11-21-2023

Examining the Interrelations between Rational Choice Inputs: Implications for Criminological Theory and Research

Benjamin Hamilton
University of Missouri-St. Louis, bchb83@umsystem.edu

Examining the Interrelations between Rational Choice Inputs: Implications for Criminological Theory and Research

Benjamin C. Hamilton

M.A., Criminology and Criminal Justice, University of Missouri – St. Louis, 2016

B.A., Criminal Justice, University of Maine at Presque Isle, 2014

A Dissertation Submitted to The Graduate School at the University of Missouri – St. Louis in partial fulfillment of the requirements for the degree:

Doctor of Philosophy in Criminology & Criminal Justice

December 2023

<u>Advisory Committee</u>

Kyle Thomas, Ph.D.
(Chairperson)

Lee Ann Slocum, Ph.D.

Chris Sullivan, Ph.D.

Thomas Loughran, Ph.D.

ABSTRACT

An essential component of any rational choice theory of criminal behavior is the notion that crime decisions are driven by an individual's expected gains and losses to illicit activities. More specifically, offenders are typically presumed to balance the pleasures of the various benefits to crime against the pains associated with crime's risks and costs, the presumption being that the offender will pursue criminal acts in the event he or she believes the expected utility to crime exceeds that which can be achieved through strictly legal means. Although criminologists have managed to test some of the more basic implications of this choice calculus, so to speak, many of the more nuanced predictions of this calculus have received relatively minimal scholarly attention. In particular, little is currently known about the extent to which a person's perceptions of the rewards, risks, and costs to crime will exert a *combined influence* on his or her criminal proclivities. Put differently, we have yet to rigorously examine whether the perceived incentives to crime will be *interdependent* on one another in shaping an offender's involvement in criminal activities. Not only is the criminological literature largely bereft of any formal discussion of any such interdependencies (insofar as what theory would seem to imply about the interrelations between rational choice inputs), but also any direct empirical examination of those interdependencies still has yet to be carried out. This dissertation sought to fill this gap in the literature by deriving three primary sets of hypotheses from prior scholarship on rational choice theories of crime (namely, Becker's [1986] model), each of which outlines the directionality, and overall functional form, of the potential interdependent relationships between the rewards, risks, and costs to crime. Each hypothesis was then examined using self-report data derived from the Pathways to Desistance Study, the results of which were

largely supportive of the existence of a generally interdependent link between rational choice variables.

Chapter 1: Introduction

Rational choice remains one of the most prominent theoretical perspectives in criminology. From the early works of Adam Smith and Jeremy Bentham, to more recent empirical tests of economic models of criminal decision-making (e.g., Thomas, Loughran, and Hamilton, 2020), scholars have long been interested in the notion that crime is, on some level, the result of a choice made by a rational actor. Of course, there has been some controversy over what is meant by the term "rational." In many cases, rationality is (broadly) defined as the pursuit of activities which are believed to satisfy a person's unique set of needs and desires to the highest degree possible (Stanovich, 2010). Two common examples of such desires include the achievement of status among one's peers through offending (i.e., the social benefits to criminal acts), and the thrill or excitement an individual experiences while engaging in illicit activities (i.e., the personal benefits to crime; Katz, 1990; Loughran et al., 2016a). In deciding whether to involve oneself in criminal activities, individuals are assumed, on some level, to weigh the anticipated pleasures of the rewards to crime against the prospective pains of crime's risks and costs (e.g., the likelihood of arrest and loss of social ties due to imprisonment; Paternoster, 2010; Matsueda, 2013). According to rational choice theorists, the sum of these pleasures and pains determines the *expected utility* of criminal acts (Becker, 1968). The higher the expected utility of engaging in criminal activities for an individual, the more likely he or she should be to engage in some quantity of criminal actions.

The idea that human behavior is shaped by a utility calculus (i.e., the act of balancing pleasures against pains) is a basic underlying assumption of nearly every normative theory of rational decision-making. Although this assumption is most clearly presented within formal mathematical models of criminal choice (e.g., Becker, 1968), traces of the utility calculus can nonetheless be seen within more "reserved" theories of bounded rationality (e.g., Kahneman & Tversky, 1979). Indeed, such theories tend to at least presume that individuals often pursue behaviors which they personally believe, however erroneously, will produce more beneficial outcomes, on the whole, relative to losses (see: Cornish & Clarke, 1986). As such, most theories of rational choice not only suggest that individuals will, generally speaking, contemplate the rewards, risks, and costs to crime prior to offending, but in doing so each individual will directly weigh the overall quantity of benefits to criminal acts *against any potential risks or losses*. In other words, theories of rational choice allow for the possibility (if not outright *predict*) that the various (dis)incentives to crime may be *interdependent* on one another when shaping a person's overall level of involvement in criminal activities.

Despite this, relatively few (if any) prior examinations of the interdependent effects of the rewards, risks, and costs to crime can be found within the criminological rational choice literature. Instead, prior tests have mostly focused on estimating the *independent* (i.e., direct) effects of the perceived outcomes to crime on criminal behavior. Traditionally, such tests rely on linear regression modeling strategies which include various measures of participants' perceived rewards, risks, and costs as predictors of self-reported offending (e.g., Matsueda et al., 2006). If the rewards to crime are significantly and positively related to offending, and the risks and costs negatively related, then the test is cited as providing supportive evidence for the criminological rational choice perspective. Whether the perceived rewards to crime will, for instance, have the same effect across all values of the perceived risks and costs is rarely a concern for these sorts of tests. Furthermore, relatively few tests even feature the rewards to crime to begin with. Aside from a few notable exceptions (namely: Grasmick & Bursik, 1990; Loughran et al., 2016a, 2016b; Matsueda et al., 2006; Nagin & Paternoster, 1993; Piliavin et al., 1986; Thomas & Vogel, 2019; Thomas, Loughran, & Hamilton, 2020; Thomas, Baumer, & Loughran, 2022), most empirical assessments in this vein have focused almost entirely on the perceived risks and costs to crime, while neglecting offenders' reward perceptions almost entirely. As such, many purported tests of rational choice can often be more accurately described as tests of perceptual deterrence (Williams & Hawkins, 1986).

An immediate consequence of these oversights is that the more nuanced dimensions of the incentives-crime link have yet to be fully explored. Although contemporary research efforts have no doubt provided a great deal of insight into the more subtle ways in which the perceived rewards, risks, and costs of crime will influence criminal behavior (e.g., Thomas et al., 2020), the discipline is far from attaining a firm grasp on the subject. More specifically, we know little about the *combined influence* (i.e., potential interaction effects) different measures of perceived (dis)incentives may have on self-reported involvement in criminal activities, let alone whether said influence can be reasonably described as a *linear* function of perceptual measures (i.e., both the direct and combined effects of the perceived rewards, risks, and costs to crime may display *nonlinear* properties; Wood, 2017). For instance, an argument could be made that some rational choice theories (e.g., Gary Becker's model) seem to suggest that individuals who hold higher perceptions of the risks and costs to crime should be less "responsive," on the whole, to the perceived social and intrinsic benefits to crime (e.g., such individuals may choose to forego the pleasures of criminal acts in an effort to avoid the *pains of capture and social losses*). However, one could also argue that such theories allow for the possibility that the benefits to crime will have a generally *weaker* influence on criminal involvement for persons whose perceptions of the risks and costs are substantially *low*. That is, individuals who believe they are unlikely to be caught and arrested—much less endure any substantial social losses if captured—may be willing to offend *regardless of any social or intrinsic benefits criminal activities may offer them* (i.e., variation in the perceived benefits to crime may matter more for individuals who hold more "middling" perceptions of arrest risk and social costs to crime, relative to those at the extreme ends of either distribution). As such, it could be the case that the perceived risks and costs to crime will have a *non-monotonic* influence on the criminogenic effect of the perceived social and personal rewards to illegal acts.[1]

Of course, it should also be mentioned that prior examinations of the direct effects of perceived rewards, risks, and costs measures on offending behavior are not without value. Such efforts have undoubtedly been an important first step in demonstrating that criminal actions are guided, at least on some level, by offenders' perceived consequences to crime. A fruitful next step, as I will seek to argue in this dissertation, is to delve deeper into the potential interrelations of the perceived social and personal benefits, expected

---

[1] By "non-monotonic," I am referring to a situation where the "shape" of some underlying relationship (e.g., a real-valued *function* of $n \geq 1$ variables) is neither fully increasing nor decreasing with respect to each element within its domain. For example, we can think of some measure of the criminogenic influence of the social and intrinsic rewards to crime as a real number (e.g., the average "slope" of the rewards effect), the value of which is assigned by an underlying function, denoted by f, defined over the set of all possible values for both perceived arrest risk (P) and social losses (C) to crime (however P and C may be defined and measured). If it is the case that, for every possible combination of values for P and C, any "marginal" (i.e., arbitrarily small) increase in either P or C leads to a *smaller* number assigned by f, then f is said to be monotonic (specifically, f is strictly decreasing with respect to P and C). If, however, some marginal increase in either P or C leads to an *increase* in the value assigned by f for some points of its domain, while also leading to a *decrease* in f at other points, then f is neither strictly increasing nor decreasing across all points of its domain and is therefore non-monotonic. Another way to think about the monotonicity of f is to consider the sign of the *partial derivatives* of f with respect to P and C, denoted by $D_P f$ and $D_C f$, respectively. If either $D_P f$ or $D_C f$ take on both positive and negative values for any two (or more) unique combinations of P and C, then f is non-monotonic (see: Loomis & Sternberg, 1968, p. 156; Rudin, 1953, p. 95).

likelihood of arrest, and anticipated social costs to crime. In doing so, I will aim to examine three primary sets of hypotheses, each derived (primarily) from the rational choice model devised by Becker (1968), along with several implications provided by prior research on criminal decision-making more broadly. The first set of hypotheses relates to the interdependency of the perceived social and intrinsic rewards to crime on perceptions of arrest risk and the social losses to criminal acts. In particular, I anticipate that offenders' responsiveness to reward perceptions is likely to vary at different points of the overall continuum of the perceived risks and costs to crime. The second set of hypotheses outlines the "reverse" scenario, wherein offenders' responsivity to perceptions of arrest likelihood and the social losses to crime are presumed to vary with respect to the perceived social and intrinsic benefits to criminality. The third and final set of hypotheses relates to the interdependency of different "types" of reward perceptions; namely, the perceived social and personal rewards to crime. Specifically, I anticipate that the general strength of the association between the perceived social (personal) rewards may well vary with respect to the perceived personal (social) rewards.

The remainder of this manuscript will proceed as follows. Chapter 2 will provide further elaboration on the concepts, mechanisms, and broader implications of the rational choice perspective. In particular, emphasis will be placed on discussing the more subtle implications of Becker's (1968) expected utility model, as well as outline the rationale behind each of this project's primary hypotheses. Chapter 3 provides an overview of the analytic strategy employed to test each hypothesis, as well as the data-frame from which self-reported perceptions of the rewards, risks, and costs to crime, as well as involvement in criminal acts, were derived (specifically, the Pathways to Desistance Study). Chapter 4 provides a summary of the results obtained from the analysis outlined in the third chapter, and Chapter 5 gives a brief discussion on the broader implications of this project's findings.

Chapter 2: Relative Perceptions of Rewards, Risks, and Costs

Rational choice theories of crime stem from the classical school of criminology, which was built upon the early works of Enlightenment scholars. While contemporary research on criminal choice often traces its roots back to the scholarship of Jeremy Bentham and Cesare Beccaria,[2] an oft overlooked figure is Adam Smith. Referred to as the "father of economics," Smith laid the foundation for the rational choice perspective within two primary works. The first is *The Theory of Moral Sentiments*, wherein Smith (1759) discusses the philosophical, moral, and behavioral implications of humankind's many passions, of which include basic drives and emotional states such as hunger, fear, love, and physical pain, as well as socially bounded feelings of fellowship, altruism, and moral guilt (Ashraf, Camerer, and Loewenstein, 2005; Paternoster, Jaynes, & Wilson, 2017). These passions, according to Smith (1759), play a crucial role in the development of a society's moral standards, in addition to shaping the beliefs and activities of its participants. In his second volume, *An Inquiry into the Nature and Causes of the Wealth of Nations*, Smith (1776) introduced the notion of the self-interested economic agent; an individual whose purchases and labor practices are driven by the desire to accumulate wealth and satisfy personal needs and desires.[3] This notion, along with the idea that passions guide human behavior, has been the crux of nearly every scholarly discussion on rational choice to date.

---

[2] Usually, in discussing the origins of deterrence theory (see: Paternoster, 2010).

[3] There is some disagreement, however, as to whether Smith's latter volume (*Wealth of Nations*) makes contradictory assumptions about human nature when compared with his former work (*Theory of Moral Sentiments*). Much of this disagreement hinges on the "moral heckler" as presented in *The Theory of Moral Sentiments*; an impartial spectator that exists within the mind of every person to dissuade him or her from blindly following one's passions (i.e., the "angel" on the shoulder; Grampp, 1948). Because of this, some argue Smith imposes a level of self-interest in *Wealth of Nations* which is not seen in his earlier work, while others suggest the two can be seen as logical extensions of one another (see: Ashraf, Camerer, and Loewenstein, 2005).

One of the earliest extensions of Smith's work can be seen in Jeremy Bentham's *An Introduction to the Principles of Morals and Legislation*.

A household name among choice-oriented criminologists, Bentham (1789) was the first to formalize Smith's concept of the "self-interested rational actor" in a theoretical context.[4] As stated by Bentham (1789:1): "Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do." By "ought to do," Bentham refers to the utilitarian principle, which dictates that any action, policy, or sanction can be described as moral and *rational* in so far as it successfully maximizes pleasure and minimizes pain for the greatest number of persons (Sidgwick, 1874). By "shall do," Bentham establishes a formal theory of individual rational choice, wherein human behavior is guided by the pursuit of pleasure and avoidance of pain; or, put differently, the maximization of *utility* through one's actions. Bentham (1789:19) defines the utility of an action thusly: "Sum up all of the values of the pleasures on one side, and those of all the pains on the other. The balance, if it be on the side of pleasure, will give the good tendency of the act upon the whole, with respect to the interests of that individual person; if on the side of pain, the bad tendency of it upon the whole." The implications of Bentham's theory are twofold. First, the utility of any activity can be evaluated to the degree at which it produces happiness *relative to the amount of suffering* it causes within persons.[5] Second, individuals will pursue activities they anticipate will produce the greatest amount of utility for themselves and avoid those which produce the least.

[4] Paternoster (2010) describes Bentham's work as delivering not just one of the first formal theories of rational choice, but of crime as well (see also: Apel, 2013; McCarthy, 2002).
[5] This is sometimes referred to as the "hedonic calculus," which features prominently in economic models of decision-making (Varian, 2009).

Since Bentham, there has been an abundance of scholarship on rationality and decision making within the social sciences, nearly all of which is derived, on some level, from Bentham's original framework (Stigler, 1950a, 1950b, 1972). Although many scholars disagree on whether human decisions can truly be described as rational, however defined, it is rare to find a subset of this literature which portrays individuals as being non-agentic, thoughtless, or otherwise insensitive to the utility of their actions. Indeed, even perspectives which relax some of the stronger assumptions about rational behavior (e.g., theories of "bounded" rationality; Simon, 1959) still adhere to many of the basic notions outlined by Bentham and, by extension, Adam Smith. Criminological rational choice can be considered as one such perspective, under which falls the deterrence, offender decision making, and situational crime prevention literatures (Cornish and Clarke, 1986). Before delving into this body of work, I will first cover some early advancements on Bentham's theoretical framework. Namely, the development of utility theories in the disciplines of economics and cognitive psychology.

Expected Utility Theory and Criminal Decision Making

The expected utility model was developed by von Neumann and Morgenstern (1944) in their book *Theory of Games and Economic Behavior*. Similar to Bentham's theory, von Neumann and Morgenstern's model assumes that individuals act as utility maximizers and will thus pursue activities which provide the greatest rewards relative to the lowest overall level of risk and fewest costs. Under this model, the expected utility to a particular choice option—which I will define as any *course of action* a person is able to pursue, among some set of alternatives, from this point onward—can be represented by the following equation:

$$EU(A) = \sum_{k=1}^{K} P_{rk} U(R_k) - \sum_{l=1}^{L} P_{cl} U(C_l) \tag{1}$$

where $EU(A)$ denotes the expected utility of action A, R_k and C_l respectively denote the quantities of some reward type k and cost type l (e.g., the "amount" of thrill or excitement one experiences from A), and P_{rk} and P_{cl} denote the respective probabilities of R_k and C_l occurring given the individual chooses to pursue A (Matsueda et al., 2006). The $U(*)$ term can be thought of as a measure of the *utility value* of any specific reward or cost, such that the more pleasurable the reward, or more painful the cost, the greater the amount of (dis)utility an individual would derive from experiencing it (Fishburn, 1989). Within this model, the (dis)utility values of the prospective rewards and costs are each weighted by their probabilities of occurring, whereby the more likely some reward k (or cost l) is to occur, the more utility the individual can expect to gain (or lose) *on average* from choosing A. By extension, any type of reward or cost that has virtually zero chance of occurring will exert *little to no influence* on the value of $EU(A)$ (Mehlkop & Graeff, 2010). Thus, any pleasurable (or painful) outcome should only influence the choices made by an individual if there it at least *some chance of that outcome occurring*.

To get a sense of how this model would work in practice, consider the prospect of choosing between two lotteries: *A* and *B*. Lottery *A* offers a 50% chance to gain $100 ($0 otherwise), while lottery *B* offers the same chance to gain $110 (again, $0 otherwise). Since both lotteries share a similar outcome – namely, a 50% chance to gain nothing – we need only to compare the utility values of the non-shared outcome (gain of $110 versus $100) to determine the "better" option of the two.[6] Assuming the individual prefers a gain of $110

[6] This follows directly from the notion that a 50% chance to gain $0 should, theoretically, produce the same utility value regardless of the option it is associated with (i.e., $.5 * U(A|\$0) = .5 * U(B|\$0)$).

over \$100 (i.e., $U[\$110] > U[\$100]$),[7] then they should also prefer lottery B over A.[8] Of course, we can just as easily add a third lottery, *C*, which offers a 50% chance for a \$120 gain. Given the previously established preference of earning more money over less, lottery *C* should now be the preferred option over both *A* and *B*. Hypothetically, we could continue to add lotteries in \$10 increments for however many iterations desired (say, lotteries *D* and *E*, which offer a %50 chance of \$130 and \$140, respectively), and, assuming the quantity of lotteries remains finite, we can develop a clear and simple preference relation between each of them (e.g., lottery *E* should be preferred over lotteries *D*, *C*, *B*, and *A*). Of course, there is no reason to restrict ourselves to strictly monetary outcomes, as we can just as easily set up an analogous "lottery" for any given consequence, or set of consequences, the individual may be concerned with (e.g., engaging in an act of shoplifting, which could offer a great deal of excitement for a person, but also comes with a chance of being caught and subsequently punished). Because of this, many decision theorists—particularly, economists (see: Kreps, 2013)—have often relied on the expected utility model to explain a broad range of human activities including, though not limited to, involvement in criminal acts (e.g., Becker, 1968).

INTUITION BEHIND THE EXPECTED UTILITY MODEL:

---

[7] I have yet to meet someone who doesn't.

[8] Note that, in this example, we can identify the "preferred" lottery simply by comparing each lottery according to its expected monetary gain. For lottery *A*, the expected gain is equal to \$50 (\$55 for lottery *B*). We can compute this in a similar mannerism to computing the expected utility of both lotteries; that is, by summing the values of the monetary "gain" of each outcome, multiplied by its respective probability of occurring. More specifically, let M_s represent the monetary gain per state "*s*." The expected gain of lottery *A* would thus be: $\sum_{s=1}^{n(s)} P_s M_s$, which would give us: $.5 * \$100 + .5 * \$0 = \$50$. In economics, this is typically referred to this as the expected *value* of lottery *A*. The expected *utility* of lottery *A*, however, refers to the *subjective valuation of each outcome* (i.e., $U[*]$) multiplied by that outcome's likelihood of occurrence, the value of which could be very different from the expected monetary gain of a given lottery (for a more detailed discussion, see Kreps, 2013).

While prior discussions of the expected utility model in the field of economics are often heavily reliant on mathematical theorems and statistical reasoning,[9] the basic premise of the model is fairly straightforward and has an intuitive appeal that has allowed the model to maintain its popularity as a normative theory of human behavior in the social sciences to this day (Nagin, 2007; Thomas & Vogel, 2019; Tversky & Kahnemann, 1979). At its core, the model is little more than a mathematical representation of how individuals should make decisions, given they behave as "rational" human beings. Specifically, this process is modeled through the anticipation of future consequences to one's actions (McCarthy, 2002). By consequences, I am referring to any event (or, in economic terms, "state" of the world) that produces some level of pleasure (reward), or pain (cost), to the individual, and follows from a given course of action (Fishburn, 1981). Since consequences do not occur until after a decision to act is made, such decisions are formed on the basis of how pleasurable or painful (as well as likely) an individual *anticipates* the consequences of that action to be. As a result, current conceptualizations of utility theories (including most variants of rational choice theory) are mostly perceptual, wherein the expected utility to an action is captured via a person's subjective *beliefs* about how pleasurable or painful the consequences associated with that action will be (Savage, 1954). According to Piliavin and colleagues (1986, p.102), any rational theory of criminal decision-making "must consider those expectations [rewards and costs] as subjectively perceived by the actor, not as inhering in the actions." The expected utility model, then, can be thought of as a theory of how individuals should make decisions based on their perceptions of future pleasures and pains, regardless of whether those perceptions accurately reflect the true consequences a

[9] For a more in-depth discussion of said theorems, as well as the more technical aspects of most standard economic models of human decision-making, see Appendix A.

person would experience if they engaged in a particular action (Geerken & Gove, 1975; Pogarsky & Loughran, 2016).[10]

Within the expected utility framework, individuals are believed to harbor perceptions of multiple types of consequences—that is, rewards, risks, and costs—prior to choosing whether or not to pursue a particular activity. Two prominent examples include the potential social and psychic consequences one expects an action to produce. Social consequences can take the form of reputational gains or losses, such as respect by one's peers, as well as support or approval from key institutional figures, such as one's parents, teachers, or employers (Anderson, 1999; Matsueda et al., 2006; Paternoster, 1989). Psychic consequences refer to the amount of thrill or excitement an individual would experience while engaging in an activity, as well as the potential discomforts associated with formal punishments (Apel, 2013; Ehrlich, 1973; McCarthy, 2002). In deciding whether to engage in a specific action or not, an individual will consider all relevant consequences and directly weigh their associated pleasures against the pains (as shown in Equation 1).[11] All else equal, the higher the "net gain" from an action, the more the agent serves to benefit from pursuing that action. If, from the agent's perspective, there do not appear to be any better alternative actions available (e.g., a higher utility value to abstaining from illicit activities

[10] Becker's model, for instance, assumes by default that individuals form their own preferences toward certain activities, as well as expectations of the consequences associated with those activities, before making the decision to offend. What Becker seems to be interested in explaining, then, is less the origins of such expectations, and more of how individuals will *respond* to them.

[11] It should be mentioned that this is a bit of an oversimplification (one that criminologists often make when discussing economic theories of rational choice). For most standard economic models—particularly, *microeconomic* theories—the question of whether the agent is actively "thinking" about the consequences prior to acting is (mostly) left unanswered. Instead, the focus of such models is largely on determining how an individual *should behave* in the event he or she acts like a utility maximizer. Consequently, however "rash" or "thoughtless" an agent's actions are—or at least *appear as such* to any outside observer—is almost entirely irrelevant. The only question which matters, under this perspective, is whether the agent's actions can be accurately described as *optimal*, regardless of however much "deliberation" the agent exerts prior to making a decision. For further discussion, see: Kreps (2013).

in favor of strictly legal behaviors), then the agent should, more often than not, pursue that particular action over any given alternative (Eide, 1999). Put differently, the probability of selecting some action hinges *not* on the basis of whether that activity offers large individual quantities of rewards, or poses minimal risks or costs, in a general sense, but rather on the degree to which that action offers the largest *bang for one's buck* (i.e., the greatest sum of benefits *relative* to the risks and costs), compared to any alternative course of action. This notion has received some attention from various choice-oriented scholars, one of which being Gary Becker, who was the first to apply the expected utility model to criminal decision making.

APPLICATION OF EXPECTED UTILITY THEORY TO CRIME:

In his seminal piece, *Crime and Punishment: An Economic Approach*, Becker (1968) proposes a rational choice approach to modeling criminal behavior that is derived almost entirely from the expected utility theory developed by von Neumann and Morgenstern (1944). According to Becker (1968), criminal behavior can be thought of as having the same properties as any "economic" activity, and thus its commission can be understood purely through the various benefits and costs crime (and its legal alternatives) has to offer to the would-be criminal. Because of this, Becker (1968, p. 170) rejected many popular criminological explanations of crime of his time, stating: "… a useful theory of criminal behavior can dispense with special theories of anomie, psychological inadequacies, or inheritance of special traits and simply extend the economist's usual analysis of choice." By the "economist's usual analysis of choice," Becker (1968) refers to standard economic theory more broadly, of which the expected utility model falls underneath. As Becker (1968, p. 176) argues, individuals will offend when "the expected utility to him exceeds the utility he could get by using his time and other resources at other activities. Some persons become 'criminals,' therefore, not because their basic motivation differs from that of other persons, but because their benefits and costs differ."

This marked a considerable divergence in focus from many popular theoretical explanations of crime at the time. For many theorists, crime was to be primarily understood as a pathological characteristic of human society, one which either largely occurs under extreme circumstances (e.g., the breakdown of social norms) or is caused by some set of stable, individual differences between persons (e.g., Gottfredson & Hirschi, 1990). Under Becker's model, however, crime was neither an aberration, nor was it purely the result of a failure of one's moral character or an act of desperation; it was simply an alternative means to satiate one's needs and desires that violated society's rules (e.g., improving one's social or personal well-being through illegal acts). In other words, under this perspective there is nothing special about crime that separates it from any law-abiding activity. Understanding crime's prevalence in society, then, can be achieved simply by applying the mechanisms underlying *any form of behavior that provides benefits and costs to the individual*. As Becker (1968, p. 177) states:

> [T]here is a function relating the number of offenses by any person to his probability of conviction, to his punishment if convicted, and to other variables, such as the income available to him in legal and other illegal activities, the frequency of nuisance arrests, and his willingness to commit an illegal act. … An increase in either p_i [probability of conviction per offense] or f_i [punishment per offense] would reduce the utility expected from an offense and thus would tend to reduce the number of offenses because either the probability of 'paying' the higher 'price' or the 'price' itself would increase.

Needless to say, Becker's analysis aided in the resurgence of the classical school of criminology. Not only did this include a revitalization of research on public policy and deterrence (see: Nagin, 1998; Paternoster, 2010), but also the formation of an area of research devoted to investigating the role of perceived incentives (i.e., benefits, risks, and costs) in criminal decision-making more broadly. This latter area encompasses the bulk of research conducted within the rational choice literature, a fair portion of which consists of tests of Becker's expected utility model of offending (Paternoster, 2010). Such tests were largely interested in isolating the effects of perceived rewards, risks, and costs on offending behavior, with the general hypothesis being that expected benefits were positively associated with offending, while the risks and costs were negatively related. If individuals are truly drawn to that which they find pleasurable, and abstain from that which offers them pain, then a rational choice model of crime would predict that, holding everything else constant, the probability of crime increases with rewards and decreases with risks and costs (Becker, 1968). Several empirical examinations of this hypothesis have been conducted in criminology.

PRIOR TESTS OF BECKER'S THEORY OF RATIONAL CHOICE:

(Insert Table 1 about here)

The body of research on rational choice in criminology can be divided into studies of perceptual deterrence and broader tests of rational choice theories (Table 1 provides an overview of such studies which have utilized data derived from the Pathways to Desistance Study).[12] Both approaches draw heavily from Becker's (1968) model, but, as many choice-oriented scholars have pointed out, there are some key differences between the two. In particular, perceptual deterrence studies have focused largely on examining the effects of the perceived certainty of arrest and severity of formal punishment on offending behavior

---

[12] Early research on criminal deterrence focused on testing the influence of objective properties of sanctions (e.g., state-level clearance rates) on offending (Kleck et al., 2005). However, scholars since have accepted that deterrence is best seen as a perceptual theory, wherein a person's subjective expectations of the likelihood and severity of formal punishments influence their offending behavior (Williams & Hawkins, 1986).

(Paternoster, 1987; Williams & Hawkins, 1986). Such tests typically employ individual-level data, wherein participants' expected legal consequences for engaging in a variety of offenses are captured via survey items, and subsequently used to predict their self-reported involvement in criminal activities (Paternoster et al., 1982). If the results suggest that individuals who believe they are highly likely to be caught and arrested for engaging in crime display lower rates of offending compared to those who believe otherwise, then said results are cited as evidence favoring deterrence and rational choice theories (e.g., Bachman, Paternoster, & Ward, 1992; Erickson et al., 1977; Grasmick & Bursik, 1990; Grasmick & Milligan, 1976; Jensen, 1969; Paternoster & Simpson, 1993; Saltzman et al., 1982; Waldo & Chiricos, 1972). According to a review of the perceptual deterrence literature by Nagin (1998, p. 7), the majority of research "points overwhelmingly to the conclusion that behavior is influenced by sanction risk perceptions."

While prior research has mostly favored the existence of a negative link between perceived likelihood of arrest and offending behavior (e.g., the "certainty effect"), support for the effect of perceived severity of sanction on offending has been considerably weaker (Nagin, 1998; 2013). Some scholars, such as Grasmick and Bryjak (1980), have argued this is due to incorrect specifications of the rational choice model laid out by Becker. Namely, the effect of perceived severity of sanction should be "moderated" by expectations of the likelihood those sanctions will occur (e.g., perceived risk of arrest), such that perceived severity of sanction has a greater influence on offending behavior as the likelihood of arrest increases (Grasmick & Bryjak, 1980). While some research has shown this to be the case (e.g., Anderson et al., 1977; Teevan, 1976), other studies, such as that by Paternoster and Iovanni (1986), have been less supportive, showing that the effect of perceived severity tends to disappear when controlling for more informal means of social control (e.g., stigmatization) regardless of participants' perceptions of the likelihood of arrest. Furthermore, while most tests have shown a significant negative relationship between perceived certainty of arrest and the number of crimes a person commits, the effect size is often substantively small (Pratt et al., 2006). Because of this, some scholars, such as Paternoster (2010, p. 765), have suggested that the body of evidence supporting the deterrence doctrine "must be swallowed with a hefty dose of caution and skepticism."

Although research in deterrence has primarily focused on the roles of perceived certainty of arrest and severity of sanction in shaping criminal activity (i.e., the "risks and costs" side of the expected utility equation), some scholars have suggested this approach is incomplete. Piliavin and colleagues (1986) argued that, under a rational choice framework, offenders consider not just the risks and costs to their behavior, but the prospective *rewards* as well.[13] As such, Piliavin et al. (1986) suggest a more comprehensive approach to testing the criminological rational choice perspective would be to examine the influence of the various perceived benefits to crime, in addition to its risks and costs, on offending behavior. Using a sample derived from the National Supported Work Demonstration, the authors found participants' expectations of the potential illicit monetary rewards to crime significantly increased their involvement in crime, while their perceived likelihood of arrest had no effect. Similarly, Matsueda and colleagues (2006) tested the effect of the perceived psychic and social rewards, in addition to perceived risk of arrest,

[13] It is worth noting that some scholars believe there is a fundamental difference between the theory of criminal deterrence and broader theories of rational choice (e.g., Becker's expected utility model). Namely, the former explains crime solely through one's perceptions surrounding the criminal justice system (e.g., likelihood of arrest, pain of imprisonment, etc.), while more complete theories of rational choice tend to consider both the role of the perceived rewards and *informal* costs to crime as well (see: Loughran et al., 2016a; Paternoster, 2010).

on offending within a sample of “high risk” adolescents derived from the Denver Youth Survey. Their results suggested that, on average, persons who believed crime offered greater benefits (psychic and social), as well as lower risk (likelihood of apprehension) and fewer costs (pain of punishment), tended to commit more offenses.

Other tests have examined the effects of not just different types of rewards to crime, but more “informal” costs as well (i.e., those mostly unrelated to the criminal justice system). For example, Grasmick and Bursik (1990) investigate the role that the potential loss of respect from conventional others and feelings of moral guilt play in shaping an individual’s criminal proclivities. Using data derived from a probability sample of adults, the authors found that individuals who believed they would experience a greater degree of moral guilt or greater loss of respect from their closest peers were less likely to suggest they were willing to engage in theft, tax evasion, and drunk driving. This vignette approach to testing rational choice theories was also utilized by Nagin and Paternoster (1993), who attempted to examine the influence of informal costs and perceived thrills on willingness to engage in drunk driving, theft, and sexual assault. Using an undergraduate sample, the authors found that individual who thought they were more likely to be dismissed from their university, lose respect of their close friends or family members, and experience diminished career prospects, reported they were less willing to commit each of the three offenses. Those who believed they would receive a greater amount of “kick” or “fun” committing each crime type, however, reported overall higher levels of willingness to offend (Nagin & Paternoster, 1994).

A semi-recent test of Becker’s model of rational choice is that of Loughran and colleagues (2016a). In this test, the authors sought to determine whether Becker’s theory can be reasonably defined as a general theory of crime, capable of effectively explaining both instrumental (e.g., larceny, robbery, tax evasion, etc.) and expressive crimes (e.g., assault, fighting, vandalism, etc.) among a "high-risk" sample of serious adolescent offenders. The authors employed a series of fixed-effects estimations of the overall influence of participants' perceptions of the social and psychic benefits to crime, perceived likelihood of arrest, and perceived social costs on the variety of offenses those participants committed. The authors also included a lengthy list of controls within their model, which enabled them to control for both time-varying and time-invariant characteristics (Loughran et al., 2016a). The results of this test, overall, suggested that an individual's variety of offenses committed tended to coincide with increases in their perceived social and personal benefits, and decreases in the expected probability of arrest, over time. The expected social costs, however, were not significantly related to offending. The results of this test led the authors to conclude "that rational choice theory is as general a theory of crime as are social learning, social control, and strain theories" (Loughran et al., 2016a, p. 107).

Another test using the Pathways data was that of Thomas, Loughran, and Hamilton (2020). Here, the authors were interested in examining whether a rational choice theory of crime could predict the specific *types* of offenses individuals commit, as well as the decision to specialize in some types of crime (e.g., violent crimes) over others (property crimes). Employing a series of logistic and hierarchical latent trait models of specialization (see: Osgood & Schreck, 2007), the authors found that not only were individuals more likely to engage in crime types they perceived to be more personally rewarding and less risky (relative to other types), but also tended to specialize in either violent or property offenses more broadly on the basis of those perceptions. And finally, a recent examination by Thomas, Baumer, and Loughran (2022a) found support for the notion that individual-level perceptions of, as well as *preferences* toward,[14] the rewards, risks, and costs to crime are driven, on some level, by broader social factors. This examination also utilized data derived from the Pathways to Desistance Study, where it was shown that living in a structurally disadvantaged neighborhood was associated with lower reported perceptions of arrest probability and the social costs to crime, as well as higher perceived social and personal benefits to crime. Additionally, the authors found that higher perceptions of neighborhood disorder, as well as a lack of access to legitimate means for achieving success, were both related to a lesser sensitivity to crime's risks and costs, and a greater sensitivity to the benefits to crime.

WHERE DO WE GO FROM HERE?

Of course, this is merely a glimpse into what choice-oriented criminologists have been able to accomplish since the resurgence of the rational choice perspective. Indeed, much has been learned not just about whether rewards, risks, and costs to crime are (statistically) related to criminal acts in a purely broad sense, but also many of the more nuanced implications of normative theories of decision-making for the study of crime. For instance, scholars have shown that offenders, generally speaking, seem to "update" their perceptions of arrest risk in a way that mimics a Bayesian style of learning (see Anwar & Loughran, 2011), while others have found evidence supporting the notion that a person's "sensitivity" to perceptions of risk and reward often varies with respect to one's moral beliefs (Paternoster & Simpson, 1996), criminal propensity (Nagin & Paternoster, 1994;

[14] By "preferences," the authors are referring to the idea that individuals may respond differently to different types of incentives. In particular, some may find the risks to crime noxious (i.e., are *deterred* by risk), while others may be indifferent to the risks to crime, and others still might even *seek out* risky activities for the thrill of doing so.

Pogarsky, 2007), and overall tendency to think through the consequences of one's actions both prior to and after acting (Paternoster & Pogarsky, 2009). Others still have investigated the role that ambiguity could play in influencing responsiveness to perceived arrest risk (Loughran et al., 2011), as well as the measurement properties of probabilistic measures of said risk perceptions more broadly (Hamilton, 2023; Thomas et al., 2018). On the whole, the discipline has come a long way in terms of better understanding the criminal decision-making process, and how this understanding may translate into solutions for public policy (Apel & Nagin, 2011).

Despite these advancements in the literature, there are still many questions left unanswered with respect to the broader implications of the utility calculus (as outlined in Becker's model). For instance, criminologists have yet to rigorously examine whether the perceived rewards, risks, and costs to crime will exert a combined influence of some kind on a person's level of involvement in illegal activities. Worse still, most prior discussions within the literature seem to only hint at what scholars might be able to anticipate in terms of the *directionality* of any such interrelation(s), meaning it is not yet clear as to whether we should expect the rewards to crime, for instance, to exert an overall weaker or stronger influence on criminal behavior as perceptions of the risks and costs to crime increase (or vice versa). Instead, thus far the discipline has seemed largely content to simply toss a set of reward, risk, and cost measures within a regression model of some kind and call it day (insofar as devising empirical examinations of a rational choice theory of crime). Any potential interrelations between such measures being are all but completely ignored, despite theory suggesting that not only such interrelations are likely to exist but may also follow a particular underlying *pattern* of some variety. As such, the remainder of this chapter will focus on delving into what theory (as well as prior research more broadly) might have to say about the *interactive effects* between the rewards, risks, and costs on offending behavior, as well as devise a series of hypotheses which will be examined using the methods outlined in the next chapter.

Interdependency of Rational Choice Constructs

Much of the debate, and controversy, within the broader criminological discourse on rational choice theories of crime resides in the *degree* to which perceived incentives influence a person's level of criminality (e.g., an individual's *preferences* toward a particular rational choice input; see also: Thomas et al., 2022a, 2022b), as well as the conditions under which incentives will differentially shape a person's level of involvement in crime (Piquero et al., 2011). Such conditions include the situational characteristics surrounding criminal decisions (e.g., exposure to criminal opportunities and broader structural factors; Osgood et al., 1996; Thomas et al., 2022a), and the existence of time-stable, individual differences that influence "rational" decision making competency (Piquero & Pogarsky, 2002; Pogarsky, 2002, 2007; Wright et al., 2004). What has received far less attention, however, is whether a person's responsivity to the various (dis)incentives to crime could be explained, in part, by his or her underlying utility calculus. In other words, it is fully possible that, within a rational choice framework, perceptions of the rewards, risks, and costs to crime will have a *combined influence* on criminal acts in some capacity. As an example, individuals may find themselves more (or less) easily seduced by the pleasures to crime depending on their *perceptions of the risk of apprehension and social costs to criminal acts*.

INTERDEPENDENCY OF REWARDS ON THE RISKS AND COSTS:

Prior research on the benefits to crime largely suggests that not only are perceptions of the social and intrinsic rewards positively associated with criminal acts, but said rewards may also have a *stronger impact* on a person's criminal tendencies relative to his or her expected risks and costs to crime (e.g., Loughran et al., 2016a, p. 102). That is, variation in measures of benefit perceptions can sometimes "outperform" measures of the anticipated probability of arrest and perceived severity of sanction when placed in regression models which aim to predict sample participants' self-reported involvement in illegal behaviors (Apel, 2013; Carroll, 1978). A commonly offered explanation for this phenomenon is the notion of *time discounting* (see Loughran, Paternoster, & Weiss, 2012a), in that, since the rewards to criminal acts are reaped at the moment of commission (e.g., the offender derives *immediate* personal satisfaction from engaging in shoplifting; Katz, 1988), it may be the case that criminal actors will "weigh" the benefits to crime to a heavier degree within their choice calculi relative to the risks and costs (which, per the nature of the justice system, tend to occur sometime *after* the crime has already been committed; Nagin & Pogarsky, 2001, 2004). Others have suggested this preferential treatment offenders often extend to the pleasures of criminality could also be due to a generally "thoughtless" nature of crime and criminals, such that most offenders are simply unconcerned with the potential risks and losses to their actions and thus tend to gravitate toward whatever is most pleasing to them in the current moment (Gottfredson & Hirschi, 1990; Hirschi, 1986). In other words, the offender, by and large, is someone for which the anticipated painful outcomes to crime have *little to no influence* on his or her utility calculus, and thus we should expect the average criminal to be more easily swayed by the *seductions of criminality* (Katz, 1988).

While some scholars have argued that a general tendency for offenders to be "less responsive" to the risks and costs of crime does not contradict the basic tenets of the rational choice perspective (e.g., McCarthy, 2002; Nagin & Paternoster, 1993, 1994),[15] such arguments often neglect the role said risks and costs could potentially play in shaping the extent to which variation in the perceived rewards to crime will impact a person's criminal proclivities. If anything, much of the scholarly discourse on this subject seems to hinge on the (mostly implicit) assumption that the criminogenic influence of the rewards to criminal acts is, for the most part, *independent* of an offender's perceptions of the risks and costs to crime. At the very least, such an assumption is integral to the notion that offenders *rarely consider the possibility of capture*—let alone any negative consequences following an arrest (e.g., loss of social status)—prior to committing an offense in lieu of focusing almost entirely on reaping the immediate benefits to criminal acts (see also: De Haan & Vos, 2003). If this is indeed the case—that is, the offender rarely "balances" his or her anticipated pleasures of criminal acts against the pains of crime's risks and costs in a way which aligns with Becker's (1968) expected utility model—then we should expect to see *little to no interrelation* between the overall influence of the perceived rewards to crime on criminal behavior and variation in the perceived risks and costs to illicit activities (e.g., the "strength" of the association between measures of reward perceptions and self-reported offending should remain more or less *constant across all values of some set of perceived risk and cost measures*). By contrast, if offenders tend to behave like utility maximizers

[15] For instance, the notion of time discounting stems directly from economic theories of decision-making, and thus one can presume that criminals may "discount the future," to some degree, while *also weighing the benefits to crime against the (discounted) risks and costs*. Put differently, although a person's level of "present orientation," so to speak, may shape his or her *preferences* toward the benefits, risks, and costs to crime, he or she is still nonetheless presumed to respond to those (dis)incentives in a way which mimics standard economic conceptions of utility maximization. For further discussion, see Frederick, Loewenstein, and O'Donoghue (2002), and Gollier (2002).

