the meanings and motives for drug use develop in the context of group membership, and anomie theories that emphasize the adaption of goals to possibilities must take such group goals into account. RICHARD A. CLOWARD & LLOYD E. OHLIN, *DELINQUENCY AND OPPORTUNITY: A THEORY OF DELINQUENT GANGS* (1960); see also Trevor Bennett, *A Decision-Making Approach to Opioid Addiction*, in *THE REASONING CRIMINAL*, *supra* note 194, at 83, 89 (“There is almost complete agreement among studies that opioid use is typically initiated in the company of friends.”).

227. See ROBINSON & DARLEY, *supra* note 207, at 201 (providing empirical research showing that “discrepancies between the criminal code and the community tend to undercut the condemnation of conviction and thereby lessen the effectiveness of condemnation as a deterrent threat”); Bruno S. Frey & Karl-Dieter Opp, *Anomie, Nutzen und Kosten*, 30 *SOZIALE WELT* 275 (1979) (distinguishing anomie theory from economic approaches to crime because anomie posits a belief in preference changes).

228. See *infra* text accompanying note 256.

omie inquires into the beliefs of those actors whom a penalty does not deter and the ensuing impact on social norms.

## 2. *The Coercion of Punishment*

The preceding discussion has analyzed one way in which law contributes to the formation of social norms — information van-guards and lore. A separate point has to do with the role of *punishment* in the formation of norms. As opposed to seeing the law's influence on norms solely through the process of statutory enactment, it is also important to view its impact through the social machinery of the act of punishment. Stephen's quote about hanging murderers is illustrative of this concept — the imposition of punishment itself influences norms.<sup>229</sup> Michel Foucault has developed the underlying theory at length, and I will compress his argument here. Foucault contends that European political systems used torture to brand a victim's body with infamy. The public nature of such torture, its "spectacle," further solidified the social meaning of the punishment. The very excessiveness of the punishment was part of its glory: justice went beyond the mere confines of the body so that its victims cried out in pain for all to hear.<sup>230</sup> The guilty man openly bore his condemnation for society to see.

Foucault's work is centrally concerned with why the system of torture developed, and he concludes that it arose to fill a gap between rarified legal discourse and the people. The tradition of penal truth inherited from the Middle Ages depended on complicated distinctions between "full proof," "semi-proof," and a host of other factors.<sup>231</sup> Public punishments were designed to reconstitute the injured sovereignty of the Crown and to restore that sovereignty by making it spectacular:

Although redress of the private injury occasioned by the offence must be proportionate, although the sentence must be equitable, the punishment is carried out in such a way as to give a spectacle not of measure, but of imbalance and excess; in this liturgy of punishment, there must be an emphatic affirmation of power and of its intrinsic superiority. And this superiority is not simply that of right, but that of the physical strength of the sovereign beating down upon the body of his adversary and mastering it . . . .<sup>232</sup>

---

229. See *supra* text accompanying note 205.

230. See FOUCAULT, *supra* note 4, at 36-39. For modern analogues, see generally ELAINE SCARRY, *THE BODY IN PAIN* (1985).

231. See FOUCAULT, *supra* note 4, at 36-39.

232. *Id.* at 49.



Punishment showed both the horror of the crime and the fact that the sovereign had mastered it.

Criminals were marked through permanent scarring to make the guilty person the herald of his own condemnation.<sup>233</sup> At some point, Foucault contends, the rampant abuse of power in state executions itself created crime. Excessive sentences and arbitrary enforcement led people at the end of the eighteenth century towards violence: “[T]he terror of the public execution created centres of illegality: on execution days, work stopped, the taverns were full, the authorities were abused, insults or stones were thrown at the executioner, . . . fights broke out, and there was no better prey for thieves than the curious throng around the scaffold.”<sup>234</sup> People were threatened by a legal system that imposed violence without restraint or reason, and were led to violence themselves. This phenomenon eventually led to the growth of the prison and a more bureaucratized form of justice.

Foucault’s ideas can be applied to the transmission of norms created by modern penal systems as well. He recognizes, for instance, that punishment is about power and that it is designed to instill a version of “truth” in individuals.<sup>235</sup> Criminal justice plays the role of a “principle of transmission.”<sup>236</sup> The ways in which it metes out justice will influence the values and preferences of citizens. Though not explicit, Foucault shows how the “sign” of punishment is used across two different axes — by being extended across *time* (that is, the brand on the body that bears the didactic mark long after the punishment has lapsed) and across *space* (that is, the “spectacle” of punishment serves to spread, through the empathy of viewers, the pain of the body to other bodies).<sup>237</sup>

Punishment — not the law books — therefore becomes a mechanism to encourage a dialogue between citizens and the govern-

---

233. *See id.* at 43. This feature of deterrence could lead to variations in deterrence in communities where the guilty herald their convictions without shame or even with pride. (In this respect, the idea is very similar to anomie theory.) For one such example, see Kenji Yoshino, *Suspect Symbols: The Literacy Argument for Heightened Scrutiny for Gays*, 96 COLUM. L. REV. 1753, 1786-88 (1996) (describing how the pink triangle has been appropriated as a badge of pride for gays).

234. FOUCAULT, *supra* note 4, at 63.

235. *See id.* at 55. He also recognizes that there is still a trace of torture in modern penal systems because imprisonment operates on a physical level by shutting the body away in a cell. *See id.* at 127-31.

236. *Id.* at 282.

237. Physical torture of course is not the only means of marking people out in this way. The “Scarlet Letter” is a less physically invasive version of the impulse, which, unlike torture, has broad literal and figurative appeal today (for example, Megan’s Law).

ment.<sup>238</sup> The penalty reveals not only the power of the government, but the laws' content as well. Each instance of punishment is, in short, a kind of fable, designed not only to shape preferences of the individual who is punished, but also those of society at large. Torture and segregation are forms of universally intelligible speech, understood even by illiterates, and thus have great didactic value. This is another, related way of understanding the government's role in norm creation. What follows is that the act of punishment generally serves as a preference-shaper. If, however, the law is out-of-step with community norms, its preference-shaping ability will be circumscribed by such norms, as it was in late eighteenth-century France<sup>239</sup> or twentieth-century Nebraska.<sup>240</sup> To the extent individual communities can construct enduring narratives that do violence to the state's chosen one, the law's general preference-shaping power will suffer.<sup>241</sup>

### 3. *The Counterproductive Coercion of Stigma*

The observation that marking the offender will inculcate social norms is, however, incomplete. One must also consider the effect of stigmatization strategies on the individual criminal. By segregating such actors from mainstream America, the criminal law could reinforce a tendency towards criminal action. In economic terms, when an individual cannot get hired for lawful work because she was once an outlaw, the relative cost of illegal activity decreases.