(i.e., as individuals who pursue criminal acts when it is generally in their best interest to do so; Chalfin & Tahamont, 2018), at least on some level, then we should instead anticipate such an interrelation to not only exist, but also follow a particular *functional form* when examined using data.[16]

To get a sense of the type of relationship one might expect to observe between some underlying "rewards effect" and the perceived risk and costs to crime, we can consider the following decision rule outlined by Matsueda (2013, p. 392):

$$if\ EU(Y) > EU(N), \qquad then\ Y = 1, \tag{2}$$

where Y denotes the decision to engage in some criminal act (i.e., Y equals 1 if the agent pursues Y, and equals 0 otherwise), and N denotes some strictly legal alternative course of action. If we think of (2) as depicting the choice of whether to engage in, say, at least one act of burglary within a particular time frame (e.g., a 6-month window), then the above expression simply states that the agent will pursue a single act of burglary (or more) in the event he or she believes the *net gain* to burglary outweighs that of any legal alternative the agent could otherwise devote his or her time and energy to (for a more in-depth discussion of the technical aspects of Equation 2, see Appendix A). Since Becker's (1968) model assumes the expected utility of an illegal action is a decreasing function of

---

[16] By the phrase "functional form," I am referring to the general *shape* of some underlying joint distribution between two (or more) variables. A relatively common example of this concept would be the "line" fitted from an ordinary least squares regression model, wherein the predicted value of some response variable is determined by a linear combination of each predictor (Wonnacott & Wonnacott, 1970). By default, such models assume a *linear functional form* of the relationship between each covariate and the outcome variable (as opposed to a *curvilinear* model of some variety, for which higher values of some "x" measure are presumed to have a *smaller marginal impact* on "y;" Wood, 2017). For the current discussion, we can think of the "strength" of some underlying link between the rewards to criminality and self-reported involvement in crime (e.g., the estimated slope coefficient β for some reward measure) as the "outcome," and the perceived risks and costs as predictors. If the value of β tends to generally decrease (or increase) at higher reported values of the risks and costs to crime, then we can say the functional form of β is generally decreasing (increasing) with respect to the perceived risks and costs of illegal activities. For a more in-depth discussion, see: Loughran and colleagues (2012c), and Pogarsky (2007).

its risks and costs (as shown in Equation 1), we can imagine a scenario where the agent is willing to forego the benefits to Y (i.e., burglary) due to anticipating a particularly high risk of apprehension and a substantial loss of social standing if captured. In fact, depending on how "averse" the agent is to the negative consequences of crime (see McCarthy, 2002), he or she might even be willing to turn down a criminal opportunity that is believed to offer a *substantially large quantity of pleasure* (e.g., high intrinsic rewards).

To illustrate this notion, consider the choice of whether to accept a lottery (similar to that which was discussed near the beginning of this chapter) where the agent pays some amount of cash up front in exchange for a chance to win a prize of some kind. According to von Neumann and Morgenstern (1944), the set of "acceptable" prizes the agent is willing to risk some amount of cash for largely depends on the size of the bet (i.e., the "cost" to play). If the risk to some lottery is relatively minimal (e.g., the agent has a 50% chance to win a bet he or she pays $20 to enter), then the agent will likely pursue said lottery even if the prize is somewhat trivial (e.g., an item worth $50). However, if the agent perceives the lottery to be extremely risky (e.g., a 5% winning odds for a bet that costs $10,000 to enter), then he or she will probably only be willing to enter that lottery in the event the potential gain is believed to be worth the possibility of such a hefty loss (e.g., a payout of $1,000,000). Likewise, within the expected utility framework (under which falls Becker's model) the risks and costs to any action the agent can take are hypothesized to determine the *minimal expected returns* the agent will, more often than not, require in order to pursue it (Kreps, 2013, 2023). Put differently, the more the agent believes the act of burglary carries a high probability of arrest and substantial social losses, the *greater the sum of social and intrinsic benefits* (along with any other beneficiary incentives) the agent will demand to receive before he or she would be willing to engage in it (i.e., the pains of the risks and costs must be adequately *compensated* by the returns; De Mesquita & Cohen, 1995; Gollier, 1999). As such, a potential hypothesis related to the interdependency of the criminogenic influence of the perceived rewards to crime on the risks and costs is the following:

> ***Hypothesis 1a***: *Variation in the perceived social and intrinsic rewards to crime will exert a weaker influence on criminal behavior for individuals who harbor higher perceptions of the likelihood of arrest and social costs to crime.*

In applying Hypothesis 1a to our previous example outlined in Equation 2, we should expect it to generally be the case that the higher a person's perceived probability of arrest and social losses to the act of burglary, the *less responsive* he or she should be to the anticipated social and intrinsic rewards to burglary (i.e., the likelihood that the agent engages in at least one act of burglary should be relatively *low* even at higher reward perceptions). In other words, the strength of the association between the perceived rewards to criminality and involvement in illegal activities should, in general, be a *decreasing function* of the perceived risks and costs.[17] Of course, this is merely one possible hypothesis which can be derived from the expected utility framework (namely, Becker's [1986] model), as one could also examine a scenario where the agent believes there to exist *almost zero possibility of arrest or social losses to some illicit action*. Here, the agent's expected utility to, say, burglary might now purely be a function of the anticipated returns (see Matsueda, 2013, p. 392), meaning the agent may find it to be "worth" his or her time to engage in at least one act of burglary *even if the overall social and intrinsic benefits to*

[17] It should be mentioned that the concept of a "decreasing" function has a precise meaning in the study of mathematical analysis (see Rudin, 1953). For our purposes, we can simply think of a decreasing function as one which assigns generally lower values to some outcome variable (Y) as the value of some predictor (X) increases.

burglary are believed to be fairly minimal. That is, relatively small reward values to criminal acts may well produce a "good enough" return on investment at sufficiently minimal quantities of the risks and costs, and thus variation in perceptual measures of the benefits to crime may, paradoxically, be *less strongly related* to criminal behavior at lower values of the perceived risks and costs. This notion can be summarized by:

> ***Hypothesis 1b****: Variation in the perceived social and intrinsic rewards to crime will exert a stronger influence on criminal behavior for individuals who harbor higher perceptions of the likelihood of arrest and social costs to crime.*

Note that Hypothesis 1b is effectively the *opposite* prediction of 1a, such that the criminogenic influence of reward perceptions is now anticipated to be a generally *increasing function* of the perceived risks and costs to crime. The basic idea behind this is that individuals should be willing to accept *any quantity of benefits* at lower risk and cost perceptions, meaning higher reward values are likely to be superfluous (i.e., for as long as $EU(Y) > EU(N)$, as outlined in Equation 2, then any marginal increase in the rewards to Y will have no observable impact on the agent's choice to pursue at least one instance of Y). However, agents who harbor *higher risk and cost perceptions* must be more "discerning" over the benefits offered by any given criminal opportunity, as the sum of those benefits could mean the difference between an agent believing he or she is best off foregoing said opportunity or pursuing it. Additionally, it may also be possible that *both* hypotheses have some degree of truth to them, in that it may generally be the case that the effectiveness of the rewards on criminal acts is at its peak for some *middle-most* set of values for the perceived risks and costs to crime. That is:

> ***Hypothesis 1c****: Variation in the perceived social and intrinsic rewards to crime will exert the strongest influence on criminal behavior for individuals who harbor more middling perceptions of the likelihood of arrest and social costs to crime.*

By "middling," I am referring to any reported value of arrest risk or social losses to illegal activities which falls somewhere between the "extreme" ends of the distribution (e.g., a reported likelihood of arrest of 30%, rather than either 0% or 100%). Thus, Hypothesis 1c states that the rewards to crime are likely to have the smallest impact on a person's decision to involve his or her-self in criminal acts at both the *lowest and highest reported values of arrest probability and social costs to crime*. The argument for this is simply a combination of Hypotheses 1a and 1b, such that sufficiently high risks and costs are likely to turn the agent off from crime regardless of his or her reward perceptions, while sufficiently *low* risks and costs are likely to result in the agent being willing to accept nearly any quantity of benefits in order to offend. As such, all that is really being said by Hypothesis 1c is that the rewards are likely to matter most for individuals who perceive generally "higher" levels of risks and costs, *but not too high as to deter them completely*. Methods for examining each of these hypotheses will be discussed in greater depth in the next chapter.

INTERDEPENDENCY OF THE RISKS AND COSTS ON THE REWARDS:

Another (closely related) issue within the offender decision-making doctrine is that of differential *deterrability*. More specifically, prior research has shown that variation in the perceived risks and costs to crime can sometimes elicit a stronger influence on the criminal proclivities of some individuals compared to others (Herman & Pogarsky, 2022; Pogarsky, 2007). So-called "incorrigible" offenders are those for which no amount of (state-induced) risks and losses are likely to prevent them from offending, while "acute conformists" are individuals who will likely avoid engaging in criminal acts *regardless of the risks or costs involved* (Pogarsky, 2002). Deterrable offenders, then, are individuals who are averse to the certainty and severity of sanctions (including *informal* sanctions such as social losses; Apel & DeWitt, 2018; Williams & Hawkins, 1986), but *not so averse as to forego criminal opportunities which pose any amount of risk, however small, to the decision-maker* (Zimring & Hawkins, 1968). Traditionally, the mechanisms undergirding the deterrability—or lack thereof—of the offender are presumed to mostly exist *independently* of his or her "rational" proclivities (e.g., a general orientation toward the present; Hirschi, 1986). As stated by Jacobs (2010, p. 417):

> *If deterrence describes the perceptual process by which would-be offenders calculate risks and rewards prior to offending, then deterrability refers to the offender's capacity and/or* ***willingness to perform this calculation****. The distinction between deterrence and deterrability is critical to understanding criminality from a utilitarian perspective.* [Emphasis added in bold]

Here, we are provided with yet another example of the all-too common presumption within the criminological study of rational choice and deterrence that differential responsiveness to incentives predominantly stems from some *non-rational* mechanism. That is, individuals for which the risks and costs to crime seem to have little "sway" over their criminal activities are presumed to have some other (ostensibly non-utilitarian) motivation for either their involvement in criminal activities or lack thereof. For instance, Pogarsky (2002) discusses the role of visceral influences (i.e., "hot" emotional states; Bachman, Paternoster, & Ward, 1992), as well as a general tendency to forego thinking through the benefits and costs to one's actions prior to offending (i.e., some offenders are highly *impulsive*; Mamayek, Loughran, & Paternoster, 2015), as potential explanations for why some offenders appear to be "impervious" to sanction threats. By contrast, individuals for which there is little "need" to employ the power of the state to get them to conform to society's rules may do so because of, say, some deeply held moral belief system which forbids the commission of illegal activities (Etzioni, 1988; Grasmick & Bursik, 1990; Grasmick et al., 1993). Such moral beliefs are presumed to alter the set of available options for the individual, whereby criminal opportunities, as well as the associated benefits, risks, and costs of those opportunities, are deliberately “ignored” by the individual (i.e., illegal acts are avoided not because they fail to offer more pleasures than pains, but rather due to *never entering a person’s utility calculus to begin with*; see Paternoster & Simpson, 1996).[18] Put simply, the bulk of prior research on this subject has been dedicated to understanding the non-instrumental reasons—namely, those unrelated to perceptions of risk and reward—for why a person may appear to be “unresponsive” to the risks and costs to crime (Jacobs, 2010; Lee et al., 2018; Loughran, Paternoster, & Piquero, 2018; Pogarsky, 2007).

Similar to that which was discussed in the previous section, we can establish a set of predictions related to the interdependency of the perceived risks and costs to crime on the anticipated benefits using the decision rule outlined in Equation 2. Since the expected utility to any illicit act (e.g., burglary) is assumed to be a positive function of its associated benefits (per Equation 1), we can imagine a scenario where the anticipated social and intrinsic rewards to criminality are substantively “high” enough to where the agent may be willing to *tolerate higher levels of risk and expected losses* (De Mesquita & Cohen, 1995). In other words, the incorrigible offender need not just be an individual who is effectively “strongarmed” into offending against his or her better judgment (whether by his or her own

[18] There is, of course, a counterargument to this notion in that moral beliefs can simply be viewed as a particular manifestation of the agent’s underlying utility function (including any preferences he or she may hold toward certain outcomes). As such, any feelings of “guilt” the agent experiences by engaging in actions he or she knows to be wrong can itself be seen as just another anticipated *painful outcome* to criminality. Such pains can, in theory, be adequately offset by a large enough quantity of benefits, such that for the right “price” the agent may be willing to pursue behaviors he or she may not be entirely proud of. For further discussion, see: Kreps (2023, p. 35).

lack of impulse control, a general absence of legitimate alternatives, or provocation by delinquent peers; see also: Matthews & Agnew, 2008). Rather, it may also be the case that criminals who are more “difficult” to deter, relatively speaking, are those who believe the rewards to offending are simply worth the possibility of arrest *even when they believe an arrest is highly likely to occur*. Thus, even fairly “risky” offenses may still be a *rational* pursuit for individuals who hold particularly high reward perceptions for those offenses (see Cohen, 1970). As a consequence, variation in perceptions of arrest risk, as well as social losses, may not elicit as strong of an impact on criminal behavior for such persons *relative to those who hold lower reward perceptions on average*. This can be summarized by the following:

> ***Hypothesis 2a****: Variation in perceived arrest risk and the social losses to crime will exert a weaker influence on criminal behavior for individuals who harbor higher perceptions of the social and intrinsic rewards to crime.*

Although relatively minimal scholarship within the discipline has sought to directly examine the idea that individuals higher in reward perceptions might be less responsive to the risks and costs to crime, some studies have at least pointed in a similar direction as that which is indicated by Hypothesis 2a. In particular, Elijah Anderson (1999) discusses the role of respect—specifically, the anticipation of what a *loss of respect* would mean for the agent—in shaping a person’s choice calculus in his study on “street codes.” According to Anderson (1999), respect is earned (as well as *maintained*) by engaging in violent acts against individuals who appear to “disrespect” the agent within certain inner-city areas (i.e., neighborhoods). In many instances, individuals living within these areas believe that a failure to maintain his or her status among other residents can, and often *will*, lead to him or her becoming a victim of violence at some point in the future. As such, the anticipated social benefits to violent acts—or rather, the *prevention of a loss in status*—for such individuals can be so "enticing" that the prospect of receiving an arrest, however (un)likely, is often insufficient to dissuade them from offending (Anderson, 1999).[19] Hence, we have at least one example of a possible "weakening" influence of the perceived risks and costs to crime as the agent comes to view criminal acts as offering a greater quantity of benefits (namely, the social benefits).

Of course, one could also imagine a scenario wherein higher benefit perceptions lead to a *stronger* overall inhibitory influence of variation in the perceived risks and costs to crime on one's involvement in illegal activities. More specifically, it may be the case that individuals with *lower* reward perceptions, on the whole, are likely to *avoid engaging in crime regardless of the risks and costs associated with any given criminal opportunity*. Such individuals would align more closely with the concept of "acute conformists," as discussed by Pogarsky (2002), in that their apparent "insensitivity" to, say, variation in perceived arrest risk is not due to a general willingness to offend despite the risk, but rather to avoid criminality *even if the odds of capture seem substantively low to them*. In simple terms, the acute conformist may simply be someone for which the benefits to crime are so minimal as to not be "worth the effort" in any capacity, meaning even relatively "small" anticipated risks and losses are likely to be sufficient to successfully deter the individual entirely (see also: Chalfin & Tahamont, 2018, p. 46). Such a notion is summarized by:

---

[19] It should be noted that the avoidance of a "loss" of some kind (e.g., prevention of a reduction in social status) is usually treated as a *beneficiary* outcome by economists. More specifically, by pursuing a course of action which avoids a painful consequence, which would otherwise be produced by some alternative option (e.g., loss of respect from "walking away" from a provocateur), the agent's behavior thus mimics that of a utility maximizer. Simply put, the outcomes to one's actions do not need to be *actively pleasurable in to benefit the agent in some way*. For further discussion, see: Kreps (2013).

> ***Hypothesis 2b****: Variation in perceived arrest risk and the social losses to crime will exert a stronger influence on criminal behavior for individuals who harbor higher perceptions of the social and intrinsic rewards to crime.*

Furthermore, we can envision a situation where both hypotheses are "true" to some extent (similar to that of Hypothesis 1c discussed in the previous section). Namely, it may be the case that variation in risk and cost perceptions matter more for individuals who hold more middling perceptions of the social and intrinsic rewards to crime, while having less of an influence on criminal behavior near the lowest and highest reported levels of benefit perceptions. This leads to the following:

> ***Hypothesis 2c****: Variation in perceived arrest risk and the social losses to crime will exert the strongest influence on criminal behavior for individuals who harbor more middling perceptions of the social and intrinsic rewards to crime.*

Note the similarity of Hypotheses 2a through 2c to those of 1a to 1c, as outlined in the previous section. That is, there is assumed to exist some "tipping point," with respect to an individual's overall perceived benefits to criminality, at which variation in risk and cost perceptions (or reward perceptions, in the case of 1a through 1c) are likely to have the largest impact on one's involvement in illegal acts. If, for instance, said point exists at an extreme-most value for some set of rewards measures (e.g., the lowest reported values for the social and intrinsic rewards to crime), then all other reported values should, generally speaking, see a *generally weaker inhibitory influence of some set of measures of the perceived risks and costs to crime*. However, if for most participants this tipping point is located more toward the center of some distribution of reported rewards values, then the inhibitory effect of the risks and costs should steadily *decrease* as one begins "moving away" from the central-most values. Likewise, neither set of hypotheses (among those provided up until this point) implies the existence of a *linear* moderating relationship between some set of perceptual measures of the rewards, risks, and costs to crime on criminal behavior. More concretely, it is not only possible for participants' "responsiveness" to some set of perceived (dis)incentives to change in a *curvilinear* fashion (e.g., the deterrent effect of the risks and costs to crime may change *exponentially* as reward perceptions increase), but also for said responsiveness to *change direction* at some middle-most reported value (e.g., the deterrent effect generally increases up until a certain reported value for the social and intrinsic rewards, but then sharply decreases beyond said value). In examining such a phenomenon, many researchers will employ a *nonlinear* (i.e., nonparametric; Manski, 2003) model of some kind (Wood, 2017). A more in-depth description of such a model, along with how the model will be used to examine each hypothesis provided throughout this chapter, will be provided in Chapter 3.

INTERDEPENDENCY OF THE SOCIAL AND INTRINSIC BENEFITS:

Finally, the decision rule outlined in Equation 2, as well as prior research on the subject of the rewards-crime link, may also have implications for the existence of a set of interrelations between different *types* of perceived benefits to crime. Indeed, a common point of discussion among choice-oriented scholars is that there are multiple types of benefits a person can expect to receive by engaging in crime, two prominent examples of which being the psychic and social benefits (Loughran et al., 2016a; Nagin and Paternoster, 1993). Although relatively little scholarship has discussed the potential interactive effects between the social and intrinsic benefits to illegal actions, prior research within the discipline (as well as the study of *microeconomics*) may nonetheless provide some insight into this subject. For instance, in their book, *The Reasoning Criminal*, Cornish and Clarke (1986) discuss what they believe to be the general structure of the criminal decision-making process. Specifically, this process begins with what is referred to as the "criminal involvement" phase, wherein the agent weighs the various pros and cons to both legal and illegal actions and comes to a general conclusion as to whether he or she should become initially involved in illicit behaviors (Cornish & Clarke, 1986, pp. 2-3). After an involvement decision is made (e.g., the agent decides his or her best course of action is to commit at least one act of burglary), the agent must now evaluate the situational factors associated with any given opportunity which presents itself to him or her (e.g., an unattended home with lackluster security measures; Cornish & Clarke, 1986, pp. 4-6). Such a process constitutes the "criminal event" phase, wherein the individual decides whether or not to pursue a particular instance of a criminal act (e.g., whether to burglarize a particular home or not).

An important component of this decision-making process, according to Cornish and Clarke (1986), is the *goal-oriented* nature of crime and criminals. That is, in both the involvement and event phases the offender considers not just the *quantity* of the potential benefits (as well as the risks and costs) to some illegal activity, but also the *types of rewards* as well. If the offender believes an act of burglary will meet his or her financial needs to a higher degree in comparison to, say, working a steady job, then he or she may choose to pursue some number of burglaries *for the explicit purpose of reaping the monetary benefits to those burglaries*. Consequently, we might anticipate that variation in his or her anticipated monetary benefits to the act of burglary will carry a greater "weight" within his or her choice calculi relative to *any other benefit type* (e.g., intrinsic thrills or social benefits) the agent may be concerned with. In other words, if burglary is sought after primarily for its financial returns, then other benefit types should be seen as *secondary* pleasures which, although may serve as an additional "seduction" to burglary, will likely have less of an impact on a person's decision of whether to pursue a particular act of burglary or not. Note that the same could also be said about any other reward type, such as the social or intrinsic rewards to crime. If the agent pursues burglary primarily for the thrills it offers, then he or she is likely to be *less concerned about the social or monetary returns*.[20]

The suggestion that individuals will become involved in crime in order to achieve a specific goal (e.g., elicit a specific reward type) has been echoed among other choice-oriented scholars as well. In his book, *Seductions of Crime*, Katz (1988) provides an overview of the various attractive qualities of crime, and the visceral psychological appeal crime can have for the individual. Among these qualities include the "sneaky thrills" associated with certain forms of theft, such as shoplifting or burglary. According to Katz (1988), crime can be exciting for the individual to engage in for a number of reasons. One such reason is the violation of social rules, such that the offender, in getting away with theft, receives the psychological appeal of stepping into "forbidden" territory and navigating throughout it without others knowing (Katz, 1988). Additionally, the potential dangers of engaging in theft provides a feeling of tension for the offender, wherein the "payoff" to the offense lies in coming up with strategies to avoid detection, and then successfully carrying them out. This produces a feeling of rush or excitement for the

[20] A similar concept in microeconomics is that of the *substitutability* of two or more goods or activities. Namely, goods which are difficult to "replace" with some alternative option are often those which the consumer will continue to purchase despite, say, a price increase of said good (e.g., the price of gasoline rarely influences purchasing decisions due to gas fulfilling a specific need not offered by any potential substitute; Varian, 2009). For this section, we can think of the decision to pursue acts of, say, shoplifting for the purpose of achieving thrills as those which the offender may find difficult to "replicate" with other (legal or nonlegal) forms of behavior. As a consequence, he or she is likely to continue to engage in shoplifting regardless of any fluctuations in his or her anticipated social rewards (or vice versa if he or she is more concerned with impressing his or her friends over whatever thrills shoplifting may offer). For further discussion, see: Kreps (2013, 2023).

offender, which, for some, can be an extremely pleasurable experience (Burt & Simons, 2013). In some cases, this "rush" can be so strong that individuals will pursue certain criminal acts solely for the purpose of achieving it. As Katz (1988, p. 52) states:

> Various property crimes share an appeal to young people, independent of material gain or esteem from peers. Vandalism defaces property without satisfying a desire for acquisition. During burglaries, young people sometimes break in and exit successfully but do not try to take anything. Youthful shoplifting, especially by older youths, often is a solitary activity retained as a private memory. 'Joyriding' captures a form of auto theft in which getting away with something in celebratory style is more important than keeping anything or getting anywhere in particular.

Here, Katz (1988) highlights several criminal activities that, even in absence of any pecuniary or social gain, provide a high enough degree of excitement for offenders to warrant their commission. Such activities fulfill a unique psychological need (e.g., excitement), and, as such, the fulfillment of said need may itself provide sufficient motivation to pursue criminal acts (McCarthy, 1995). Alternatively, the individual may also primarily engage in criminal activities for the purpose of achieving a heightened social status among his or her peers. As Katz (1988, pp. 114-115) states:

> In contrast to impassioned violence among spouses and friends, adolescent attackers of other adolescents may never get close enough or stay around long enough to witness their victims' destruction. But their own 'battle scars' are sure to be reviewed repeatedly in intimate collective settings. … The animating concern is less a sadistic consumption of the suffering of others than the construction within one's circle of proof of a heroic commitment to the group's grandiose stature.

Here, violence is described as a primarily social endeavor, through which one can often achieve a great deal of respect among his or her friends and peers. Such social rewards, according to Katz (1988), can even "trump" whatever other pleasurable outcomes violent acts may bring the individual, including whatever thrills or excitement violence may produce for the individual. Indeed, this notion can also (once again) be seen in the work of Elijah Anderson (1999), as the maintenance of social standing is suggested to take precedence over virtually every other consequence to criminal acts the agent may be concerned with. This leads into our next hypothesis:

> ***Hypothesis 3a****: Variation in the perceived intrinsic (social) rewards to crime will exert a stronger influence on criminal behavior for individuals who harbor higher perceptions of the social (intrinsic) rewards to crime.*

Conversely, it may also be the case that higher perceptions of one reward type may actually *strengthen* the overall influence of other reward types on criminal behavior. That is, it is possible that individuals who perceived the act of burglary to offer a high quantity of, say, thrills or excitement may also find themselves *more susceptible* to the social rewards to crime. An example of this notion is provided by Neal Shover (1996, pp. 64-65) in his book *Great Pretenders*, wherein he discusses the roles of both the intrinsic thrills and anticipated financial returns to the act of robbery in shaping the offender's *preferences toward robbery over alternative forms of illicit acts*. More specifically, Shover (1996) highlights the power dynamics at play for those who engage in more "confrontational" criminal acts (i.e., robbery), as such acts can often elicit a level of excitement rarely achieved through more "impersonal" acts (e.g., burglary), as well as a sense of *control* over a situation that the offender may otherwise rarely experience within other areas of his or her personal life (e.g., financial instability). In addition, Shover (1998) mentions the nearly equal importance of the financial returns to robbery, in that the relatively "quick" nature of securing additional funds by taking it from someone directly meant far less effort being placed toward converting, say, stolen items from a house or car into currency. Indeed, both aspects of engaging in robbery (i.e., the thrills and the money) were highlighted as two primary reasons for why some offenders specifically preferred the act of robbery over that of burglary, as despite the heightened "danger" associated with mugging another person (e.g., the victim fighting back, higher chance of arrest, etc.), the intoxicating blend of excitement and quick cash seemed to instill a nearly insatiable drive to continue mugging within a select number of chronic offenders (Shover, 1998, p. 65). More broadly, Shover's study appears to lend some support for the notion that, for many offenders, the elicitation of a single reward type (e.g., social benefits) may be insufficient to "motivate" the offender to pursue a particular criminal act (e.g., robbery). Instead, the average offender may only opt for said criminal act in the event he or she anticipates receiving a substantively high quantity of *two or more types of perceived rewards*. Likewise, individuals who hold higher perceptions of a particular type of benefit to criminal activities, on average, could potentially be *more responsive to other types*, rather than less so (as suggested by Hypothesis 3a). That is:

> ***Hypothesis 3b****: Variation in the perceived intrinsic (social) rewards to crime will exert a stronger influence on criminal behavior for individuals who harbor higher perceptions of the social (intrinsic) rewards to crime.*

Finally, an argument could also be made that the point at which variation in one reward (e.g., social rewards) type achieves its strongest influence on a person's involvement in illegal acts may be some middling value of an alternative reward type (e.g., intrinsic rewards). The logic for this claim follows from that of the previous two sections, in that the decision rule outlined in Equation 2 suggests that variation in any given (dis)incentive type is likely to have a stronger impact the closer the agent is to some "tipping point" in his or her expected consequences to illegal acts and legal alternatives (once again, a more technical breakdown of this notion, as well as overall mathematical justification for each of the hypotheses provided through this chapter, are provided in Appendix A). In simple terms, it may be the case that for individuals with a sufficiently *low* perception of, say, the social benefits to crime may be generally less responsive to the personal benefits due to the summed "seductions" of a particular criminal act (e.g., burglary) being, on average, insufficient for the agent to truly believe he or she is better off engaging in at least one instance of said act. However, the same may also be true for those who harbor sufficiently *high* perceptions of the social (intrinsic) rewards to crime, as such individuals may find the social (intrinsic) benefits alone to be worth the risks and costs to engaging in at least one particular type of criminal act (burglary). For such individuals, the prospect of an alternative "gain" of some kind (e.g., intrinsic thrills or excitement), while certainly desirable for the agent, may have relatively little *observable impact* on his or her level of involvement in criminal activities. This idea can be summarized by:

> ***Hypothesis 3c****: Variation in the perceived intrinsic (social) rewards to crime will exert the strongest influence on criminal behavior for individuals who harbor higher more middling perceptions of the social (intrinsic) rewards to crime.*

Before ending this chapter, it should first be noted that none of the hypotheses established up to this point make any a priori assumptions about whether the moderating influence between two (or more) rational choice inputs will be *symmetrical* with respect to one another. That is, I do not assume that any distinct "pair" of perceived (dis)incentive variables will exhibit a clear, two-way moderation effect (as is usually assumed in standard regression models which employ a multiplicative interaction term of some variety; Wood, 2017). For instance, it may be the case that while an increase in the perceived social rewards to crime may lead to a generally stronger influence of the intrinsic rewards on a person's involvement in criminal activities, it is also possible that an increase in the perceived intrinsic rewards actually lead to a *weaker* overall influence of the social rewards on criminal acts. The same is also true of the predictions outlined in the previous two section, as a generally weaker influence of the social and intrinsic rewards to crime at higher values of risk and cost perceptions, for example, does not imply that perceived arrest likelihood and social costs will exert a stronger influence on crime decisions at higher values of the perceived rewards.[21] Put differently, it is possible that the moderating influence of multiple types of benefit, risk, and cost perceptions may not conform to the standard *two-way and predominantly linear interaction effect* as commonly examined using more conventional modeling strategies (e.g., the inclusion of a multiplicative interaction term within one of the many sub-variants of the generalized linear model; Wood, 2017). Rather, for each set of hypotheses it is presumed that any such moderating influence may only be meaningful in a *single direction* (e.g., higher social rewards may imply a stronger influence of the intrinsic rewards but not vice versa), as well as display some level of *nonlinearity* (e.g., a non-monotonic functional form as hypothesized by 1c, 2c, and 3c). As such, the examination of such properties is likely to necessitate the usage of a less "conventional" analytic strategy; one which is (ideally) capable of employing more flexible assumptions of both the distributional form and overall directionality of any given (possibly *one-way*) interactive effect (e.g., a generalized additive model of some variety; Wood, 2017). Details

[21] Note that such a phenomenon would naturally follow from the usage of a multiplicative interaction term to examine the moderating influence between, say, the perceived social benefits and anticipated arrest likelihood. In particular, if the slope coefficient estimated within a regression model containing said interaction term is negative, then this result would be interpreted as perceptions of social benefits having a generally weaker criminogenic effect at higher values of perceived arrest risk, while perceived arrest risk would have a *stronger* overall influence on crime at higher levels of the social benefits (Raudenbush & Bryk, 2002). This follows from the hypothesized negative relationship between arrest risk and offending behavior, as a "more negative" relationship between risk and crime at higher reward levels would imply a stronger deterrent effect of risk overall. For this project, I am assuming that such a "two-way" interaction may not exist, as it may very well be the case that higher perceptions of both reward and risk have a generally *weakening* influence on one another at higher values. Such a phenomenon can often be overlooked in more traditional examinations of moderation effects (see Wood, 2017).

as to how such an approach can be leveraged for the purpose of evaluating each primary set of hypotheses for this dissertation are provided in the next chapter.

Chapter 3: Data and Methods

The following empirical investigation will utilize data derived from the Pathways to Desistance study (Mulvey, 2012). This study is a longitudinal investigation of adolescent individuals who were previously convicted of a serious offense (typically, a felony) in either the juvenile or adult court systems located in Maricopa County, AZ, and Philadelphia County, PA. As part of the eligibility criteria for joining the study, the offender was required to have committed his or her respective offense before the age of 18 (the youngest recruits were around 14 years of age at the time of their offense). In an effort to represent a broader range of criminal backgrounds (e.g., those convicted of a robbery, assault, etc.), participants who were convicted of a drug-related offense were allowed to comprise no more than 15% of the total sample (Schubert et al., 2004). The investigators also approached any female offenders who met the eligibility requirements for the study so as to prevent the sample from consisting of only male offenders (in total, around 13.6% of participants were female). Participants were recruited within a three-year window between November 2000 and January 2003, during which approximately 80% of persons approached by the researchers agreed to partake in the study, achieving a total sample size of 1,354 individuals. Responses to a wide range of survey measures were captured within (approximately) 2-hour long interviews with participants, each of which was conducted in either the participant's own home or in a private room in whichever facility he or she might have been held at the time. To answer each question, participants were given laptops with pre-programmed question prompts, as well as any necessary skip patterns, to which participants could provide answers to via a number pad. After an initial baseline interview, each participant was interviewed in 6-month intervals for six total follow-up periods. Past this point, participants were interviewed annually for an additional four interviews, summing to eleven data waves in total over the course of 84 months. For further elaboration on the sample recruitment strategy and overall design of the study, refer to Schubert and colleagues (2004).

MEASURES:

Self-Reported Offending:

For each wave of the Pathways study, participants were inquired about the number of crimes they committed within the specified recall period. For the baseline wave, participants were asked to report on the number of offenses they engaged in within the past year, while each subsequent wave afterward they were asked to recall the number of crimes since the previous interview. Among these included violent crimes (namely: aggravated assault, getting into a fight or gang-related fight, shooting at or shooting someone), instrumental crimes (auto theft, burglary via entering a building or car, shoplifting, receiving or selling stolen property, illegal usage of a credit card, selling marijuana or other illicit substances), crimes that are both violent and instrumental in nature (carjacking and robbery either with or without a weapon), and miscellaneous property offenses (vandalism and arson).[22] A generalized offending measure was created by assigning each participant a variety score equal to the quantity of crime *types* he or she reported engaging in at least one instance of per data wave. Variety scores were preferred over a count measure for two

[22] Crime types which seemed less relevant to the measures of perceived rewards, risk, and costs were dropped from the analysis. Among these included: engaging in illicit sexual activities for money, carrying a gun, driving while drunk or high, and joyriding. Additionally, self-report measures related to the commission of either a homicide or sexual assault were dropped as well. The reason for this is two-fold. First, both measures are masked within the publicly available dataset due to confidentiality concerns, and, as a consequence, a fair portion of studies which utilize the Pathways data do not employ these measures (e.g., Hamilton, 2023; Thomas & Vogel, 2018; Thomas et al., 2020). Second, among studies that have utilized these measures (e.g., Anwar & Loughran, 2011), the reported base rates for the homicide and rape variables are often extremely small, especially when compared to other reported crime types. As such, the inclusion of these measures in any generalized offending variable (e.g., a variety score) will, more than likely, have little to no impact on the substantive findings obtained within the present analysis. For further discussion, see: Hinkle et al. (2013).

reasons. First, variety scores can help normalize what is often a heavily right-skewed distribution of responses in raw count measures (e.g., the "sold marijuana" item had at least one response of 995) by capping each participant's reported number of offenses per crime type at a value of 1. Second, each of the hypotheses established in the previous chapter were largely derived from applying the decision rule outlined in Equation 2 to the decision to engage in at least one offense of a given type for some arbitrary time period (e.g., that which occurs between two adjacent data waves). Because of this, a variety score outcome measure seemed more appropriate for this analysis.[23]

Perceived Social Benefits:

Participants' expectations of the potential social rewards to offending are captured via three scales, each of which inquires about the degree to which the participant believed he or she would receive various social benefits for engaging in either fighting, theft, or robbery. Among these benefits include the amount of respect gained from one's peers, other adults, or their significant other, and the satisfaction associated with "getting back" at others. To give an example of how each set of social rewards measures was presented to participants, the social benefits to robbery scale consisted of the following:

> ***Now think about how people would react if you robbed someone of their money, clothes, or other goods. How much do you agree or disagree with the following statements about how people might react?***
>
> *If I rob someone, other people my age will respect me more.*
> *If I rob someone, I'll get more respect from adults in my neighborhood.*
> *If I rob someone, people my age will be afraid to mess with me.*
> *If rob someone, I'll impress my boyfriend (or girlfriend).*
> *I can get back at someone who messes with me if I rob him (or her) or someone close to him (or her).*

[23] It should also be mentioned that prior examinations of rational choice and deterrence theories using the Pathways data (e.g., Hamilton, 2023; Loughran et al., 2016a).

Responses for each of these items were coded on a 1 to 4 Likert scale (1 being "strongly disagree" and 4 being "strongly agree"), with the total expected social benefits for an individual, at a given time period, computed via the mean of all items per crime type.[24] Additionally, participants' responses to the social benefits scale seemed to be fairly consistent, eliciting a Cronbach's alpha (α) score of .89 at the baseline wave.

Perceived Personal Benefits:

The anticipated personal benefits from offending were measured by prompting each participant to report on the amount of excitement they expected to receive from engaging in criminal activities. Specifically, participants were provided the following:

> ***How much 'thrill' or 'rush' is it to do any of the following things?*** *[If you have never done any of these things, give your rating for how much 'thrill' or 'rush' you think it would be for you]*
>
> *Fighting*
> *Robbery with gun*
> *Stabbing someone*
> *Breaking into a store or home*
> *Stealing clothes from a store*
> *Vandalism*
> *Auto theft*

For each crime type, participants were instructed to provide a numerical response ranging from 0, which denoted "no fun or kick at all," to 10 denoting "a great deal of fun or kick." Similar to the perceived social rewards measure, the total quantity of the perceived personal benefits to crime for an individual was determined by computing that individual's

[24] Curiously, the codebook for the individualized items within the pathways study states that, at the baseline wave, social reward items were placed on a 1 – 5 scale, with the middle-most category being "neither agree nor disagree." At each subsequent wave, however, each social rewards item is presented without this category, instead taking the 1 – 4 form mentioned here. It is not mentioned on the official website that a change in the scale occurred between the baseline wave and wave 1 (see: https://www.pathwaysstudy.pitt.edu/index.html), and not a single individual is presented as answering "Strongly Agree" for any item within the baseline wave. It is possible that this may simply be a codebook error.

mean response across each of the 7 items at a given data wave. In addition, responses to this variable seemed to display a fair level of consistency ($\alpha = .88$ at the baseline wave).

Perceived Likelihood of Arrest:

Like the perceived personal rewards to offending measure, the perceived risk of arrest scale utilizes participants' numerical responses to several different crime types to compute a mean score denoting one's expected likelihood of arrest at a given data wave. More specifically, participants were given the following prompt:

How likely is it that you would be caught and arrested for the following crimes?

Fighting
Robbery with gun
Stabbing someone
Breaking into a store or home
Stealing clothes from a store
Vandalism
Auto theft

Note that these crime types are the same as those used in the personal rewards scale. Once again, each participant was asked to provide a numerical response to each crime type which ranged from 0 to 10. A response of 0, in this case, meant the individual believed there was "no chance" they would be caught and arrested, and a response of 10 meant they believed they were "absolutely certain to get caught." This scale also showed a fair degree of consistency ($\alpha = .89$ at the baseline wave).

Perceived Social Costs:

Finally, participants' expected social costs were captured by asking them how likely they believed they were to experience a variety of negative social outcomes in the event they were caught and arrested for engaging in a criminal offense. In particular, respondents were prompted with the following:

If the police catch me doing something that breaks the law, how likely is it that:

I would be suspended from school.
I would lose respect from my close friends.
I would lose respect from my family members.
I would lose respect from neighbors or other adults.
I would lose respect from my girlfriend/boyfriend.
It would make it harder to find a job.

Response options were placed on a Likert scale ranging from a score of 1 (“very unlikely it would happen”) to 5 (“very likely it would happen”), the mean of which determined a participant’s total expected social costs at a specific data wave. This scale also showed a decent level of consistency, eliciting an alpha of .76 at the baseline wave.

Controls:

To reduce concerns over the potential influence of unobserved heterogeneity, each model will contain a set of control variables as covariates, among which will include demographic measures capturing each participant’s race, sex, and age (in years) at the time he or she was first interviewed within the baseline data wave. Additionally, prior research has suggested that stable individual differences (e.g., impulsivity) might potentially influence perceptions of utility, as well as condition the effect of perceptions of reward and risk on offending behavior (Nagin & Paternoster, 1993; Piquero et al., 2011; Thomas & McGloin, 2013). Likewise, several measures of stable individual differences—namely, impulse control (measured via the “impulse control” scale contained in the Weinberger Adjustment Inventory), psychosocial maturity (Psychosocial Maturity Inventory), and future orientation (Future Outlook Inventory)—will also be included in each regression model. Furthermore, a key component of Becker’s model is that of a person’s expectations of the returns and losses to legal alternatives to crime (e.g., legal work). To account for this, a dichotomous variable denoting each participant’s reported employment status (1 =

employed, and 0 = not employed) at a particular data wave will also be included within each of the models outlined throughout the remainder of this chapter.

ANALYTIC STRATEGY:

To examine the combined effect of the perceived rewards, risks, and costs on offending behavior, the following model will be estimated:

$$g\big(\phi(Y_{it+1})\big) = \alpha + T_e(\boldsymbol{I}_{it}) + \sum_{k=1}^{K} \beta_k\big(\boldsymbol{X}_{it}^{k}\big) + \varepsilon_{it} \qquad (3)$$

where Y_{it+1} denotes individual i's self-reported offending variety score at time $t + 1$, $\boldsymbol{I}_{it}$ is a vector containing each of the perceived incentive variables (social and personal rewards, certainty of arrest, and social costs), and $\boldsymbol{X}_{it}^{k}$ is a vector of controls.[25] By setting the outcome variable to "lead" each predictor in this way (i.e., the value of Y for each participant is measured at $t + 1$, while the value of $\boldsymbol{I}$ for said participants is measured at time t), the risk of achieving a *reverse causal* estimate is minimized.[26] The g term refers to a generalized additive structure of the model, and ϕ assigns a Tweedie distribution to

[25] Note that the value of the subscript "t" refers to any data wave which can feasibly be captured within the model (e.g., if $t = 1$, then the model utilizes perceptions reported in the first data to predict each participant's self-reported offending in the second data wave). Likewise, the model outlined in Equation 3 utilizes a "pooled" data structure, wherein each row of the data frame contains a given participant's reported perceptions (and criminal involvement) at a specific data wave (e.g., person 1 at time period 2). Because of this, the model effectively treats a person's answers at two (or more) different time periods as *independent* observations (e.g., person A's reported reward, risk, and cost perceptions at data wave 1 is treated as a separate case from that of A's perceptions at wave 2). Additionally, the model is set to exclude any participant i, for some time period t, for which at least one of the variables is assigned a missing value (e.g., a scale for which the individual did not provide an answer to one or more items). Consequently, only participants who *provided answers to every item per variable for a given data wave* were included within each of the models estimated throughout this project.

[26] Such a possibility is often an issue when examining the influence of *attitudes* of any kind on criminal behavior. In most cases, survey measures of the rewards, risks, and costs to crime capture a person's *current* beliefs about the consequences to criminality, while self-reported offending measures capture *prior* involvement in criminal activities (e.g., how many acts of burglary a person commits since previously being surveyed). If we were to estimate a model utilizing measures of participants' current perceptions ($\boldsymbol{I}_{it}$) to predict prior offending (Y_{it}), our results may simply reflect an *experiential effect*. That is, a person's experiences with criminal acts directly shape his or her perceptions of the benefits and costs to crime, rather than the other way around (see: Saltzman et al., 1982).

Y_{it+1}, which is analogous to a Negative Binomial model (see Wood, 2017, p. 115). The T_e term denotes a tensor product smooth, the "surface" of which corresponds to the model's estimated value of Y_{it+1} for any unique combination of values provided for each of the perceived reward, risk, and cost measures, while holding each control variable ($\boldsymbol{X}_{it}^{k}$) constant at its respective sample mean. Like any generalized additive model, the tensor product makes no a priori assumptions about the shape of the underlying function to be estimated (aside from those related to the mean and variance structure of the outcome variable; Wood, 2017), and instead generates a model manifold (i.e., a set of predicted values of Y_{it+1} for each participant i at time t) from the data directly. In contrast to more conventional (i.e., *linear*) regression procedures, such an approach allows the model to provide a direct approximation of the *functional form* of the relationship between the outcome variable (Y_{it+1}) and each primary predictor of interest ($\boldsymbol{I}_{it}$).

Additionally, since the tensor product is fitted using *every possible combination* of values from multiple predictors, the researcher can also identify any potential interaction effects between those predictors by observing any changes the overall shape—more specifically, the *curvature*—of the estimated model manifold across different values of each moderating variable (details for such a process will be provided throughout the remainder of this chapter). Furthermore, a primary benefit of employing a tensor product smooth over some alternative smoothing method is its insensitivity to the "scale" of the predictors. More specifically, the tensor product is *invariant* under re-scaling of the independent variables, and thus will always produce the same distributional shape regardless of the scale each predictor variable is measured on (e.g., a 1 to 4 Likert scale versus a numerical measure from 0 to 10). As such, the tensor product is often useful for observing the overall *shape* of the relationship between two or more predictor variables which are not placed on the same scale as one another. Since each of the perceived rewards, risks, and costs measures found in the Pathways data are (mostly) measured on different scales, a tensor product smooth seemed the most appropriate for this analysis.

While a full technical breakdown of the statistical machinery undergirding tensor product smooths—much less generalized additive models more broadly—is well beyond the scope of this project, a brief summary of its inner workings will nonetheless be provided here (for a more in-depth discussion, see: Wood, 2017, pp. 227 – 237). Denote the set of all possible combinations of reported values for each of the perceived social and intrinsic rewards, arrest risk, and social cost measures by $\mathcal{U}$, such that:

$$\mathcal{U} \subset \mathbb{R}^m, \qquad m < n,$$

where $\mathcal{U} \subset \mathbb{R}^m$ is read as "$\mathcal{U}$ is a subset of some m-dimensional Euclidean vector space,"[27] and n is the number of cases—that is, the *sample size*—of the model's data frame. Note that $m = 4$ in this instance since each "point" contained in $\mathcal{U}$ can be represented as a (unique) linear combination of four variables (namely, the primary reward, risk, and cost measures supplied within the Pathways data). Consequently, any tensor product derived from $\mathcal{U}$ will take the following form:

$$T_e(\mathcal{U}) = \sum_{j=1}^{J} \gamma_j \lambda_j(\boldsymbol{I}_{it}), \qquad for\ any\ \boldsymbol{I}_{it} \in \mathcal{U},$$

where λ_j denotes a 4-dimensional basis function defined over $\mathcal{U}$, and γ_j is that function's respective coefficient. Similar to a standard linear regression model, the value of γ_j can be interpreted as the "effect" λ_j has on Y_{it+1} with respect to each point in $\mathcal{U}$ (i.e.,

[27] For further discussion on Euclidean spaces, as well as vector spaces more broadly, see: Shilov (1971).

every combination of the perceived reward, risk, and cost measures). The difference, however, is that in this instance the value of any particular basis function λ_j does not have any meaningful interpretation with regard to how any specific perception variable is measured (i.e., the basis functions of the model do not correspond with however "low" or "high" participants might rank across each predictor variable). Rather, the role of the basis functions is purely instrumental, as each λ_j can be uniquely combined in some mannerism (e.g., via the "best fitting" set of γ_j) to produce a model manifold which, theoretically, should "fit" the observed joint distribution(s) of each of the perceived incentive and offending variables to the highest degree possible (for further discussion on basis functions, see: Wood, 2017, p. 162).

Of course, a potential pitfall to allowing the model to fit a more "flexible" set of predictions to the response variable (Y_{it+1}), with respect to each predictor ($\boldsymbol{I}_{it}$), is the possibility of accidentally overfitting the data. Put differently, the model may overcompensate for random fluctuations in Y_{it+1}, such that while the model might achieve a greater overall "fit" to the observed distribution of Y_{it+1} (e.g., a higher r-squared value), the model may also struggle to make predictions for any *unobserved* values of Y_{it+1} (e.g., those which the model is not explicitly trained on). As such, most statistical packages, such as the "mgcv" package in R, will, by default, *penalize* the overall flexibility of any generalized additive model. For instance, in constructing a tensor product smooth, the "gam" function in mgcv employs a set of cubic regression splines for the predictor variables, each of which encourages the model to prefer smoother functional form estimates over "better fitting," more noisy ones (for details, see: Wood, 2017, p. 232). In an effort to conform to standard practices in generalized additive modeling, each of the models conducted for this project will be estimated using the mgcv package in R.

Furthermore, to estimate the model outlined in Equation 3, each predicted value of Y_{it+1} will be fitted using a restricted maximum likelihood (REML) procedure (Patterson & Thompson, 1971). The benefits of REML for this analysis, over that of a more standard maximum likelihood (ML) method, are twofold. First, the nested structure of the Pathways data can often lead to downwardly biased estimates of any variance and co-variance components for (non-hierarchical) models which employ ML methods. This occurs due to such models treating every observation of the predictor, as well as outcome, variables as being truly *independent* of one another, rather than as observations nested within persons over time (Raudenbush & Bryk, 2002). Such an approach tends to overlook time-stable characteristics of individuals (i.e., fixed effects), which can sometimes give the appearance of a greater degree of similarity between observed values of predictors (namely, those nested within persons) relative to what would be observed within a *non-nested* data structure. The REML method corrects for this by appropriately restricting the degrees of freedom of the model, thus ensuring any within-person similarities over time will not influence variance estimates (see: Raudenbush & Bryk, 2002, pp. 53-54). Second, REML allows for direct comparisons in the overall "performance" of different models; particularly, that which is achieved by allowing for a more flexible (i.e., *nonlinear*) fit to the observed data points.

Testing for Nonlinearity:

A good first step in any analysis employing a generalized additive model is to determine whether the desired smoothing procedures are necessary for achieving a better model fit. Namely, could the overall influence of the perceived social and personal rewards, arrest likelihood, and social costs on criminal behavior be just as well approximated using a more standard linear modeling procedure? Perhaps more importantly, even if a nonlinear model achieves a better fit for each of the perception variables, will the *interactive effects* of those variables be pronounced enough to warrant a tensor product smooth? Put differently, if the *marginal* influence of each perception variable (e.g., social rewards) remains constant across all possible combinations of values for every other perception variable (e.g., personal benefits, arrest likelihood, and social costs), then one can achieve a good fit simply by assigning a single (i.e., *independent*) smooth term to each of the perceived rewards, risks, and costs to crime. In such a scenario, the outcome is said to be an *additive* function of each predictor, in that any combined (i.e., moderating) influence between those predictors can safely be ignored without jeopardizing the ability of the model to produce accurate predictions (Wood, 2017). If, however, there do appear to be substantial interdependencies between each main predictor of interest, then a tensor product smooth will achieve a better fit relative to any model which fits a set of independent smooths for the primary variables.

To assess whether the perceived social and personal rewards, probability of arrest, and social costs display a nonlinear influence on offending variety—including both the marginal (i.e., independent) and combined effects of each perceived incentive—the following two models will be estimated in addition to the tensor product smooth model as defined in Equation 3:

$$\phi(Y_{it+1}) = \alpha + \sum_{j=1}^{J} \beta_j\left(\boldsymbol{I}_{it}^{j}\right) + \sum_{k=1}^{K} \beta_k\left(\boldsymbol{X}_{it}^{k}\right) + \varepsilon_{it}, \qquad (4)$$

where β_j denotes a (static) coefficient for the j-th perception variable, and:

$$g\big(\phi(Y_{it+1})\big) = \alpha + \sum_{j=1}^{J} f_j\big(\boldsymbol{I}_{it}^{j}\big) + \sum_{k=1}^{K} \beta_k\big(\boldsymbol{X}_{it}^{k}\big) + \varepsilon_{it}, \qquad (5)$$

where f_j denotes a marginal smooth function assigned to the j-th incentive. Notice the lack of the g term in the first model (i.e., Equation 4), which would normally denote a smooth estimate of some variety for the outcome variable (i.e., g refers to a generalized additive model; Wood, 2017). Likewise, Equation 4 denotes a generalized *linear* regression model which assigns a Tweedie distribution to Y_{it+1}. To determine whether the second model achieves a better fit than the first, a common practice in generalized additive modeling is to compare each model's respective Akaike Information Criterion (AIC) score. If the second model outlined in Equation 5 achieves a lower AIC value than the first, then it is said to achieve a better fit overall (see: Wood, 2017, p. 335).[28] The same comparison can also be made with respect to the AIC value of the tensor product model, which, ideally, will be smaller than those of each of the additional two models established in this section. Additionally, one can also examine the restricted maximum likelihood (REML) scores for each model,[29] for which lower values, once again, imply a better overall model fit (Wood,

---

[28] It should be mentioned that AIC measures also tend to prefer simpler (i.e., linear) models over more complicated (i.e., nonlinear) ones. Although some have suggested the comparison of AIC scores between "smooth" and "non-smooth" models will often provide an overly conservative assessment of the degree to which any nonlinear model achieves a better fit to the data, such an assessment will nonetheless favor the usage of a nonlinear model if it achieves a lower AIC value relative to its linear counterpart. For further discussion, see: Wood (2017, pp. 301-304).

[29] Note that REML methods are often preferred over alternative smoothing procedures, such as generalized cross validation (GCV), for the purpose of producing a model which is (relatively) insensitive to sampling variance. In other words, a generalized additive model fitted via REML procedures tends to be less "swayed" by random fluctuations in data, compared to other methods, and is thus less prone to over-fitting. Of course, this also means such models are at a greater risk of "under-fitting" the data, as REML methods tend to prefer more rigid smooths over more flexible ones (i.e., REML will typically fit a "less wiggly" functional form for the model). Such a risk is rarely of much concern for social scientists, who typically prefer achieving a more rigorous set of estimates over those which, although potentially capable of producing more "accurate" predictions of some response variable, are more error-prone (Hamilton, 2023). Because of this, REML

2017, p. 362). Furthermore, each model will also provide a measure of the statistical significance of the perceptions variables. Specifically, every generalized additive model assigns a p-value to each of its respective smooth terms (including any tensor products), which denotes the probability of achieving an equivalent fit to the data via the zero function (e.g., $f_j = 0$ for all values of j). The lower the p-value for a given smooth term, the smaller the probability that a particular variable (or set of variables, in the case of a tensor product smooth) exhibits no influence on the outcome variable (see: Wood, 2017, pp. 304-305).

Examining the Interdependency of the Rewards on the Risks and Costs:

A distinct advantage that generalized additive models—particularly, tensor product smooths—have over more standard linear regression procedures is that the predicted values of the outcome variable in the former are derived almost entirely from observed distributions in data. Because of this, generalized additive models are often useful to researchers looking to get a sense of the “true” underlying distributional form of the relationship between some response variable (e.g., Y) and a set of predictors (e.g., Loughran et al., 2012). It is perhaps unsurprising, then, that the set of predicted values (i.e., the model manifold) produced by such models tend to lend themselves reasonably well to standard calculus operations (e.g., the computation of derivatives and integrals; Wood, 2017). For tensor product smooths specifically, the researcher is supplied with an estimated model surface which is derived from *multiple* variables, upon which a variety of useful procedures can be conducted. Since a primary goal of this project is to investigate the degree to which the overall influence of the perceived rewards to crime will be

methods were employed in fitting each of the models for this project. For further discussion, see: Wood (2017, p. 267).

interdependent on the perceived risks and costs, a potentially useful approach would be to estimate the following:

$$\iint_S \left(\frac{\partial Y}{\partial x_1} + \frac{\partial Y}{\partial x_2}\right) dx_1 dx_2,$$

where S denotes the set of estimated values of Y_{it+1} for all possible combinations of values for two given predictor variables (i.e., x_1 and x_2), holding all other variables constant. The contents of the double integral in the above expression refer to the *differential of* Y with respect to x_1 and x_2, the value of which can be thought of as the overall steepness of the curvature of S while traveling in a "positive" direction along both x_1 and x_2.[30] Put simply, a value of $\iint_S(*)$ greater than zero corresponds to an overall positive influence of x_1 and x_2 on Y, while values of $\iint_S(*)$ less than zero denote an overall negative influence of x_1 and x_2 on Y. More importantly, the further the value of $\iint_S(*)$ strays from zero, the *stronger the overall effect* of x_1 and x_2 on Y. In the event x_1 and x_2 refer to the perceived social and personal benefits of crime, then the value of $\iint_S(*)$ denotes the overall "rewards effect" for a *specific combination of values for each of the perceived arrest risk and social costs measures* (denoted by x_3 and x_4, respectively).

To estimate the value of $\iint_S(*)$, given some set of fixed values assigned to x_3 and x_4, I will employ a common "trick" in vector calculus which relies solely on the characteristics of the boundary of S. Let ω be the following 1-form:

$$\omega = -Y(\boldsymbol{I})dx_1 + Y(\boldsymbol{I})dx_2,$$

---

[30] More formally, the value of $\frac{\partial Y}{\partial x_1} + \frac{\partial Y}{\partial x_2}$ denotes the *directional derivative of* Y with respect to $1 \in \mathbb{R}^2$, such that 1 is a vector whose components are all equal to 1. For further discussion, see: Edwards Jr. (1973) and Spivak (1965).

where $Y(\boldsymbol{I})$ denotes the estimated value of Y_{it+1} with respect to $\boldsymbol{I}_{it}$, and dx_1 refers to some marginal increase in x_1 (dx_2 is defined analogously).[31] By taking the exterior derivative of ω (as defined by Spivak [1965]), we get the following 2-form:

$$d(\omega) = \left(\frac{\partial Y}{\partial x_1} + \frac{\partial Y}{\partial x_2}\right) dx_1 \wedge dx_2,$$

where $\wedge$ denotes the wedge product of dx_1 and dx_2 (Edwards Jr., 1973). By integrating $d(\omega)$ over S, we now get:

$$\iint\limits_S \left(\frac{\partial Y}{\partial x_1} + \frac{\partial Y}{\partial x_2}\right) dx_1 \wedge dx_2.$$

Note that the above is simply a restatement of our previous expression for $\iint_S(*)$, the only difference being that we are now explicitly referencing the underlying 2-form which belongs to the double integral (i.e., $dx_1 \wedge dx_2$). It follows from the generalized Stokes' Theorem[32] (see: Spivak, 1965, pp. 124, 135) that:

$$\iint\limits_S \left(\frac{\partial Y}{\partial x_1} + \frac{\partial Y}{\partial x_2}\right) dx_1 \wedge dx_2 = \oint\limits_{\partial S} -Y(\boldsymbol{I})dx_1 + Y(\boldsymbol{I})dx_2,$$

where $\oint_{\partial S}(*)$ denotes a line integral on ∂S (i.e., the *oriented boundary* of S). More specifically:

$$\oint\limits_{\partial S} \omega = \oint\limits_{C^1} \omega + \oint\limits_{C^2} \omega - \oint\limits_{C^3} \omega - \oint\limits_{C^4} \omega,$$

[31] This is, of course, an oversimplification, as terms such as "dx_1" have an extremely precise interpretation in the study of differential forms (particularly, the "exterior differential calculus;" Lovelock & Rund, 1975, p. 130), the technical details of which are well beyond the scope of this project. For an in-depth breakdown, I highly recommend Michael Spivak's (1965) *Calculus on Manifolds* (specifically, the discussion provided between pages 86 and 95).

[32] Technically, the following example more closely resembles the contents of Green's Theorem, which itself is merely a special case of both the classical and generalized Stokes' Theorems. For a more detailed breakdown of each of these theorems, see: Spivak (1965).

where each C^k can be thought of as a "link" in a chain of line integrals for each of the four "sides" of the domain spanned by x_1 and x_2 (see: Flanders, 1963, pp. 57-63). For instance, C^1 denotes the full range of values of x_1 while holding x_2 constant at its respective lowest value, and C^3 denotes the same except x_2 is now held constant at its highest value. Likewise, C^2 and C^4 refer to the range of x_2 while holding x_1 constant at its highest and lowest values, respectively. Since ∂S is presumed to be an oriented boundary around S,[33] each of the integrals for C^3 and C^4 return a "negative" value (see: Spivak, 1965, pp. 97-100). Additionally, since C^1 and C^3 vary only with respect to x_1 (i.e., x_2 is held constant), while C^2 and C^4 vary with respect to x_2 (x_1 constant), it follows that:

$$\oint_{\partial S} \omega = \oint_{C^1} -Y(\boldsymbol{I})dx_1 + \oint_{C^2} Y(\boldsymbol{I})dx_2 - \oint_{C^3} -Y(\boldsymbol{I})dx_1 - \oint_{C^4} Y(\boldsymbol{I})dx_2,$$

which can alternatively be written as:

$$\oint_{\partial S} \omega = \oint_{C^3} Y(\boldsymbol{I})dx_1 - \oint_{C^1} Y(\boldsymbol{I})dx_1 + \oint_{C^2} Y(\boldsymbol{I})dx_2 - \oint_{C^4} Y(\boldsymbol{I})dx_2.$$

Put differently, the value of $\oint_{\partial S} \omega$ is equal to the total area under Y on C^2 and C^3 deducted by the total area under Y on C^1 and C^4. Finally, we can think of $\oint_{\partial S} \omega$ as itself a *function* of x_3 and x_4 (i.e., arrest likelihood and social losses), the surface of which depicts the overall rewards effect at *each possible combination of values for the perceived arrest risk and the social costs measures* (as will be discussed in the next section, we can just as easily do this same process in *reverse*, such that S now denotes the surface of predicted values for Y_{it+1} for all x_3 and x_4, holding x_1 and x_2 fixed). Such a surface can be approximated by computing the value of $\oint_{\partial S} \omega$ at several points of x_3 and x_4, and then

[33] As is the case for any surface fitted over a closed and bounded subset of $\mathbb{R}^n$ (Munkres, 1991).

fitting a (second) tensor product smooth through those computed values with respect to x_3 and x_4.[34] Doing so would not only help "visualize" the overall influence of the perceived social and personal benefits on self-reported criminality, with respect to perceived arrest risk and social losses, but would also allow for the identification of the *exact values of the risk and cost measures* for which the rewards effect is strongest, as well as weakest. Such an approach would also capture any *nonlinear* moderation effect(s), which can help determine whether the criminogenic influence of the perceived rewards to crime appears to adhere to either Hypothesis 1a, 1b, or 1c, as previously outlined in chapter 2.[35]

---

[34] For this project, my approach to computing these values is, admittedly, a bit crude. Namely, I opted to use the "predict.gam" function in the mgcv package to determine the estimated value of Y_{it+1} for a finite set of values (usually, 10) within a given subsection of the boundary of S (i.e., some C^k), and then computed the sum of those estimated values to determine the "line integral" for that subsection (i.e., the value of $\oint_{C^k} Y(\boldsymbol{I})$). I then repeated this process for each other subsection, and then deducted the line integrals for the "lower" portions of the boundary of S (i.e., C^1 and C^4) from those of the "upper" portions of said boundary (i.e., C^2 and C^3). Such a process provides an estimate of $\oint_{\partial S} Y(\boldsymbol{I})$, as defined previously, for a given set of fixed values for perceived arrest risk (x_3) and social costs (x_4). To estimate the degree to which the value of $\oint_{\partial S} Y(\boldsymbol{I})$ changed with respect to x_3 and x_4, I programmed an algorithm to conduct the aforementioned computation process for over 10,000 unique combinations of values for x_3 and x_4, each of which was evenly "spaced" from one another and, in totality, covered the full range of values for both x_3 and x_4. I then estimated a tensor product smooth through this set of estimated values of $\oint_{\partial S} Y(\boldsymbol{I})$, with respect to x_3 and x_4, which produced a smooth, 2-dimensional surface that can be visualized using the "vis.gam" function in mgcv. For further discussion on these functions, and mgcv more broadly, see: Wood (2017).

[35] An argument could be made that much of what is being discussed here could just as effectively be achieved through a much "simpler" approach. Namely, rather than going through the trouble of estimating a generalized additive model and subsequently performing some number of highly technical calculus operations, one could instead just divide the sample across different conditions of each moderating variable (e.g., arrest risk and social costs) and observe any changes in the "size" of the estimated coefficients for each main predictor (e.g., social and intrinsic rewards) derived from a far less complex (e.g., linear) modeling strategy. For example, one could estimate the overall influence of the social and personal rewards variables on offending variety specifically for sample participants who report "lower" values of arrest risk and social costs, and then do the same for those who report "higher" values of the risks and costs. If the value of β for each reward variable is lower in the latter group, relative to the former, then such a result would suggest that higher values of risk and cost perceptions are related to a generally lower overall effectiveness of the perceived reward variables. Of course, this approach also places a particularly heavy burden on the researcher, who must now decide the appropriate number of "sub-groups" to examine for each of the moderating variables. Too few sub-groups and the researcher risks achieving an overly simplistic view of any underlying interactive effect, whereas too many sub-groups will limit the statistical power of each respective model. A generalized additive model bypasses such concerns by not only using the responses of *every participant* to generate its findings, but also achieves a far more *detailed* (i.e., less simplistic) set of estimates relative to those which would otherwise be derived from simply examining a linear model for different subgroups. Nonetheless, a supplementary set of results utilizing the "subgroup" strategy are provided in Appendix B.

Furthermore, since tensor product smooths are invariant with respect to re-scaling of the predictors, it will also be assumed that:

$$\oint_{C^1} f(\boldsymbol{I})dx_1 = \oint_{C^2} f(\boldsymbol{I})dx_2 = \oint_{C^3} -f(\boldsymbol{I})dx_1 = \oint_{C^4} -f(\boldsymbol{I})dx_2,$$

for some (constant) function f. That is, the "area" under f is presumed to solely be determined by the *shape* of f itself (i.e., the "width" of the range of values for either x_1 or x_2 do not contribute to the value of their respective line integrals).

Examining the Interdependency of the Risks and Costs on the Rewards:

In a similar fashion, we can also examine the overall degree to which perceptions of arrest risk and social costs exert an "inhibitory" (i.e., *deterrent*; Nagin, 1998, 2013) influence on self-reported offending at different levels of the perceived social and personal rewards to crime. That is, does variation in x_3 and x_4 appear to have a stronger influence on a person's offending level for different combinations of values for x_1 and x_2? In this scenario, we are now interested in estimating the following:

$$\iint_S \left(\frac{\partial Y}{\partial x_3} + \frac{\partial Y}{\partial x_4}\right) dx_3 dx_4,$$

where S is defined similarly to the previous section, with the exception that the domain of S is now determined by the range of possible values for x_3 and x_4, holding x_1 and x_2 constant for some fixed set of values. Once again, we can achieve a "simpler" estimate of $\iint_S(*)$ using the basic properties of the boundary of S (again denoted by ∂S). Let ψ denote the following 1-form:

$$\psi = -Y(\boldsymbol{I})dx_3 + Y(\boldsymbol{I})dx_4.$$

Note that ψ is identical to ω in every respect, except for the replacement of x_1 and x_2 with x_3 and x_4, respectively. Taking the exterior derivative of ψ gives us the following:

$$d(\psi) = \left(\frac{\partial Y}{\partial x_3} + \frac{\partial Y}{\partial x_4}\right) dx_3 \wedge dx_4,$$

from which we can define the value of $\iint_S (d(\psi))$ by:

$$\iint_S \left(\frac{\partial Y}{\partial x_3} + \frac{\partial Y}{\partial x_4}\right) dx_3 \wedge dx_4 = \oint_{\partial S} -Y(\boldsymbol{I})dx_3 + Y(\boldsymbol{I})dx_4,$$

which, again, follows from the generalized Stokes' Theorem. Same as before, we can compute the value of $\oint_{\partial S} \psi$ by:

$$\oint_{\partial S} \psi = \oint_{C^3} Y(\boldsymbol{I})dx_3 - \oint_{C^1} Y(\boldsymbol{I})dx_3 + \oint_{C^2} Y(\boldsymbol{I})dx_4 - \oint_{C^4} Y(\boldsymbol{I})dx_4,$$

where each C^k is defined equivalently to how they were in the previous section, with the only difference being the substitution of x_3 in place of x_1, and x_4 in place of x_2. If the predicted value of Y_{it+1} sees an overall decrease with respect to x_3 and x_4 (given some fixed x_1 and x_2), then the value of $\oint_{\partial S} \psi$ will be less than zero (i.e., $\oint_{\partial S} \psi$ will denote a *negative* effect of x_3 and x_4 on Y_{it+1}). Additionally, the greater the overall decrease in Y_{it+1}, the further the value of $\oint_{\partial S} \psi$ will *stay from zero*. This implies that any estimated surface representing the value of $\oint_{\partial S} \psi$, with respect to x_1 and x_2, will achieve its *lowest* points for any subset of its domain where x_3 and x_4 exert an *overall stronger influence on* Y_{it+1}. More specifically, if x_3 and x_4 have a stronger inhibitory effect on Y_{it+1} for, say, lower values of x_1 and x_2, then the value of $\oint_{\partial S} \psi$ will also be lower at these points. Therefore, any geometric depiction of $\oint_{\partial S} \psi$ will represent the "weakest" overall influence of x_3 and x_4 by *higher points* of some visualized surface embedded in $\mathbb{R}^3$.[36] Such a surface

---

[36] Note that this is the opposite of what we would expect to see if were to instead examine the overall influence of the social and intrinsic benefits to crime as a function of the risks and costs. That is, any surface estimated using the methods outlined in the previous section will denote a stronger overall "rewards effect" by *higher* points of some fitted surface, rather than lower points. This is, of course, an immediate consequence of how

can be approximated in a similar mannerism to the previous section, in that one can compute the value of $\oint_{\partial S} \psi$ for several combinations of points for x_1 and x_2, and then fit a tensor product smooth through those computed values with respect to x_1 and x_2, which will allow for an overall evaluation of Hypotheses 2a through 2c.[37] Finally, I will once again assume:

$$\oint_{C^1} f(\boldsymbol{I})dx_3 = \oint_{C^2} f(\boldsymbol{I})dx_4 = \oint_{C^3} -f(\boldsymbol{I})dx_3 = \oint_{C^4} -f(\boldsymbol{I})dx_4,$$

for some fixed function f. Thus, the value of any given line integral (i.e., that which is defined for some C^k) will purely be a function of the overall "shape" of the predicted value of Y_{it+1} over its respective domain.

Examining the Interdependency of Multiple Reward Types:

To assess the degree to which multiple reward types will exhibit a "moderating" influence on one another, one can employ a similar approach to that which was discussed in the previous section. More specifically, one can estimate the following:

$$\iint_{R(X)} \Delta Y(x_1) dx_3 dx_4\,, \qquad given\ any\ fixed\ x_2,$$

where $R(X)$ denotes the set of all possible combinations of values for x_3 and x_4 (i.e., the "range" of the perceived likelihood of arrest and social costs measures, respectively). Here, the contents of $\iint_{R(X)}(*)$ include the overall change in the estimated value of Y_{it+1} with respect to x_1, denoted by $\Delta Y(x_1)$. That is:

---

the values of $\iint_S(d(\psi))$ and $\iint_S(d(\omega))$ are interpreted, wherein the former denotes a stronger influence for lower estimated values, while the latter denotes a stronger influence at higher values.

[37] This was achieved using nearly identical methods to those outlined in the previous section, with the exception of "swapping" the places of x_1 and x_2 with x_3 and x_4, respectively.

$$\Delta Y(x_1) = \int_L^U \frac{\partial Y}{\partial x_1} dx_1 = Y(U) - Y(L),$$

where L and U respectively denote the lowest and highest possible values for some measure of the perceived social benefits to crime (i.e., x_1), and the right-most portion of the above expression follows directly from the fundamental theorem of calculus (see: Spivak, 1965, pp. 100-102). Hence, $\Delta Y(x_1)$ simply denotes the difference between the estimated value of Y_{it+1} at the highest and lowest points of x_1. Since the value of $\Delta Y(x_1)$ could also be a function of the perceived risks and costs to crime (as suggested in the previous section), it is fully possible that any particular selection of values for x_3 and x_4 (e.g., their respective sample means) could potentially influence any estimate of the moderating influence of x_2 on x_1 (or vice versa). As such, by computing the value of $\iint_{R(X)}(*)$, given some fixed value of x_2 (i.e., the personal benefits to crime), we can achieve an average estimate of the overall influence of x_1 on Y *across all possible values of* x_3 *and* x_4. Such an approach not only effectively "dodges" the problem of having to select particular values of x_3 and x_4 to hold constant (as is often done in linear modelling procedures), but also provides a more complete depiction of how the criminogenic influence of x_1 varies with respect to x_2. More concretely, we can directly observe the degree to which x_2 shapes the "effectiveness" of x_1 (and vice versa) *regardless of any possible combination of values for* x_3 *and* x_4.

Similar to the previous section, the value of $\iint_{R(X)}(*)$ can be computed across several points of x_2, through which a smooth curve can be fitted.[38] In doing so, one can

[38] As with the previous section (see the previous footnote), this was also done in a somewhat crude mannerism. Specifically, I used the "predict.gam" function to retrieve the estimated values of Y_{it+1} for both the lowest (i.e., L) and highest (U) values of x_1. I then simply deducted the former estimate (i.e., the predicted

both visualize the overall degree to which the criminogenic influence of x_1 is a function of x_2, as well as identify the general "shape" of the moderating influence of x_2 on x_1, from which Hypotheses 3a through 3c (specifically, in applying these hypotheses to the moderating influence of personal benefit perceptions on the criminogenic influence of the social rewards to crime) can be evaluated. The same can also be done in reverse, in that we can instead estimate:

$$\iint_{R(X)} \Delta Y(x_2) dx_3 dx_4 , \qquad given\ any\ fixed\ x_1,$$

the process of which is nearly identical to the previous example (except, in this case, $\iint_{R(X)}(*)$ now varies with respect to x_1). In doing so, the general impact of the *personal* rewards to crime on self-reported involvement in criminal activities can be evaluated with respect to the perceived social benefits. Similarly, this would generate an additional curve depicting the effectiveness of the intrinsic benefits to crime as a function of the social rewards, which would subsequently allow for an evaluation of Hypotheses 3a through 3c (applied specifically to the influence of the perceived personal rewards with respect to the social rewards).

SIGNIFICANCE TESTING:

Y value for L) from the latter (U), which determined the overall influence of x_1 on Y_{it+1} at a given combination of values for x_2, x_3, and x_4. I then employed an algorithm to repeat this process for around 8,000 unique combinations of values for x_2, x_3, and x_4 (once again, evenly spaced and covering the full range of each variable), and then estimated a tensor product smooth through these computed values with respect to x_2, x_3, and x_4. To compute the value of $\iint_{R(X)}(*)$ with respect to some fixed value of x_2, I once again employed the predict.gam function to retrieve the estimated value of $\Delta Y(x_1)$ for approximately 100 unique combinations of values for x_3 and x_4, and then computed the sum of each estimate of $\Delta Y(x_1)$. I then repeated this process for approximately 1,000 (evenly spaced) values of x_2, and then fit a smooth term through each estimated "surface integral" (i.e., $\iint_{R(X)}(*)$). This produced a smooth curve depicting the overall influence of x_1 on Y for each unique value of x_2. The same process was conducted in reverse to assess the overall effect of x_2 on Y as a function of x_1.

A challenging aspect of conducting analytical procedures on any set of predicted values (i.e., the "surface" of a model manifold) is that there rarely exists standardized methods for testing the statistical significance of one's findings. Indeed, such challenges are often present in any "unconventional" modeling strategy, particularly those which seek to employ weaker assumptions, in some form or another, relative to more standard modeling practices (see also: Hamilton, 2023). Nonetheless, such tests are by no means impossible to achieve, and, more often than not, simply require researchers to exercise a bit more "creativity" in designing them. As such, I will now discuss my plan for examining the robustness of the findings produced via the methods outlined within the previous two sections. Particularly, I will seek to determine the degree to which such results could be produced by chance variation in the data (which, in most cases, is that which is produced by sampling variance; Manski, 2003). In such a scenario, we would expect the true value of any computed surface integral (e.g., $\iint_S (*)$) to be equal to zero across all possible values of each "fixed" perception variable. That is:

$$\iint\limits_S \left(\frac{\partial Y}{\partial x_1} + \frac{\partial Y}{\partial x_2}\right) dx_1 dx_2 = 0, \qquad for\ all\ x_3\ and\ x_4,$$

as well as:

$$\iint\limits_S \left(\frac{\partial Y}{\partial x_3} + \frac{\partial Y}{\partial x_4}\right) dx_3 dx_4 = 0, \qquad for\ all\ x_1\ and\ x_{2,}$$

and:

$$\iint\limits_{R(X)} \Delta Y(x_1) dx_3 dx_4 = 0, \qquad for\ all\ x_2.$$

For the first two expressions, a "surface integral" value of zero can be achieved in the event the estimated value of Y_{it+1} remains constant across the boundary of S. Recall that, for any constant function f, the respective values of each line integral, with respect to any C^k, will be equivalent to one another. It follows that, for the first expression:

$$\oint_{C^3} f(\boldsymbol{I})dx_1 - \oint_{C^1} f(\boldsymbol{I})dx_1 + \oint_{C^2} f(\boldsymbol{I})dx_2 - \oint_{C^4} f(\boldsymbol{I})dx_2 = 2(I_f) - 2(I_f) = 0,$$

where I_f denotes the value of any line integral in the above expression with respect to f (where f is equal to some real number K across all "boundary" values of x_1 and x_2). With this in mind, we could potentially test the "significance" of any estimated value of $\iint_S(*)$, for some fixed x_3 and x_4, by the following:

$$\oint_{\partial S} -Y_{CI}(\boldsymbol{I})dx_1 + Y_{CI}(\boldsymbol{I})dx_2,$$

such that:

$$Y_{CI} = \min_{Y \in CI} |Y - \bar{Y}(\partial S)|,$$

where $\bar{Y}(\partial S)$ denotes the "mean" value of Y across the boundary of S, and:

$$CI = \{Y \in [L, U] : L = Y(\boldsymbol{I}) - Z * \sigma;\ U = Y(\boldsymbol{I}) + Z * \sigma, \}.$$

That is, CI denotes a *confidence interval* fitted to the estimated valued of Y_{it+1} at the point $\boldsymbol{I_{it}}$, which is determined by multiplying the standard error (denoted by σ) of said estimate by some real number Z (in the event CI denotes a 95% confidence interval, then $Z = 1.96$). Thus, Y_{CI} refers to the *closest possible estimate of* Y_{it+1} *to the "boundary mean"* (i.e., $\bar{Y}(\partial S)$) contained within some confidence interval fitted to $Y(\boldsymbol{I})$. If it is the case that $\bar{Y}(\partial S)$ is contained within this confidence interval for every point in ∂S, it follows that:

$$\oint_{\partial S} -Y_{CI}(\boldsymbol{I})dx_1 + Y_{CI}(\boldsymbol{I})dx_2 = \oint_{\partial S} -\bar{Y}(\partial S)dx_1 + \bar{Y}(\partial S)dx_2,$$

from which follows:

$$\oint_{C^1} \bar{Y}(\partial S)dx_1 = \oint_{C^2} \bar{Y}(\partial S)dx_2 = \oint_{C^3} \bar{Y}(\partial S)dx_1 = \oint_{C^4} \bar{Y}(\partial S)dx_2,$$

and therefore:

$$\oint_{\partial S} -\bar{Y}(\partial S)dx_1 + \bar{Y}(\partial S)dx_2 = 0.$$

Thus, two additional surfaces over x_3 and x_4 (as defined previously) will be estimated using this approach, one for which CI denotes a confidence interval of 95% (i.e., $Z = 1.96$), and one for a confidence interval of 99% ($Z = 2.576$).[39] In addition, an identical process will be employed to examine the robustness of the estimated surface depicting the value of $\oint_{\partial S} \psi$ with respect to x_1 and x_2 (i.e., that which was employed to evaluate Hypothesis 2). Finally, a similar approach can be taken with respect to testing the significance of any fitted set of estimates of $\iint_{R(X)}(*)$ with respect to x_2 (as well as x_1). Namely, a confidence interval can be fitted around the estimated value of Y_{it+1} for each of the lowest and highest possible values of x_1 (e.g., L and U), with respect to some set of fixed values for x_2, x_3, and x_4. In the event the boundary mean of Y_{it+1} (e.g., $\frac{Y(L)+Y(U)}{2}$) is contained in both confidence intervals, we can simply assign said average to L and U, in which case the value of $\Delta Y(x_1)$ will equal zero for some unique combination of values for

[39] The method I used for this approach simply involves extracting the value of the standard error of a single "point" of $\boldsymbol{I}_{it}$, which can be done using the predict.gam command for the mgcv package, with the value of "se.fit" set to "TRUE." I then programmed an algorithm to re-run each of the previous model algorithms, with one exception: the predicted value of Y_{it+1} was set to be equal to Y_{CI} (as defined above), the "width" of which is determined by multiplying the extracted standard error value by either 1.96 (95% interval) or 2.576 (99% interval). If the surface of the model manifold (i.e., the predicted value of Y_{it+1} as estimated via Equation 2) falls within every confidence interval established per point of $\boldsymbol{I}_{it}$, then every "interaction" surface fitted using the methods established in this chapter will be completely flat (i.e., assigned a value of zero across all points of each "fixed" variable). Additionally, if it is the case that some subareas of each of these surfaces were fitted to a "noisier" predicted value of Y_{it+1}, compared to other areas, then those sections will be "penalized" more heavily (i.e., assigned a value closer to zero). This can allow for one to monitor any changes in the overall *shape* of each surface (e.g., some sections of S which initially were assigned a "strong" estimated influence of x_1 and x_2 on Y_{it+1} may now assign a *weak* effect).

x_2, x_3, and x_4. If $\Delta Y(x_1) = 0$ for all x_3 and x_4, given some fixed x_2, then so too will $\iint_{R(X)}(*) = 0$ at that particular value of x_2. Hence, two additional curves will be fitted for the moderating influence of multiple reward types on criminal behavior, employing confidence intervals of 95% ($Z = 1.96$) and 99% ($Z = 2.576$).

A NOTE ON BETWEEN VERSUS WITHIN-PERSON COMPARISONS:

A potential criticism of this dissertation's analytic strategy relates to the employment of between-person comparisons with respect to reported values of the rewards, risks, and costs to crime. In particular, recent scholarship has suggested that measures of reported probabilities of arrest may, on some level, be "arbitrary" insofar as comparing the scores of one participant against those of another (Thomas et al., 2018). Because of this, some scholars have advocated for moving away from making between-person comparisons of reported arrest probabilities entirely, and to instead make *within-person* comparisons whenever possible (e.g., a fixed effects model of some kind; Thomas et al., 2020). Regardless, I have decided to employ between-person comparisons within each of the models estimated within this project for two reasons.