---

238. This idea also suggests that the types of punishment may be important. *See id.* at 45, 104-05. If one goal of punishment is to create an immediate link between the punishment and the crime, then the punishment should replicate the crime. For example, eighteenth-century torture was often imposed at the very site of the crime, and sometimes the torture would theatrically reproduce the crime. Such measures were believed to enhance deterrence. As Beccaria put it, "[t]he punishment should, as far as possible, fit the nature of the crime . . . so that [fear of punishment] removes and redirects the mind to ends other than those which the enticing idea of breaking the law would wish to point it." BECCARIA, *supra* note 4, at 49. Marat put the point a similar way:

To derive the punishment from the offence is therefore the best means of proportioning punishment to crime.

If this is the triumph of justice, it is also the triumph of liberty, for then penalties no longer proceed from the will of the legislator, but from the nature of things; one no longer sees man committing violence on man.

JEAN-PAUL MARAT, *PLAN DE LÉGISLATION CRIMINELLE 72-73* (Aubier Montaigne 1974) (1790), cited in FOUCAULT, *supra* note 4, at 105..

This view of deterrence often led to rather ingenious means of punishment. French legal reformer Vermeil stated that abusers of public liberty should be deprived of their own liberty, that speculation and usury should be punished by fines, theft by confiscation, murder by death, fire-raising by the stake, poisoners by poison, and so on. *See* FOUCAULT, *supra* note 4, at 105.

239. *See supra* text accompanying note 234.

240. *See supra* text accompanying note 213.

241. *See* Cover, *supra* note 189, at 35-39, 47-50. *See generally* Butler, *supra* note 212.

Moreover, from a psychological perspective, those branded outlaws may begin to internalize such labels and fulfill the expectation that they believe the criminal system and society have for them. Instead of reducing crime, stigmatization strategies may increase the criminal activity of particular actors.

This analysis flows quite naturally from Erving Goffman's important work on stigma.<sup>242</sup> Goffman explains that "normal" society shuns stigmatized individuals — those that deviate from the norm. Such individuals may choose either to correct the stigma (for example, a physically deformed person who elects plastic surgery), devote effort to overcome the stigma's effect and thus open doors that appear closed (for example, work hard in school to compensate for the deformity), or join with others who face the same stigma.<sup>243</sup>

Those who cannot remove the stigma, whose identities are spoiled, will often arrange their lives to avoid contact with normal — that is, unstigmatized — people.<sup>244</sup> Even if the deformity can be hidden, the risk of being exposed will often serve as an inducement to avoid such contacts. As one unemployed man put it during the Depression:

"When I go out, I cast down my eyes because I feel myself wholly inferior. When I go along the street, it seems to me that I can't be compared with an average citizen, that everybody is pointing at me with his finger. I instinctively avoid meeting anyone. Former acquaintances and friends of better times are no longer so cordial."<sup>245</sup>

This comment is revealing, as it shows that isolation is a product of the subjective internalization of stigma. The comment raises, moreover, the possibility that normal people avoid those who are stigmatized. Such avoidance stems in part from the importance of the "with" relationship in society — the social reality that when *A* is "with" *B*, people learn about *A*'s social identity from his presence with *B*.<sup>246</sup> "Normal" people do not want to be seen with the stigmatized. Finally, even if a stigmatized person surmounts these hurdles, and normal people treat her with respect, she is left wondering whether the normal people are faking it and whether they are defining her only in terms of her stigma.<sup>247</sup>

---

242. ERVING GOFFMAN, *STIGMA: NOTES ON THE MANAGEMENT OF SPOILED IDENTITY* (1963).

243. *See id.* at 8–10, 23–25.

244. *See id.* at 12.

245. Bohan Zawadzki & Paul Lazarsfeld, *The Psychological Consequences of Unemployment*, 6 *J. SOC. PSYCHOL.* 224, 239 (1935).

246. *See GOFFMAN, supra* note 242, at 47.

247. *See id.* at 14.

Both internal and external avoidance prompts those with stigmata to find sympathetic others. Those with the similar stigmata can provide the individual with moral support and the comfort of feeling at ease.<sup>248</sup> Illiterates, for example, will behave remarkably differently when surrounded by others like them:

The existence of a different value system among these persons is evinced by the communality of behavior which occurs when illiterates interact among themselves. Not only do they change from unexpressive and confused individuals, as they frequently appear in larger society, to expressive and understanding persons within their own group, but moreover they express themselves in institutionalized terms. Among themselves they have a universe of response. They form and recognize symbols of prestige and disgrace; evaluate relevant situations in terms of their own norms and in their own idiom; and in their interrelations with one another, the mask of accommodative adjustment drops.<sup>249</sup>

Illiterates do not only band together with each other, they develop a set of beliefs and "evaluate" "situations in terms of their own norms."<sup>250</sup> Sometimes, when stigmata are invisible, some individuals will try to "pass" as unstigmatized people.<sup>251</sup> Yet members of a group or those outside the group may threaten to expose a person's stigma to the outside world.<sup>252</sup> A modern-day example might be "outing" in gay politics.

Goffman concentrates primarily on physical and social handicaps, but his conceptualization of stigma provides several useful insights into criminal punishment. First, criminal punishment imposes a stigma on individuals that is sometimes visible — recall Foucault's similar description of marking victims. This stigma may lead criminals to avoid contact with law-abiding people. Outside contact becomes problematic because of the risk that normal people will disapprove or define a criminal only in terms of his stigma.<sup>253</sup> Outsiders, for their part, will avoid criminals because of the possibility that being seen with one will contaminate them both

---

248. *See id.* at 20.

249. Howard E. Freeman & Gene G. Kassenbaum, *The Illiterate in American Society: Some General Hypotheses*, 34 *SOC. FORCES* 371, 374 (1956).

250. *Id.*

251. *See* GOFFMAN, *supra* note 242, at 135.

252. *See id.* at 85–86.