First, between-person comparisons allow for a relatively straightforward—and possibly more *meaningful*—interpretation of the results produced by each of the models described throughout this chapter. For instance, the model manifold estimated via Equation 2 can be thought of as the predicted value of Y_{it+1} for any given person i who reports a particular set of values for each element of $\boldsymbol{I}_{it}$. Thus, any estimation of the overall "change" in Y_{it+1} with respect to, say, x_1 and x_2 can be interpreted as the *average influence* that variation in x_1 and x_2 will exert on self-reported involvement in criminal activities more broadly (given some set of fixed values assigned to x_3 and x_4). By contrast, a within-person model is derived from comparing each participant's reported values for each time-varying predictor (e.g., $\boldsymbol{I}_{it}$) at a specific time period relative to his or her *person-specific* mean (e.g., $\bar{\boldsymbol{I}}_i$) across all time periods of the study (we can denote such comparisons by $\ddot{\boldsymbol{I}}_{it}$, which is obtained via $\boldsymbol{I}_{it} - \bar{\boldsymbol{I}}_i$). As a consequence, the "space" of possible combinations of values for any fixed x_3 and x_4 (i.e., $R(X)$), for example, now refers to the *degree of deviance* of each participant's reported values for x_3 and x_4, at time t, from his or her respective sample average(s). As such, the boundary of $R(X)$ is no longer determined by the range of possible values of x_3 and x_4 any participant might potentially give at some data wave, but rather the *range of deviance scores* supplied by sample participants.[40] Because of this, some participants' feasible set of values for x_3 and x_4 may only occupy a *subset of* $R(X)$ (e.g., any person who provides a mean response of 8 for the x_3 measure has a set of possible deviance scores bounded between -8 and 2 for any time period t). Likewise, if one were to hold participants' time-demeaned responses for x_3 constant at -7, for instance, then one must also assume such a score will be meaningful for participants who *do not provide a mean response of 7 or greater*.[41] Such an assumption can be avoided through the usage of

[40] This implies the value of x_3 now belongs to some subset of $(-10,10)$ rather than [0,10] (i.e., the boundary determined by the measure for x_3 provided within the Pathways data). Note the "open" parentheticals used for the $+/-10$ interval, which simply state that a "within-person" measure of x_3 can never truly achieve a value of either plus or minus 10. This, of course, follows from the computation of the mean, as in order to achieve a deviance score of -10, for instance, the individual must simultaneously report a value of 0 for a given data wave while being assigned a person-specific mean score of 10. However, since the survey measure for x_3 is bounded from above by 10, it follows that a person-specific mean of 10 is achieved if and only if the participant reports a value of 10 *across all data waves*. Therefore, a deviance score of -10 is impossible (the same also applies to a deviance score of 10).

[41] Note that most linear models which employ time-demeaned predictors *do not make this assumption*. Since the predicted value of Y in such models is determined by a linear combination of the predictor variables, each of which is held constant at its respective sample mean, it follows that any coefficient β_k in such models can be interpreted as the average influence predictor k exerts on Y given each control variable is held constant at zero (which follows from the sum of every participant's person-specific deviance scores being equal to zero, and therefore any "global" mean of k will equal zero as well). Since a value of zero will always fall within the range of possible values provided by participants (i.e., it is always true that $0 \in [L_i, U_i]$, where L_i and U_i respectively denote the minimum and maximum values of predictor k provided by participant i), such models

a between-person model (e.g., it is always possible, in theory at least, for participant i to provide a response anywhere between 0 and 10 for x_3 at any t), and thus the utilization of between-person comparisons for each predictor seemed more appropriate for this analysis.

Second, recent scholarship has suggested that within-person comparisons—particularly, those made for self-reported arrest probabilities—may be insufficient to fully address the identification problems induced by measurement error. For instance, a study by Hamilton (2023) highlighted several potential *time-varying* sources of "arbitrariness" in answers provided for probabilistic survey measures, including the degree of confidence a person feels in his or her responses at a given time period (Bruine de Bruin et al., 2002), as well as his or her propensity to provide satisficed answers between data waves. As an example, a participant may "round" his or her numerical responses to a higher degree—if not provide meaningless answers altogether (Manski & Molinari, 2010)—during later periods wherein his or her interest in the study has waned (see: Roberts et al., 2019). Additionally, the results of the Hamilton (2023) study suggested that, even under considerably weaker measurement assumptions, supportive conclusions for the existence of a negative link between risk perceptions and criminal behavior can still be drawn from participants' reported probabilities of arrest (specifically, by comparing the self-reported offending levels of participants who reported a higher likelihood of capture against those who reported a lower perceived arrest risk). Such findings call into question the supposed meaninglessness of comparing reported probabilities of arrest between persons, as one would normally expect a relatively "arbitrary" set of responses to produce a *more fragile* estimate of some underlying relationship between perceived arrest risk and criminal

---

are able to avoid making any problematic assumptions about the possible values participants are (theoretically) capable of reporting.

behavior (that is, any between-person model should elicit a “noisier” estimate of β; Braga & Apel, 2016). As such, one could argue that any “loss” in credibility of inference (see: Manski, 2003) induced through the utilization of between-person comparisons may be sufficiently compensated by, say, an improvement in the *interpretability* of one’s results (namely, those achieved within each of the models outlined throughout this chapter). Such a “gamble” was one I was more than willing to take, and thus I opted to forego a strictly within-person analysis for this project.

Chapter 4: Results

(Insert Table 2 about here)

Table 2 provides a list of summary statistics for each of the variables discussed in Chapter 3. Specifically, Table 2 presents the mean (or percentage), standard deviation, median, range, and IQR for each of the perceived reward, risk, and cost variables, as well as each control and outcome measure, captured during the first wave (post-baseline) of the Pathways to Desistance Study. Since the outcome variable is set to "lead" the independent variables at $t + 1$, each summary measure presented in Table 2 for offending variety is captured at the second data wave. Here, evidence of an overdispersed outcome measure is clearly seen, as the median value for self-reported offending variety is zero (implying at least half of the participants reported committing no offenses during wave 2), with a mean of 1.4, a standard deviation of 2.32, and a maximum reported value of 17. Both the social and personal reward measures display a similar (though less extreme) downward bias, with mean scores of 1.94 (S.D. = .49) and 2.27 (S.D. = 2.49), respectively. Additionally, the third quartile score for the social rewards is 2.27, which falls below the "middle-most" value of the social rewards scale (i.e., 2.5), while the third quartile for the personal rewards is 3.86 (which falls below 5). By contrast, the perceived arrest risk and social cost variables display fewer signs of skewness (though other distributional problems can often arise in perceived probability measures; see: Hamilton, 2023), with respective means of 5.2 (S.D. = 2.99) and 2.97 (S.D. = .91). About 13.28% of participants reported being actively employed during wave 1, as well as provided a mean response of 3.07 (S.D. = .94), 3.06 (S.D. = .46), and 2.48 (S.D. = .58) for the impulse control, psychosocial maturity, and future orientation measures, respectively. The mean age (in years) of the sample, as measured during the baseline wave, is 16.04, with a standard deviation of 1.14 and a range of 14 to 19 (note that participants were only required to have *committed* an offense prior to the age of 18 in order to be included in the study, meaning those released *after* the age of 18 could still qualify for as long as their crime occurred during adolescence; Schubert et al., 2004). Around 13.6% of participants were coded as female (86.4% male), 20.2% were coded as white, 41.4% black, 33.5% Hispanic, and 4.8% were coded as "other race."

(Insert Table 3 about here)

Table 3 provides a summary of the linear model outlined in Equation 4. Overall, the perceived social and personal rewards variables are estimated to have a positive influence on offending variety, while the perceived arrest risk and social costs measures are estimated to be negatively related to self-reported criminality. The coefficients fitted for the social and personal benefits variables are .26 (S.E. = .03) and .09 (.006), respectively, while the respective coefficients for the arrest risk and social costs measure are -.06 (.005) and -.08 (.016). Each perceived incentive variable is significantly related to offending variety, with corresponding p-values less than .001. For the control variables, each of the coefficients for impulse control (-.26), psychosocial maturity (-.06), future orientation (-.13), and employment (-.006) are negative; however, only the impulse control ($p < .001$) and future orientation measures ($p < .001$) are significantly related to self-reported offending (maturity is marginally significant at $p < .10$). For the demographic measures, offending variety is estimated to decrease with age (β = -.053; $p < .001$), female participants reported fewer offenses relative to males (β = -.52; $p < .001$), and black (β = -.29; $p < .001$) and Hispanic (β = -.18; $p < .001$) participants, along with participants coded as "other race" (β = -.017; not significant), reported fewer offenses relative to whites.

Furthermore, since the model is fitted using a Tweedie distribution, the mgcv package in R returns an estimate for the variance parameter fitted to the outcome variable, denoted by "p." As discussed by Wood (2017, p. 115), the value of p is (in part) a function of the degree of overdispersion of the outcome measure, such that p takes on a value closer to 2 as the level of overdispersion increases (closer to 1 otherwise). For this model, the estimated value of p is 1.21, suggesting that some level of correction for overdispersion is necessary (see: Dunn & Smyth, 2005). Table 3 also provides the number of cases utilized within each of the models (i.e., the pooled sample size, denoted by N), which is 11,529.

(Insert Table 4 about here)

Table 4 summarizes the marginal smooth model outlined in Equation 5. For this model, the perceived incentive measures are now individually fitted with a smoothing parameter (i.e., f_k), each of which is constructed by a "linear" combination of 5 basis functions (i.e., $f_k = \sum \gamma_j \lambda_j(x_k)$; $j = 1, 2, \dots, 5$). Note these basis functions take the form of cubic regression splines (as discussed in Chapter 3) and can thus be used to estimate a nonlinear association between any predictor (i.e., x_k) and the response variable (Y_{it+1}). To get a sense of the overall influence of each incentive variable on offending, the model provides a set of estimates for the *effective degrees of freedom* (EDF) of each smoothing parameter f_k.[42] It should be noted, however, that generalized additive models often employ a different definition for "degrees of freedom" compared to more standard modeling practices (e.g., strictly linear models). More specifically, the EDF value of a smooth term f_k, for some predictor x_k, is an approximation of the degree to which f_k seems to "fit" the observed data points above and beyond any functional form of x_k contained in the *null*

[42] Note that EDF values are computed for any model fitted using REML methods (see: Wood, 2017, p. 251).

space of the model (see: Wood, 2017, pp. 251-252). For this analysis, said null space contains the "zero" function (i.e., $H_0 : f_k = 0$ for all values of x_k; Wood, 2017, p. 305), and thus the EDF value can be interpreted as the extent to which f_k appears to "explain" the variance of Y_{it+1} relative to a scenario where x_k has *no influence* on the response variable. As such, the p-value estimated from the model for each smooth f_k takes on a similar meaning to how it is usually interpreted in linear models (e.g., the probability that some relationship between x_k and Y_{it+1} was generated by chance; Wood, 2017, p. 304). Furthermore, the upper bound of any estimated EDF value is determined by the number of basis functions (i.e., k) deducted by the number of "free" parameters (i.e., the number of predictors) which, in this case, is equal to 1 for all f_k. Thus, the EDF values for each f_k in this model cannot exceed a value of 4.

Each of the estimated relationships between the control variables and offending variety, for the marginal smooth model (Equation 5), display nearly identical properties to those estimated in the purely linear model (i.e., both sets of variables are assigned similar coefficients and p-values). For the smooth terms (f_k), the estimated EDF values for the social and personal rewards measures are 2.71 and 3.69, respectively, while the EDF values for the arrest risk and social costs measures are 2.72 and 2.63, respectively. Note that the EDF values are positive for each of the perceived incentive variables, including the risk and cost measures. This is to be expected, since EDF is a measure of the overall "fit" of a predictor and thus tells us nothing about the *directionality* of the association (getting a sense of the overall "direction" of f_k is mostly achieved through a direct visualization of f_k). Additionally, all four EDF values for the perceived incentive measures achieve a p-value less than .001, and hence the model suggests that offending variety can be meaningfully explained, on some level, by the perceived social and personal rewards, arrest likelihood, and social costs to crime.

(Insert Table 5 about here)

Table 5 outlines the estimates of the model represented by Equation 3. Here, each of the marginal smooth terms (i.e., f_k) are replaced by a single smooth—specifically, the *tensor product smooth*—denoted by T_e, and, once again, the control variables are estimated to have a similar influence on offending variety as the previous two models (with the exception of the maturity variable, which is no longer marginally significant). Notice the number of basis functions (i.e., the value of k) for T_e is now 625, whereas the marginal smooths were estimated using only 5 basis functions each. This is due to the way in which the tensor product is constructed; namely, T_e denotes a smooth function of *multiple variables*, the values of which are treated analogously to components of an m-vector contained in some subset of $\mathbb{R}^m$ (i.e., the set $\mathcal{U}$ as previously defined in Chapter 3). Consequently, the tensor product essentially "views" each input variable not as a distinct entity, but rather *different parts of a single underlying object* (i.e., T_e fits a predicted value of Y_{it+1} to *every combination of values* for its predictor variables, as opposed to fitting a separate estimate of Y_{it+1} to the values of each individual predictor; Wood, 2017, pp. 227-228). Likewise, each basis function for T_e denotes a unique combination of "marginal" basis functions (i.e., λ_j) assigned to every predictor (x_k), the shape of which can be thought of as an m-dimensional surface (i.e., *manifold*) embedded in $\mathbb{R}^{m+1}$ (see: Spivak, 1965). Since each perceived incentive variable is fitted with 5 basis functions (as determined in the previous model), we can identify the total number of basis functions for T_e simply by computing the total number of combinations of basis functions for each x_k (i.e., we take 5 times itself for 4 iterations, or $5^4 = 625$). Furthermore, since T_e treats any given combination of values for each x_k as a single data point, the number of free parameters in the model remains equal to 1, and hence the upper limit of any estimated EDF value for T_e is 624. Per Table 5, the EDF value for the tensor product term is 91.82, which is statistically significant at $p < .001$.

(Insert Table 6 about here)

To compare the overall performance of each model (i.e., the degree to which each model's predictions for Y_{it+1} appear to "fit" the observed distribution), the respective AIC, REML, and adjusted r^2 values per model were estimated. Table 6 provides a summary of these estimates. Overall, the tensor product smooth model seems to perform the best of the three, achieving simultaneously the lowest AIC (31,051) and REML (15,568) scores, as well as the highest adjusted r^2 value (.148), relative to the marginal smooth and linear models. The marginal smooth model performed second best, achieving an approximate AIC score of 31,094, an REML score of 15,574, and an adjusted r^2 of .131. The linear model exhibited the "worst" performance of the three, with an AIC of 31,199, an REML of 15,625, and an adjusted r^2 of 0.127. In summary, the overall fitness measures seem to favor the more "complicated" models over that of the strictly linear model, implying that not only are the effects of the perception variables likely to be nonlinear, but there may also exist substantial *interdependencies* between them.

(Insert Figure 1 about here)

Figure 1 provides a sequence of plots for each of the marginal smooths fitted in the model outlined in Equation 4. The top-left plot provides a visualization of the estimated marginal smooth for the perceived social rewards (note that the solid curve embedded in this figure denotes the estimated value of Y_{it+1} for each value of the social rewards measure, while the dotted curves define a 95% confidence interval fitted for each estimate), and the top-right plot visualizes the marginal smooth with respect to the perceived personal rewards. The bottom-left plot visualizes the marginal smooth for perceived arrest risk, and the bottom-right plot captures the smooth fitted for the social costs measure. For the most part, each of the functional forms depicted in these plots are (relatively) nonlinear, as the social rewards appear to display a non-monotonic function (i.e., the "effect" is not strictly increasing for all values of x_1), while the personal benefits show signs of exhibiting a "diminishing" influence on self-reported criminality at higher values. The function fitted to the perceived arrest risk measure mimics that of a negative "s-curve," while the social costs measure seems to have little influence on offending variety at lower values. Of course, such "marginal" curves are only meaningful insofar as the relationships between each perceived incentive variable and self-reported offending are *independent* of one another. To get a sense of their potential interdependencies, a more detailed look at their *combined functional form* is needed.

(Insert Figures 2.1 and 2.2 about here)

Figures 2.1 and 2.2 provide visualizations of the estimated value of $\oint_{\partial S} \omega$ (as previously defined in Chapter 3) for each combination of values for the perceived arrest risk and social cost measures. More concretely, both figures represent the degree to which the overall influence of the perceived social and personal benefits on self-reported offending *varies with respect to perceived arrest likelihood and social costs*. Such a function takes the form of a 2-dimensional surface embedded in $\mathbb{R}^3$, for which Figure 2.1 provides a depiction of (note the "I(Omega)" label for the z-axes of this figure refers to

$\oint_{\partial S} \omega$, and thus can be read as "the strength of the overall influence of the social and intrinsic rewards on offending variety"). On the whole, the rewards to crime seem to have the strongest effect on offending variety at lower values of perceived arrest risk (x_3) and social costs (x_4), while exhibiting a much weaker overall influence at both middling and higher values. Because of this, the results shown in 2.1 align closest with the expected distributional form outlined in Hypothesis 1a (i.e., the prediction that the general influence of the rewards to crime will diminish as the risks and costs to crime increase).

It should be noted that the surface described in Figure 2.1 is drawn to scale (i.e., the "height" of the surface is always fitted to the highest and lowest estimated values for the z-axis), and thus only summarizes the general *shape* of the surface. As such, Figure 2.2 provides as a means for observing the *values* of the general strength of the association between the rewards to crime and offending variety (i.e., $\oint_{\partial S} \omega$) computed at each combined value of the perceived arrest risk and social costs variables (i.e., a topological map). Overall, the estimated strength of the social and intrinsic rewards measures achieved a value anywhere between 15 and 55 (possibly higher) between the lowest and middle-most values of the perceived arrest risk and social costs variables, as well as a value of roughly 10 or less at each higher value of the risk and cost measures. Once again, the results shown in 2.2 exhibit a general decrease in the effectiveness of the rewards measures as the perceived risk and cost variables approach their respective highest values (as anticipated by Hypothesis 1a).

(Insert Figures 2.3, 2.4, 2.5, and 2.6 about here)

Figures 2.3 and 2.4 provide additional estimates of the "surface" depicted in Figures 2.1 and 2.2. Specifically, each figure denotes the "worst-case" (i.e., closest to zero) estimate of the rewards effect ($\oint_{\partial S} \omega$) within a 95% confidence interval fitted across the perceived arrest risk and social cost measures. Overall, the shape of the distribution remains (mostly) intact, although the estimated value of the rewards effect is considerably smaller relative to those shown in Figure 2.2 (as to be expected). The same can also be seen for the surface fitted by a 99% confidence interval, as shown in Figures 2.5 and 2.6. Despite the computed value of the rewards effect being even closer to zero, the general shape of the surface remains relatively static. As such, the overall functional form outlined in Hypothesis 1a can still be seen in 2.3 and 2.5, despite employing considerably weaker identifiability assumptions, thus providing more robust support for Hypothesis 1a.

(Insert Figures 3.1 and 3.2 about here)

The surface and topological plots respectively provided in Figures 3.1 and 3.2 visualizes the estimated strength of the association between the perceived arrest risk and social cost measures (i.e., the value of $\oint_{\partial S} \psi$, as defined in Chapter 3) with respect to the perceived social (x_1) and intrinsic (x_2) benefits to crime (note the "I(Psi)" label on the z-axis in Figure 3.1 refers to computed value of $\oint_{\partial S} \psi$). Similar to before, the overall influence of perceptions of arrest risk and the social costs to crime appears to "strengthen" as the social and personal rewards measures approach their respective highest values (note that *lower* values of $\oint_{\partial S} \psi$ correspond to a stronger estimated effect of the risks and costs to crime, since both are hypothesized to be negatively related to a person's involvement in criminal acts). Because of this Figure 3.1 aligns most with the predicted functional form outlined in Hypothesis 2b. Furthermore, as shown in Figure 3.2, estimates of the general inhibitory influence of the risk and cost variables reach a value of -30 near the highest values of the perceived social and intrinsic reward measures, and approximately a value of

-4 near the lowest values of the rewards. Such estimates imply that offenders who harbor higher reward perceptions are, generally speaking, substantially more responsive to variation in their expectations of arrest likelihood and the social losses to illegal activities (as suggested by Hypothesis 2b).

(Insert Figures 3.3, 3.4, 3.5, and 3.6 about here)

However, this generally decreasing (i.e., *strengthening*) pattern of the overall influence of perceived arrest risk and social costs on offending variety, with respect to perception of the social and personal rewards to crime, is not observed for the surface fitted within a 95% confidence interval. As shown in Figures 3.3 and 3.4, the lowest computed values of the inhibitory effect of the risks and costs ($\oint_{\partial S} \psi$) are now estimated to fall somewhere within the intersection of "lower-mid" values of the perceived social rewards measure and highest values of the personal rewards variable. Additionally, Figure 3.4 suggests the range of estimated values for said inhibitory effect now fall somewhere between 0 and -3, as opposed to the range of -4 and -30 seen in Figure 3.2. Unsurprisingly, the results of the surface fitted within a 99% confidence interval, as shown in Figures 3.5 and 3.6, are nearly identical to those found using a 95% confidence interval (the primary difference, however, is that the range of estimated values for the general influence of the risk and cost measures now fall within a range of 0 and -1.4). Relative to the primary surface visualized in Figures 3.1 and 3.2, each of the worst-case manifolds depict a far "noisier" relationship between the overall inhibitory effect of the perceived risks and costs to crime with respect to variation in participants' reported reward perceptions. Furthermore, both manifolds are considerably "flatter" than the primary surface as well, as their respective minimum estimated values for the effectiveness of the risks and costs are closer to zero than almost every point of the surface shown in 3.2. In particular, the surface fitted within a 99% confidence interval achieves an approximate minimum value of -1.4, while the primary surface achieves a *maximum* of approximately -4. Likewise, the results of the worst-case models suggest it is possible that the perceived social and intrinsic rewards to criminal acts have *zero influence* on an offender's overall responsivity to his or her perceptions of arrest risk and social costs (or, if any such influence exists, it does not mimic the distributional form anticipated by any of the hypotheses outlined in Chapter 2).

(Insert Figures 4.1 and 4.2 about here)

Figure 4.1 provides a set of plots visualizing the estimated value of the overall influence of the perceived social rewards to crime (i.e., $\iint_{R(X)}\left(\Delta Y(x_1)\right)$, as defined in Chapter 3) with respect to the personal benefits to crime. The uppermost plot depicts the overall change in the influence of the social rewards to crime on offending variety for each value of the personal rewards variable, while the two additional plots depict the same curve as the top plot except fitted with 95% and 99% confidence intervals, respectively. Overall, the criminogenic influence of the social benefits to crime seems to be weakest at the highest value of the personal rewards measure, and strongest at the lowest value of said measure as anticipated by Hypothesis 3a. However, one also observes a fair degree of non-monotonicity in the functional form of the underlying moderating relationship, as the curve depicted in the top plot of 4.1 appears to reverse direction at two distinct points. Although one could argue the curve seems to depict a generally downward trend overall, the sudden "upswing" certainly is strange when considering that each hypothesis derived within the previous chapter was based upon the idea of there existing a *tipping point* of some kind (i.e., a specific point at which an effect is strongest, and tends to diminish the further one "moves away" from it along the set of values for some set of moderating variables; see Appendix A for further elaboration).[43] Because of this, an argument could also be made that the results of Figure 4.1 do not support any of the hypothesis established in Chapter 2. This argument is further supported by each of the "worst-case" curves fitted within the 95% and 99% confidence interval models displaying a similar non-monotonic pattern as the primary curve (although the distribution considerably "flattens" for any value of the reported value of the personal rewards measure which is greater than 2). As such, while the results shown in Figure 4.1 seem to support the notion that the personal benefits may exhibit at least some moderating influence on the criminogenic effect of the social rewards, it is not entirely clear whether said moderating influence truly aligns with Hypothesis 3a.

Finally, Figure 4.2 depicts the overall influence of the personal rewards on self-reported offending variety with respective to the perceived social rewards. On the whole, the curve shown in the upper-most plot in 4.2 clearly depicts a non-monotonic functional form of the estimated strength of the intrinsic rewards variable (i.e., $\iint_{R(X)}\big(\Delta Y(x_2)\big)$), with respect to the social rewards measure, except in this instance it appears there exists only a *single* change in directionality; one which occurs at some middling value of the social rewards variable as anticipated by Hypothesis 3c. In particular, the overall shape of the curve shows a general *increase* in the strength of the personal rewards influence between

---

[43] Another peculiar observation shown in Figure 4.1 is that at higher values of the perceived personal rewards to crime, the overall influence of the social rewards is suggested to have a *negative effect* on self-reported offending. Such a finding certainly contradicts the expectation that the rewards will always be positively linked to crime in some capacity. Of course, such a pattern may also just be a product of the way in which the tensor product smooth is computed, as the primary goal of any generalized additive model is to fit a functional form that is "closer to reality," but not *so close* as to generate meaningless predictions. Likewise, in attempting to "smooth out" the functional form of some underlying empirical association, it is fully possible for the model to "overcorrect" for chance variation. That is, such a model may fit a more rigid set of curves (or a higher dimensional surface) than might otherwise be appropriate using "real world" data, which could lead to an overestimation (or underestimation) of the predicted values of the outcome variable at certain points of each of the predictor variables (see Wood, 2017).

the lowest reported value of the social rewards to crime (i.e., 1) up until roughly a value of 2, while also showing a general *decrease* between in the strength of the personal rewards effect between a reported value of 2 and the highest value for the social benefits measure (i.e., 4). A similar pattern is reflected in the sub plots fitted with 95% and 99% confidence intervals, although both also depict a slight "upswing" in the effect of the personal rewards measure on offending variety around between a reported value of 3.5 and 4 on the social benefits variable. However, said upswing is also somewhat negligible, and could simply be a product of a generally higher quantity of "reporting error" near the highest value of the social rewards measure (given the curve actually *approaches zero* between a social rewards value of 3.5 and 4, this could very well be the case; see footnote 43). Likewise, each of the plots provided in 4.2 seem to show at least some support for the predicted distributional form outlined in Hypothesis 3c.

Chapter 5: Discussion and Conclusions

Skepticism toward the rational choice perspective has permeated the criminological discourse since the resurgence of the deterrence and offender decision-making doctrines. Indeed, a commonly held belief among criminologists has been that criminal acts can rarely, if ever, be reasonably described as rational (e.g., Gottfredson & Hirschi, 1990). If such acts do appear to display some level of rationality (e.g., offenders seem to respond to perceptions of incentives to some extent; Loughran et al., 2016a), then it is often suggested that it must be of a decidedly "bounded" nature (e.g., Cornish & Clarke, 1986). In particular, the notion that criminals engage in any calculative process prior to offending, as commonly outlined in normative theories of decision-making, has received a fair amount of pushback within the discipline (e.g., De Hann & Vos, 2003). Some scholars have even argued that to truly understand the rational dimensions of crime and criminals, one must first dispense of the belief that criminals are willing—much less *capable*—of making good decisions for themselves (e.g., Hirschi, 1986). Put differently, much of the scholarly debate on rational offending up to this point has centered on the question of whether criminal actions can be realistically described as a means of *maximizing utility* for the offender. More specifically, do the illicit activities of persons appear to follow the prescriptions of an underlying utility calculus, as is often defined within normative theories of rational choice (e.g., Becker, 1986)? This dissertation sought to provide at least a cursory answer to this question by investigating the potential interdependencies between rational choice inputs; particularly, those which seem to align with the rational choice calculus as it is defined in Becker's (1986) expected utility model of crime.

The results of this investigation appear to provide some level of support for the existence of such interdependencies between measures of the perceived rewards, risks, and costs to crime and self-reported offending. More specifically, the offending patterns of participants who reported relatively “low” perceptions of arrest risk and social costs seemed to be more responsive to variation in their perceived social and personal benefits to crime (as outlined in Hypothesis 1a). By contrast, participants who suggested they were more likely to be caught and arrested for engaging in criminal acts, and would (consequently) endure greater social losses, appeared to be much less sensitive to the perceived benefits of criminal actions. In addition, persons who reported lower perceptions of the social and intrinsic benefits to crime appeared to be less responsive, on the whole, to variation in reported arrest risk and social costs (as suggested in Hypothesis 2b). However, this result was not robust to the possible influence of sampling variation, as both surfaces depicted in Figures 3.3 through 3.6 not only achieved a relatively “flat” distribution of the overall estimated influence of perceived arrest likelihood and social losses on criminal involvement, with respect to the perceived social and personal benefits to crime, but neither exhibited a similar overall pattern to that which was found in the primary model (as shown in Figures 3.1 and 3.2). This analysis also suggested that both the social and personal benefits to crime exerted an overall *negative* moderating influence on one another, as higher reported values of the social rewards to crime were typically associated with a lesser influence of the perceived personal rewards on self-reported offending, and vice versa. Furthermore, the results also seemed to suggest that each interrelation between the perceived incentive variables exhibited a fair degree of nonlinearity. Namely, the criminogenic influence of the rewards variables saw a nearly *exponential* increase in strength as both the perceived arrest risk and social cost variables approached their respective lowest values, while the interactive effects of the perceived social and personal rewards measures displayed some level of non-monotonicity (i.e., increasing for a subset of values for each variable while decreasing in others). Overall, the results of the previous chapter seem to at least imply that the strength of the rewards-crime link is likely to be a function of perceived arrest likelihood and the social losses to crime, as well as alternative types of benefit perceptions.

What is perhaps the most compelling implication of the results produced throughout this dissertation is the challenge those results pose to the notion that criminality does not conform to the ostensibly unrealistic depiction of offenders provided by rational choice theorists (particularly, Gary Becker). Indeed, a common refrain even among choice-oriented criminologists is that the "hyper-rationalized" image of the average criminal, as is often put forward by economists, rarely conforms to real world decision-making, and thus the discipline is likely better off parsing out the myriad of ways in which the offender's choices will *deviate* from normative models of rational decision-making, rather than conform to them (e.g., the application of the many disparate empirical findings of behavioral economists; Pogarsky, Roche, & Pickett, 2018). Although efforts toward examining the supposedly "irrational" characteristics of crime decisions are likely to be fruitful in their own right, one cannot help but wonder if foregoing a more in-depth examination of the many (yet to be discovered) implications of such models may result in inadvertently throwing the baby out with the bath water (a similar concern is raised by Thomas and colleagues [2020]). If the results outlined in the previous chapter are anything to go by, it would seem the purported "uselessness" of normative theories of rational choice for the study of crime might have been overstated. If anything, it is likely we have only begun to scratch the surface of what can potentially be learned from such theories.

Although the results of this study seem to support the existence of an overall interdependent relationship between rational choice constructs, further research will likely be necessary in order to tease out the underlying mechanisms of this relationship. In particular, efforts should be taken to see if there are any conditions under which the distributional form of any specific interrelation (e.g., the interdependency of the rewards on the risks and costs) may shift in directionality, as well as monotonicity (e.g., are there some offenders for which the rewards to crime exert a stronger influences on criminal acts at *higher* values of risk and cost perceptions, as suggested by Hypothesis 1b?). Given the relatively "flexible" nature of the hypotheses provided in the second chapter, such a shift may well align with the predictions outlined in Equation 1. Another potential avenue for further exploration would be to investigate whether the interdependent relations observed in this dissertation are similar to those found in alternative datasets. Particularly, datasets capturing the perceived benefits, risks, and costs to crime held by *non-serious* offenders. Such efforts would not only help to evaluate the generalizability of this dissertation's results, but might also provide additional insight into *why* the particular distributional patterns observed in the previous chapter were found.[44] Additionally, examinations employing a wider range of perceived rewards, risks, and cost measures would likely be beneficial as well. Namely, a broader range of reward *types* (e.g., the expected monetary

---

[44] As an example, it may be the case that serious offenders who hold higher perceptions of, say, the social rewards to crime will often appear to be less sensitive to the personal rewards due to finding the social benefits to crime to be generally sufficient to justify being involved in crime on some level, while less serious offenders may find the social benefits alone to be generally *insufficient* to justify criminal acts, and thus may appear to be *more* sensitive to the personal rewards to crime as their perceptions of the social benefits increase.

rewards to crime; Loughran et al., 2016b) could help determine whether the generally diminishing influence of multiple reward types, as shown in chapter 4, are also seen for anticipated benefit types beyond the social and personal rewards to criminal acts.

As for the potential policy implications of this project, I can confidently (and enthusiastically) say there are none. The notion of somehow making criminal acts less "fun" or socially rewarding for chronic offenders borders on absurdity, and the argument that crime can potentially be deterred by increasing sanction risks and costs has been around since the inception of the discipline (Nagin, 1998). As such, I do not see any possible application of the results provided throughout this dissertation for the inhibition of crime by state practitioners, nor do I feel it would be appropriate to try and stretch said results beyond their logical limits simply for the purpose of having something to say about "solving" the crime problem. The efforts undertaken here were purely and unapologetically conducted for the purpose of examining some of the more nuanced implications of a rational choice theory of crime. Whether these efforts eventually translate into meaningful policy creation I leave to future research endeavors.

While I firmly believe the decision to employ between-person comparisons within the main body of results of this dissertation was the appropriate choice, future research is likely to benefit from examining whether similar interrelations can be observed using a within-person model of some variety (e.g., a fixed-effects model). Such an examination could provide some insight into whether an individual seems to be more (or less) responsive to some types of (dis)incentives during time periods wherein he or she holds higher perceptions of other types *relative to his or her person-specific mean for those types*. At minimum, such efforts would allow for a more in-depth examination of the *within-*

person implications of the expected utility model, particularly those related to the interrelations of different types of reward, risk, and cost variables. Furthermore, one of the dangers of employing a generalized additive model of any kind is the possibility of the estimated model manifold being less “well behaved” than the researcher would like it to be (e.g., a general tendency toward over or under-estimating some predicted value at specific points of the distribution of responses for the predictor variables; Wood, 2017). Likewise, the utilization of an alternative modeling strategy to that which was outlined in Chapter 3 is likely to be beneficial as well (insofar as examining the *robustness* of the interrelations shown in Chapter 4).

Finally, an even deeper dive into the theoretical underpinnings of the interdependencies of rational choice inputs will likely be immensely beneficial for the discipline. Although I have done my best to try and tease out the broader implications of the rational choice perspective—particularly, the rational choice model developed by Becker (1968)—for those interdependencies, I also found it immensely difficult to achieve a precise set of predictions related to any particular interdependent relationship (for example, it was not clear to me whether Hypothesis 1a should be preferred over 1b or 1c, or vice versa), and thus I ended up settling on the relatively “forgiving” range of predictions outlined in Chapter 2 (for a full breakdown of the extent of my efforts on this front, see Appendix A). However, there could still be reason to believe that a subset of the hypotheses provided in the second chapter should be more “probable” than some number of alternative predictions (e.g., Hypothesis 1a may more closely align with what we would expect to see, given the contents of the expected utility model supplied by Becker). Likewise, a much more in-depth examination of what theory might have to say about the interrelations between rational choice constructs (or possibly even the development of a more detailed rational choice theory of crime) is likely to offer a wealth of insight into both the findings of this dissertation, as well as the underlying nature of crime decisions more broadly.

References

Anderson, E. (1999). *Code of the street: Decency, violence, and the moral life of the inner city*. WW Norton & Company.

Anderson, L. S., Chiricos, T. G., & Waldo, G. P. (1977). Formal and informal sanctions: A comparison of deterrent effects. *Social problems*, *25*(1), 103-114.

Anwar, S., & Loughran, T. A. (2011). Testing a Bayesian learning theory of deterrence among serious juvenile offenders. *Criminology*, *49*(3), 667-698.

Apel, R. (2013). Sanctions, perceptions, and crime: Implications for criminal deterrence. *Journal of Quantitative Criminology*, *29*(1), 67-101.

Apel, R., & DeWitt, S. E. (2018). Informal and formal sanctions. *Deterrence, Choice and Crime: Contemporary Perspectives. NY, New York: Routledge*, 139-156.

Ashraf, N., Camerer, C. F., & Loewenstein, G. (2005). Adam Smith, behavioral economist. *Journal of Economic Perspectives*, *19*(3), 131-145.

Bachman, R., Paternoster, R., & Ward, S. (1992). The rationality of sexual offending: Testing a deterrence/rational choice conception of sexual assault. *Law and Society Review*, *26*(2), 343-372.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime* (pp. 13-68). Palgrave Macmillan, London.

Bender, E. A. (1978). *An introduction to mathematical modeling*. New York: Wiley.

Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation*.

Boyd, S. & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Braga, A. A., & Apel, R. (2016). And we wonder why criminology is sometimes considered irrelevant in real-world policy conversations. *Criminology & Public Policy*, *15*, 813.

Bruine de Bruin, W., Fischbeck, P. S., Stiber, N. A., & Fischhoff, B. (2002). What number is “fifty-fifty”? Redistributing excessive 50% responses in elicited probabilities. *Risk Analysis: An International Journal*, *22*(4), 713-723.

Burt, C. H., & Simons, R. L. (2013). Self-control, thrill seeking, and crime: Motivation matters. *Criminal Justice and Behavior*, *40*(11), 1326-1348.

Carroll. J. S. (1978). A psychological approach to deterrence: The evaluation of crime opportunities. *J Pers Soc Psychol*, *36*, 512–1520.

Chalfin, A. J., & Tahamont, S. (2018). The economics of deterrence: A review of the theory and evidence. *Deterrence, Choice, and Crime, Volume 23*, 29-79.

Cohen, S. B. (1970). The suitability rule and economic theory. *Yale Law Journal*, *80*, 1604-1635.

Cornish, D. B., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending*. Transaction Publishers.

De Haan, W., & Vos, J. (2003). A crying shame: The over-rationalized conception of man in the rational choice perspective. *Theoretical Criminology*, *7*(1), 29-54.

De Mesquita, B. B., & Cohen, L. E. (1995). Self-interest, equity, and crime control: a game-theoretic analysis of criminal decision making. *Criminology*, *33*(4), 483-518.

Dunn, P. K., & Smyth, G. K. (2005). Series evaluation of Tweedie exponential dispersion model densities. *Statistics and Computing*, *15*, 267-280.

Edwards Jr., C. H. (1973). *Advanced calculus of several variables*. Dover Publications.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of political Economy*, *81*(3), 521-565.

Eide, E. (1999). Economics of criminal behavior. In A. Marciano & G. Ramello (Eds.), *Encyclopedia of law and economics* (pp. 345-389). New York: Springer-Verlag.

Erickson, M. L., Gibbs, J. P., & Jensen, G. F. (1977). The deterrence doctrine and the perceived certainty of legal punishments. *American Sociological Review*, *42*(2), 305-317.

Etzioni, A. (1988). *The moral dimension: Toward new economics.* New York: Free Press.

Fishburn, P. C. (1981). Subjective expected utility: A review of normative theories. *Theory and Decision 13*(2), 139-199.

Fishburn, P. C. (1989). Generalizations of expected utility theories: A survey of recent proposals. *Annals of Operations Research*, *19*(1), 1-28.

Flanders, H. (1963). *Differential forms with applications to the physical sciences*. Courier Corporation.

Geerken, M. R., & W. R. Gove (1975). Deterrence: Some theoretical considerations. *Law & Society Review 9*(3), 497-513.

Gollier, C. (1999). *The economics of risk and time*. MIT press.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.

Grampp, W. D. (1948). Adam Smith and the economic man. *Journal of Political Economy*, *56*(4), 315-336.

Grasmick, H. G., & Bryjak, G. J. (1980). The deterrent effect of perceived severity of punishment. *Social forces*, *59*(2), 471-491.

Grasmick, H. G., & Bursik Jr, R. J. (1990). Conscience, significant others, and rational choice: Extending the deterrence model. *Law and Society Review*, *24*(3), 837-861.

Grasmick, H. G., & Milligan, H. (1976). Deterrence theory approach to socioeconomic/demographic correlates of crime. *Social Science Quarterly*, *57*(3), 608-617.

Grasmick, H. G., Blackwell, B. S., Bursik Jr, R. J., & Mitchell, S. (1993). Changes in perceived threats of shame, embarrassment, and legal sanctions for interpersonal violence, 1982–1992. *Violence and Victims*, *8*(4), 313-325.

Halmos, P. R (1960). *Naïve set theory*. van Nostrand.

Hamilton, B. C. (2023). Identification problems in probabilistic measures of perceived arrest risk: Estimating a partially-identified certainty effect. *Journal of Quantitative Criminology*, 1-26.

Herman, S., & Pogarsky, G. (2022). Morality, deterrability, and offender decision making. *Justice Quarterly*, *39*(1), 1-25.

Hirschi, T. (1986). On the compatibility of rational choice and social control theories of crime. In *The reasoning criminal* (pp. 105-118). Routledge.

Hirschi, T. (2004). Self-control and crime. *Handbook of self-regulation*, 537-552.

Jacobs, B. A. (2010). Deterrence and deterrability. *Criminology*, *48*(2), 417-441.

Jensen, G. F. (1969). “Crime doesn't pay”: Correlates of a shared misunderstanding. *Social Problems*, *17*(2), 189-201.

Katz, J. (1990). Seductions of crime: Moral and sensual attractions in doing evil.

Kellett, C. M., & Braun, P. (2023). *Introduction to nonlinear control: Stability, control design, and estimation*. Princeton University Press.

Kolmogorov, A. N., & Fomin, S. V. (1975). *Introductory real analysis*. Courier Corporation.

Kreps, D. M. (2013). *Microeconomic foundations I: Choice and competitive markets.* Princeton University Press.

Kreps, D. M. (2023). *Microeconomic foundations II: Imperfect competition, information, and strategic interaction.* Princeton University Press.

Lee, H., Sullivan, C. J., & Barnes, J. C. (2018). Maturity of judgment and perceptual deterrence. *Criminal Justice and Behavior*, *45*(11), 1762-1781.

Loughran, T. A., Nguyen, H., Piquero, A. R., & Fagan, J. (2013). The returns to criminal capital. *American Sociological Review*, *78*(6), 925-948.

Loughran, T. A., Paternoster, R., Chalfin, A., & Wilson, T. (2016a). Can rational choice be considered a general theory of crime? Evidence from individual-level panel data. *Criminology*, *54*(1), 86-112.

Loughran, T. A., Paternoster, R., Piquero, A. R., & Pogarsky, G. (2011). On ambiguity in perceptions of risk: Implications for criminal decision making and deterrence. *Criminology*, *49*(4), 1029-1061.

Loughran, T. A., Paternoster, R., & Piquero, A. R. (2018). Individual difference and deterrence. In *Deterrence, Choice, and Crime, Volume 23* (pp. 211-236). Routledge.