253. Consider what one criminal said:

"And I always feel this with straight people — that whenever they're being nice to me, pleasant to me, all the time really, underneath they're only assessing me as a criminal and nothing else. It's too late for me to be any different now to what I am, but I still feel this keenly, that that's their only approach, and they're quite incapable of accepting me as anything else."

TONY PARKER & ROBERT ALLERTON, *THE COURAGE OF HIS CONVICTIONS* 111 (1962).

socially and legally. Stigma also becomes a way to justify one's career choices, much the way those with scars and harelips may justify their decisions.<sup>254</sup> Thus, those who have already committed crime may feel that other options are closed to them and continue their criminal activity — sour grapes, again.<sup>255</sup>

Second, stigmatization from the law-abiding world will prompt criminals to band together with others like them. The stigma imposed from outsiders is celebrated within this group, and their norms differ from the world of the nonstigmatized. They develop *subnorms* that may be antithetical to those of the law-abiding world. This may become both an inducement to further crime, as law breaking is seen as a socially positive act within the group, and a disincentive to noncriminal alternatives. As one criminal describes it, "I can remember . . . on more than one occasion . . . going into a public library near where I was living, and looking over my shoulder a couple of times before I actually went in, just to make sure no one who knew me was standing about and seeing me do it."<sup>256</sup> We can now understand the explanatory power of adaptive preference and anomie theories. The former theory explains how the preferences of lawbreakers develop — as an adaptation to a world where crime is a more realistic option than lawful employment. The latter explains how such attitudes become entrenched within a social group and the formation of subnorms.

In this fashion, crime may spiral upward. The youth who is caught for selling one vial of crack emerges from confinement as a social pariah. He internalizes that belief and avoids contact with the law-abiding world. His isolation from the lawful world leads him to keep company with other pariahs. The subnorms of this group reward the criminal activity that the law-abiding world punishes, and devalues the lawful alternatives that the law-abiding world celebrates. The punishment, then, produces the crime it was

---

254. Dr. Baker and Dr. Smith explain:

For years the scar, harelip or misshapen nose has been looked on as a handicap, and its importance in the social and emotional adjustment is unconsciously all embracing. It is the "hook" on which the patient has hung all inadequacies, all dissatisfactions, all procrastinations and all unpleasant duties of social life, and he has come to depend on it not only as a reasonable escape from competition but as a protection from social responsibility.

William Y. Baker & Lauren H. Smith, *Facial Disfigurement and Personality*, 112 JAMA 301, 303 (1939).

255. Some support exists for the claim that the stigma imposed by criminal sentences precludes lawful employment. See HERBERT S. MILLER, *THE CLOSED DOOR: THE EFFECT OF A CRIMINAL RECORD ON EMPLOYMENT WITH STATE AND LOCAL PUBLIC AGENCIES* (Manpower Admin., U.S. Dept. of Labor Contract No. 81-09-70-02, 1972).

256. PARKER & ALLERTON, *supra* note 253, at 109 (internal quotation marks omitted).

intended to prevent. What is more, it may even produce other types of crime, substitutions of sorts, both because stigmatized individuals avoid the law-abiding world and because they may learn new ways of earning money from members of the stigmatized group.

Even the very names that criminals develop — aliases and nicknames — expose the subnorm problem.<sup>257</sup> Criminals may use aliases to avoid the effects of stigma — both legal and social — by forming a new identity. But often the stigma catches up with and exposes them. At this point, stigma has the counterproductive effects noted above. Similarly, they take nicknames known only to the underworld in part to enhance their reputation for lawbreaking, to revel in the stigma amongst sympathizers. So, the perhaps fleeting taste that once led them to crime may soon land them in a social milieu where crime is the norm, and preferences will again need to adjust to fit this situation.

None of this undermines the earlier point about taste-shaping and the importance of punishment in deterring crime. The point is that the same mechanisms that deter one group — punishment, stigma, lore, and so on — can work to increase crime in another group. The degree to which one trend will outweigh the other is a difficult statistical issue. But whatever the extent is, the penalty scheme has distributional effects that lead some people to become entrenched in crime, and others to be free of it. Particularly when these effects replicate preexisting social inequalities, the criminal law may impose drastic costs.

### C. *Preference and Reference*

Microeconomics understands preexisting preferences to be exogenous. I have argued in this article that this view misstates the impact that legal systems may have on preferences, and that an understanding of criminal law divorced from its taste-shaping impact can be a poor one. As an alternative, I have suggested that the law has a taste-shaping function and that tastes may be shaped by a legal system that operates indirectly and unconsciously on consumers. In this section I examine another way of examining preference formation and substitution: psychology and cognitive bias.

Psychological preference formation theory stresses the way people process information about alternatives. A choice among options creates conflict, particularly when one must consider several

---

257. Cf. GOFFMAN, *supra* note 242, at 59.

different and often competing attributes. Economists use the assumption of value-maximization to analyze such choices. This economic theory predicts that preference ordering between two options should not change when additional alternatives are introduced. One example of this assumption is the "regularity condition," which states that a nonpreferred option cannot become preferred when other options are added. An increasing body of psychological literature, however, demonstrates that these economic assumptions do not accurately describe the reality of human decisionmaking.<sup>258</sup> Among other things, this research has shown how the addition of technically irrelevant alternatives can influence preference.

The simplest point is status quo bias. Eldar Shafir has demonstrated that conflict between attributes leads people to seek new options or to refrain from choice altogether.<sup>259</sup> A person faced with choosing between attractive *X* and the less-attractive status quo will choose *X*. But if the choice is between the less-preferred status quo and equally attractive *X* and *Y*, the person may defer the choice and select the status quo. The fact that each alternative is better than the status quo may not be enough to induce a person to pick one of them because the choice between the two alternatives is hard to justify.<sup>260</sup>

A more complicated point has to do with decisionmaking bias created by the introduction of supposedly irrelevant alternatives. Imagine, for example, two products of equal value to a consumer; Product *A* is high quality with a high price, while Product *B* is low quality with a low price. A consumer is indifferent between Product *A* and Product *B* because Product *B*'s low price compensates

---

258. See Daniel Kahneman & Amos Tversky, *The Psychology of Preferences*, *SCI. AM.*, Jan. 1982, at 160; Mark Kelman et al., *Context-Dependence in Legal Decision Making*, 25 *J. LEGAL STUD.* 287 (1996); Amos Tversky & Daniel Kahneman, *Causal Schemas in Judgments Under Uncertainty*, in 1 *PROGRESS IN SOCIAL PSYCHOLOGY* 49, 61-70 (Martin Fishbein ed., 1980); Lee Ross & Craig A. Anderson, *Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 129, 149 (Daniel Kahneman et al. eds., 1982); Amos Tversky & Daniel Kahneman, *Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment*, 90 *PSYCHOL. REV.* 293 (1983); Amos Tversky & Daniel Kahneman, *Rational Choice and the Framing of Decisions*, 59 *J. BUS.* S251, S260-62 (1986).