Loughran, T. A., Paternoster, R., & Weiss, D. (2012a). Hyperbolic time discounting, offender time preferences and deterrence. *Journal of Quantitative Criminology*, *28*, 607-628.

Loughran, T. A., Piquero, A. R., Fagan, J., & Mulvey, E. P. (2012b). Differential deterrence: Studying heterogeneity and changes in perceptual deterrence among serious youthful offenders. *Crime & Delinquency*, *58*(1), 3-27.

Loughran, T. A., Pogarsky, G., Piquero, A. R., & Paternoster, R. (2012c). Re-examining the functional form of the certainty effect in deterrence theory. *Justice Quarterly*, *29*(5), 712-741.

Loughran, T. A., Reid, J. A., Collins, M. E., & Mulvey, E. P. (2016b). Effect of gun carrying on perceptions of risk among adolescent offenders. *American Journal of Public Health*, *106*(2), 350-352.

Lovelock, D., & Rund, H. (1975). *Tensors, differential forms, and variational principles*. Courier Corporation.

Mamayek, C., Loughran, T., & Paternoster, R. (2015). Reason taking the reins from impulsivity: The promise of dual-systems thinking for criminology. *Journal of Contemporary Criminal Justice*, *31*(4), 426-448.

Manski, C. F. (2003). *Partial identification of probability distributions*. Springer-Verlag.

Manski, C. F., & Molinari, F. (2010). Rounding probabilistic expectations in surveys. *Journal of Business & Economic Statistics*, *28*(2), 219-231.

Matsueda, R. L. (2013). Rational choice research in criminology: A multi-level framework. In R. Wittek, T. A. B. Snijders, & V. Nee (Eds.), *Handbook of rational choice social research* (pp. 283-321). California: Stanford University Press.

Matsueda, R. L., Kreager, D. A., & Huizinga, D. (2006). Deterring delinquents: A rational choice model of theft and violence. *American Sociological Review*, *71*(1), 95-122.

Matthews, S. K., & Agnew, R. (2008). Extending deterrence theory: Do delinquent peers condition the relationship between perceptions of getting caught and offending?. *Journal of research in crime and delinquency*, *45*(2), 91-118.

McCarthy, B. (1995). Not just “For The Thrill Of It”: An instrumentalist elaboration of Katz' explanation of Sneaky Thrill Property Crimes. *Criminology*, *33*(4), 519-538.

McCarthy, B. (2002). New economics of sociological criminology. *Annual Review of Sociology*, *28*(1), 417-442.

Mehlkop, G., & Graeff, P. (2010). Modelling a rational choice theory of criminal action: Subjective expected utilities, norms, and interactions. *Rationality and Society*, *22*(2), 189-222.

Mulvey, E. P. (2012). Research on Pathways to Desistance [Maricopa County, AZ and Philadelphia County, PA]: Subject measures, 2000–2010. ICPSR29961-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Mulvey, E. P., Steinberg, L., Fagan, J., Cauffman, E., Piquero, A. R., Chassin, L., Knight, G. P., Brame, R., Schubert, C. A., Hecker, T., & Losoya, S. H. (2004). Theory and research on desistance from antisocial activity among serious adolescent offenders. *Youth Violence and Juvenile Justice*, *2*(3), 213-236.

Munkres, J. R. (1991). *Analysis on manifolds*. CRC Press.

Munkres, J. R. (2021). *Topology (2nd edition)*. Pearson Education.

Nagin, D. S. (1998). Criminal deterrence research at the outset of the twenty-first century. *Crime and justice*, *23*, 1-42.

Nagin, D. S. (2007). Moving choice to center stage in criminological research and theory. *Criminology*, *45*, 259.

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice*, *42*(1), 199-263.

Nagin, D. S., & Paternoster, R. (1993). Enduring individual differences and rational choice theories of crime. *Law and Society Review*, *27*, 467-496.

Nagin, D. S., & Paternoster, R. (1994). Personal capital and social control: The deterrence implications of a theory of individual differences in criminal offending. *Criminology*, *32*(4), 581-606.

Nagin, D. S., & Pogarsky, G. (2001). Integrating celerity, impulsivity, and extralegal sanction threats into a model of general deterrence: Theory and evidence. *Criminology*, *39*(4), 865-892.

Nagin, D. S., & Pogarsky, G. (2004). Time and punishment: Delayed consequences and criminal behavior. *Journal of Quantitative Criminology*, *20*, 295-317.

Osgood, D. W., & Schreck, C. J. (2007). A new method for studying the extent, stability, and predictors of individual specialization in violence. *Criminology*, *45*(2), 273-312.

Osgood, D. W., Wilson, J. K., O'malley, P. M., Bachman, J. G., & Johnston, L. D. (1996). Routine activities and individual deviant behavior. *American Sociological Review*, *61*(4), 635-655.

Paternoster, R. (1987). The deterrent effect of the perceived certainty and severity of punishment: A review of the evidence and issues. *Justice Quarterly*, *4*(2), 173-217.

Paternoster, R. (1989). Absolute and restrictive deterrence in a panel of youth: Explaining the onset, persistence/desistance, and frequency of delinquent offending. *Social Problems*, *36*(3), 289-309.

Paternoster, R. (2010). How much do we really know about criminal deterrence? *The Journal of Criminal Law and Criminology*, *100*(3), 765-824.

Paternoster, R., & Iovanni, L. (1986). The deterrent effect of perceived severity: A reexamination. *Social Forces*, *64*(3), 751-777.

Paternoster, R., & Pogarsky, G. (2009). Rational choice, agency and thoughtfully reflective decision making: The short and long-term consequences of making good choices. *Journal of Quantitative Criminology*, *25*, 103-127.

Paternoster, R., & Simpson, S. (1993). A rational choice theory of corporate crime. In R. V. G. Clarke & M. Felson (eds.), *Routine Activity and Rational Choice* (pp. 37-56). Transaction Publishers.

Paternoster, R., & Simpson, S. (1996). Sanction threats and appeals to morality: Testing a rational choice model of corporate crime. *Law and Society Review*, 549-583.

Paternoster, R., Jaynes, C. M., & Wilson, T. (2017). Rational choice theory and interest in the "fortune of others". *Journal of research in crime and delinquency*, *54*(6), 847-868.

Paternoster, R., Saltzman, L. E., Chiricos, T. G., & Waldo, G. P. (1982). Perceived risk and deterrence: Methodological artifacts in perceptual deterrence research. *The Journal of Criminal Law and Criminology*, *73*(3), 1238.

Patterson, H. D., & Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, *58*(3), 545–554.

Pickett, J., Cullen, F., Bushway, S. D., Chiricos, T., & Alpert, G. (2018). The response rate test: Nonresponse bias and the future of survey research in criminology and criminal justice.

Piliavin, I., Gartner, R., Thornton, C., & Matsueda, R. L. (1986). Crime, deterrence, and rational choice. *American Sociological Review*, *51*(1), 101-119.

Piquero, A. R., & Pogarsky, G. (2002). Beyond Stafford and Warr's reconceptualization of deterrence: Personal and vicarious experiences, impulsivity, and offending behavior. *Journal of Research in Crime and Delinquency*, *39*(2), 153-186.

Piquero, A. R., Paternoster, R., Pogarsky, G., & Loughran, T. (2011). Elaborating the individual difference component in deterrence theory. *Annual Review of Law and Social Science*, *7*, 335-360.

Pogarsky, G. (2002). Identifying "deterrable" offenders: Implications for research on deterrence. *Justice Quarterly*, *19*(3), 431-452.

Pogarsky, G. (2007). Deterrence and individual differences among convicted offenders. *Journal of Quantitative Criminology*, *23*(1), 59-74.

Pogarsky, G., & Loughran, T. A. (2016). The policy-to-perceptions link in deterrence: Time to retire the clearance rate. *Criminology & Pubic Policy 15*(3), 777-791.

Pogarsky, G., Roche, S. P., & Pickett, J. T. (2018). Offender decision-making in criminology: Contributions from behavioral economics. *Annual Review of Criminology*, *1*, 379-400.

Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), *Taking stock: The status of criminological theory* (pp. 367-96). New Brunswick, NJ: Transaction.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd Edition)*. Sage Publications.

Roberts, C., Gilbert, E., Allum, N., & Eisner, L. (2019). Research synthesis: Satisficing in surveys: A systematic review of the literature. *Public Opinion Quarterly*, *83*(3), 598-626.

Rudin, W. (1953). *Principles of mathematical analysis*.

Saltzman, L., Paternoster, R., Waldo, G. P., & Chiricos, T. G. (1982). Deterrent and experiential effects: The problem of causal order in perceptual deterrence research. *Journal of Research in Crime and Delinquency*, *19*(2), 172-189.

Savage, L. J. (1954). *The foundations of statistics*.

Schubert, C. A., Mulvey, E. P., Steinberg, L., Cauffman, E., Losoya, S. H., Hecker, T., ... & Knight, G. P. (2004). Operational lessons from the pathways to desistance project. *Youth violence and juvenile justice*, *2*(3), 237-255.

Shilov, G. (1971). *Linear Algebra*. Dover Publications.

Shover, N. (1996). *Great pretenders: Pursuits and careers of persistent thieves*. Routledge.

Sidgwick, H. (1874). *The methods of ethics.*

Simon, H. (1957). *Models of man*. New York; John Wiley.

Smith, A. (1759). *The theory of moral sentiments*.

Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations*.

Spivak, M. (1965). *Calculus on manifolds: A modern approach to classical theorems of advanced calculus*. CRC press.

Stanovich, K. E. (2010). Dual-process theory and the great rationality debate. *Rationality and the Reflective Mind*, 3-28.

Stigler, G. J. (1950a). The development of utility theory. I. *Journal of political economy*, *58*(4), 307-327.

Stigler, G. J. (1950b). The development of utility theory. II. *Journal of political economy*, *58*(5), 373-396.

Stigler, G. J. (1972). The adoption of the marginal utility theory. *Quarterly Journal of Business and Economics*, *11*(4), 188.

Strogatz, H. S. (2015). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering (2nd edition)*. CRC Press.

Teevan Jr, J. J. (1976). Deterrent effects of punishment: Subjective measures continued. *Canadian Journal of Criminology and Corrections*, *18*(2), 152-160.

Thomas, K. J., & Marie McGloin, J. (2013). A dual-systems approach for understanding differential susceptibility to processes of peer influence. *Criminology*, *51*(2), 435-474.

Thomas, K. J., & Vogel, M. (2019). Testing a rational choice model of "desistance:" Decomposing changing expectations and changing utilities. *Criminology*, *57*(4), 687-714.

Thomas, K. J., Baumer, E. P., & Loughran, T. A. (2022a). Structural predictors of choice: Testing a multilevel rational choice theory of crime. *Criminology*, *60*(4), 606-636.

Thomas, K. J., Hamilton, B. C., & Loughran, T. A. (2018). Testing the transitivity of reported risk perceptions: Evidence of coherent arbitrariness. *Criminology*, *56*(1), 59-86.

Thomas, K. J., Loughran, T. A., & Hamilton, B. C. (2020). Perceived arrest risk, psychic rewards, and offense specialization: A partial test of rational choice theory. Forthcoming in *Criminology*.

Thomas, K. J., O'Neill, J., & Loughran, T. A. (2022b). Estimating latent preferences for crime: Implications for rational choice, identity, and desistance theories. *Justice Quarterly*, 1-31.

Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263-291.

Varian (2009). *Microeconomic analysis.* W. W. Norton & Company.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Waldo, G. P., & Chiricos, T. G. (1972). Perceived penal sanction and self-reported criminality: A neglected approach to deterrence research. *Social Problems*, *19*(4), 522-540.

Williams, K. R., & Hawkins, R. (1986). Perceptual research on general deterrence: A critical review. *Law and Society Review 20*(4), 545-572.

Wonnacott, R. J., & Wonnacott, T. H. (1970). *Econometrics*. John Wiley & Sons.

Wood, S. N. (2017). *Generalized additive models: An introduction with R*. CRC press.

Wright, B. R., Caspi, A., Moffitt, T. E., & Paternoster, R. (2004). Does the perceived risk of punishment deter criminally prone individuals? Rational choice, self-control, and crime. *Journal of Research in Crime and Delinquency*, *41*(2), 180-213.

Yandl, A. L., & Bowers, A. (2016). *Elementary Point-set Topology: A Transition to Advanced Mathematics*. Courier Dover Publications.

Zimring, F., & Hawkins, G. (1968). Deterrence and marginal groups. *Journal of Research in Crime and Delinquency*, *5*(2), 100-114.

Table 1. Rational Choice and Deterrence Studies using the Pathways to Desistance Data

Citation:	Study Goal:	Variables Used:	Model:	Results:
Anwar & Loughran (2011)	Investigate the process of risk perception updating.	Risk of Apprehension	Ordinary Least Squares	Supportive
Loughran et al. (2011)	Investigate ambiguity in perceived risk of apprehension	Risk of Apprehension	Linear Probability Model	Supportive (ambiguity seems to influence offending)
Loughran et al. (2012b)	Determine if perceptions of rewards, risks, and costs change over time, and whether those changes are due to offending.	Personal Rewards Risk of Apprehension Social Costs	Difference of Means Test (F-Statistic)	Mostly Supportive
Loughran et al. (2012c)	Examine the functional form of the relationship between perceived risk and offending.	Risk of Apprehension	Generalized Additive Model	Supportive (non-linear risk effect)
Loughran et al. (2013)	Examine the influence of prior offending experience, specialization, and embeddedness on self-reported illegal earnings.	Illegal Wages	Ordinary Least Squares	Supportive
Thomas et al. (2013)	Assess if individual differences influence sanction risk updating.	Risk of Apprehension	Random Coefficients Model	Mixed
Loughran et al. (2016a)	Test the generality of a rational choice theory of crime.	Social Rewards Personal Rewards Risk of Apprehension Social Costs	Fixed Effects Poisson	Mixed (Social costs not sig.)
Loughran et al. (2016b)	Estimate effect of gun carrying on perceptions of crime.	Social Rewards Personal Rewards Risk of Apprehension Social Costs	Fixed Effects	Supportive (gun carry influenced beliefs)
Wilson et al. (2016)	See if low impulse control and vicarious arrests influence sanction risk updating.	Risk of Apprehension	Random Coefficients Model	Supportive
Lee et al. (2018)	Investigate whether maturity of judgment can influence perceived rewards, risks, and costs.	Social Rewards Personal Rewards Risk of Apprehension Social Costs	Hybrid Hierarchical Fixed Effects Model	Supportive (maturity influences beliefs)
Thomas & Vogel (2019)	Examine the age-graded effects of rewards, risks, and costs on crime.	Social Rewards Personal Rewards Risk of Apprehension Social Costs	Nonlinear Decomposition Model	Mostly Supportive
Thomas, Loughran, & Hamilton (2020)	Test the crime-specific dimensions of Beckerian rational choice theory.	Personal Rewards Risk of Apprehension	Hierarchical Latent Trait Model	Mostly Supportive
Thomas et al. (2022a)	Examine the link between structural factors and preferences toward RC inputs.	Social Rewards Personal Rewards Risk of Apprehension Social Costs	Hierarchical Latent Trait Model	Supportive (structural factors influence pref.)

Thomas et al. (2022b)	Examine changes in preferences toward risk over time	Risk of Apprehension	Hierarchical Latent Trait Model	Supportive (pref. vary with age)
Hamilton (2023)	Examine the robustness of an estimated "certainty effect" under weaker measurement assumptions.	Risk of Apprehension	Algorithmic Quasi-Poisson Models	Mostly Supportive

Table 2: Summary Statistics at Wave 1

Variable Name	Mean/%	S.D.	Min	1st Q.	Med.	3rd Q.	Max	N
Outcome Variable (Wave 2)								
Offending Variety	1.397	2.32	0.00	0.00	0.00	2.00	17	1,258
Perceived Incentives								
Social Rewards	1.942	0.49	1.00	1.67	2.00	2.27	3.87	1,261
Personal Rewards	2.274	2.49	0.00	0.00	1.43	3.86	10	1,261
Arrest Risk	5.194	2.99	0.00	2.86	5.00	7.71	10	1,247
Social Costs	2.969	0.91	1.00	2.33	2.97	3.67	5	1,260
Controls								
Impulse Control	3.071	0.94	1.00	2.38	3.00	3.75	5.00	1,261
Maturity	3.055	0.46	1.10	2.77	3.03	3.40	4.00	1,258
Future Orient.	2.480	0.58	1.00	2.13	2.41	2.88	4.00	1,258
Employment	13.28%	0.34	0				1	1,265
Age (baseline)	16.04	1.14	14	15	16	17	19	1,354
Female	13.59%	0.34	0				1	1,354
White	20.24%	0.40	0				1	1,354
Black	41.43%	0.49	0				1	1,354
Hispanic	33.53%	0.47	0				1	1,354
Other Race	4.8%	0.21	0				1	1,354

Table 3: Linear Model (Tweedie: p = 1.212; N = 11,529)

Variable Name	β_k	S.E.	T-Value
Intercept (β_0)	2.232	0.245	9.110
Perceived Incentives			
Social Rewards	0.259***	0.031	8.281
Personal Rewards	0.094***	0.006	16.168
Arrest Risk	-0.058***	0.005	-11.451
Social Costs	-0.078***	0.016	-4.893
Controls			
Impulse Control	-0.263***	0.017	-15.767
Maturity	-0.059†	0.033	-1.778
Future Orient.	-0.131***	0.026	-4.965
Employment	-0.006	0.038	-0.160
Age	-0.053***	0.012	-4.369
Female	-0.516***	0.050	-10.337
Black	-0.287***	0.039	-7.406
Hispanic	-0.176***	0.036	-4.853
Other	-0.017	0.066	-0.263

† $p < .10$; *** $p < .001$

Table 4: Marginal Smooths Model (Tweedie: p = 1.209; N = 11,529)

Variable Name	β_k/EDF	S.E./$(k-1)$	T-Value/F
Intercept (β_0)	2.272	0.222	10.229
Perceived Incentives			
f_1(Social Rewards)	2.708***	4	15.818
f_2(Personal Rewards)	3.688***	4	81.013
f_3(Arrest Risk)	2.719***	4	27.517
f_4(Social Costs)	2.626***	4	7.852
Controls			
Impulse Control	-0.248***	0.017	-14.821
Maturity	-0.055†	0.033	-1.667
Future Orient.	-0.114***	0.027	-4.289
Employment	-0.012	0.038	-0.313
Age	-0.057***	0.012	-4.711
Female	-0.498***	0.050	-9.937
Black	-0.267***	0.039	-6.852
Hispanic	-0.178***	0.036	-4.920
Other	-0.044	0.066	-0.667

EDF = Effective Degrees of Freedom; k = Number of Basis Functions.

† $p < .10$; *** $p < .001$

Table 5: Tensor Product Model (Tweedie: p = 1.204; N = 11,529)

Variable Name	β_k/EDF	S.E./$(k-1)$	T-Value/F
Intercept (β_0)	2.272	0.222	10.229
Tensor Product $(T_e(\boldsymbol{I}_{it}))$	91.82***	624	1.506
Controls			
Impulse Control	-0.247***	0.017	-14.765
Maturity	-0.050	0.033	-1.512
Future Orient.	-0.113***	0.027	-4.254
Employment	-0.007	0.038	-0.171
Age	-0.055***	0.012	-4.568
Female	-0.490***	0.050	-9.796
Black	-0.261***	0.039	-6.709
Hispanic	-0.170***	0.036	-4.690
Other	-0.031	0.066	-0.465

EDF = Effective Degrees of Freedom; k = Number of Basis Functions.
*** p < .001

Table 6: Model Comparisons

Model	AIC	REML	Adj. r^2
Linear Model	31,199.07	15,625	0.127
Marginal Smooths Model	31,093.84	15,574	0.131
Tensor Product Model	31,051.12	15,568	0.148

Figure 1: Marginal Smooth Visualizations

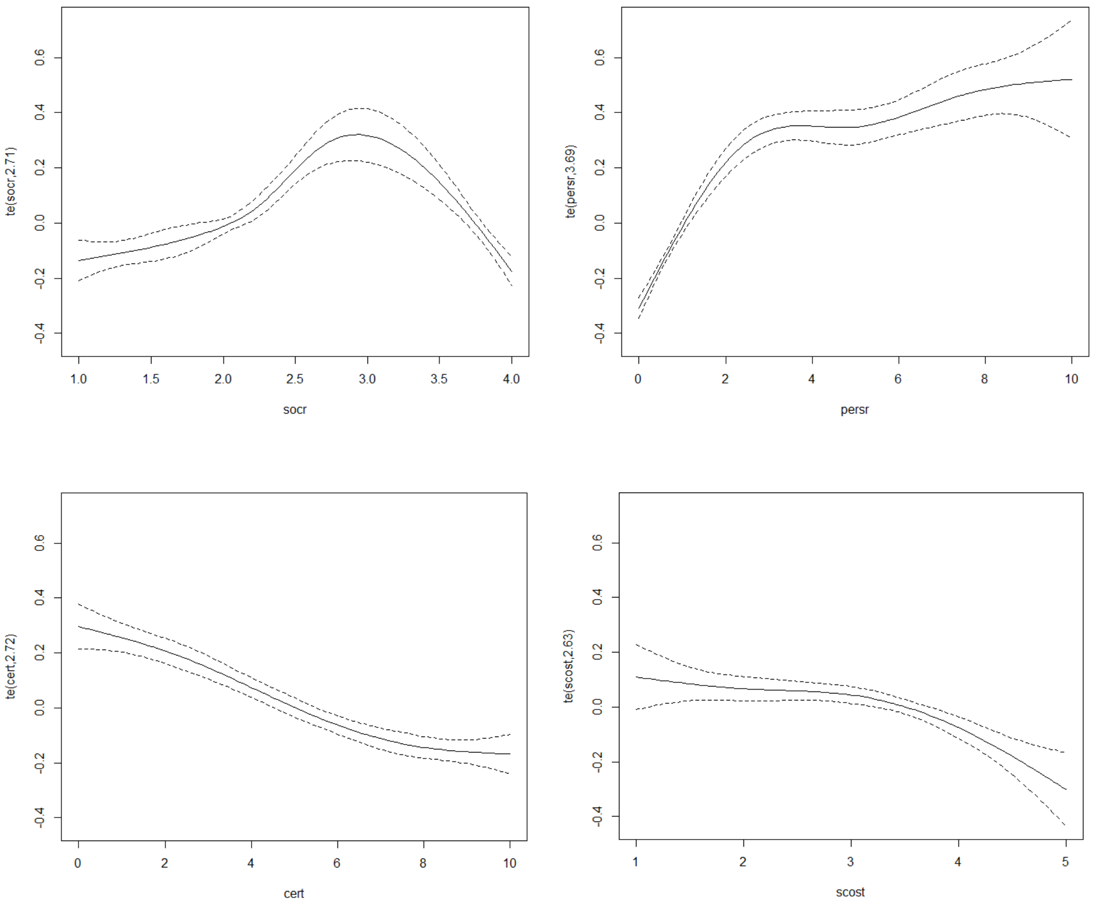


Figure 2.1: Surface Plot of the Interdependency of the Rewards on the Risks and Costs.

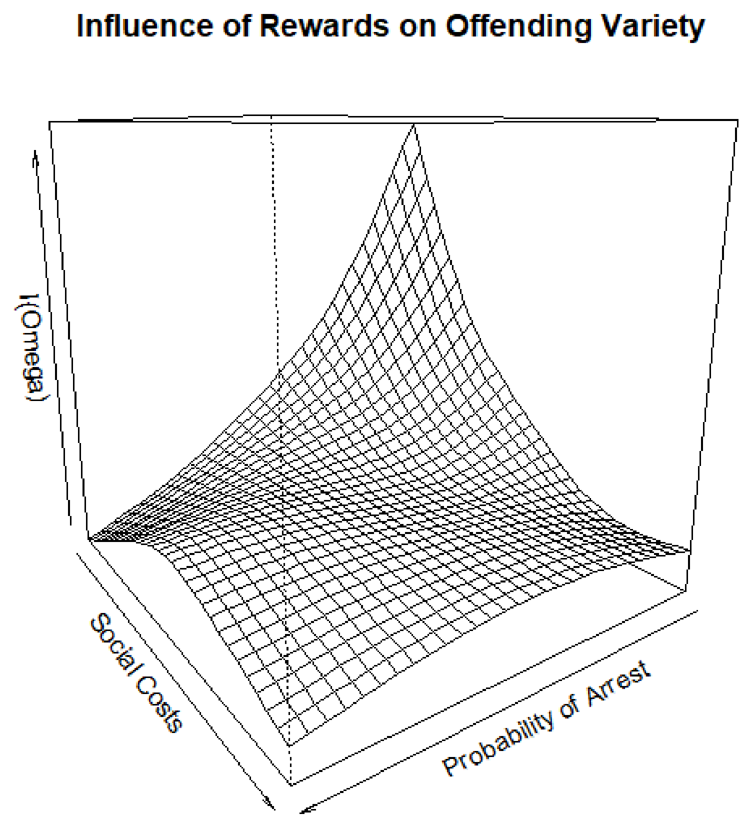


Figure 2.2: Topological Plot of the Interdependency of the Rewards on the Risks and Costs.

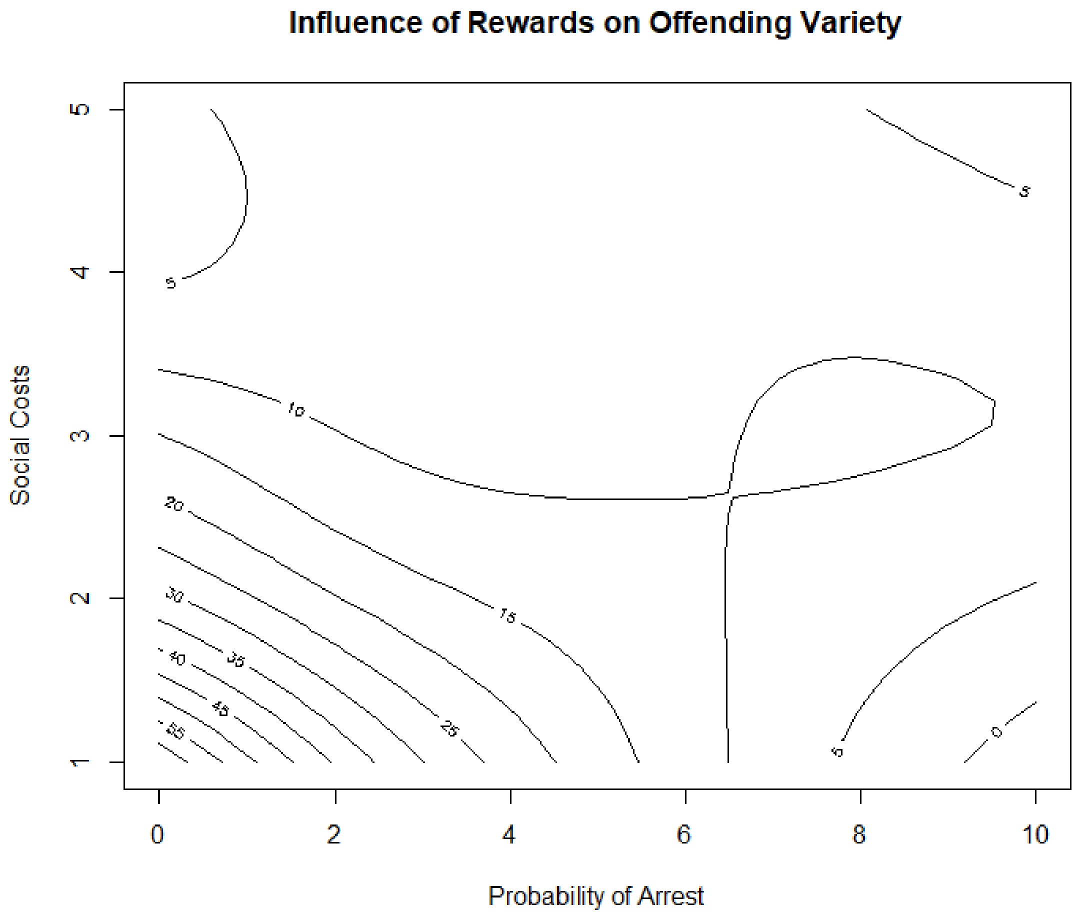


Figure 2.3: Surface Plot of the Interdependency of the Rewards on the Risks and Costs. (95% Confidence Interval)

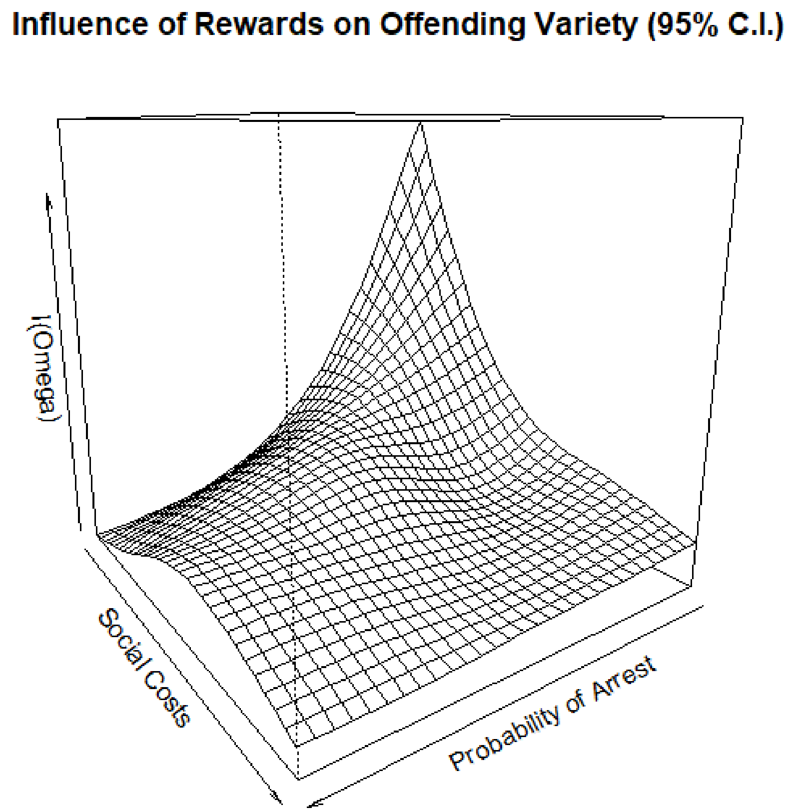


Figure 2.4: Topological Plot of the Interdependency of the Rewards on the Risks and Costs. (95% Confidence Interval)

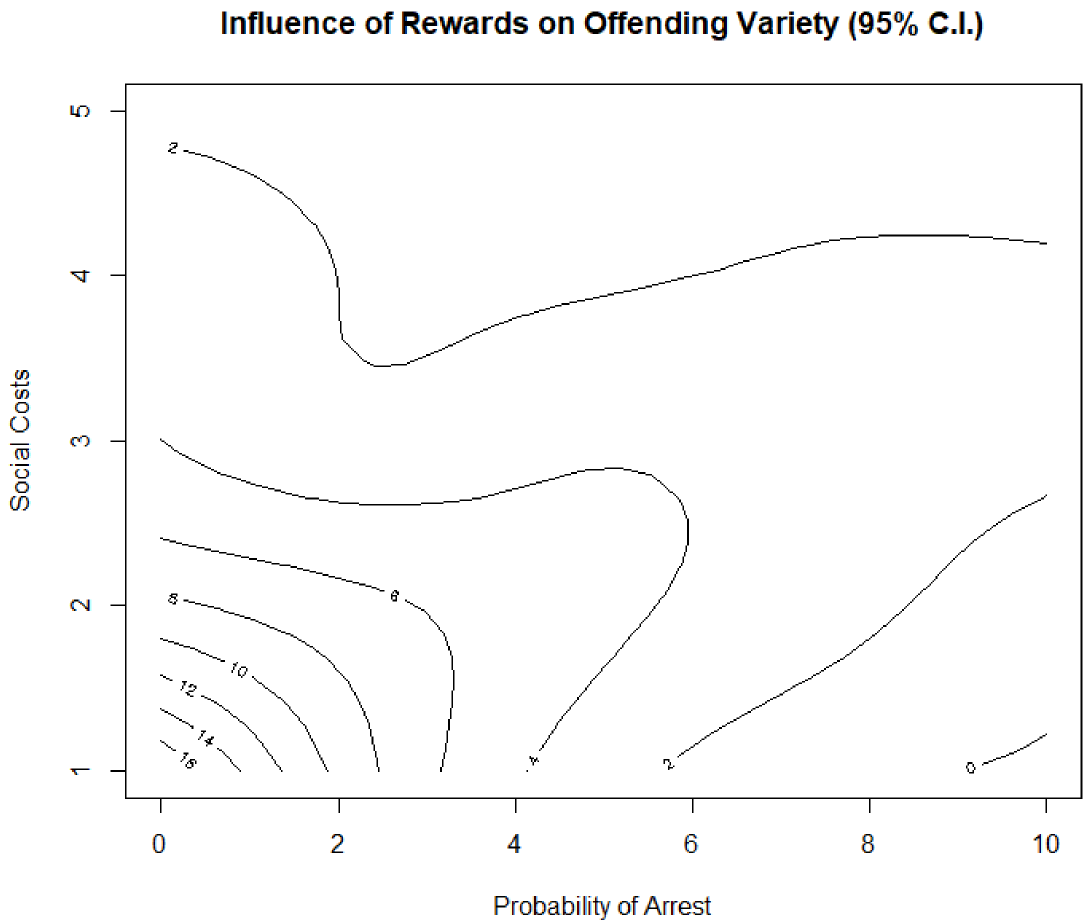


Figure 2.5: Surface Plot of the Interdependency of the Rewards on the Risks and Costs. (99% Confidence Interval)

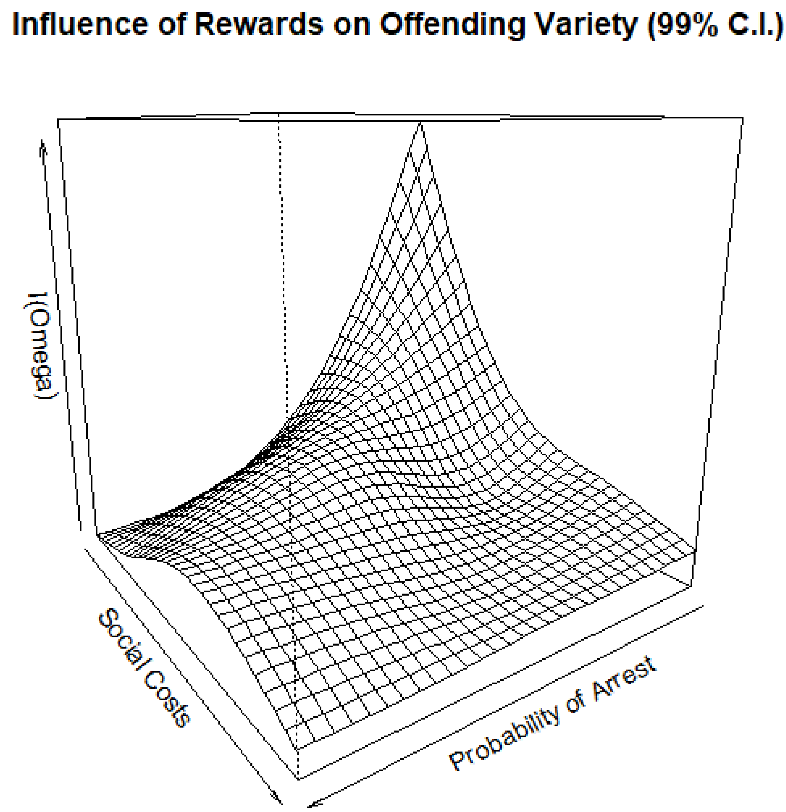


Figure 2.6: Topological Plot of the Interdependency of the Rewards on the Risks and Costs. (99% Confidence Interval)

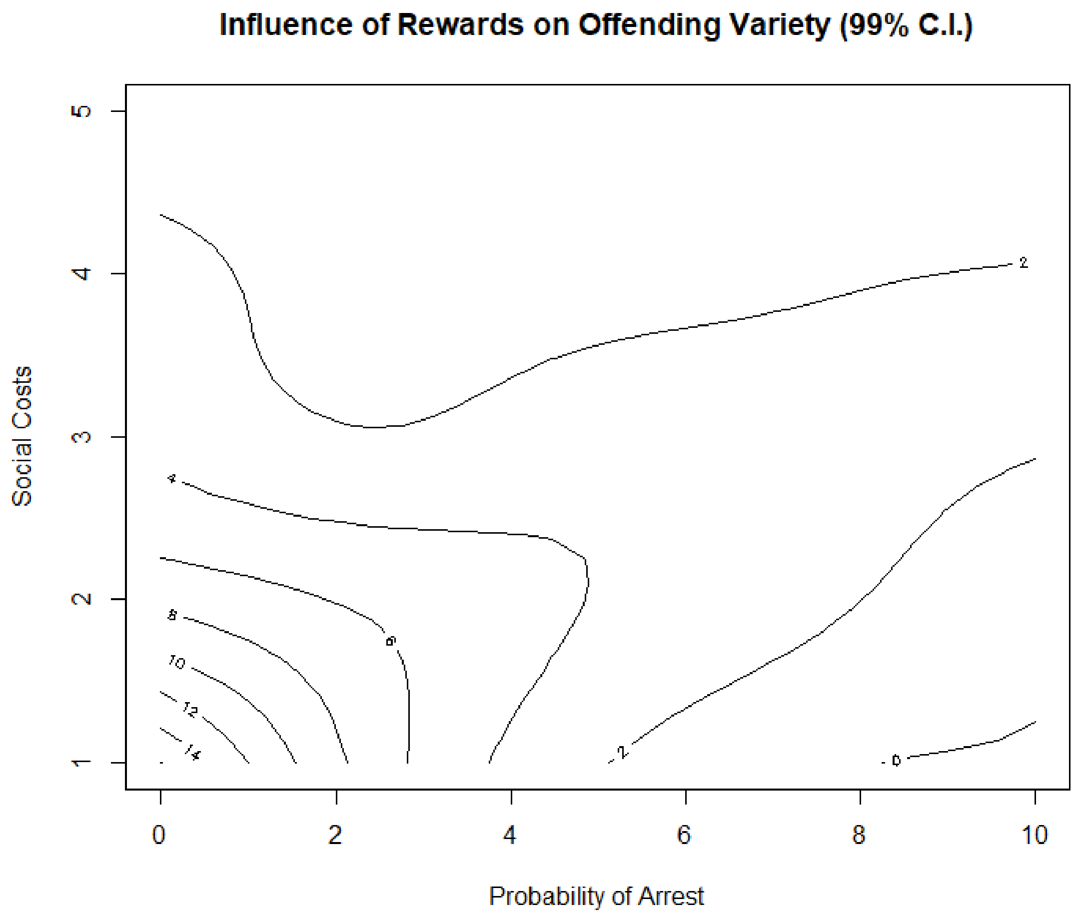


Figure 3.1: Surface Plot of the Interdependency of the Risks and Costs on the Rewards.

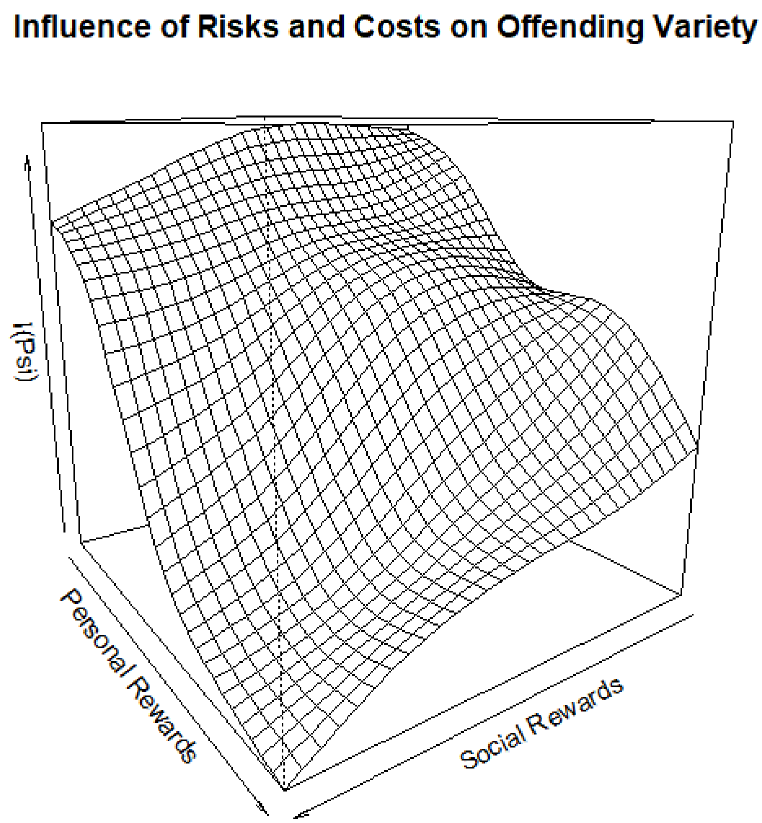


Figure 3.2: Topological Plot of the Interdependency of the Risks and Costs on the Rewards.