259. Eldar Shafir et al., *Reason-Based Choice*, 49 *COGNITION* 11, 18, 21 (1993).

260. In one study, students were told to imagine that they had planned to spend the night working on a paper due the next day. Half of them were told that they then discovered that an admired author was going to give a public lecture, and half were told about the lecture and also about a screening of a movie that they wanted to see. Many more students in the second half — with both choices — chose to go to the library (40% versus 21%). See Donald A. Redelmeier & Eldar Shafir, *Medical Decision Making in Situations That Offer Multiple Alternatives*, 273 *JAMA* 302, 302-03 (1995).

for its low quality. If a third option, Product *C*, is introduced, with the same low price as Product *B* but even lower quality, people may begin to favor Product *B* over Product *A*, because Product *C* makes Product *B* look like a good value. Conversely, if Product *C* has the same high price as Product *A* but with less quality, people will buy Product *A*. Even though people are not receiving additional information, the extraneous information skews their choices. In other words, a particular option can become more desirable simply because the options are presented or framed along with irrelevant information.<sup>261</sup>

The addition of an inferior alternative may thus enhance the desirability of an option. Cognitive psychologists dub this the asymmetric dominance effect — the tendency to prefer *x* over *y* increases by the addition of alternative *z* that is inferior to *x* but superior to *y*.<sup>262</sup> Here is a clear example: Simonson and Tversky offered subjects a choice between \$6 and “an elegant Cross pen.”<sup>263</sup> 36% chose the pen and 64% took the cash. A second group was given the choice between \$6, the Cross pen, and “a second pen that was distinctly less attractive.” This time, 46% took the Cross and 52% chose the cash. Again, this shows that the regularity condition does not hold: the cheap pen option should not influence the choice between the Cross pen and the cash. Yet whether the tradeoffs within the new set are favorable or unfavorable can influence the tendency to prefer a particular alternative — what Simonson and Tversky call a “tradeoff contrast.”

Another exception to the regularity condition is extremeness aversion. Extremeness aversion predicts that within an offered set,

---

261. Shafir et al. tested this prediction by offering groups of students hypothetical choices between two CD players on a one-day clearance sale. The “popular SONY player” costs only \$99 while a “top-of-the-line AIWA player” costs \$169. An equal number of students, 27%, picked each brand, and 46% chose to wait until they learned more about the various models. See Shafir et al., *supra* note 259, at 22. On the other hand, a second group of students was posed the hypothetical without the AIWA model. This time, 66% of people picked the SONY, and only 34% selected the deferment choice. The third group of students was presented with a choice between the SONY, “an inferior AIWA player for the regular list price of \$105,” or the deferment choice. This time, 73% picked the SONY player and only 24% picked deferment. The introduction of the cheap AIWA should not have influenced the choice between deferment and the SONY, but it did. More people were willing to buy the SONY, which looked like a better deal once the cheap AIWA was shown. See *id.* at 22-23; see also Amos Tversky & E. Shafir, *Choice Under Conflict: The Dynamics of Deferred Decision*, 3 *PSYCHOL. SCI.* 358, 360-61 (1992).

262. See Joel Huber et al., *Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis*, 9 *J. CONSUMER RES.* 90 (1982); Douglas H. Wedell, *Distinguishing Among Models of Contextually Induced Preference Reversals*, 17 *J. EXPERIMENTAL PSYCHOL.* 767 (1991).

263. Itamar Simonson & Amos Tversky, *Choice in Context: Tradeoff Contrast and Extremeness Aversion*, 29 *J. MARKETING RES.* 281, 287 (1992).



options with extreme values are relatively less attractive than those with intermediate values.<sup>264</sup>

Both asymmetric dominance and extremeness aversion are explanations of why reference points influence choices between options. These ideas are very useful in analyzing criminal deterrence and substitution of alternatives. The substitution perspective predicts that individuals do not view the costs and benefits of a particular crime in a vacuum. Rather, they examine them in light of the costs and benefits of other crimes. The psychological addendum to substitution suggests that people evaluate the relative harms and benefits of a particular crime by using reference points. Consequently, when the law proclaims, through a harsh penalty, that the cost of a particular activity is very high, it might make other crimes appear more attractive than they were before the penalty.

A harsh penalty on an activity might, therefore, invert Andenaes's idea of general deterrence. Andenaes, as we saw, argued that the criminal law creates deterrence by educating people about those acts that should not be done.<sup>265</sup> But Andenaes's educational effect can be stood on its head. High penalties on crime *X* may not only educate people about the particular danger of *X*, but also about the comparably less dangerous — that is, less punished — crimes *Y* and *Z*, even if *Y* and *Z* are in reality more dangerous. *Y* and *Z* may then look more attractive than they did before. Theories about the educational impact of criminal law should incorporate these reference points in order to account for these substitution-like effects.

To illustrate, let us return to the crack cocaine example. While the new crack penalties did not directly say anything about heroin, they may have influenced people's perception of heroin's utility. The penalty increase could have changed the way people thought about the cost of consuming crack. The law — through Andenaes-like and lore effects — might have shaped people's view of crack as the most evil drug. This would give rise to asymmetric dominance and extremeness aversion effects. By penalizing crack as an ex-

---

264. See Itamar Simonson, *Choice Based on Reasons: The Case of Attraction and Compromise Effects*, 16 J. CONSUMER RES. 158 (1989). For example, subjects were shown five cameras varying in quality and price. One group was given a choice between a \$170 Minolta and a higher quality \$240 Minolta. The second group was given the additional option of an even higher quality \$470 Minolta. In the first group, subjects were split between the two cameras, but in the second group, 57% chose the middle option and the remaining subjects were equally divided between the two extremes. The introduction of an extreme option reduced the desirability of the other extreme option, but not of the option in the middle. See Shafir et al., *supra* note 259, at 25.

265. See *supra* note 204 and accompanying text.

treme drug, the high crack penalties, via extremeness aversion, made heroin look better than it did before and thereby increased the taste for it. For some consumers, heroin became a middle choice — not as bad as either crack or total abstinence from drugs.<sup>266</sup>

In the case of drugs, the psychology of addiction exacerbates the reference point effect. As anyone close to an alcoholic or drug abuser knows, one of the most powerful facets of addiction is denial.<sup>267</sup> Harsh crack penalties may have the subtle effect of playing into denial. People do not want to think they have reached rock bottom. The existence of crack, and its treatment under the law, may provide an excuse for these people. By concentrating on crack, the law has indirectly told people that other drugs are safer when, in fact, the reverse could be true when it comes to heroin.<sup>268</sup> The media exacerbated this problem by devoting constant attention to the one “new terrible drug” on the streets and neglecting older — and perhaps more dangerous — drugs.<sup>269</sup> One reason that heroin use is increasing at such an alarming rate could be that the contemporary understanding of crack may have influenced views about heroin.<sup>270</sup>

The substitution effect, therefore, might work in tandem with general principles from the psychology of addiction and from the effects of reference points more generally to increase consumption of heroin. By concentrating on crack, the law may indirectly suggest that heroin is safer than it earlier appeared to be. Calling this a

---

266. Similarly, the asymmetric dominance effect suggests that when crack and heroin were considered equals, users may have abstained due to conflict about which drug to try. (Recall that people will often defer choice even when that deferment is not “rational,” see *supra* note 261.) But when the equality between crack and heroin was severed by the mandatory minimums, it may have increased the consumption of drugs — indeed, even of both drugs. Extremeness aversion, however, would temper the increase in crack and would channel much of the increase in drug use to heroin.

267. Cf. Robert L. DuPont, *Nicotine*, 272 JAMA 1221 (1994) (reviewing NICOTINE ADDICTION: PRINCIPLES AND MANAGEMENT (C. Tracy Orleans & John Slade eds., 1993) and stating that “[d]enial is a hallmark of addiction to alcohol and other drugs”).

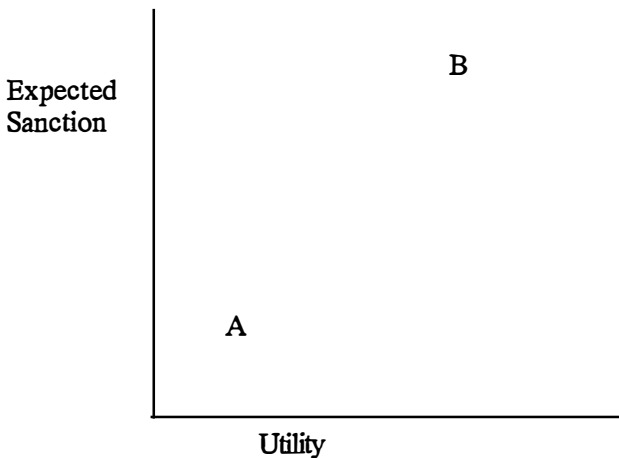
268. See *supra* note 66.

269. See *Heroin Comes Back*, *supra* note 67, at 63 (“If crack didn’t have the attention of the media . . . heroin would have been on the front pages of every newspaper in America.” (quoting Robert Stutman, head of the Drug Enforcement Administration’s New York field office)).

270. Heroin-related emergency room visits increased by 65% from 1988 to 1993. See Lupo, *supra* note 59, at 1. One source reports that in 1992, for the first time in recent years, there was a significant drop in the perceived risk of harm associated with heroin use among high school seniors in the Monitoring the Future Study. See DRUG ENFORCEMENT ADMIN., U.S. DEPT. OF JUSTICE, U.S. DRUG THREAT ASSESSMENT 44 (1993). But the drops, which occurred in the class of 1991, were not that significant. See OFFICE OF NATL. DRUG CONTROL POLY., *supra* note 62, at 12 tbl.7.

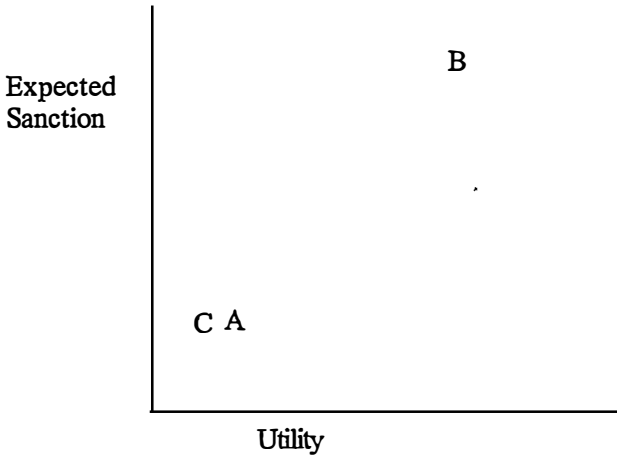
substitution effect is somewhat of a misnomer, since it is based on two noneconomic ideas — the inability of value-maximization to predict choice in the real world and the sociological conception of penalty messages. This interplay nicely demonstrates my claim that incorporation of these concepts, gathered from several different disciplines, may yield a more realistic picture about the deterrent effect of a criminal penalty than a simple and elegant prediction from any one particular method alone.

If extremeness aversion and asymmetric dominance turn out to be powerful real-world forces, they may help governments devise solutions to minimize harmful acts. By introducing alternatives that are superior in one aspect and relatively equivalent in another, the law may influence the choices that people make. The introduction of technically irrelevant options may also make a previous option look more attractive.<sup>271</sup> For example, imagine that individuals face choices between the following options:



Now, assume that a set of would-be criminals is indifferent between the two choices because the high expected sanction compensates for its high utility. The introduction of a third option that is close to option A will induce people to pick option A.

271. See Joel Huber & Christopher Puto, *Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects*, 10 J. CONSUMER RES. 31 (1983).



The government could, therefore, essentially manufacture option *C* in order to induce people to pick option *A*. People may not pick *C*, but may be more likely to pick *A* than before. This strategy could enhance the law's substitution-creating role.<sup>272</sup>

The number of people who will be influenced by *C* will depend on the degree of extremeness aversion and asymmetric dominance present in the options. This analysis is very difficult to do in the abstract, and must await empirical testing. Until now, however, no one has tested such a proposition, in part because the whole notion of what crimes are "close" to each other depends on a sociological notion of grouped offenses that economists shy away from, and in part because the economic idea that criminals choose among offenses has not yet permeated the legal landscape.