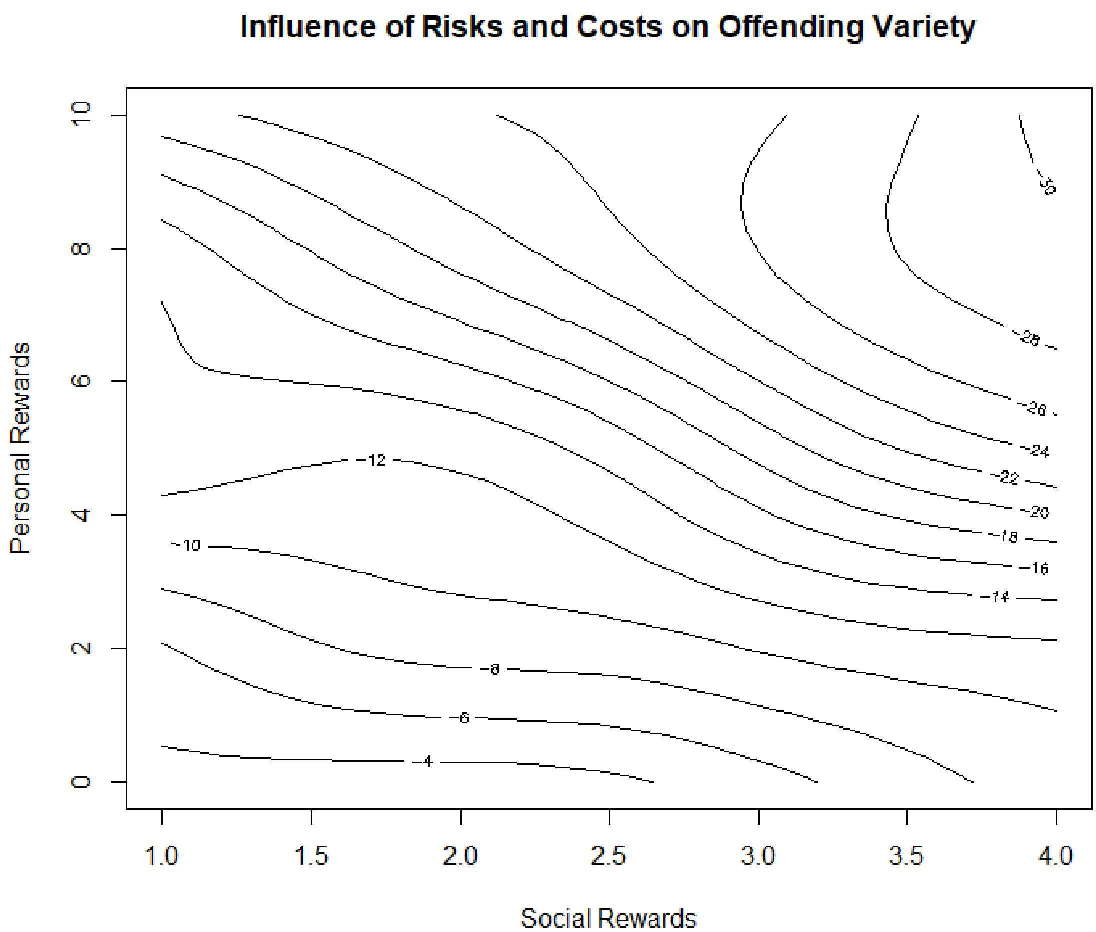


Figure 3.3: Surface Plot of the Interdependency of the Risks and Costs on the Rewards. (95% Confidence Interval)

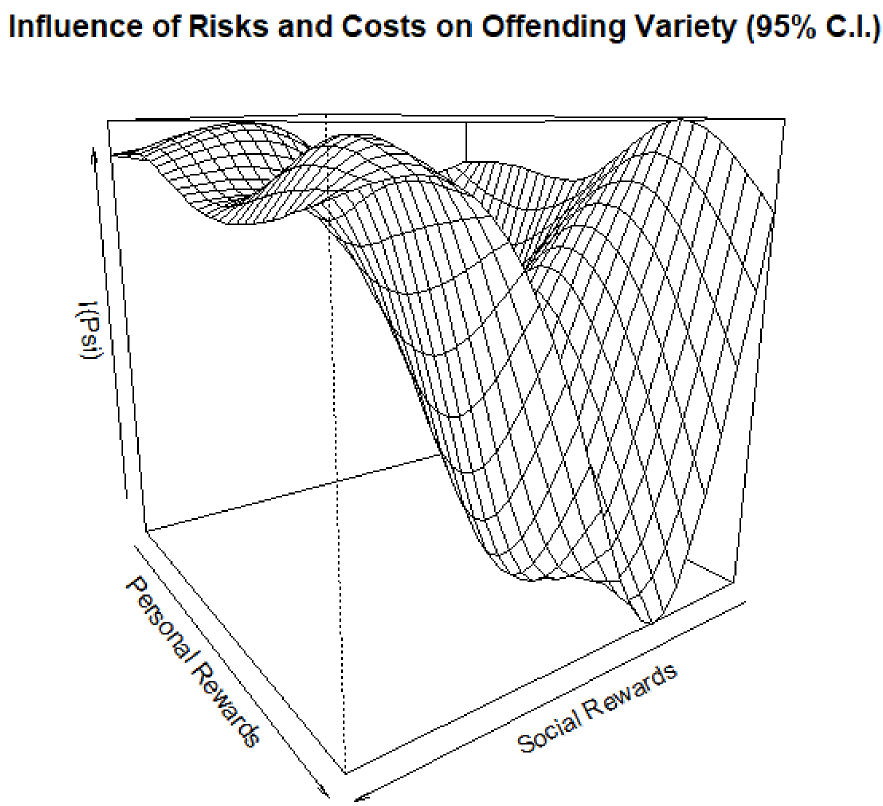


Figure 3.4: Topological Plot of the Interdependency of the Risks and Costs on the Rewards. (95% Confidence Interval)

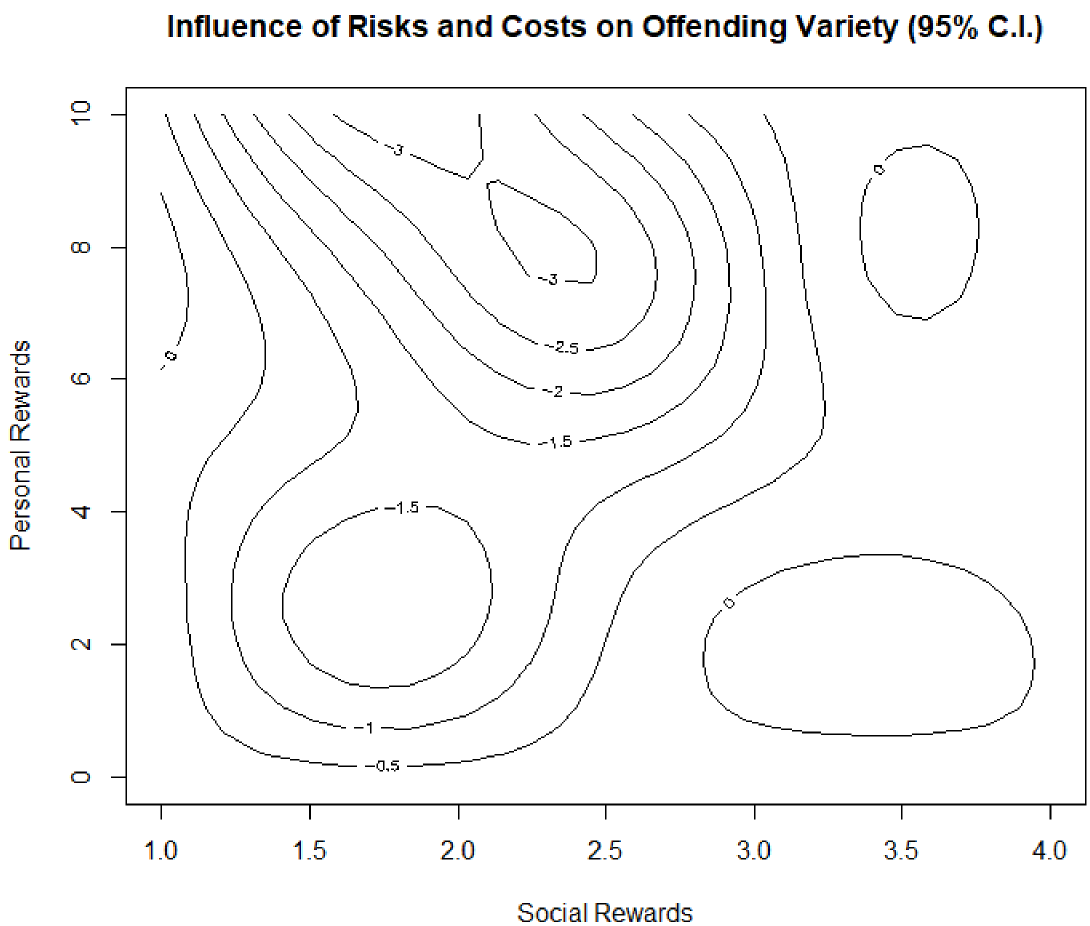


Figure 3.5: Surface Plot of the Interdependency of the Risks and Costs on the Rewards. (99% Confidence Interval)

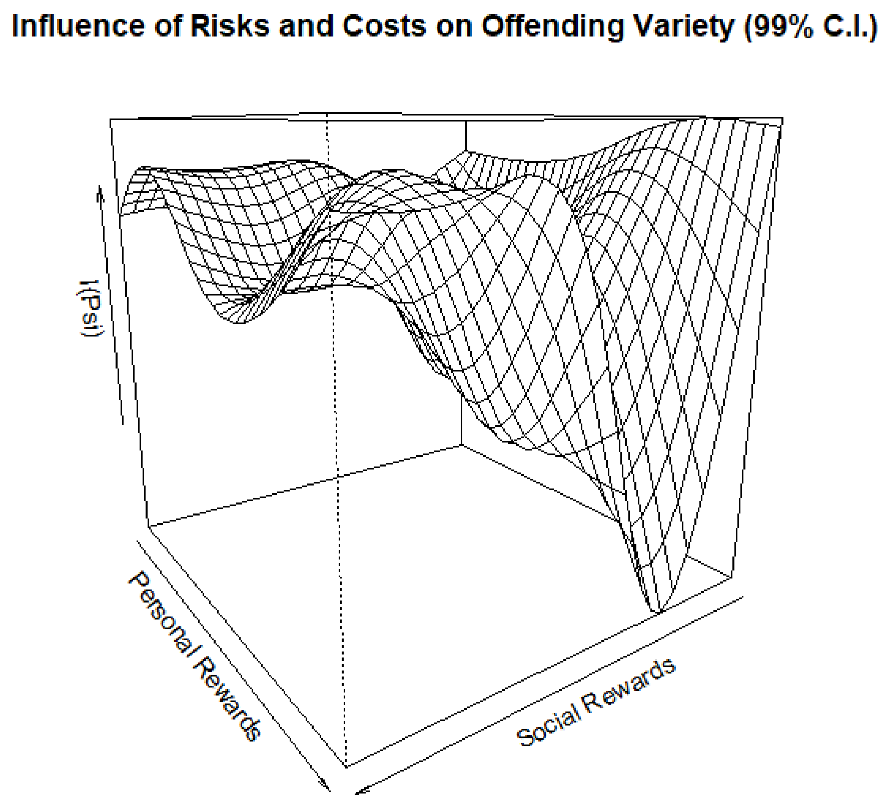


Figure 3.6: Topological Plot of the Interdependency of the Risks and Costs on the Rewards. (99% Confidence Interval)

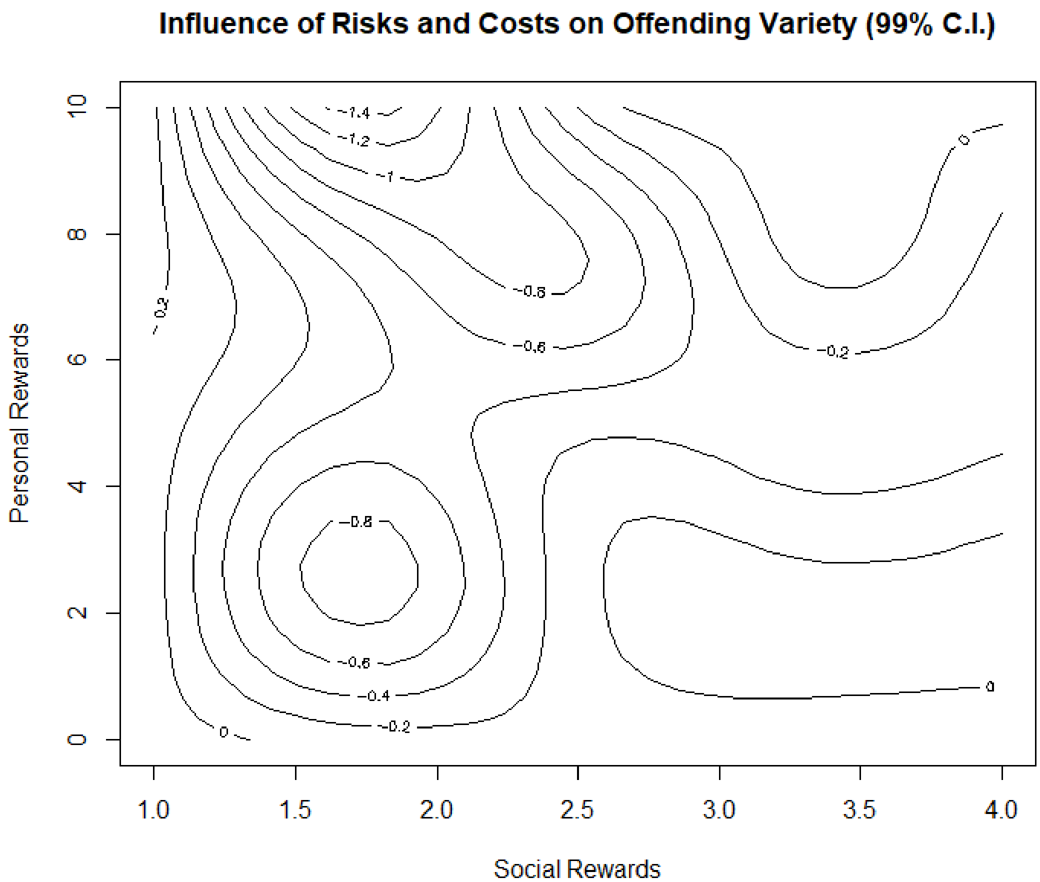


Figure 4.1: Smooth Plots of the Interdependency of the Social Rewards on the Personal Rewards.

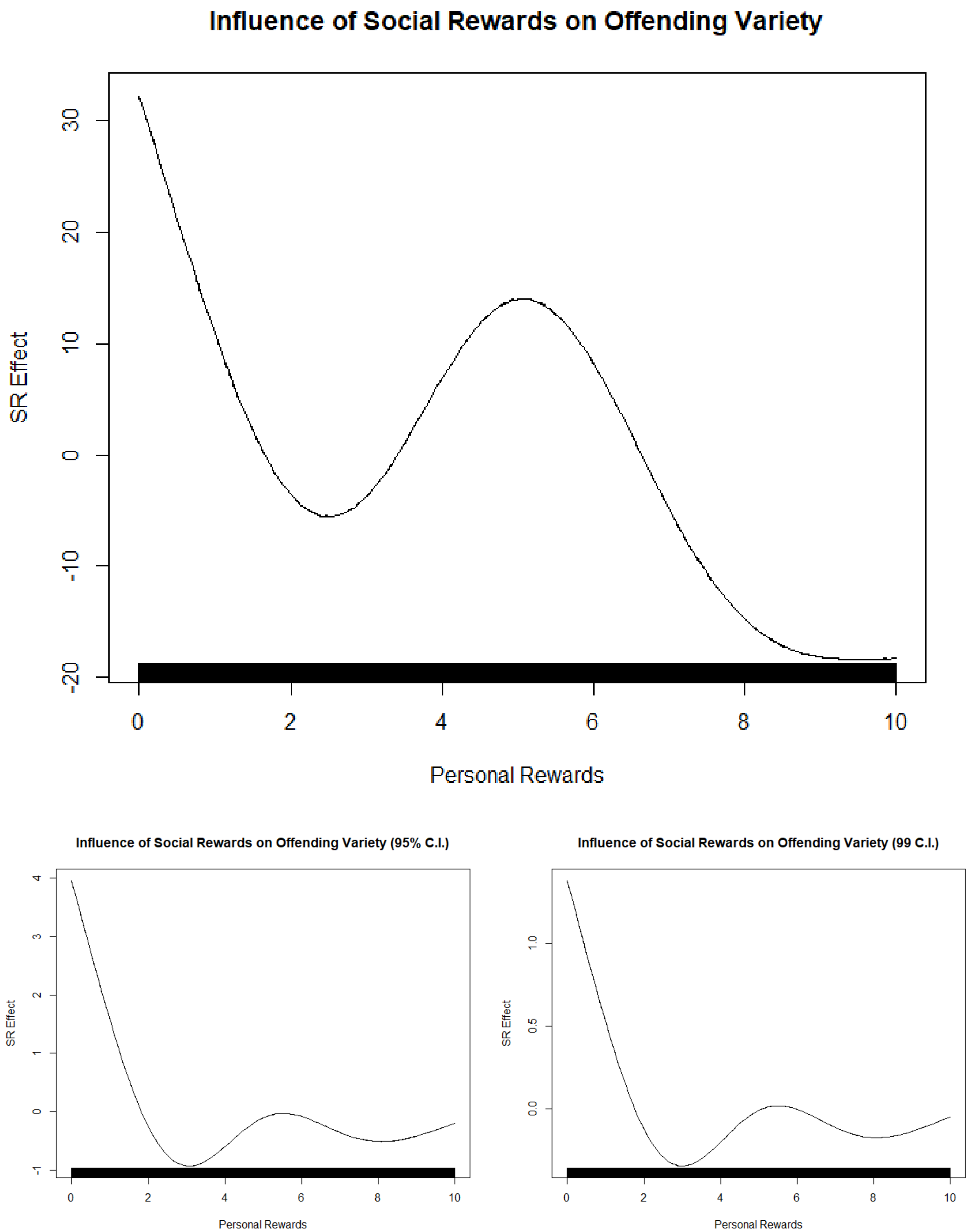
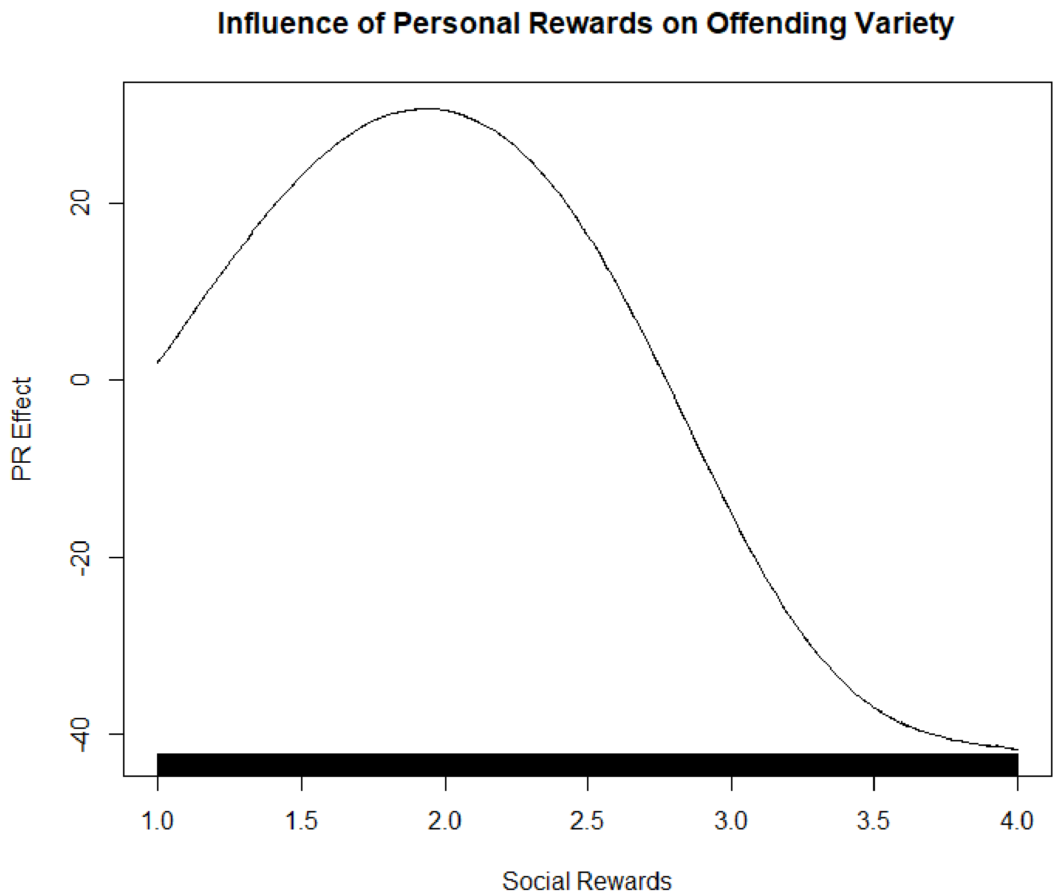
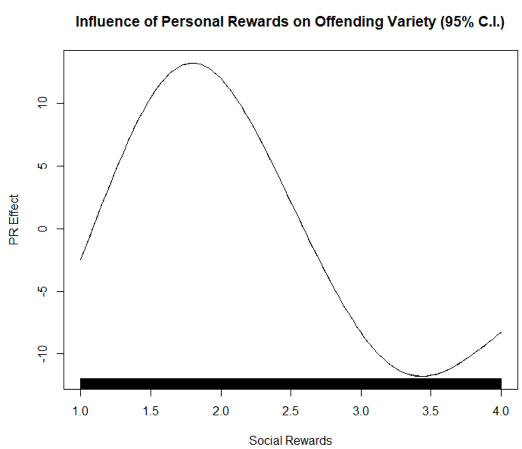
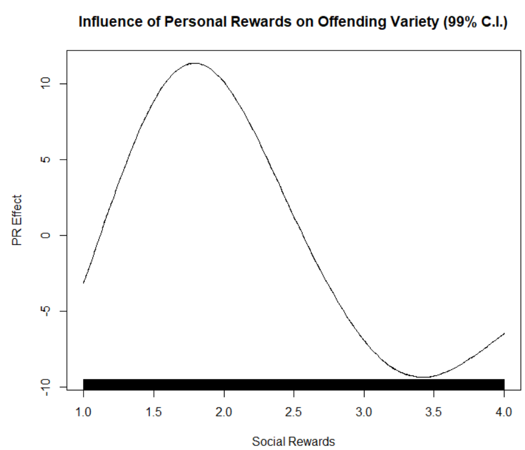


Figure 4.2: Smooth Plots of the Interdependency of the Personal Rewards on the Social Rewards.







APPENDIX A: MATHEMATICAL DISCUSSION

The following discussion is intended to provide an overview of the more technical aspects of this dissertation. Namely, I seek to delve much deeper into the mathematical framework which from this project's primary hypotheses were derived, as well as (hopefully) give the reader a clearer sense of why these particular hypotheses were chosen. In doing so, my primary aim is to draw heavy inspiration from microeconomic conceptions of choice and utility maximization (specifically those which are covered in Kreps [2013, 2023]), much of which can also be seen, in some capacity, in Gary Becker's (1968) seminal piece. Furthermore, I do not intend for this discussion to be an "ultra-rigorous" breakdown of the core concepts of a microeconomic approach to consumer choice and decision-making, nor will I seek to formally prove every conjecture throughout. Rather, I simply wish to provide a "good enough" coverage of those concepts without sacrificing mathematical rigor altogether. As such, while I will rely heavily on theorems provided in several subfields of mathematics—namely, linear algebra, convex optimization, and elementary real analysis—I leave the proofs of those theorems to scholars whose mathematical maturity far surpasses my own (as can be found in each of the citations I will provide throughout this discussion). Without further delay, I will now begin laying the foundation upon which this dissertation's core concepts are built, starting with the agent's utility function as it pertains to illegal and legal actions.

THE UTILITY FUNCTION:

To begin, I will assume the agent has, at his or her disposal, a finite set of behaviors which he or she may devote some amount of attention (e.g., time and energy) to. Denote this set by the symbol $\mathcal{A}$, such that:

$$\mathcal{A} = \{A : A \in \mathbb{R}_+^k\}$$

where A is a vector contained within the non-negative orthant of k-dimensional Euclidean space.[45] That is:

$$A_j \geq 0, \qquad j = 1,2, \dots, k$$

where A_j is the jth component of the vector A, the value of which can be any non-negative real number. Each component A_j can be thought of as a person's level of involvement in some activity (e.g., committing a burglary, going to school or work, seeing a movie, etc.), such that higher values of A_j represent greater levels of involvement within that activity, and a value of zero for A_j means "no involvement." For each activity set, define the following function:

$$EU : \mathbb{R}_+^k \longrightarrow \mathbb{R}$$

where EU denotes the expected utility of a given activity set (i.e., some $A \in \mathbb{R}_+^k$), and $\mathbb{R}$ denotes the set of real numbers (put differently, $EU(A)$ is a *real-valued* function). For the sake of brevity, I will assume EU is continuous and differentiable everywhere over its domain (that is, $\mathbb{R}_+^k$) up to at least a second level (i.e., second-order derivatives are assumed to exist; see: Loomis and Sternberg, 1968), thus allowing for standard calculus operations to be performed. Additionally, I assume EU is concave over $\mathbb{R}_+^k$, such that, for all $A, B \in \mathbb{R}_+^k$ and $\theta \in [0,1]$:

$$EU(\theta * A \ + [1 - \theta] * B) \geq \theta * EU(A) + (1 - \theta) * EU(B).$$

Put simply, if the expected utility to option A is greater than that of option B, then every option which falls directly "between" A and B will provide at least as much utility

[45] The previous equation is written in set notation, which can be read as follows: $\mathcal{A}$ is the set of all A, such that A is any element (in this case, A is a vector) which belongs to $\mathbb{R}_+^k$.

as B (i.e., any "adjustment" the agent makes to his or her activity set in the direction of A, starting from B, can only potentially improve his or her odds of achieving a better outcome; see: Kreps, 2013). The motivation behind this assumption is twofold. First, the concavity of EU guarantees the existence of a solution set for any (appropriately) constrained optimization problem (see: Boyd and Vandenberghe, 2004). Second, it allows for the possibility that the agent will derive less satisfaction as his or her level of involvement in a particular activity increases (i.e., activities may provide *diminishing marginal utility* at higher values of A_j; Becker, 1968). Thus, it is assumed that, on some level, the more a person's desires become sated by engaging in a specific activity (or collection of activities), the *less pleasure* that person will likely derive from devoting additional time and energy to it.

Applying this notion to involvement in criminal (and legal) actions, the agent is thus incentivized to both 1) satiate a *variety* of needs and wants by engaging in behaviors (criminal or otherwise) which address them, and 2) only pursue behaviors which address those needs to the highest degree possible among all alternatives. In other words, the agent is assumed to prefer to spend his or her resources (time, energy, etc.) on precisely those behaviors which offer the greatest "bang for one's buck" (Kreps, 2013). A common microeconomic approach to modeling this idea involves the following: assume the agent cannot devote an infinite amount of attention to any particular activity (or set of activities) and must instead choose which behaviors he or she will spend some amount of time and energy on and those which he or she will avoid entirely (if any). The agent can devote varying quantities of his or her time and attention to each activity, but in doing so will leave fewer resources available to distribute to behavioral alternatives (i.e., every action has an *opportunity cost*). Put differently, the agent is presumed to have a limited quantity of resources of any kind to work with, and is thus tasked with budgeting his or her efforts in a way which, generally speaking, promotes the best outcomes for him or her (Kreps, 2013).

I will denote the agent's "resource pool" by E, where E is any positive real number (i.e., $E > 0$). Denote the "energy cost" for each activity A_j by the vector $e \in \mathbb{R}^k_{++}$,[46] such that:

$$e_j > 0, \qquad e^T A \leq E$$

where e_j is the jth component of e, and $e^T A$ denotes the inner product of e and A (that is: $e^T A = \sum_{j=1}^{k} e_j A_j$).[47] Here, we can think of e_j as the amount of E required in order to consume a single unit of A_j (e.g., if the subscript j refers to burglary, then e_j is the total quantity of time and energy the agent must expend in order to pull off a single act of burglary). Note that e_j is assumed to be strictly positive, meaning there are no "free" actions; any activity the agent can involve his or her self in comes at some opportunity cost, with some actions incurring a greater (or lesser) cost than others. Since the agent's chosen activity set A must not incur a total resource expenditure greater than E, the set of all "feasible" activity sets the agent may select from includes every value of A whose inner product with e is less than or equal to E (as shown in the previous equation).

Before going any further, I first want to clarify that the value of e_j (and, by extension, E) is not meant to be a concrete quantity which can be directly measured. Rather,

---

[46] The symbol $\mathbb{R}^k_{++}$ refers to the strictly positive orthant of $\mathbb{R}^k$, meaning every component of e must be greater than zero (as specified in the following equation).

[47] Note that the inner product is defined similarly to the so-called "dot product" as commonly seen in linear algebra. The equation for the inner product can be read as follows: the value of $e^T A$ is determined by multiplying each "row-wise" component of the vectors e and A and subsequently summing each result (it is assumed that e and A are written as "column" vectors, with e^T denoting the transpose of e; Shilov, 1971).

it is a purely abstract entity whose primary purpose is simply to *restrict* the agent's actions in a way which reflects real-world decision making on some level. This serves both the technical goal of establishing a precisely defined (though very much abstract) set of boundaries on the agent's choice set, as well as the theoretical goal of providing a broader framework through which further insight into rational decision-making (including criminal decisions) can hopefully be achieved. Additionally, by allowing e_j to vary between activity types it is not only implied that some activities may require less exertion than others, but also that said exertion is an explicitly *rational* component of the decision to offend. Thus, the "ease" with which (some) criminal actions might be conducted suggests even the most prudent of utility maximizers may be willing to engage in some quantity of criminal acts despite the potential risks and costs (as an example, while stealing a $20 bill from a stranger's wallet might seem like a "petty" reward to some, to others it could represent well over an hour's worth of labor obtained in mere seconds).

Additionally, I do not assume the agent is under any obligation to utilize his or her entire resource pool, so to speak (e.g., the agent selects A only if $e^T A = E$), nor do I presume the agent has unlimited "access" to his or her preferred activities at all times. Rather, I simply assume the agent would prefer to be primarily involved in activities which provide an equivalent (or higher) anticipated "return on investment" to any given alternative, and will generally aim to behave accordingly whenever possible (depending on the opportunities available to the agent). As such, the reader is advised to think of the approach taken here as merely an attempt to provide a somewhat reasonable—though abstract—*approximation* of how an individual's involvement in criminal behaviors might be influenced by his or her perceptions of the rewards, risks, and costs to crime (particularly, the extent to which the influence of those incentives might be interdependent on one another). With this all having been said, I will now begin working through some of the more technical aspects of the agent's task of maximizing utility.

UTILITY MAXIMIZATION AND CONSUMER CHOICE:

Define the utility maximization problem by:

$$Maximize\ EU(A), \qquad subject\ to\ A_j \geq 0\ and\ e^T A \leq E.$$

Since E is finite and every $e_j A_j \geq 0$, it follows that the set of all feasible activity sets, which I will denote from this point onward by $\mathcal{F}$, is bounded from above.[48] Because $\mathcal{F}$ is a subset of $\mathbb{R}^k_+$, it is clear that $\mathcal{F}$ is also bounded from below (i.e., there are no negative components of A). It follows that $\mathcal{F}$ is convex and compact in $\mathbb{R}^k_+$,[49] which implies the existence of a unique set which contains all "maximal" elements of $\mathcal{F}$ (see: Kreps, 2013). That is:

$$\mathcal{S} = \{A : EU(A) \geq EU(B)\}$$

where $\mathcal{S}$ denotes the *solution set* of $\mathcal{F}$. That is, $\mathcal{S}$ is defined as the set which contains all choice options (i.e., activity sets) which yield the highest expected utility value relative to any possible alternative.[50] Thus, the agent is tasked with identifying this set of options, and subsequently pursuing (at least one of) those options to the best of his or her ability.

---

[48] In other words, for each A_j there exists a large enough real number N such that $A_j < N$ for all $e^T A \leq E$. Note that such a number can be found simply by selecting any $N > e^T A = E$, such that all components of A, with the exception of A_j, are equal to zero.

[49] The convexity of $\mathcal{F}$ is an immediate consequence of the linearity of $e^T A$. That is, for any $A, B \in \mathcal{F}$, the set of all points of $\mathbb{R}^k_+$ which fall between A and B (i.e., any point which can be represented by $\theta * A + [1 - \theta] * B$ for some $\theta \in [0,1]$) is also contained in $\mathcal{F}$ (see: Yandl and Bowers, 2016). The compactness of $\mathcal{F}$ follows from both the topological structure of Euclidean space (i.e., there are no "holes" in any bounded region of $\mathbb{R}^k$), and the fact that $\mathcal{F}$ contains its own boundary (i.e., there are no points in $\mathbb{R}^k_+$ which "touch" $\mathcal{F}$ that are themselves not contained in $\mathcal{F}$). For a more in-depth discussion on the topological properties of Euclidean space(s), see Munkres (2021).

[50] It should be noted that since EU is continuous and concave over $\mathbb{R}^k_+$, the solution set $\mathcal{S}$ is itself convex. That is, if A and B are elements of $\mathcal{S}$ (which implies $EU(A) = EU(B)$), then $\mathcal{S}$ also contains all points which

Before we can get a sense of which options are most (or least) likely to feature in $\mathcal{S}$, we first need to consider the *component-specific* influence of each A_j on the expected utility value of A (i.e., how might some *marginal* increase in A_j influence the value of $EU(A)$?). Since EU is concave over $\mathcal{F}$, it follows that if the agent would benefit in some way from increasing his or her involvement in A_j (for some fixed $A \in \mathcal{F}$), then:

$$EU(A) > EU(B), \qquad given\ A_j > B_j$$

where B_j is the rate of consumption of activity j for activity set B (note that all other components of A and B are assumed equal). Thus, if both A and B are contained in $\mathcal{F}$, then B is, by definition, a suboptimal option and therefore not contained in $\mathcal{S}$ (A, on the other hand, may be contained in $\mathcal{S}$, but only if no better options are available). However, if A is not contained in $\mathcal{F}$ (e.g., the agent exceeds his "resource limit," such that $e^T A > E$), then although the agent may prefer A over B, he or she is nonetheless forced to "settle" for the inferior option (i.e., although $EU(A) > EU(B)$, only option B can potentially belong to $\mathcal{S}$).[51]

Of course, another implication of the concavity of EU is that the *rate* at which EU changes may decrease at higher values of A_j (e.g., as the desires addressed by A_j become

---

fall between A and B (see previous footnote). This follows from the notion that the agent will be "just as well off" in shifting from any given option toward some equivalent or better alternative.

[51] It should be noted that, under this conception, it is fully possible for the agent to possess needs and desires which are *never fully satisfied*. That is, it could be the case that the agent will continue to derive additional amounts of pleasure (i.e., marginal utility) from continually increasing his or satiation levels for some specific set of desires (e.g., the agent may always be happier having more money over less). Likewise, rather than assuming the agent avoids, say, criminal acts because he or she has managed to "completely satisfy" each of his or her respective needs and wants through legal alternatives, we can instead assume the agent simply derives a greater *relative satisfaction* through conformity (i.e., even if some number of criminal activities would be pleasurable, on some level, for the agent to commit, he or she is simply *better off* pursuing only legal actions).

sated). Economists typically refer to this rate of change as the *marginal utility* of a particular good or action (see: Kreps, 2013), which I will denote by:

$$MU(A_j) = \frac{\partial EU}{\partial A_j}$$

where MU is the marginal utility function of A_j, which, as defined above, is equal to the partial derivative of EU with respect to A_j. Given any two options A and B which differ only in the j-th component, the concavity of EU implies the following:

$$MU(A_j) \geq MU(B_j), \qquad A_j < B_j.$$

Thus, the marginal utility of any activity A_j can only (potentially) decrease as the value of A_j increases. An important question for economists—and rational choice theorists more broadly—is what happens if two or more *different* goods (or activities) are “swapped out” with one another to some degree? For instance, if the agent increases his or her involvement in A_j while simultaneously *decreasing* his or her involvement in A_i (given $e^T A$ remains constant and $i \neq j$), what conclusions can we draw about the respective marginal utilities of those actions? Assume the following occurs:

$$EU(A) > EU(B), \qquad A_j < B_j \; and \; A_i > B_i, \qquad e^T A = e^T B.$$

Given all other components of A and B are equal (i.e., $A_k = B_k$ for all $k \neq i \neq j$), the above suggests that, relative to B, the agent benefits more by trading some (proportional) quantity of activity j for activity i. As such, we can conclude that, for option B:

$$MU_e(B_j) < MU_e(B_i), \qquad where \; MU_e(B_j) = \frac{MU(B_j)}{e_j}.$$

That is, the *change in marginal utility* of activity j relative to its "price" e_j, denoted by $MU_e(B_j)$, is less than that of i (note that the value of $MU_e(B_j)$ is always defined since $e_j > 0$ for every j; see: Kreps, 2013). Put simply, the agent has more to gain from some proportional increase in activity B_i relative to activity B_j. Because of this, any (arbitrarily small) adjustment made from B_j toward B_i, holding $e^T B$ constant, not only benefits the agent in general, but it also ensures that option B is *not contained in* $\mathcal{S}$. This naturally follows from the convexity of $\mathcal{F}$, as any "movement" from one option to another which preserves the value of $e^T A$—assuming, of course, both options contain only non-negative components—guarantees that the agent is still within the bounds of $\mathcal{F}$. Likewise, we can conclude two things: 1) generally speaking, "exchanging" inferior actions (e.g., A_j) for better ones (A_i), in some amount, will produce better outcomes for the agent (i.e., doing so will increase the value of EU), and 2) in order for A to be contained in $\mathcal{S}$, each selected quantity of A_j must either be *just as good as any alternative action or otherwise equal to zero.* That is, $A \in \mathcal{S}$ only if:

$$MU_e\left(A_j^*\right) \geq MU_e(A_i^*), \qquad for\ all\ A_j^* > 0\ and\ i \neq j,$$

where A_j^* denotes the value of A_j for some fixed $A \in \mathcal{S}$.[52] Note that this can be demonstrated in a straightforward manner via the following: assume there exists some $A_j^* > 0$ and A_i^* such that $MU_e(A_j^*) < MU_e(A_i^*)$ for any fixed option A contained in $\mathcal{S}$. We know that in any instance where the change in marginal utility for two different actions is

[52] Some scholars refer to A_j^* as the *optimal* value of A_j. That is, out of all possible activity sets, only those which have a value of A_j equal to A_j^* are those which maximize utility. For this discussion, however, I do not assume A_j^* corresponds to any specific value (or possibly a set of values). Rather, my use of A_j^* is purely in reference to the consumption rate of activity j for some fixed $A \in \mathcal{S}$ (i.e., select any A in $\mathcal{S}$; whatever the value of A_j, for this specific option A, is also the value of A_j^*).

"desynched" in this manner we can always reduce A_j^* by some arbitrarily small amount, in exchange for some increase in A_i^* (again, proportional to e), and thus achieve a greater expected utility value (i.e., EU increases with $A_j^* - \frac{\delta}{e_j}$ and $A_j^* + \frac{\delta}{e_i}$ for some small enough $\delta > 0$). However, since we know any proportional exchange of the "consumption rate" of two actions always yields an activity set contained in $\mathcal{F}$, it follows that there is at least one superior option to A the agent could select instead. This contradicts the assumption that A belongs to $\mathcal{S}$.[53]

MARGINAL UTILITY AS A FUNCTION OF REWARDS AND COSTS:

Of course, all of this technical wizardry does not tell us much about the underlying *mechanisms* of what might encourage the agent to shift his or her level of involvement in any action or set of actions, much less to what *degree* he or she will do so. To this end, we will have to employ some more (hopefully not too egregious) assumptions. Namely, I assume the marginal utility value of any action is a direct function of the rewards, risks, and costs of that action, such that:

$$MU(A_j) \geq MU(A_i), \qquad if\ R^j \geq R^i\ and\ C^j \leq C^i$$

where R^j and C^j respectively denote the marginal reward and cost vectors for activity j (note that the usage of superscripts here does *not* refer to exponentiation). I will also assume the reward and cost vectors both belong to the positive orthant of their respective Euclidean spaces. More specifically:

$$R^j \in \mathbb{R}_+^k\ and\ C^j \in \mathbb{R}_+^l, \qquad for\ any\ j.$$

[53] A common approach in microeconomics to modeling this notion is through the application of Lagrangian mechanics, similar to that which is often employed in the physical sciences (e.g., physics). For those interested in a more in-depth discussion of this approach, see Kreps (2013).

Additionally, I assume that each component of R^j and C^j is "weighted" by its probability of occurring, such that:

$$R_k^j = p_k^j r_k^j, \qquad C_l^j = p_l^j c_l^j$$

where r_k^j denotes the value of the reward "type" r_k for activity j, and p_k^j denotes the probability of that reward type occurring given some increase in A_j (the cost types are defined similarly).[54] Thus, the previous equations imply that the marginal utility of activity A_j will be at least as good as A_i in the event the expected value of each reward type of the former is at least as high as the latter, and the expected value of each cost type is lower.[55] Hence, if the value of, say, r_k^j increases for some reward type k (given $p_k^j > 0$), we can conclude that the value of $MU(A_j)$ will generally increase (or rather, will at least *not decrease*) as well for all $A \in \mathcal{A}$.

It should be noted, however, that I do not assume the converse of this statement is true. Rather, it could be the case that $MU(A_j) \geq MU(A_i)$ in the event some component of R^j is *less than* that of R^i (the same is true if some component of C^j is greater than C^i). Put differently, while I assume it is always true that the marginal utility of A_j is at least as good as A_i if it is the case that each of the rewards to activity j are greater than activity i, and the

[54] It should be mentioned that this is a highly reductive view of the role that the "likelihood" of an outcome occurring will play throughout the agent's decision-making process. As with any problem involving random outcomes, one can model the utility maximization problem in a number of ways, including that of a "finite" number of outcomes for a single choice at one point in time, or as a *dynamic* choice problem where the agent seeks to maximize the value of some utility function over multiple points in time (or possibly continuous time). The latter is the subject of the study of dynamic programming and optimal control theory in mathematics, both of which could easily fill an entire dissertation's worth of material in applying them to any rational choice problem (let alone the implications of such a problem for crime). As such, I chose a much simpler approach here for the sake of brevity, convenience, and ultimately to foster a far more intuitive discussion of the subject. For any reader interested in a more in-depth breakdown, I recommend Strogatz (2015) and Kellett and Braun (2023).