#### D. *Adaptive Preferences Revisited*

We now return to analyzing the question of where preferences originate. The previous sections have argued that the law has an important role in shaping preferences. But, naturally, other variables influence preference as well. Adaptive preference theory helps answer these questions by abstracting away current preferences, which inevitably will be shaped by the legal system. This project of assessing preferences by divorcing them from the taste-shaping effect of law has an impressive pedigree in both political

272. *See supra* section II.B. If the law makes legal substitutes out of less harmful activities, it may also reduce the number of crimes that are thought of as nondeterrable. Tastes are shaped within the context of a legal system, and a system that provides people with an incentive to commit the less harmful offense may inculcate a system of beliefs to value that offense more highly than the more harmful ones. If the taste-shaping impact of the criminal law is strong enough, it could wipe out most nondeterrable crimes.

science and philosophy. Even works such as John Rawls's *A Theory of Justice*<sup>273</sup> can be read as presenting a way of understanding preferences in the absence of a legal regime.<sup>274</sup>

Adaptive preferences may provide another explanation of taste formation, most obviously by explaining the preference differentials between criminals and noncriminals. Because people value only accessible options, criminals tend to value drugs in a way that non-users do not. This does not, however, answer the question, from where does the preference to commit crimes initially come? If people value what they have, then why would people commit crime? The full answer to this question is beyond the scope of this article. Accordingly, I will only set out a few preliminary thoughts about the role of counteradaptive preferences in driving criminal behavior. A counteradaptive preference refers to the tendency of humans to want what they cannot have. In lay parlance, the grass is always greener on the other side.<sup>275</sup> While the economist stresses the opportunity for profit illegality engenders, the counteradaptive preference theorist stresses the psychological temptations prohibitions create.

In the drug context, adaptive and counteradaptive preferences provide policy experts with a new path for examining traditional deterrence questions. The adaptive preference question examines whether a legal penalty will have a taste-shaping effect. Such an inquiry may proceed on several levels, and we have encountered them throughout this Part. If, for example, the law keeps a drug away from a group of people, will their preferences adapt in such a way that they will not have a taste for the drug? Conversely, if the distributional effect of a drug law is to flood a particular area with the drug (such as inner cities or schools), will preferences adapt or be shaped in a way that encourages consumption?

The counteradaptive question, on the other hand, asks whether a legal penalty will exacerbate criminal activity by tempting people to commit crimes. Here again, there are several lines of inquiry. Does the temptation to get what one cannot easily have explain, say, drug use among affluent teens? Does this yearning for the unavailable explain the fierce "straight-edge" teenagers in high crime areas who do not drink, smoke, and so on? More generally, is perceived availability shaped by the physical presence of the drug (as

---

273. JOHN RAWLS, *A THEORY OF JUSTICE* (1971).

274. See Sunstein, *supra* note 182, at 1146 n.63.

275. See ELSTER, *supra* note 180, at 111.

the economist would see it), by the law (as the taste-shaping economist would have it), by norms (as the sociologist would put it), by the information skew created by particular choices (as the cognitive psychologist would say), or by some combination of these? An answer to this question would fill a book by itself, and I only mean to raise the question now.

In the case of drugs, adaptive preferences introduce another complication. Unlike preferences for other items, a preference for an addictive narcotic, by its very nature, may change future preferences. Because such a drug is highly addictive upon merely trying it, a user's preferences will change because of its properties. That is, current consumption is a complement to future consumption. Whereas counteradaptive preferences may explain why a user first tries a drug, the drug itself may be responsible for further consumption and addiction, itself a form of adaptive preference.

Cass Sunstein has placed addiction in a category separate from adaptive preference. In Sunstein's schema, addiction is classified as Category 3 — where the “preferences are endogenous to the act of consumption.”<sup>276</sup> Category 2, on the other hand, “includes preferences that are a product of legal rules allocating entitlements and wealth.”<sup>277</sup> By now, my point will be obvious to the reader: addiction belongs in both of these categories. One cannot fully understand addiction without understanding the preferences that give rise to addiction — preferences that the legal regime in part creates. This article has suggested several possible ways in which legal rules that allocate wealth may influence tastes and behavior, and how they may encourage illegal activity.<sup>278</sup> By emphasizing the ways in which the legal system shapes taste, therefore, adaptive preference theory provides a mechanism to help grasp the advances and limits

---

276. Sunstein, *supra* note 182, at 1139. He defines addiction as “a process in which the subjective costs of not consuming a particular good increase dramatically over time, while the subjective benefits decrease or remain stable.” *Id.* at 1158. The preference for a good is not, in his view, static, but changes as a result of increased use. Elster similarly believed that addiction “is much more specific than sour grapes: it is to be explained more by the nature of the object of addiction than by the tendency of the human mind to adapt to whatever objects are available.” ELSTER, *supra* note 180, at 121; see also Gary Becker et al., *Rational Addiction and the Effect of Price on Consumption*, AM. ECON. REV., May 1991, at 237; Gary S. Becker & Kevin M. Murphy, *A Theory of Rational Addiction*, 96 J. POL. ECON. 675 (1988).

277. Sunstein, *supra* note 182, at 1138.

278. See *supra* text accompanying notes 32, 103, 242, 257 (arguing that once the legal regime calls certain people lawbreakers, those people are likely to value their status as lawbreakers more than they did when only “others” were the lawbreakers); *supra* text accompanying notes 109-12 (arguing that wealth effects might mean that criminalization of drugs only shifts consumption among certain socioeconomic groups).

of substitution and a way to understand where preferences for illegal activity are formed.

IV. SYNTHESIZING THE EFFECTS OF DETERRENCE

The above analysis paints a complicated picture and raises a host of questions. By now, readers can see how concepts such as substitution, taste-shaping, norms, asymmetric dominance/extremeness aversion, and adaptive/counteradaptive preferences apply to a host of activities. Does a law against embezzlement lead to an increase in people taking kickbacks? Does it influence social norms, or are such norms already entrenched? Will a doubling of the kickback penalty to compensate for any increase have an effect on tastes? Would it make kickbacks look more or less attractive? Does a market for such offenses exist? These are all pieces of one large puzzle — a project that, though difficult, might turn out not to be impossible.