[55] Note that this relation is assumed to remain true *regardless of the values chosen for A_j and A_i*. This means we can select any possible $A \in \mathcal{A}$ and it will always be the case that $MU(A_j) \geq MU(A_i)$, given the rewards to j are higher, and the costs lower, relative to i. This will be relevant further along in this discussion.

costs lesser, I do not assume that this is the *only* scenario in which $MU(A_j)$ may be greater than or equal to $MU(A_i)$. More concretely, I allow for the possibility that some actions may *make up* for a deficiency in some reward type(s) by offering a *greater reward of a different type* (or, alternatively, offers a lower overall cost to the agent, and thus a smaller set of rewards may still yield a modest utility value). As such, even if, say, the social benefits to activity j are less than that of activity i, in the event j also offers a high enough personal reward (e.g., thrills or excitement), then it is possible that $MU(A_j) > MU(A_i)$ despite offering relatively fewer social rewards.

OPTIMAL CHOICES AND STRICT DOMINANCE:

We are now in a position to begin thinking about the relative effectiveness of the rewards, risks, and costs to crime (as well as legal alternatives) on the agent's overall level of involvement in criminal activities. Assume the agent is "drawn" to the solution set $\mathcal{S}$, in that he or she tends to pursue activity sets which, generally speaking, are believed to achieve outcomes similar to those of any $A \in \mathcal{S}$ (i.e., he or she will "gravitate" most toward any option A which minimizes the value of $EU(B \in \mathcal{S}) - EU(A)$). Put differently, however "imperfect" the agent's choices might be, he or she is nonetheless assumed to be interested in maximizing personal utility and should thus be prone to *altering his or her behavior* whenever he or she believes doing so will produce a more ideal (i.e., *desirable*) set of outcomes. Our goal is to now determine the circumstances under which changes in R^j and C^j (for some activity j or possibly a set of activities, such as criminal behaviors) will shape the solution set $\mathcal{S}$, which, in turn, could impact the agent's behavior to some observable degree.

Since we know that $A_j^* > 0$ only if $MU_e(A_j^*) \geq MU_e(A_i^*)$ for any A contained in $\mathcal{S}$ (and all $i \neq j$), it follows that, for any $A_i^* > 0$:

$$MU_e(A_i^*) \geq MU_e\left(A_j^*\right).$$

Therefore, for all $A_j^*, A_i^* > 0$:

$$MU_e(A_i^*) = MU_e(A_j^*).$$

In this example, assume we (arbitrarily) increase at least one component of R^j and, by doing so, increase the marginal utility of A_j^* (i.e., the value of $MU(A_j^*)$ for some $A \in \mathcal{S}'$, where $\mathcal{S}'$ denotes the previous solution set; that is, the set of all optimal options A *before* any changes to R^j or C^j occur). Since this change would increase the numerator of $MU_e(A_j^*)$ (recall: $MU_e\left(A_j\right) = MU(A_j)/e_j$), it follows that:

$$MU_e(A_i^*) < MU_e(A_j^*)$$

and hence the agent can exchange some amount of A_i^* for A_j^* (starting at the point A) to achieve a greater expected utility value (the agent might also be able to simply increase his or her involvement in A_j^* if $e^T A < E$). Thus, by increasing the rewards—and hence also the *change in marginal utility*—to some activity j, there exists at least one alternative option $B \in \mathcal{F}$ such that $EU(B) > EU(A)$ (given $A_j^*, A_i^* > 0$, $i \neq j$). Consequently, not only is A not contained in the "updated" solution set $\mathcal{S}$ (i.e., A belongs to $\mathcal{S}'$ but not $\mathcal{S}$), but there also must exist some alternative option $B \in \mathcal{S}$ which assigns a *higher consumption rate* to activity j. More specifically, for any (previously optimal) A which contains at least two non-zero activities i and j, by increasing the value of $MU(A_j)$ we can always find some (now optimal) B such that:

$$B_j^* > A_j^*, \qquad given\ R_B^j \geq R_A^j\ and\ C_B^j \leq C_A^j,$$

where R_A^j and C_A^j respectively denote the reward and cost vectors assigned to A_J (i.e., the "prior" values of the rewards and costs), and R_B^j and C_B^j likewise denote the same vectors assigned to B_j (i.e., the "altered" rewards and costs). Of course, it should be noted that B is not guaranteed to exist in the event A contains only one non-zero element (i.e., $A_i^* = 0$ for all $i \neq j$). In such a scenario, B exists only if $e^T A < E$, as the agent can simply increase his or her consumption of A_j to achieve a greater expected utility value (note that this implies $MU_e\left(A_j^*\right) = 0$ for any $A \in \mathcal{S}'$; Kreps, 2013). However, if $e^T A = E$, then the solution set will remain unchanged since the agent cannot increase his or her consumption of A_j without exceeding the value of E (this is obviously true since all $A_i = 0$, and thus cannot be "exchanged" for any amount of A_j). This is, of course, an extraordinary scenario, as it requires the agent to only be involved in one specific type of activity at any given moment. Likewise, it is probably safe to assume that any (previously) optimal option A likely features, at minimum, more than one non-zero activity for any given agent.[56]

Another situation where an increase in $MU(A_j^*)$ might not influence the solution set is if $A_j^* = 0$ for all $A \in \mathcal{S}'$. Since A_j can equal zero for any $A \in \mathcal{S}'$ for as long as $MU_e\left(A_j^*\right) \leq MU_e(A_i^*)$ for all $i \neq j$, it follows:

$$MU_e\left(A_j^*\right) < MU_e(A_i^*), \qquad only\ if\ A_j^* = 0.$$

That is, the only situation in which the change in marginal utility to activity j can be less than that of at least one alternative i, given A belongs to $\mathcal{S}$, is if $A_j^* = 0$ (note that this also implies that $MU_e\left(A_j^*\right) < MU_e(A_i^*)$ for all $A_i^* > 0$, since the change in marginal

[56] Note that everything discussed so far also applies to any decrease in the costs to a given activity j, as this could also lead to a higher value of $MU(A_j)$, in which case there exists some $B \in \mathcal{S}$ such that $B_j > A_j$.

utility of all non-zero actions must be equivalent to one another). Here, we can always find some arbitrarily small increase (decrease) in R^j (C^j) such that:

$$MU_e\left(A_j^*\right) < MU_e\left(A_j^*\right) + \delta < MU_e(A_i^*), \qquad for\ some\ A_i^* > 0$$

where δ denotes a (small) positive real number equal to the marginal utility "gain" of some change to R^j (or C^j). In other words, it is possible that, for some $A_j^* = 0$, changes in R^j and C^j (and, by extension, $MU_e(A_j^*)$) may occur without altering the solution set (i.e., $\mathcal{S}' = \mathcal{S}$). If we think of j as referring to, say, the act of burglary, then the agent may very well continue to abstain from burglary even if he or she comes to perceive a greater reward (or lower cost) to engaging in burglary. Put differently, marginal changes in R^j and C^j should have relatively little—if any—effect on the agent's involvement in activity j for as long as j is *strictly dominated* by at least one other activity i for all $A \in \mathcal{S}$ (i.e., whenever $MU_e\left(A_j^*\right) < MU_e(A_i^*)$ for some $A_i^* > 0$, regardless of whichever previously optimal A we might initially select).[57]

An important question, however, remains: based on our discussion so far, can we identify a situation where some activity j is strictly dominated by an alternative activity i? To answer this question we first need to find an instance where the agent might be completely *indifferent* to two different types of activities. Recall our assumption that

[57] Although very much implied, it should nonetheless be mentioned that this particular conception of the agent's choice problem allows for the possibility that changes in R^j and C^j may influence the "optimal" value of A_j^* for only a subset of $A \in \mathcal{S}'$. That is, it could be the case that some number of previously optimal options *remain optimal* even after an adjustment to the rewards and costs to activity j are made, while other options do not. This can occur if, say, an increase to R^j leads to an increase in $MU(A_j^*)$ for some choices of $A \in S'$, while remaining the same for others (note that such a situation is allowable only if it preserves the concavity of EU). Regardless, it remains true in general that the agent can never benefit from *decreasing* his or her involvement in j for any given increase in R^j or decrease in C^j (i.e., $EU\left(B_j\right) \leq EU\left(B_j^*\right)$ for any $B_j < B_j^* = A_j^*$, given $R_B^j \geq R_A^j$ and $C_B^j \leq C_A^j$).

$MU(A_j) \geq MU(A_i)$ for all possible $A \in \mathcal{A}$ whenever it is the case that $R^j \geq R^i$ and $C^j \leq C^i$. By extension:

$$MU(A_j) \leq MU(A_i), \qquad if\ R^j \leq R^i\ and\ C^j \geq C^i,$$

and therefore:

$$MU(A_j) = MU(A_i), \qquad if\ R^j = R^i\ and\ C^j = C^i.$$

Thus, in the event the respective reward and cost vectors for activities i and j are equivalent to one another, then so too are the *marginal utilities* of i and j for any given option A. If it is also the case that both i and j invoke the same "energy" cost (i.e., $e_j = e_i$), then:

$$MU_e(A_j) = MU_e(A_i), \qquad if\ R^j = R^i;\ C^j = C^i;\ \ and\ e_j = e_i.$$

In this scenario, we can freely exchange any proportional quantity of A_j for A_i (and vice versa), at any "starting" point $A \in \mathcal{A}$, *without altering the expected utility of* A. That is:

$$EU(\theta * A^{j=0} + [1-\theta] * A^{i=0}) = K, \qquad for\ all\ \theta \in [0,1],$$

where K is a (constant) real number, $A^{j=0}$ is any $A \in \mathcal{A}$ with the j-th component equal to zero, and $A^{i=0}$ is the (unique) A equivalent to $A^{j=0}$ in all components except for i and j, the values of which are "swapped" (i.e., $A_i^{i=0} = A_j^{j=0} = 0$ and $A_j^{i=0} = A_i^{j=0} \geq 0$). Because of this, if either $A_j^* > 0$ or $A_i^* > 0$ (or both) for any $A \in \mathcal{S}$, then every alternative activity set B produced via some (allowable) equivalent exchange between j and i will *also be contained in* $\mathcal{S}$ (this, of course, follows from EU remaining constant for any such exchange, which implies every B achieved this way will also be a "maximal" element in $\mathcal{F}$). Economists refer to this type of situation as *perfect substitution*, in that the agent is completely indifferent to whether he or she engages solely in i, j, or any (proportional) combination of the two (Kreps, 2013). For this discussion, I assume such a situation occurs in the event the agent believes he or she could achieve the *exact same outcomes for the same amount of effort* from engaging in two or more different activity types (i.e., when $R^j = R^i$, $C^j = C^i$, and $e_j = e_i$ for any combination of $i \neq j$).

Now assume we increase some component(s) of R^j, which results in:

$$MU(A_j) > MU(A_i), \qquad for\ some\ i \neq j\ and\ all\ A \in \mathcal{A}.$$

Given $e_j = e_i$, this implies:

$$MU_e(A_j) > MU_e(A_i).$$

Consequently, activity i is now strictly dominated by j, as any possible selection of A with a positive i-th component (i.e., $A_i > 0$) will always yield a scenario where the change in marginal utility of i is *less than that of* j. Since $A \in \mathcal{S}$ only if $MU_e(A_i) \geq MU_e(A_j)$ for all $A_i > 0$ (and all $j \neq i$), it follows that $\mathcal{S}$ does not contain any A with a non-zero i-th component. Simply put, in this situation the agent will always be better off avoiding i altogether in favor of pursuing (some amount of) j instead, and thus it is in the agent's best interest to always opt for engaging in j over i whenever possible.

SUBSTITUTABILITY OF LEGAL AND ILLEGAL ACTIONS:

An intuitive, though by no means obvious, implication of the "strict dominance" notion is that the overall *substitutability* of two or more activities might have something to do with the degree of overlap between the respective rewards and costs of those actions. Namely, the more "similar" the types of needs addressed (or pains caused) by any two activities, the *less of an impact* any proportional exchange of the two might have on the expected utility of A (i.e., the agent may still achieve a "high" utility value in exchanging some quantity of i for j, regardless of whether said exchange is technically suboptimal). The key idea behind this is that the agent is often best served by satiating a *variety* of needs and desires through his or her actions (while also minimizing as many different types of pains as possible), and hence could potentially benefit more from some $A \in \mathcal{F}$ which provides a "modest" degree of satiation for several desires, relative to an alternative option (i.e., $B \in \mathcal{F}$) which satiates fewer desires to a higher degree at the expense of the rest. As such, the agent may well be incentivized, on some level, to *minimize* his or her involvement in activities which produce similar reward (and cost) types as one another, as doing so allows the agent to devote greater attention to activities which produce *different types of rewards and costs*.

The foundations for this idea have (mostly) been established in the discussion leading up to this point. More specifically, recall that the concavity of EU implies that the marginal utility of any activity j can only *decrease* as the agent's level of involvement in j increases. Put differently, for any $A \in \mathcal{A}$ which varies only in the j-th component and any (finite or countably infinite) sequence of A_j^k where:

$$A_j^1 < A_j^2 < \cdots < A_j^k,$$

it follows:

$$MU\left(A_j^1\right) \geq MU\left(A_j^2\right) \geq \cdots \geq MU\left(A_j^k\right).$$

Likewise, since any relation between R^j (as well as C^j) and $MU(A_j)$ must also preserve the concavity of EU, it follows that, for any activity i which provides identical rewards and costs to j (i.e., $R^j = R^i$ and $C^j = C^i$):

$$MU(A_i) = MU\left(A_j^k\right) \, for \, all \, k = 1,2,\ldots,k,$$

and therefore:

$$MU(A_i^1) \geq MU(A_i^2) \geq \cdots \geq MU(A_i^k),$$

where A_i^k denotes the value of A_i at the k-th iteration of the sequence A_j^k. Since all components of any select $A \in \mathcal{A}$ other than j are held constant throughout this sequence (i.e., $A_i^1 = A_i^2 = \cdots = A_i^k$), the above equation states that the marginal utility of activity i is itself a *decreasing function of* A_j (i.e., $MU(A_i)$ diminishes with A_j as well as A_i). In simple terms, all this is really saying is that since activity j satiates the same desires (and causes the same pains) as those addressed by i, it is possible that activity i may become *less appealing* for the agent in the event his or her involvement in j increases. In such a scenario, we might even expect the agent to reduce his or her involvement in i, since any opportunity losses associated with i are likely to be compensated, to some degree, by j.[58]

The influence of activity j on $MU(A_i)$ (and vice versa) can be summarized by the second-order partial derivative of EU with respect to A_j and A_i (note that this is also sometimes referred to as the *cross-partial* derivative between i and j; see: Kreps, 2013). Since it is assumed that the marginal utilities of any i and j which produce identical outcomes are equivalent to one another for any $A \in \mathcal{A}$, it follows that:

$$\frac{\partial^2 EU}{\partial {A_j}^2} = \frac{\partial^2 EU}{\partial A_j \partial A_i} = \frac{\partial^2 EU}{\partial {A_i}^2}, \qquad if\ R^j = R^i\ and\ C^j = C^i,$$

[58] Although presented in a fairly abstract way, one can often see concrete examples of this phenomenon in day-to-day life. For instance, wealthy individuals are typically less willing to engage in various "penny pinching" behaviors more commonly seen among the less wealthy (e.g., clipping coupons, purchasing second-hand goods, etc.). Individuals who work desk jobs tend to benefit more from increasing their level of physical exercise compared to those involved in intensive manual labor. Persons experiencing a great deal of work (or life) related stress usually benefit most from engaging in activities which promote relaxation (e.g., meditation), while avoiding those which may cause their stress levels to elevate (e.g., working overtime). For further elaboration, see chapters 11 and 19 in Kreps' (2013, 2023) *Microeconomic Foundations* series.

where $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ denotes the second (i.e., cross) partial derivative with respect to i and j, and $\frac{\partial^2 EU}{\partial {A_j}^2}$ is the second partial derivative with respect to j, and likewise for i (note that the above statement is true in a trivial sense if $i = j$). Combined with the concavity of EU, it follows:

$$\frac{\partial^2 EU}{\partial {A_j}^2} \leq 0, \qquad for\ any\ j.$$

That is, it must always be the case that any second partial derivative (including cross partials) must be less than or equal to zero for any choice of A (Boyd and Vandenberghe, 2004). As such, for every pair of activities i and j, it can only be the case that an increase in j (or i) leads to a potential *decrease* in the attractiveness of i (j) to the rational actor.

It should be noted that this conception does not tell us much about how any two (or more) second partial derivatives might compare with one another in the event R^j (C^j) is not identical to R^i (C^i).[59] As such, I will also assume the following:

$$\frac{\partial^2 EU}{\partial {A_j}^2} \leq \frac{\partial^2 EU}{\partial A_j \partial A_i} \leq \frac{\partial^2 EU}{\partial {A_i}^2}, \qquad if\ R^j \geq R^i\ and\ C^j \geq C^i.$$

In other words, I assume the *rate* at which the marginal utility diminishes for any activity j is itself a function of the rewards and costs to j (for the sake of brevity, I will also assume this function is both continuous and monotonic with respect to R^j and C^j). Specifically, any second partial derivative with respect to j can only *decrease* as the rewards and costs of j increase. Furthermore, I assume the cross-partial derivative with

[59] For instance, the second partial derivative with respect to any activity j could either increase or decrease with R^j (as well as C^j), given the value of said derivative remains less than zero.

respect to any two activities i and j must fall somewhere between their respective second partial derivatives whenever it is the case that every type of reward and cost of j is respectively greater than or equal to those of i, or vice versa.[60] Likewise, for as long as $R^j \geq R^i$ and $C^j \geq C^i$ (or vice versa), the cross-partial derivative of i and j should also be a decreasing function of the rewards and costs to i and j. More specifically, consider any two activities i and j which the agent is completely indifferent between (i.e., $R^j = R^i$ and $C^j = C^i$). If we decrease any component of either R^i or C^i, it follows that:

$$\frac{\partial^2 EU}{\partial {A_j}^2} \leq \frac{\partial^2 EU}{\partial A_j \partial A_i} \leq 0.$$

Suppose $\frac{\partial^2 EU}{\partial {A_j}^2} < 0$. Given some decrease in R^i or C^i which increases the value of $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ (assuming such an increase exists for some change in R^i or C^i), we now have:

$$\frac{\partial^2 EU}{\partial {A_j}^2} < \frac{\partial^2 EU}{\partial A_j \partial A_i} \leq 0.$$

We can easily generalize this to any decreasing sequence of R^i or C^i. For example:

$$R_1^i > R_2^i > \cdots > R_n^i \text{ and } C_1^i > C_2^i > \cdots > C_m^i,$$

---

[60] To give a more concrete example of this assumption, consider the function $\sqrt{ax + by}$. If we compute the second partial derivatives with respect to x and y, as well as their cross-partial derivative, we respectively get $-\frac{a^2}{4(ax+by)^{3/2}}$, $-\frac{b^2}{4(ax+by)^{3/2}}$, and $-\frac{ab}{4(ax+by)^{3/2}}$ (this can be independently verified by the reader, if desired). If we think of a and b as the respective values of some reward type for x and y, it can be easily shown that, for any $a \geq b \geq 0$: $-\frac{a^2}{4(ax+by)^{3/2}} \leq -\frac{ab}{4(ax+by)^{3/2}} \leq -\frac{b^2}{4(ax+by)^{3/2}}$. More specifically, since the denominator is the same for each second partial derivative, one only needs to compare the values of their respective numerators to determine their order for any non-negative a and b. Since $a^2 > ab > b^2$ whenever $a > b > 0$, it follows that $-(a^2) < -(ab) < -(b^2)$, and thus this particular function fulfills the requirements of our (most recent) assumption. For an analogous example of the respective costs to x and y, consider the function $-(ax + by)^2$ (the details are left to the reader).

where R_n^i denotes the n-th iteration of some decreasing sequence of reward vectors for activity i, and C_m^i denotes the same for some sequence of cost vectors.[61] Here, the value of $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ should be an increasing function of the terms in both sequences (note this is merely the inverse of the statement that $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ generally *decreases* as the rewards and costs to i increase). That is:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_1^i\right) \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_2^i\right) \leq \cdots \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_n^i\right) \leq 0,$$

and:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(C_1^i\right) \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(C_2^i\right) \leq \cdots \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(C_m^i\right) \leq 0.$$

Hence, the "distance" between $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ and zero can only decrease as either R^i or C^i approaches its respective zero vector for some $i \neq j$, assuming it is always the case that $R^j \geq R^i$ and $C^j \geq C^i$. Of course, this raises an important question: does this relationship hold even if we do *not* assume that $R^j \geq R^i$ and $C^j \geq C^i$, or vice versa? In other words, is it the case that the cross-partial derivative of i and j always gets "closer" to zero as the rewards and costs to i decrease *regardless of the consequences of* j? To answer this question, I will impose one more (and final) assumption:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i} = 0, \qquad if\ R^{jT} R^i = 0\ and\ C^{jT} C^i = 0,$$

[61] It should be noted that none of the terms in either sequence references any particular component(s) of either R^i or C^i. More specifically, by writing $R_1^i > R_2^i$, I am merely implying that at least one component of R_1^i is greater than that of R_2^i, while all other components of R_1^i are either greater than or equal to those of R_2^i. A similar definition can be found in Kreps (2013).

where $R^{jT}R^{i}$ and $C^{jT}C^{i}$ respectively denote the inner (dot) products of the reward and cost vectors to j and i. Simply put, if activities i and j produce *completely dissimilar consequences*, then any increase in j will have no impact on the value of $MU(A_i)$ and vice versa. As such, for any (increasing) sequence of reward vectors:

$$R_1^i < R_2^i < \cdots < R_n^i, \qquad given\ R_n^{iT}R^j = 0\ for\ all\ R_n^i,$$

it follows:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_1^i\right) = \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_2^i\right) = \cdots = \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_n^i\right) = 0.$$

Suppose we fix some overlapping components of R^j and R^i, such that:

$$R_1^i < R_2^i < \cdots < R_n^i, \qquad given\ R_n^{iT}R^j = K\ for\ all\ R_n^i,$$

where K is some non-negative real number. Since the value of $R^{jT}R_n^i$ is assumed constant for every R_n^i, we can deduce that:

$$R_n^i > R_{n-1}^i\ only\ if\ \left(R_n^i - R_{n-1}^i\right)^T R^j = 0\ for\ all\ n,$$

where $R_n^i - R_{n-1}^i$ is the vector "sum" of R_n^i and $-1 * R_{n-1}^i$ for any choice of n in the above sequence (e.g., $R_2^i > R_1^i$ only if $\left(R_2^i - R_1^i\right)^T R^j = 0$). Likewise, it follows that the value of $R^{jT}R_n^i$ remains constant only in the event each positive component of $R_n^i - R_{n-1}^i$, for every possible value of n, is equal to *zero* in R^j (i.e., any and all reward types altered throughout the sequence R_n^i are precisely those types which are *not addressed by* R^j). If we were to compute the cross-partial derivatives for each of these "difference" vectors, we would get:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_2^i - R_1^i\right) = \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_3^i - R_2^i\right) = \cdots = \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_n^i - R_{n-1}^i\right) = 0.$$

This, of course, follows from $\left(R_n^i - R_{n-1}^i\right)^T R^j$ being equal to zero for all selections of n. Notice that by "first-differencing" each R_n^i in this manner we are effectively removing the influence that any potential overlapping reward type for i and j might have on their cross-partial derivative. In other words, we can think of the above expression as what we would expect to happen to $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ in the event we only altered some number of reward types of i which are completely independent of those which are addressed by j. Specifically, this expression states that every change made to any reward type uniquely addressed by i will *not influence the value of* $\frac{\partial^2 EU}{\partial A_j \partial A_i}$. Additionally, since the cross-partial derivative with respect to i and j can only potentially decrease with any overlapping component(s) of R^j and R^i, it follows that:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_1^i\right) = \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_2^i\right) = \cdots = \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_n^i\right) \leq 0,$$

$$given\ R_n^{iT} R^j = K\ for\ all\ R_n^i.$$

Note that this expression remains valid if we were to replace each R_n^i with some (also increasing) sequence of C_n^i instead. Furthermore, since the inner product between any R^j (or C^j) with its corresponding zero vector always yields a value of zero (see: Shilov, 1971), it follows that, for any $j \neq i$:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i} = 0, \qquad if\ R^i = 0\ and\ C^i = 0.$$

That is, the cross-partial derivative between any activities j and i will *always be equal to zero* if it is the case that i (or j) produces no rewards or costs for the agent. Consider, then, a countably infinite, strictly decreasing sequence of R^i and C^i which converges zero (i.e., the respective zero vectors for R^i and C^i). More specifically:

$$\left(R_1^i, C_1^i\right) > \left(R_2^i, C_2^i\right) > \cdots > \left(R_n^i, C_n^i\right) > \cdots ,$$

$$such\ that\ \left(R_n^i, C_n^i\right) \longrightarrow (0,0)\ as\ n \longrightarrow \infty,$$

where $\left(R_n^i, C_n^i\right)$ is the n-th pair of vectors in the above sequence, and $\left(R_{n-1}^i, C_{n-1}^i\right) > \left(R_n^i, C_n^i\right)$ only if $R_{n-1}^i > R_n^i$ and $C_{n-1}^i > C_n^i$ in all components (i.e., every reward and cost type of the first term is greater than the second). The expression "$\left(R_n^i, C_n^i\right) \longrightarrow (0,0)$ as $n \longrightarrow \infty$" can be read: "the pair of vectors R_n^i and C_n^i approach their respective pair of zero vectors as the value of n approaches infinity." Since the value of $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ is not affected by changes in any rewards and costs unique to i, we can deduce:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_{n-1}^i, C_{n-1}^i\right) \neq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_n^i, C_n^i\right),$$

$$only\ if\ \left(R_{n-1}^i - R_n^i\right)^T R^j \neq 0\ or\ \left(C_{n-1}^i - C_n^i\right)^T C^j \neq 0\ for\ any\ n.$$

Therefore, we can ignore every component of R_n^i and C_n^i respectively equal to zero in R^j and C^j. Suppose $\left(R_1^i, C_1^i\right) > \left(R^j, C^j\right)$ for all non-zero components of R^j and C^j. Per the basic principles of mathematical analysis (see: Rudin, 1953), there exists some $k > 1$ such that, for all $n < k$:

$$\left(R_1^i, C_1^i\right) > \left(R_2^i, C_2^i\right) > \cdots > \left(R_{k-1}^i, C_{k-1}^i\right) > \left(R^j, C^j\right).$$

Since the cross-partial derivative of i and j is bound from below by $\frac{\partial^2 EU}{\partial {A_i}^2}$, which itself is a decreasing function of the rewards and costs to i (as previously established), it follows that:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_1^i, C_1^i\right) \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_2^i, C_2^i\right) \leq \cdots \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_{k-1}^i, C_{k-1}^i\right).$$

Additionally, since the sequence $\left(R_n^i, C_n^i\right)$ converges to (0,0), for any pair of R^j and C^j we can also find some $l > k$ such that, for all $n \geq l$:

$$\left(R^j, C^j\right) > \left(R_l^i, C_l^i\right) > \left(R_{l+1}^i, C_{l+1}^i\right) > \cdots.$$

Since the distance between $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ and zero can only decrease as each of the rewards and costs to i approach zero for all $\left(R_n^i, C_n^i\right) < \left(R^j, C^j\right)$ (as also previously established), it follows:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_l^i, C_l^i\right) \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_{l+1}^i, C_{l+1}^i\right) \leq \cdots.$$

Furthermore, since every second partial derivative (including any cross-partial derivatives) is assumed to be a continuous and monotonic function of the rewards and costs to any i and j (which, recall, are elements of $\mathbb{R}_+^k$ and $\mathbb{R}_+^l$, respectively), it follows from the intermediate value theorem (see: Munkres, 2021, p. 154) that,[62] for all $k \leq n < l$:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_{k-1}^i, C_{k-1}^i\right) \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_k^i, C_k^i\right) \leq \cdots \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_{l-1}^i, C_{l-1}^i\right) \leq \frac{\partial^2 EU}{\partial A_j \partial A_i}\left(R_l^i, C_l^i\right).$$

Finally, since $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ is equal to zero in the event $\left(R_n^i, C_n^i\right) = (0,0)$, the value of $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ with respect to $\left(R_n^i, C_n^i\right)$ converges to zero as n approaches infinity. That is:

[62] Note that this is a direct result of the connectedness of $\mathbb{R}_+^k$ and $\mathbb{R}_+^l$. More specifically, since every closed and convex subset of any finite-dimensional Euclidean vector space is connected, as per the principles of elementary point-set topology, any continuous, real-valued function over such a subset (e.g., $\mathbb{R}_+^k$ or $\mathbb{R}_+^l$) has the intermediate value property. A summary of this property is as follows: given $f: X \longrightarrow \mathbb{R}$ is continuous over some connected space X, for any $x, y \in X$ such that $f(x) < f(y)$, we can always find some $z \in X$ such that $f(x) < f(z) < f(y)$ for every possible value of f contained in the open interval $(f(x), f(y))$. Since each second-order derivative with respect to j is a decreasing function of R^j and C^j, it follows that, for any two distinct terms in the sequence (R_n^i, C_n^i), there exists some (R^i, C^i) such that $(R_n^i, C_n^i) > (R^i, C^i) > (R_{n+m}^i, C_{n+m}^i)$ for every value of $\frac{\partial^2 EU}{\partial A_j \partial A_i}$ which falls between $\frac{\partial^2 EU}{\partial A_j \partial A_i}(R_n^i, C_n^i)$ and $\frac{\partial^2 EU}{\partial A_j \partial A_i}(R_{n+m}^i, C_{n+m}^i)$. For further discussion, see Yandl and Bowers (2016).

$$\lim_{(R_n^i, C_n^i) \to (0,0)} \frac{\partial^2 EU}{\partial A_j \partial A_i} = 0, \qquad for\ any\ j\ and\ i.$$

Therefore, the cross-partial derivative of i and j is a direct function of the extent to which the rewards and costs to i are also addressed by j. The greater the degree of overlap between the consequences to both activities, the "stronger" the effect any increase in i will have on the marginal utility of j (or vice versa). With this in mind, the question now becomes: how might changes in the rewards and costs to i influence the *optimal consumption rate* of j under varying levels of overlap in their respective rewards and costs?

OPTIMAL CHOICE AS A FUNCTION OF REWARDS AND COSTS:

Let's consider a (relatively) simple example to start. Assume each $A \in \mathcal{A}$ contains only two activities (i.e., $A \in \mathbb{R}_+^2$), denoted by i and j, both of which produce completely dissimilar consequences (i.e., $R^{jT}R^i = 0$ and $C^{jT}C^i = 0$). Suppose we increase any number of non-zero components of R^i, as well as decrease some non-zero components of C^i, and, by doing so, increase the value of $MU(A_i)$ for all $A \in \mathcal{F}$. As before, I will denote the set of "prior" solutions by $\mathcal{S}'$ (i.e., the set of all optimal selections of $A \in \mathcal{F}$ before altering any components of R^i and C^i) and the "posterior" solution set by $\mathcal{S}$. Say we select some $A \in \mathcal{S}'$ such that $A_i^* > 0$ and $A_j^* > 0$. As previously established, every $A \in \mathcal{S}$ assigns a value of A_i^* which is at least as high as that which is assigned by any $A \in \mathcal{S}'$ (i.e., the value of A^i can only potentially increase as its marginal utility increases). Since the cross-partial derivative of any i and j with non-overlapping consequences is equal to zero, it follows that the value of $MU(A_j)$ remains constant regardless of our selection of A in $\mathcal{S}$ (i.e., the marginal utility of j is not altered through any changes in A_i). Thus, for any $A \in \mathcal{S}$ we can always find some $A \in \mathcal{S}'$ such that:

$$A_j^*(\mathcal{S}') \leq A_j^*(\mathcal{S}),$$

where $A_j^*(\mathcal{S}')$ denotes the value of A_j^* for some $A \in \mathcal{S}'$ (same for $A_j^*(\mathcal{S})$). If the value of $e^T A$ is less than E for any selection of A in both $\mathcal{S}$ and $\mathcal{S}'$, it follows:

$$A_j^*(\mathcal{S}') = A_j^*(\mathcal{S}), \qquad for\ some\ A \in \mathcal{S}'\ and\ A \in \mathcal{S}.$$

The rationale behind this expression is as follows: since changes in A_i do not influence the value of $MU(A_j)$, we can simply "increase" the value of A_i to achieve equilibrium between $MU(A_i)$ and $MU(A_j)$, as required in order for A to belong to $\mathcal{S}$, without altering the value of A_j (note that such an equilibrium must exist since it is assumed that $e^T A < E$ for each element in $\mathcal{S}$, as otherwise we could simply keep "increasing" A_i until $e^T A = E$, at which point equilibrium may only be achievable by *decreasing* some quantity of A_j; see: Boyd and Vandenberghe, 2004). Additionally, if there exists only one optimal value of A_j^* for every A in both $\mathcal{S}'$ and $\mathcal{S}$, then:

$$A_j^*(\mathcal{S}') = A_j^*(\mathcal{S}), \qquad for\ all\ A \in \mathcal{S}'\ and\ A \in \mathcal{S},$$

which follows from:

$$MU\left(A_j^*\right) = MU(A_i^*) = 0, \qquad given\ e^T A < 0\ for\ all\ A \in \mathcal{S}'\ and\ A \in \mathcal{S}.$$

However, if the value of $e^T A = E$ for each optimized selection of A (in both $\mathcal{S}'$ and $\mathcal{S}$), then, by increasing the value of $MU(A_i)$, we are now "forced" to *lower* the value of A_j in order to achieve equilibrium (see: Kreps, 2013). In fact, since $MU(A_j)$ is a decreasing function of A_j, by selecting a lower quantity of A_j the value of $MU(A_j)$ can only *increase*. Thus, to achieve equilibrium we need only continually make (minor) adjustments to both A_i and A_j until their marginal utilities are equal (while, of course, preserving the value of $e^T A$).

Now let's consider an identical scenario with one alteration: the cross-partial derivative of both i and j is now less than zero, although neither i nor j strictly dominates the other (e.g., there is at least some overlap between their respective rewards and costs, but each activity also addresses some set of needs the other does not). If we were to once again raise the rewards to i and lower its costs by the same amount (thus achieving the same upward shift in marginal utility), we would, obviously, need to raise the value of A_i. However, in this scenario whenever we raise the value of A_i, we are simultaneously *lowering the value of* $MU(A_j)$. As a consequence, to achieve equilibrium in the case where $e^T A < E$ for each of our optimized selections of A, we must, in addition to increasing A_i, also *decrease* the value of A_j. That is:

$$MU(A_j) < MU(A_i) = 0, \qquad for\ some\ increase\ in\ A_i,$$

and hence our selection of any $A \in \mathcal{S}$ in our previous example will not belong to $\mathcal{S}$ in this situation (this, of course, follows from A only being a solution to our utility maximization problem if it is the case that the marginal utilities of all non-zero actions are equal to zero, given $e^T A < E$). Additionally, any decrease in A_j will *increase* the value of $MU(A_i)$ (since, again, the cross-partial derivative is negative), and thus any selection of $A \in \mathcal{S}$ in this scenario will involve both a lowering of A_j, as well as an *even larger* increase in A_i. Furthermore, the same is also true in the event $e^T A = E$ for each of our optimized selections of A, as any proportional exchange between A_i and A_j will have a *greater impact* on their respective marginal utilities. In other words, the greater the overlap in the rewards and costs to i and j, the *more their respective optimal values are swayed* by marginal changes in the consequences to i (or j).

Of course, these are fairly simple examples, and naturally things become considerably more complicated when three or more activities are involved. As an example, although i and j might have a cross-partial derivative of zero, there could exist some third behavior k which has a negative cross-partial derivative with both i and j (e.g., k addresses some number of needs addressed by both i and j, while i and j address completely dissimilar needs). In this situation, an increase in A_i may require a decrease in A_k, which may then require (or at least allow for) some *increase* in A_j in order to achieve equilibrium. Regardless, generally speaking it should be the case that changes in the rewards and costs to some activity i should tend to have a greater impact on any optimized selection of A_i whenever there exists some set of alternative actions which address similar needs to i to, roughly, the same degree.[63]

IMPLICATIONS FOR THEORY AND RESEARCH:

Before delving more into the theoretical, as well as empirical, implications of this discussion, I first want to reiterate that the particular approach I have chosen to take here is by no means the end-all, be-all of rational choice theorizing (or even microeconomic reasoning, for that matter). Arguments have been made for a less mathematical conception of criminal choice (see: Cornish and Clarke, 1986), and, as such, I do not wish to suggest that the only worthwhile direction of the criminological rational choice perspective is of a purely "Beckerian" nature. I do, however, firmly believe that such efforts have nonetheless been immensely fruitful for the study of criminal decision-making, and that we have succeeded thus far in only scratching the surface of the implications of Becker's model.

[63] The proof of this "conjecture," however, is well beyond the scope of this project, and as such the reader is recommended to refer to the mathematical optimization literature for a more in-depth look into this subject. A good place to start, in my view, is Boyd and Vandenberghe (2004).

Likewise, the technical discussion I provide throughout this appendix (along with this dissertation as a whole) should be seen as merely one attempt among many to understand what those implications are, and to submit them to empirical testing. Hopefully, I have succeeded in establishing a reasonable (enough) conceptual foundation which both aligns with the spirit of the Beckerian approach and can help guide other scholars' thinking about crime decisions. With that out of the way, I will now dive into the primary implications of this foundation, among which include the primary hypotheses for this dissertation.

Criminal Behavior as a Nonlinear Function of Rewards, Risks, and Costs:

Perhaps the "simplest" implication of the preceding discussion is that the degree to which marginal changes in any particular incentive (i.e., some unique type of reward, risk, or cost) influences criminal behavior is itself a function of *each other incentive the individual is concerned with*. In other words, any estimated empirical link between, say, the social benefits and crime may itself depend on an individual's perceptions of, say, the personal benefits to crime, probability of arrest, and the social costs to offending. On a more technical level, we can think of criminal involvement as being a *nonlinear function* of the perceived rewards, risks, and costs to crime (Strogatz, 2015). By a nonlinear function, I am simply referring to a situation where involvement in criminal activities, which I will denote as Y, cannot be (appropriately) represented by the usual linear regression model. That is:

$$E(Y) = \alpha + \sum_{k=1}^{K} \beta_k r_k + \sum_{l=1}^{L} \beta_l c_l + \varepsilon,$$

where r_k and c_l respectively denote some specific type of reward and cost (once again, I will assume the values of each type are "weighted" by their respective probabilities of occurring), and β_k and β_l denote any real number (i.e., "coefficient") assigned to each reward and cost. Here, the expected value of Y is represented as a *linear combination* of each of the model's "inputs," meaning any particular person's level of involvement in Y can be approximated by adding the values of each of his or her perceived rewards, risks, and costs weighted by their respective coefficients. This is, of course, a fairly optimistic view of estimation, as it not only assumes that each unique type of reward, risk, and cost will exhibit a "linear" effect on Y, but that they will also do so *independently of one another*. That is, every unique incentive type will exhibit the same influence on Y *regardless of the values of each other incentive type*. The question now becomes: how reasonable is this approach for examining the influence of perceived incentives on criminal behavior?

Consider first the "shape" of Y as a function of a particular incentive (say, r_k), given all other incentive types are held constant. Let Y be vector whose components are equal to the "optimal" consumption rate(s) of a set of criminal activities (e.g., burglary, robbery, shoplifting, and so on), and r_k is a vector denoting the value of, say, the personal benefits to each activity contained in Y. For Y to be a linear function of r_k, it must be the case that:

$$\beta_k^j = \frac{\partial Y}{\partial r_k^j}, \qquad for\ all\ R^j \in \mathbb{R}_+^k.$$

That is, for every possible "change" in any given r_k^j contained in r_k, the rate of change in Y with respect to each r_k^j (i.e., $\frac{\partial Y}{\partial r_k^j}$) must remain constant (i.e., equal to some real number β_k^j). Assume there exists some $r_k > 0$ (denoted by r_k^0) for which every "lesser" set of personal benefits leads to an optimal level of involvement in each (illegal) activity j

equal to zero. Put differently, the agent is best off avoiding all activities contained in Y for any r_k whose components are all less than r_k^0. It follows that:

$$\nabla_Y r_k = 0, \qquad for\ all\ r_k < r_k^0,$$

where $\nabla_Y r_k$ denotes the gradient vector of r_k (i.e., the set of partial derivatives assigned to each $r_k^j \in r_k$). As such, any increase in r_k which results in a rewards vector whose components remain less than those of r_k^0 will lead to no change in Y. Obviously, such an association is best described by a value of β_k equal to zero. However, if there also exists a value of r_k (denoted by r_k^1) for which every "greater" set of personal rewards leads to a value of Y greater than zero, we can identify a countably infinite sequence of r_k (say, r_{kn}, where n denotes a positive integer) for which:

$$r_{k1} < \cdots < r_{km} < r_k^0 < r_{km+1} < \cdots < r_k^1 < r_{kn} < \cdots, \qquad m < n.$$

Here, any element in the sequence r_{kn} whose subscript is less than or equal to m achieves a gradient vector of zero, while each element whose subscript is greater than or equal to n may well achieve a non-zero gradient vector (i.e., it is possible that some increase in r_k beyond r_{kn} may lead to an *increase in* Y). In such a scenario, any selected value of β_k will lead to some level of error in the predicted value of Y for a subset of possible values of r_k (if, for instance, $\beta_k = 0$, then $\beta_k \neq \nabla_Y r_k$ for some $r_k > r_k^1$). Thus, even for a single reward type, representing its overall influence on Y with a static coefficient (i.e., any β_k) may be insufficient (the same is true for any type of risk or cost as well). A better approximation may be achieved if were to instead allow the value of β_k to *vary* with respect to r_k. For example:

$$\beta_k = \|\nabla_Y r_k\|, \qquad given\ all\ else\ equal.$$

That is, β_k is no longer static and is instead equal to the value of the (Euclidean) *length of the gradient of* r_k. In theory, such an approach should lead to a better approximation of the value of Y for all possible values of r_k (given we hold each other incentive type constant), as β_k now denotes the *average* influence of r_k on Y (e.g., if the overall effect of some marginal increase in any component of r_k is "stronger" for some values of r_k, then each component of $\nabla_Y r_k$ will be greater as well, which will lead to a higher value of $\beta_k = \|\nabla_Y r_k\|$). Such an approach is often seen in generalized additive models (see: Wood, 2017), where the "effect" of any given predictor variable is represented by a *function* of said predictor, whose shape can take a number of different (possibly nonlinear) forms. However, even this approach has some problems. Namely, it assumes that the influence of r_k on Y can be captured by a single "curve" embedded in $\mathbb{R}^2$, whose shape remains static across all possible values of each other incentive type.