The following box illustrates some of the possible results from criminalization of an activity:

**FACTORS TO BE CONSIDERED  
WHEN IMPOSING A PENALTY**

	PARTICULAR ACTIVITY		OTHER ACTIVITIES
DECREASES	a) Opportunity set decreases b) Taste decreases c) May unconsciously decrease taste (lore)	<b>1</b>	a) Y-Optimality b) Complementarity from penalized activity and escalator effects
INCREASES	a) Marginal deterrence & cliffs b) Monopolization/true profit c) Inverse sentencing d) Distributional price effects e) Geographic substitution f) Stigma/labeling g) Giffen goods h) Horizontal and vertical tradeoffs	<b>3</b>	<b>4</b> a) Substitution b) Shapes taste towards substitutes c) Extremeness aversion/asymmetric dominance d) Monopolization/true profit e) Complementarity from law-breaking f) Income effects/two types of deterrence

To illustrate all of these effects, let us return to the example of crack cocaine. With respect to Box 1, a strong penalty on crack cocaine possession and dealing could decrease dealing by decreasing a criminal’s opportunity set, because the “price” — legal and perhaps social — of dealing is higher.<sup>279</sup> Furthermore, the penalty

279. See *supra* Part I.

could weaken a taste for dealing or prevent one from being formed in a youth.<sup>280</sup> This effect may be so even if the actor is unaware of the law's role, as lore may develop that crack dealing is an evil activity.<sup>281</sup> Within Box 2, when the crack penalty is set at a Y-optimal point, a penalty on crack may prevent people from committing other sorts of crimes even if it does not prevent crack dealing. This is because the crack penalty has an income effect that makes other crimes more costly. A simpler income effect is that a high monetary price for crack will also mean that users will not have money to consume other drugs.<sup>282</sup> If crack consumption has complementary relationships with other crimes — such as consumption of other drugs and physical abuse — the crack penalty will also reduce these other types of crime if it succeeds in reducing consumption.<sup>283</sup>

Box 3 illustrates how the crack penalty may not actually decrease consumption. If the penalty for crack dealing has cliffs, there will be marginal deterrence problems — if five grams and eight grams are both punished with five years in jail, dealers will deal in eight-gram amounts.<sup>284</sup> Furthermore, criminalization creates the possibility for entrepreneurship and true profit, further increasing the returns from illegal activity.<sup>285</sup> If the crack penalty is increased at a time when lore about crack is entrenched, an inverse sentencing effect is possible, and the expected sanction may drop.<sup>286</sup> In addition, if the law induces price-setters in affluent communities to decrease their consumption, thereby reducing the price of the drug, it may increase the horizontal consumption of less affluent users, whom monetary price deters more than social price.<sup>287</sup>

Box 3 includes several other effects. For example, if the crack penalty is lower in other jurisdictions, the penalty may simply externalize the problem as dealers pack up and move to those locations. Such geographic substitution could mitigate much of the benefit of a penalty.<sup>288</sup> The penalty may also stigmatize those arrested under it, making it harder for them to find lawful employment and easier to find criminal employment.<sup>289</sup> In other words, even if the overall

---

280. *See supra* section III.A.

281. *See supra* section III.B.1.

282. *See supra* section II.C.

283. *See supra* notes 35, 67, 197 and accompanying text.

284. *See supra* section I.A.

285. *See supra* text accompanying note 97.

286. *See supra* text accompanying notes 212-22.

287. *See supra* section I.D.3.

288. *See supra* text accompanying notes 118-19.

289. *See supra* section III.B.3.



number of people dealing crack (horizontal consumption) is lower, those who deal crack will deal more of it (vertical consumption). Finally, if crack is a Giffen good, increasing the penalty could impoverish users of the drug and increase the vertical consumption of crack.<sup>290</sup> In certain circumstances, such as with drug *consumption*, the real evil may be excessive vertical consumption, in which case the tradeoff between types of consumptions can be counterproductive.<sup>291</sup>

Box 4 explains how the crack penalty may increase other crimes. Dealers and consumers may substitute another drug, such as heroin, instead of crack. This may also increase people's taste for the substitute and contribute to an anomie effect whereby subnorms in a community erode the taboo against it.<sup>292</sup> In addition, those stigmatized by their association with the drug may act in other illegal ways because legal employment is no longer an option and because they encounter other persons who dabble in a variety of crimes.<sup>293</sup> Furthermore, if the crack penalty makes crack appear to be the worst drug, consumption of other drugs may increase as extreme-ness aversion causes them to look better than they did before.<sup>294</sup>

To the extent the law creates the potential for monopolization, it may encourage organized crime and further law-breaking by such organizations.<sup>295</sup> If the law raises the monetary price of a drug, it may also have a straight income effect that will either lead to additional crime to earn money to buy drugs, or, perhaps, to a decrease in consumption of other drugs to save money to buy the penalized drug.<sup>296</sup>

Finally, the deterrent effect of the crack penalty is likely to vary for different users. For users with high income, the monetary price effect will not matter much, but the stigma of a possible conviction will.<sup>297</sup> The reverse may be true for users with less wealth. Accordingly, those high income users that face the stigma of being caught may find themselves thrust into an environment where crime is a means of survival. This could increase the commission of all sorts of crimes. And for those low-income criminals whose incomes drop

---

290. See *supra* section II.D.1.

291. See *supra* text accompanying note 166-70.

292. See *supra* text accompanying notes 224-28.

293. See *supra* text accompanying notes 32, 103, 242, 257.

294. See *supra* text accompanying notes 264-66.

295. See *supra* text accompanying note 107.

296. See *supra* text accompanying note 155.

297. See *supra* section I.D.3.

further because of a penalty — either because they are consumers of the drug or because the penalty requires them to pay more in bribes and thus cuts into their dealing profits — a penalty may increase the commission of other crimes in order to return the criminal to her prepenalty wage.<sup>298</sup>

Now consider the nonmarket crime of assault. A penalty on assault will increase its price and thus reduce a criminal's opportunity set. It may also decrease the taste for assault by inculcating a norm against it — an effect that may be either direct or indirect through the vehicle of lore. *Y*-optimality is a more complicated matter — a high penalty on assault, even if it does not reduce the commission of assault, may prevent other crimes because the actor is not willing to take what is now a larger aggregate risk. If the actor wants to assault someone and rob a bank, penalizing the assault at a very high level may deter the bank robbery.<sup>299</sup> This result holds even if the two crimes, assault and bank robbery, are entirely separate. Additionally, a penalty on assault may prevent complementary crimes. So if assault tends to lead to battery or other forms of violence, a high assault penalty may reduce the commission of those offenses.

Conversely, a high assault penalty may — though it deters many assaults — prompt those assaults that do occur to be very harmful, or even encourage assaulters to assault many people at once. This happens when the penalty does not draw a heavy distinction between very harmful and less harmful assaults, or between one assault and many. That is the insight, once again, offered by marginal deterrence. The effect might also be phrased in horizontal and vertical terms: the number of people who assault could be lower, but the assaults may be much worse. A very strong penalty may also create an inverse sentencing effect because juries and judges will not want to impose high sentences on those who commit assault. For those convicted of assault, a stigmatization effect is possible, which may increase the commission of assault. Those who have assaulted may find themselves separated from society, even if they are not punished, and this phenomenon may have anomie and subnorm effects that could increase crime.<sup>300</sup>

---

298. See *supra* text accompanying notes 155, 167.

299. This example, as with all discussions of *Y*-optimality, assumes that the actor has a high enough taste for the assault that it exceeds the expected cost.

300. Two other ideas discussed above, monopolization and Giffen goods, are not really applicable to this nonmarket crime.

The application of the concepts in Box 4 is more difficult for nonmarket crimes. Depending on an individual's taste for assault, a high penalty may prompt substitution to a different crime. The range of substitution will be a function of the individual's preference in relation to the price of the competing options. Substitution is, therefore, theoretically possible. The real question concerns the range of crimes that may be substitutes. The penalty structure, which might group offenses together, may influence this range. To the extent that a high penalty for assault makes other crimes, such as robbery, look better than they did before, cognitive bias may further encourage substitution. Finally, high assault penalties cause stigmatization effects that could drive people to further crime, just as in the crack example. There is, again, no real "income effect."<sup>301</sup>

The above discussion highlights just how many complex effects are at work in deterrence. Which of these results will happen in a given situation is, of course, difficult to predict. The extent to which they will happen is even more difficult to determine. These tasks, naturally, demand empirical work. My only goals here have been to defend the complexity of the question and to demonstrate just how much work is left to do.

I began this article by promising that the deterrence hypothesis is much more complicated than has been assumed. In the end, scholars might decide that the questions about elasticity of substitution, norms, framing, and so on are just too difficult to figure out.<sup>302</sup> If so, that itself would be a useful finding. If the criminal law cannot really tell whether it is accomplishing its goal of deterrence and cannot tell whether the deterrence it does accomplish comes at the cost of creating crimes that may be much worse, it may be time to rethink our deterrence-based consequentialist premises.<sup>303</sup> If such premises rely on intuition and guesswork — guesswork that is itself motivated by our natural gravitation towards simplicity and ele-

---

301. In the abstract, I imagine that one could stretch the concept to argue that if an assault penalty made it more difficult to satiate one's taste for an assault because assaulters had to tone down their threats to avoid heavy punishment, it might increase the number of assaults so that criminals could reach their target "income."

302. For one incisive example showing that elasticities of substitution can be measured, see Ian Ayres & Steven D. Levitt, *Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack* (Aug. 1996) (unpublished manuscript, on file with author). Ayres's work measures, using interrupted time-series analysis, whether the LoJack auto-theft device increases the commission of other crimes, such as robbery, through statistical analysis.

303. See Norval Morris, *Impediments to Penal Reform*, 33 U. CHL L. REV. 627, 631 (1966) ("[E]very criminal law system in the world, except one [Greenland], has deterrence as its primary and essential postulate. It figures most prominently throughout our punishing and sentencing decisions — legislative, judicial and administrative." (footnote omitted)).

gance — other nondeterrent-based approaches to criminal punishment should be explored.

### TO THE FUTURE

Consider the words of the New Hampshire Constitution of 1784: All penalties ought to be proportioned to the nature of the offense. No wise legislature will affix the same punishment to the crimes of theft, forgery, and the like, which they do to those of murder and treason. Where the same undistinguishing severity is exerted against all offenses, the people are led to forget the real distinction in the crimes themselves, and to commit the most flagrant with as little compunction as they do those of the lightest dye.<sup>304</sup>

My own work has emphasized this document as central to an understanding of our Federal Constitution.<sup>305</sup> What is interesting here is the insight it also offers into contemporary criminal law. Over two hundred years ago, those who drafted the New Hampshire Constitution realized the importance of substitution effects in criminal law — that “theft” and “murder” could be substitutes. What’s more, the drafters also realized that the criminal law performs a taste-shaping function; that the “people” would be “led to forget” the distinctions and commit “the most flagrant” crimes.

Criminal lawyers have not examined whether our two-hundred-year experience bears out the predictions of the New Hampshire Constitution. This failure is symptomatic of more general problems in contemporary criminal law. After all, criminal law, for much of the nineteenth century and the beginning of the twentieth, was at the forefront of interdisciplinary studies in law. Criminologists borrowed heavily from psychiatry and philosophy in an attempt to understand why people act the way they do and how government should punish them. Yet recently, a movement inward has dominated criminal law scholarship. This is unfortunate, and not only because the implosion may lead to incomplete answers to age-old questions in criminal law. The narrowness of conventional criminal law has also prompted criminal lawyers not even to ask questions that are now commonplace in other areas of law. Advances in the one area of major interdisciplinary progress, economic analysis of criminal law, have slowed because the insights gleaned from a model with a rational, static-preference, statute-reading actor have reached their logical stopping point.

---

304. N.H. CONST. art. XVIII (amended 1793).

305. See, e.g., Bruce Ackerman & Neal Katyal, *Our Unconventional Founding*, 62 U. CHI. L. REV. 475, 484-86, 561-62 (1995).

The beauty of substitution analysis — in both its economic and noneconomic incarnations — is that it opens up criminal law to these inquiries. What different incentives does criminalization create? What is the role of the criminal law when actors do not have perfect knowledge about penalties? How do degrees of risk-aversion play out in criminal law? What are the effects of excessive penalties? Can the criminal law shape tastes as well as imposing opportunity constraints? Can crimes of passion be substitutes? How does income correlate with crime? How does the psychology of framing relate to substitution? Substitution theory raises all of these questions, and each requires extensive individualized consideration. The answers will determine the role of deterrence in our criminal law.