Consider the sequence of r_{kn} defined previously. Here, we assumed that the value of all other incentive types never changed throughout this sequence, which allowed us to identify a fixed set of r_k^0 and r_k^1 (i.e., fixed vectors for which all r_k below r_k^0, or above r_k^1, lead to a restricted set of possible values for Y). Let c_{lm} be an increasing, countably infinite sequence of vectors denoting the value of, say, the social costs to each j in Y. Since the costs to each j can only lower their respective marginal utility values, it follows that:

$$MU_Y(r_{kn} \mid c_{l1}) \geq MU_Y(r_{kn} \mid c_{l2}) \geq \cdots \geq MU_Y(r_{kn} \mid c_{lm}) \geq \cdots,$$

where MU_Y denotes the set of marginal utilities for each member of Y for any fixed element of r_{kn}, given the value of c_l is held constant at some term in the sequence c_{lm}. Because of this, it is possible that, for some high enough term in c_{lm}:

$$\|r_k^1(c_{l1})\| < \inf(\|r_k^1(c_{ln})\|), \qquad for\ some\ n > 1,$$

where $\inf(\|r_k^1(c_{ln})\|)$ denotes the "shortest" possible length of any given r_k^1 with respect to some $c_{ln} > c_{l1}$. In other words, it is possible that any previously selected r_k^1 (holding c_l constant at c_{l1}) may *no longer produce only positive values of* Y *for all* $r_k > r_k^1$ at higher values of c_l. Consequently, any static function representing the association between r_k and Y may fail to properly capture the "true" underlying relationship for different values of other incentive types (including other types of rewards as well as risks and costs). Likewise, a more reasonable approach might be to think of Y as a *multivariable function* of different incentives. For instance:

$$Y(r_k, c_l) : \mathbb{R}^n \longrightarrow \mathbb{R}^m, \qquad n = k + l,$$

where $Y(r_k, c_l)$ denotes a function carrying each element of $\mathbb{R}^n$ into a subset of $\mathbb{R}^m$ (i.e., the "image set" of Y). Note that m is assumed to be equal to the number of components of Y, while n is merely the sum of all components for some set of rewards and costs vectors. Of course, we can simplify things even further by denoting the value of Y as, say, the total quantity (or *variety*) of offenses committed with respect to each j. If we also denote the vectors r_k and c_l by their average (i.e., mean) values across all offense types (as commonly done in prior empirical examinations of criminological rational choice), we get:

$$Y(r^k, c^l) : \mathbb{R}^n \longrightarrow \mathbb{R}, \qquad n = k + l,$$

where r^k and c^l denote vectors containing the average "gain" or "loss" for any possible number of reward and cost types (e.g., r^k is a vector with two components, each of which respectively denotes the average social and personal benefits to all activities in Y). Since the value of Y is now treated as a real-valued function with respect to several types of rewards, risks (e.g., the possibility of apprehension and associated formal punishments by the state), and costs, it is now possible for the functional form (i.e., the fitted "curve") of the relationship between, say, the social benefits and Y to take on a *variety of shapes* depending on the values of each other incentive variable. That is, for any given combination of (fixed) values for each other incentive type (e.g., the personal benefits, probability of arrest, and social costs), it may be possible to find some other combination of values such that:

$$f_Y(r_k|\boldsymbol{I}_1) \neq f_Y(r_k|\boldsymbol{I}_2), \qquad \boldsymbol{I}_m \in \mathbb{R}^{n-1},$$

where f_Y denotes the functional form of the association between r_k and Y, holding each other incentive type, denoted by $\boldsymbol{I}_m$, constant (note that since $\boldsymbol{I}_m$ denotes a fixed point for all non-r_k incentives, every possible value of $\boldsymbol{I}_m$ thus belongs to an $n-1$ dimensional Euclidean space). Here, the above expression merely states that the set of points contained in the image set of f_Y, for some fixed $\boldsymbol{I}_1$, may not include all points contained in the corresponding image set of f_Y with respect to $\boldsymbol{I}_2$ (i.e., it is possible that $f_Y(r_k|\boldsymbol{I}_1) - f_Y(r_k|\boldsymbol{I}_2)$ may not return the "zero function;" Kolmogorov & Fomin, 1970). Such "multivariable" functions can also be estimated through generalized additive modeling procedures. Particularly, one can fit a *tensor product smooth* over any number of predictor variables in a data frame, which produces an estimated model manifold (i.e., an n-dimesional "surface" embedded in $\mathbb{R}^{n+1}$) which can potentially capture any nonlinear direct, as well as *moderating*, influence of each predictor on Y. Given there seems to be at least some reason to believe that the functional form of any given incentive type may, itself, be a function of other incentives, such a method could be useful for examining the more nuanced dimensions of the incentives-to-crime relationship (details are provided in Chapter 3).

Criminogenic Effect of the Rewards as a Function of the Risks and Costs:

Next, let's consider whether the agent's *responsiveness* to the perceived social and personal rewards to crime could be a function of crime's risks and costs, at least on some level. By "responsiveness," I am referring to the overall influence changes in the rewards to criminal activities will exert on the optimized value of Y (as defined in the previous section). Say we have some measure of the social benefits to crime, which (ideally) can be represented by a closed interval on the real number line. That is:

$$r_k \in [L, U],$$

where $[L, U]$ denotes the range of possible values observable for r_k (in this case, L denotes the lowest value of r_k, while U denotes the highest). In estimating the overall influence of r_k on Y, for some fixed $\boldsymbol{I}_m$ (again, as defined in the previous section), we simply wish to get a general sense of how much Y tends to *change* as the value of r_k "moves" from L to U. One such estimate would be the following:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_m) = \int_L^U \frac{\partial Y}{\partial r_k} dx, \qquad given\ some\ fixed\ \boldsymbol{I}_m,$$

where $\beta_{\boldsymbol{I}}$ is meant to capture the general influence of r_k on Y. Here, the value of $\beta_{\boldsymbol{I}}$ is determined by the definite integral of the partial derivative of r_k (defined over $[L, U]$), given each other incentive type is held constant at the point $\boldsymbol{I}_m$. It follows from the fundamental theorem of single-variable calculus (see: Spivak, 1971) that:

$$\int_L^U \frac{\partial Y}{\partial r_k} dx = Y(U) - Y(L),$$

where $Y(U)$ is simply the value of Y when $r_k = U$ (the value of $Y(L)$ is defined analogously). Put simply, the value of $\beta_{\boldsymbol{I}}$ is just the *difference in* Y along the boundary of $[L, U]$ (i.e., the degree to which Y changes between the lowest and highest possible values of r_k). Since it is possible that the shape of f_Y (see the previous section) may differ between distinct values of $\boldsymbol{I}_m$, we can think of $\beta_{\boldsymbol{I}}$ as a *function of* $\boldsymbol{I}_m$. In particular, $\beta_{\boldsymbol{I}}$ may achieve higher (or lower) values at distinct points of the distribution of some measured set of the risks and costs to crime. Let $\boldsymbol{I}_C$ be a point contained in the set of all combinations of values for some measure of the perceived probability of arrest, along with a measure of the perceived social costs to crime (i.e., each possible $\boldsymbol{I}_C$ belongs to a closed and bounded subset of $\mathbb{R}^2$). It follows that:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C) : \mathbb{I}_C \longrightarrow \mathbb{R}, \qquad \mathbb{I}_C \subset \mathbb{R}^2,$$

where $\mathbb{I}_C$ is the unique subset of $\mathbb{R}^2$ which contains all possible $\boldsymbol{I}_C$ (note I am also assuming the complement of $\mathbb{I}_C$ only contains points of $\mathbb{R}^2$ which fall outside of the "boundary" of every possible value of $\boldsymbol{I}_C$). Define some increasing sequence of $\boldsymbol{I}_C$ by $\boldsymbol{I}_{Cn}$, such that:

$$\boldsymbol{I}_{C1} < \boldsymbol{I}_{C2} < \cdots < \boldsymbol{I}_{Cn}.$$

For the above expression, a statement such as $\boldsymbol{I}_{C1} < \boldsymbol{I}_{C2}$ simply means that each component of $\boldsymbol{I}_{C1}$ is, respectively, less than those of $\boldsymbol{I}_{C2}$. Since any increase in the risks and costs to crime can only decrease the marginal utility of all activities contained in Y (as previously shown), it may be the case that higher values of $\boldsymbol{I}_C$ lead to lower values of $\beta_{\boldsymbol{I}}$. Such a phenomenon was used as an example in the previous section, where some sequence of "cost" vectors was presumed to lead to an increase in the "minimum" quantity of rewards necessary for Y to achieve an optimal value greater than zero. For this example, let's presume there exists some element of $\boldsymbol{I}_{Cn}$ for which:

$$\inf(r_k^1 | \boldsymbol{I}_{Cm}) > U, \qquad m \leq n.$$

That is, the "smallest" possible value of r_k for which $Y > 0$, given some $\boldsymbol{I}_{Cm}$ contained in $\mathbb{I}_C$, is greater than the maximum possible value of r_k contained in $[L, U]$. Since the infimum (i.e., the "greatest lower bound;" Rudin, 1953, p. 4) of r_k^1 can only increase with $\boldsymbol{I}_C$, it follows that:

$$Y(r_k \in [L, U] | \boldsymbol{I}_{Cn}) = 0, \qquad for\ all\ n > m,$$

and therefore:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C) = 0, \qquad for\ all\ \boldsymbol{I}_C \geq \boldsymbol{I}_{Cn}.$$

Note that follows from $Y(U) = Y(L) = 0$ for every $\boldsymbol{I}_C$ whose components are respectively greater than or equal to those of $\boldsymbol{I}_{Cn}$. Hence, this particular scenario seems to imply that there may exist some combination of values for the perceived probability of arrest and social costs to crime for which observed changes in the perceived social benefits to criminal activities (i.e., r_k) may have *zero influence* on the optimized value of Y (i.e., $Y = 0$ regardless of r_k). Of course, we could just as easily establish a similar relationship between the overall influence of *multiple types of rewards* and $\boldsymbol{I}_C$. Say we have a measure of the perceived thrills and excitement for engaging in criminal acts, which I will denote by $r_t \in [L_t, U_t]$ (note that I will now also denote the value of the social rewards by $r_k \in [L_k, U_k]$). Similar to $\mathbb{I}_C$, the set of all possible combinations of values for the social and personal benefits to crime can be defined by:

$$\mathbb{I}_R = [L_k, U_k] \times [L_t, U_t] \subset \mathbb{R}^2,$$

where $\mathbb{I}_R$ is a closed and bounded subset of $\mathbb{R}^2$, the boundary of which separates all possible values of r_t and r_k from those which lie outside of the "rectangle" defined by $[L_k, U_k] \times [L_t, U_t]$ (note that each element of $\mathbb{I}_C$ can be said to be contained within a "rectangle" embedded in $\mathbb{R}^2$ as well). In this instance, we can estimate the overall influence of r_t and r_k by:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C) = \iint\limits_{\mathbb{I}_R} \left(\frac{\partial Y}{\partial r_k} + \frac{\partial Y}{\partial r_t}\right), \qquad for\ some\ fixed\ \boldsymbol{I}_C,$$

where $\beta_{\boldsymbol{I}}$ is now a function of two variables, the value of which is equal to the double integral (defined over $\mathbb{I}_R$) of the *differential* of Y with respect to r_t and r_k. That is, $\beta_{\boldsymbol{I}}$ captures the overall degree to which Y appears to change with both r_t and r_k, except in this instance we have two partial derivatives to consider instead of one. Specifically, we compute the sum of the partial derivatives of r_t and r_k for each point in $\mathbb{I}_R$, which is equivalent to taking the inner product of $\nabla_Y(r_t,r_k)$ and the "1" vector (i.e., the differential of Y is merely its *directional derivative* with respect to 1, such that 1 is a vector contained in $\mathbb{R}^2$ whose components are each equal to 1; Edwards Jr., 1973). Because of this, the fundamental theorem of single-variable calculus (as employed previously) is insufficient for computing the value of $\beta_{\boldsymbol{I}}(\boldsymbol{I}_C)$. Instead, we can employ what is sometimes referred to as the fundamental theorem of *vector* calculus, otherwise known as Stokes' Theorem (Spivak, 1971), which makes use of the properties of the *boundary* of $\mathbb{I}_R$ to compute $\beta_{\boldsymbol{I}}$ (details of how, exactly, this theorem can be employed are provided in Chapter 3).

Once again, we can consider a situation for which some value of $\boldsymbol{I}_C$ elicits an optimal Y value of zero for all elements of $\mathbb{I}_R$. That is:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C) = 0, \qquad for\ all\ \boldsymbol{I}_C \geq \boldsymbol{I}_{Cn}.$$

Thus, regardless of which (observable) values we select for r_t and r_k, their respective partial derivatives will be equal to zero (and so too will the value of $\beta_{\boldsymbol{I}}$) for every $\boldsymbol{I}_C \geq \boldsymbol{I}_{Cn}$. As such, this scenario (once again) appears to imply that higher values of each component of $\boldsymbol{I}_C$ (i.e., higher perceived probability of arrest and social costs) may lead to lower values of $\beta_{\boldsymbol{I}}$. It should be noted, however, that it may also be the case that some "higher" values of $\boldsymbol{I}_C$ may actually lead to a *stronger* overall influence of r_t and r_k on Y. Consider, for example, a scenario where the "smallest" possible value of $\boldsymbol{I}_C$ (say $\boldsymbol{I}_C^0$) is equal to the zero vector. More concretely, this "minimum" value of $\boldsymbol{I}_C^0$ corresponds to the agent believing there exists a zero chance of arrest, as well as no social losses acquired, for engaging in any amount of Y. Likewise, since there exists very little "disincentive" for engaging in criminal activities, the agent may happily take advantage of any criminal opportunity which comes his or her way, even if said opportunities offer relatively "small" benefits. Because of this, the agent may, paradoxically, seem relatively "insensitive" to changes in r_t and r_k, despite the fact that he or she is heavily involved in criminal activities. Alternatively, we can consider the following:

$$D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R) = \lim_{h \to 0} \frac{EU(N + hy) - EU(N)}{\|hy\|}, \qquad given\, h > 0.$$

where $D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R)$ denotes the *directional derivative* of EU at the point N, with respect to some (infinitesimally small) proportional exchange between N and Y, denoted by the vector y (see: Spivak, 1965), the value of which is a function of $\boldsymbol{I}_C$ and $\boldsymbol{I}_R$. Let N denote the optimal strictly legal course of action (i.e., EU can only potentially increase at the point N if the agent were to involve his or her-self in some amount of Y). It follows that, for some sufficiently small value of $h > 0$:

$$D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R) > 0, \qquad only\ if\ EU(N + hy) - EU(N) > 0,$$

and:

$$D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R) < 0, \qquad only\ if\ EU(N + hy) - EU(N) < 0.$$

Suppose we want to get a sense of how much of an overall change in the value of EU, for some exchange of N for Y, we can expect to see *on average* with respect to all $\boldsymbol{I}_R \in \mathbb{I}_R$. Denote this average by ΔEU_R, for which:

$$\Delta EU_R(\boldsymbol{I}_C) = \iint_{\mathbb{I}_R} |D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R)|, \qquad for\ some\ fixed\ \boldsymbol{I}_C,$$

where $|D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R)|$ denotes the absolute value of $D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R)$. Consider the following scenario for which, given $\boldsymbol{I}_C = \boldsymbol{I}_C^0$ and some sufficiently small $h > 0$:

$$EU(N + hy) - EU(N) > 0, \qquad for\ all\ \boldsymbol{I}_R \in \mathbb{I}_R.$$

Suppose we select some $\boldsymbol{I}_C > \boldsymbol{I}_C^0$ for which the above expression remains true. Since the value of EU with respect to every $Y > 0$ is a decreasing function of $\boldsymbol{I}_C$ (as previously established), it follows that, for any fixed $Y = N + hy$:

$$EU(Y|\boldsymbol{I}_C) < EU(Y|\boldsymbol{I}_C^0), \qquad for\ all\ \boldsymbol{I}_R \in \mathbb{I}_R,$$

thus:

$$D_Y(\boldsymbol{I}_C, \boldsymbol{I}_R) < D_Y(\boldsymbol{I}_C^0, \boldsymbol{I}_R), \qquad for\ all\ \boldsymbol{I}_R \in \mathbb{I}_R,$$

and hence:

$$\Delta EU_R(\boldsymbol{I}_C) < \Delta EU_R(\boldsymbol{I}_C^0).$$

Note that we can also establish a similar derivative in the "opposite" direction:

$$D_N(\boldsymbol{I}_C, \boldsymbol{I}_R) = \lim_{h \to 0} \frac{EU(Y - hy) - EU(Y)}{||hy||}, \qquad given\ h > 0.$$

In this instance, Y denotes some fixed $Y > 0$, and $D_N(\boldsymbol{I}_C, \boldsymbol{I}_R)$ denotes the directional derivative of EU moving *toward* N (i.e., away from Y). Since the marginal utility of Y is also a decreasing function of $\boldsymbol{I}_C$, it follows that, given $EU(Y) > EU(N)$:

$$D_N(\boldsymbol{I}_C^0, \boldsymbol{I}_R) < D_N(\boldsymbol{I}_C, \boldsymbol{I}_R) < 0.$$

That is, the agent has *more to lose* from trading some amount of Y for N for lower values of $\boldsymbol{I}_C$ (i.e., those closer to $\boldsymbol{I}_C^0$). Likewise, some marginal change in $\boldsymbol{I}_R$ (i.e., r_t or r_k) is likely to have *less impact*, generally speaking, on the optimized value of Y at the point $\boldsymbol{I}_C^0$ relative to any $\boldsymbol{I}_C$ for which $\Delta EU_R(\boldsymbol{I}_C) < \Delta EU_R(\boldsymbol{I}_C^0)$. Put differently:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C^0) < \beta_{\boldsymbol{I}}(\boldsymbol{I}_C), \qquad given\ \Delta EU_R(\boldsymbol{I}_C^0) > \Delta EU_R(\boldsymbol{I}_C).$$

Hence, at low enough values of $\boldsymbol{I}_C$, marginal increases in the risks and costs to crime may, paradoxically, lead to *higher* values of $\beta_{\boldsymbol{I}}$, as the agent has more to "gain" from pursuing some set of behavioral alternatives (i.e., since he or she is able to avoid some quantity of losses by pursuing less of Y, the agent is likely to be *more responsive* to the benefits of Y). In addition, the agent is now forced to weigh the social and personals benefit to crime against some (albeit relatively minimal) quantity of losses, and thus show greater "discernment" over which criminal opportunities he or she will take advantage of (e.g., he or she will only opt for those which produce a high enough quantity of benefits, on the whole). Given the above, it seems unlikely that any precise statements can be made with respect to the *directionality* of the overall "moderating" influence of the risks and costs to crime on responsivity to rewards. However, predictions can certainly be made about the general *shape* of some geometric depiction of the value of $\beta_{\boldsymbol{I}}$ with respect to all $\boldsymbol{I}_C \in \mathbb{I}_C$. Namely, we can "restate" Hypotheses 1a through 1c (as outlined in Chapter 2) by the following:

Proposition 1*: For every $\boldsymbol{I}_C \in \mathbb{I}_C$ at which $\beta_{\boldsymbol{I}}$ achieves its global maximum value, it follows that $\beta_{\boldsymbol{I}}(\boldsymbol{I}_C^i) \geq \beta_{\boldsymbol{I}}(\boldsymbol{I}_C^j)$ for all $\boldsymbol{I}_C^i, \boldsymbol{I}_C^j \in \mathbb{I}_C$ for which:*

$$d(\boldsymbol{I}_C, \boldsymbol{I}_C^j) = d(\boldsymbol{I}_C, \boldsymbol{I}_C^i) + d(\boldsymbol{I}_C^i, \boldsymbol{I}_C^j).$$

PROOF: Note that d refers to the Euclidean distance between any two points contained in $\mathbb{I}_C$, as defined in the usual mannerism by:

$$d\left(\boldsymbol{I}_C^i, \boldsymbol{I}_C^j\right) = \sqrt{\sum_{l=1}^{L} \left(I_l^i - I_l^j\right)^2},$$

where I_l^i denotes the l-th component of $\boldsymbol{I}_C^i$ (I_l^j is defined analogously). The expression provided in Proposition 1 refers to the *triangle inequality*, for which "equality" holds if and only if $\boldsymbol{I}_C$ and $\boldsymbol{I}_C^j$ denote the endpoints of some line segment (contained in $\mathbb{I}_C$) which also contains $\boldsymbol{I}_C^i$ (see: Kolmogorov & Fomin, 1970, pp. 37-38). Put simply, if one can draw a straight line through the points $\boldsymbol{I}_C$, $\boldsymbol{I}_C^i$, and $\boldsymbol{I}_C^j$, and it is the case that $\boldsymbol{I}_C^j$ is at least as far away from $\boldsymbol{I}_C$ as $\boldsymbol{I}_C^i$, then the distance between $\boldsymbol{I}_C$ and $\boldsymbol{I}_C^j$ is the sum of the piecewise distances between the sub-segments $\left(\boldsymbol{I}_C, \boldsymbol{I}_C^i\right)$ and $\left(\boldsymbol{I}_C^i, \boldsymbol{I}_C^j\right)$. Thus, Proposition 1 simply states that if one were to "start" at some $\boldsymbol{I}_C$ for which:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C) = \sup\left(\beta_{\boldsymbol{I}}(\boldsymbol{I}_C \in \mathbb{I}_C)\right),$$

and began "moving away" from this point ($\boldsymbol{I}_C$) in a given direction, then, for every $\boldsymbol{I}_C^i$ and $\boldsymbol{I}_C^j$ encountered in this mannerism:

$$\beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^i\right) \geq \beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^j\right), \qquad given\ d\left(\boldsymbol{I}_C, \boldsymbol{I}_C^i\right) \leq d\left(\boldsymbol{I}_C, \boldsymbol{I}_C^j\right).$$

Assume for the purpose of contradiction there exists some $\boldsymbol{I}_C^i$ for which:

$$\beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^i\right) < \beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^j\right), \qquad given\ d\left(\boldsymbol{I}_C, \boldsymbol{I}_C^i\right) \leq d\left(\boldsymbol{I}_C, \boldsymbol{I}_C^j\right).$$

Such a phenomenon implies:

$$\Delta EU_R(\boldsymbol{I}_C) < \Delta EU_R\left(\boldsymbol{I}_C^i\right) > \Delta EU_R\left(\boldsymbol{I}_C^j\right), \qquad for\ all\ \boldsymbol{I}_R \in \mathbb{I}_R.$$

Note that for any (straight) line segment connecting each of the three points in the above expression, we can deduce:

$$d\left(I_l, I_l^j\right) \leq d\left(I_l, I_l^j\right), \qquad for\ all\ l,$$

where I_l denotes the l-th component of $\boldsymbol{I}_C$. If it is the case that I_l increases between $\boldsymbol{I}_C$ and $\boldsymbol{I}_C^j$, it follows that:

$$I_l < I_l^i < I_l^j,$$

and thus:

$$\Delta EU_l(I_l) < \Delta EU_l\left(I_l^i\right) > \Delta EU_l\left(I_l^j\right),$$

where ΔEU_l denotes some *marginal change* in ΔEU_R with respect to the l-th component of $\boldsymbol{I}_C$. Put simply, it must be generally the case that any "substitution" of some non-criminal course of action (N) in place of crime (Y) must have a larger impact on the agent's expected utility value (with respect to all $\boldsymbol{I}_R \in \mathbb{I}_R$) in moving from some marginal "cost" value of I_l to that of I_l^i, while simultaneously having a *smaller impact on EU* moving from I_l^i to I_l^j. Given all else is equal, such a phenomenon can only be observed if (assuming all other components of $\boldsymbol{I}_C$ are held constant):

$$0 \leq \iint\limits_{\mathbb{I}_R} |D_Y(I_l, \boldsymbol{I}_R)| < \iint\limits_{\mathbb{I}_R} \left|D_Y\left(I_l^j, \boldsymbol{I}_R\right)\right| < \iint\limits_{\mathbb{I}_R} \left|D_Y\left(I_l^i, \boldsymbol{I}_R\right)\right|.$$

Since $EU(Y > 0)$ is a decreasing function of I_l, it follows that:

$$\left|D_Y\left(I_l^j, \boldsymbol{I}_R\right)\right| < \left|D_Y\left(I_l^i, \boldsymbol{I}_R\right)\right|, \qquad for\ some\ fixed\ \boldsymbol{I}_R,$$

only if, for some sufficiently small $h > 0$:

$$EU\left(N + hy|I_l^j, \boldsymbol{I}_R\right) + K = EU(N) = EU\left(N + hy|I_l^i, \boldsymbol{I}_R\right) - J, \qquad given\ K < J.$$

If it is the case that:

$$EU(N) \leq EU\left(N + hy | I_l^j, \boldsymbol{I}_R\right) < EU\left(N + hy | I_l^i, \boldsymbol{I}_R\right),$$

then the value of K in the previous expression will always be less than J. If, however:

$$EU\left(N + hy | I_l^j, \boldsymbol{I}_R\right) < EU\left(N + hy | I_l^i, \boldsymbol{I}_R\right) \leq EU(N),$$

then $K > J$. Thus, for the sake of simplicity we can assume that:

$$EU(N) < EU\left(N + hy | I_l^j, \boldsymbol{I}_R\right) < EU\left(N + hy | I_l^i, \boldsymbol{I}_R\right), \qquad for\ all\ \boldsymbol{I}_R \in \mathbb{I}_R,$$

from which follows:

$$\left|D_Y\left(I_l^j, \boldsymbol{I}_R\right)\right| < \left|D_Y\left(I_l^i, \boldsymbol{I}_R\right)\right|, \qquad for\ all\ \boldsymbol{I}_R \in \mathbb{I}_R,$$

and thus:

$$\iint\limits_{\mathbb{I}_R} \left|D_Y\left(I_l^j, \boldsymbol{I}_R\right)\right| < \iint\limits_{\mathbb{I}_R} \left|D_Y\left(I_l^i, \boldsymbol{I}_R\right)\right|.$$

In other words, since by increasing the value of I_l we achieve a “closer” value of $EU(N + hy)$ to that of $EU(N)$ for all $\boldsymbol{I}_R$, it follows that ΔEU_l approaches zero with I_l and therefore $\Delta EU_l\left(I_l^i\right) > \Delta EU_l\left(I_l^j\right)$. However, since $I_l < I_l^i$, this also implies:

$$\Delta EU_l(I_l) > \Delta EU_l\left(I_l^i\right) > \Delta EU_l\left(I_l^j\right).$$

Note we can apply this same process to each other component of $\boldsymbol{I}_C$, in which case:

$$\Delta EU(\boldsymbol{I}_C) > \Delta EU\left(\boldsymbol{I}_C^i\right) > \Delta EU\left(\boldsymbol{I}_C^j\right),$$

which implies:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_C) < \beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^i\right) < \beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^j\right), \qquad given\ d\left(\boldsymbol{I}_C, \boldsymbol{I}_C^i\right) \leq d\left(\boldsymbol{I}_C, \boldsymbol{I}_C^j\right).$$

Hence, $\beta_{\boldsymbol{I}}$ can never achieve its global maximum value at any point $\boldsymbol{I}_C$ for which $\beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^i\right) < \beta_{\boldsymbol{I}}\left(\boldsymbol{I}_C^j\right)$ for some $\boldsymbol{I}_C^i$ which falls directly between $\boldsymbol{I}_C$ and $\boldsymbol{I}_C^j$. This contradicts our initial assumption, thus proving Proposition 1.■

Inhibitory Effect of the Risks and Costs as a Function of the Rewards:

In a similar fashion, we can also examine the overall influence that variation in any given (closed and bounded) subset of the risks and costs to crime (e.g., p_l and c_l) will exert on the optimized value of Y with respect to some unique set of rewards to Y, which I will denote by $\boldsymbol{I}_R$ (note that $\boldsymbol{I}_R$ is defined analogously to $\boldsymbol{I}_C$ in the previous section, with the exception that $\boldsymbol{I}_R$ refers to a fixed point in $\mathbb{R}_+^k$ instead of $\mathbb{R}_+^l$). As before, we can consider the following:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_R) = \iint\limits_{\mathbb{I}_C} \left(\frac{\partial Y}{\partial p_l} + \frac{\partial Y}{\partial c_l}\right) dxdy\,, \qquad for\ some\ fixed\ \boldsymbol{I}_R,$$

where p_l denotes the probability of arrest for engaging in Y, and c_l denotes the severity of social losses if captured (note that in this example I am once again treating p_l, c_l, and Y as scalar entities; i.e., each denotes a real number determined by, say, the respective means of p_l^j, c_l^j, and A_j for all j criminal activities contained in the vector $\vec{Y}$). The value of $\beta_{\boldsymbol{I}}$ is defined equivalently to that of the previous section, and $\mathbb{I}_C$ is defined as the cartesian product of the range of values for p_l and c_l. That is:

$$\mathbb{I}_C = \left[L_p, U_p\right] \times \left[L_c, U_c\right] \subset \mathbb{R}^2,$$

where L_p and U_p respectively denote the lower and upper bounds of p_l, and L_c and U_c respectively denote the same bounds for c_l. Note that since the (marginal) influence of both p_l and c_l on Y is presumed to be negative, it follows that $\beta_{\boldsymbol{I}}$ denotes a "stronger" inhibitory effect, with respect to some fixed $\boldsymbol{I}_R$, if and only if:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_R) < \beta_{\boldsymbol{I}}(\boldsymbol{I}'_R), \qquad given\ \boldsymbol{I}_R \neq \boldsymbol{I}'_R.$$

In other words, the greater the overall change in the optimized value of Y, with respect to p_l and c_l, the *lower the value of* $\boldsymbol{\beta_I}$. The question now becomes: how might the value of $\beta_{\boldsymbol{I}}$ change with respect to $\boldsymbol{I}_R$? Consider the following sequence:

$$\boldsymbol{I}_{R1} < \boldsymbol{I}_{R2} < \cdots < \boldsymbol{I}_{Rn}.$$

Since the marginal utility of each (criminal) activity j contained in the vector $\vec{Y}$ (note that each component of $\vec{Y}$ is simply the value of A_j for all $j \in \vec{Y}$) is an increasing function of the rewards vector, it follows that, given all else is equal:

$$MU_Y(\boldsymbol{I}_{R1}) \leq MU_Y(\boldsymbol{I}_{R2}) \leq \cdots \leq MU_Y(\boldsymbol{I}_{Rn}),$$

where MU_Y denotes a vector, the components of which are equivalent to the change in marginal utility with respect to each j contained in $\vec{Y}$ (more specifically, the j-th component of MU_Y is simply $MU_e(A_j)$). Denote the value of $MU_e(A_i^*|\boldsymbol{I}_{R1})$, for some (optimized) $A_i^* > 0$, by λ (as previously discussed, we can select any activity i which is assigned an optimized "consumption rate" greater than zero to determine the value of λ, since $MU_e(A_i^*) = MU_e(A_j^*)$ for all $A_i^*, A_k^* > 0$ and $i \neq j$). That is, λ denotes the optimized marginal utility value of any non-zero activity (or set of activities) the agent may feasibly pursue, with respect to some fixed set of rewards for $\vec{Y}$ (specifically, $\boldsymbol{I}_{R1}$). Assume the following:

$$MU_Y(p_l, c_l|\boldsymbol{I}_{R1}) < \mathbf{1}_j * \lambda,$$

where $MU_Y(p_l, c_l|\boldsymbol{I}_{R1})$ denotes the (vector) value of MU_Y as a function of p_l and c_l (recall the values of p_l and c_l respectively denote the *average* probability of arrest and social losses for all $j \in \vec{Y}$), and $\mathbf{1}_j$ denotes a 1-vector of length j.[64] Likewise, the above

[64] More precisely, $\mathbf{1}_j$ is a column vector with j components, all of which are equal to 1 (see: Shilov, 1971).

statement can be read as: "$MU_e(A_j^*) < \lambda$ for all $j \in \vec{Y}$ and $p_l, c_l \in \mathbb{I}_C$, given $\boldsymbol{I}_R = \boldsymbol{I}_{R1}$." It follows that:

$$\left\|\vec{Y}^*\right\| = 0, \qquad for\ all\ p_l, c_l \in \mathbb{I}_C,$$

where $\left\|\vec{Y}^*\right\|$ denotes the Euclidean length of any optimized value of $\vec{Y}$. Since the marginal utility of each $j \in \vec{Y}$ is less than λ for all possible values of p_l and c_l, it follows that the optimized value of each component of $\vec{Y}$ is *equal to zero* at every point within the domain of $\mathbb{I}_C$ (given all else equal). As a consequence:

$$\frac{\partial Y}{\partial p_l} = \frac{\partial Y}{\partial c_l} = 0, \qquad for\ all\ p_l, c_l \in \mathbb{I}_C,$$

from which follows:

$$\iint\limits_{\mathbb{I}_C} \left(\frac{\partial Y}{\partial p_l} + \frac{\partial Y}{\partial c_l}\right) dxdy = 0, \qquad given\ \boldsymbol{I}_R = \boldsymbol{I}_{R1},$$

and thus:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_{R1}) = 0.$$

Such an outcome should seem almost obvious (both mathematically and intuitively), given the discussion leading up to this point. After all, if the benefits to each $j \in \vec{Y}$ fail to exceed those achieved through any proportional level of involvement in legal alternatives (i.e., the agent achieves a greater quantity of benefits from legal actions in exchange for the same amount of "effort" he or she would need to expend on some level of involvement in criminal behaviors), then, by definition, the agent is best off avoiding Y entirely *regardless of the probability of arrest or social losses associated with* Y (note I am also assuming that each of the risks and costs to legal actions are no greater than those produced by illegal acts). Consider now the following:

$$MU_Y(p_l, c_l|\boldsymbol{I}_{Rk}) > \boldsymbol{I}_j^0 * \lambda, \qquad \boldsymbol{I}_{Rk} < \boldsymbol{I}_{Rn},$$

where $\boldsymbol{I}_j^0$ denotes a column vector with j components, such that:

$$j = \begin{cases} 1, & MU_e(A_j^*) > \lambda \\ 0, & MU_e(A_j^*) \leq \lambda \end{cases}$$

That is, if the change in marginal utility for illegal act j is greater than λ (which, recall, is determined by $\boldsymbol{I}_{R1}$), then $\boldsymbol{I}_j^0$ is assigned a value of 1 for component j. If the change in marginal utility of j is less than or equal to λ, then $\boldsymbol{I}_j^0$ is assigned a j-th component of 0. Since the marginal utility of any $j \in \vec{Y}$ can only decrease with p_l and c_l, it follows that:

$$\frac{\partial Y}{\partial p_l} \leq 0 \; and \; \frac{\partial Y}{\partial c_l} \leq 0, \qquad for\ all\ p_l, c_l \in \mathbb{I}_C,$$

and therefore:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_{Rk}) \leq 0.$$

Furthermore, since:

$$MU_Y(\boldsymbol{I}_{Rk}) \leq MU_Y(\boldsymbol{I}_{Rk+1}) \leq \cdots \leq MU_Y(\boldsymbol{I}_{Rn}),$$

it follows that:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_{Rj}) \leq 0, \qquad for\ all\ k < j < n.$$

Put differently, if the value of $\beta_{\boldsymbol{I}}$ is potentially less than zero at the k-th iteration of the sequence $\boldsymbol{I}_{Rn}$, it follows that $\beta_{\boldsymbol{I}}$ may also be less than or equal to zero *for all iterations of this sequence which fall between* $\boldsymbol{I}_{Rk}$ *and* $\boldsymbol{I}_{Rn}$. As such, a potential implication of the above expression is that individuals should, generally speaking, be *less responsive* to the risks and costs to crime at lower reward values for criminal actions. Such a notion follows from the logic that individuals who have relatively little to gain from offending may be "undeterrable" not because they do not care about the risks and costs to crime, but rather because they are likely to *avoid crime altogether* regardless of any variation in perceptions of arrest risk and social costs.

Before moving on to the next section, we also need to consider the possibility of individuals becoming less responsive to the risks and costs at *higher* values of $\boldsymbol{I}_R$. Let's assume there exists no "cap" to the sequence $\boldsymbol{I}_{Rn}$, such that:

$$\boldsymbol{I}_{R1} < \boldsymbol{I}_{R2} < \cdots < \boldsymbol{I}_{Rn} < \cdots.$$

Additionally, assume there exists some $\boldsymbol{I}_{Rk}$ for which:

$$\boldsymbol{I}_{Rk} > \boldsymbol{I}_R^0,$$

where $\boldsymbol{I}_R^0$ denotes a (fixed) set of rewards for all legal alternatives to crime (note that I also assume every benefit type not addressed by illegal acts is equal to zero in $\boldsymbol{I}_R^0$). Here, the above expression states that every (non-zero) benefit type offered by illegal acts exceeds the values of those addressable by legal behaviors. Since $\boldsymbol{I}_{Rn}$ is an infinitely increasing sequence, it follows that:

$$\boldsymbol{I}_R^0 < \boldsymbol{I}_{Rk} < \boldsymbol{I}_{Rk+1} < \cdots,$$

Since the value of MU_Y increases with $\boldsymbol{I}_{Rn}$, it follows that any "proportional exchange" of some level of involvement in criminal acts with that of legal alternatives will result in an overall *greater loss in expected utility* as $\boldsymbol{I}_{Rn}$ approaches infinity (i.e., the value of $\Delta EU_C(\boldsymbol{I}_R)$, defined analogously to that of $\Delta EU_R(\boldsymbol{I}_C)$ in the previous section, increases with $\boldsymbol{I}_R$). Likewise, any overall decrease in the marginal utility of illegal actions imposed by p_l and c_l will be "offset," on some level, by the rewards to crime (particularly, any social and personal benefits to criminality). In fact, as $\boldsymbol{I}_{Rn}$ continues to increase, the greater the overall degree to which the "disutility" produced by p_l and c_l will be sufficiently "compensated" by the rewards to crime. As such, it may also be the case that, for high enough reward values, individuals will typically display *less sensitivity* to the risks and costs to crime. Hence, it (once again) cannot be determined whether the overall influence of some empirical measures of p_l and c_l will be a monotonic function of $\boldsymbol{I}_R$, as there may exist some $\boldsymbol{I}_{Rk}$ for which:

$$\beta_{\boldsymbol{I}}(\boldsymbol{I}_{R1}) > \cdots > \beta_{\boldsymbol{I}}(\boldsymbol{I}_{Rk-1}) > \beta_{\boldsymbol{I}}(\boldsymbol{I}_{Rk}) < \beta_{\boldsymbol{I}}(\boldsymbol{I}_{Rk+1}) < \cdots.$$

As such, we should at least expect the value of $\beta_{\boldsymbol{I}}$ to achieve no more than one "local minimum" at some point in either the interior or on the boundary of some rewards space, denoted by $\mathbb{I}_R$ (i.e., there should not exist multiple "dips" in any surface depicting the value of $\beta_{\boldsymbol{I}}$ with respect to $\mathbb{I}_R$; Boyd & Vandenberghe, 2004). Such a notion can be summarized by the following restatement of Hypotheses 2a through 2c:

Proposition 2: *For every $\boldsymbol{I}_R \in \mathbb{I}_R$ at which $\beta_{\boldsymbol{I}}$ achieves its global minimum value, it follows that $\beta_{\boldsymbol{I}}(\boldsymbol{I}_R^i) \leq \beta_{\boldsymbol{I}}(\boldsymbol{I}_R^j)$ for all $\boldsymbol{I}_R^i, \boldsymbol{I}_R^j \in \mathbb{I}_R$ for which:*

$$d(\boldsymbol{I}_R, \boldsymbol{I}_R^j) = d(\boldsymbol{I}_R, \boldsymbol{I}_R^i) + d(\boldsymbol{I}_R^i, \boldsymbol{I}_R^j).$$

PROOF: Apply the same process as the proof of Proposition 1, except that the value of $\beta_{\boldsymbol{I}}$ is now assumed to increase with respect to some $\boldsymbol{I}_R^i$ and $\boldsymbol{I}_R^j$.■

Moderating Influence of Multiple Types of Rewards:

Finally, let's examine the potential interactive effect between multiple types of rewards to criminal behaviors. Namely, should we expect the overall influence of one reward type (e.g., social benefits) to vary across different values of a different reward type (e.g., thrills or excitement)? Consider first a situation wherein some increase in a given reward type tends to weaken a person's responsiveness toward other types of reward (i.e., his or her level of involvement in criminal activities tends to be less "swayed" by changes in alternative benefit types as his or her perceptions of some other type increases). Say we identify a sequence of values for, say, the social benefits to crime such that:

$$r_{k1} < r_{k2} < \cdots < r_{kn},$$

Note that the above sequence is defined similarly to that which was used as an example in the previous two sections. One scenario where such a sequence may lead to a general "insensitivity" to other reward types could be:

$$\frac{\partial^2 EU}{\partial A_j \partial A_i}(r_{k1}) = \frac{\partial^2 EU}{\partial A_j \partial A_i}(r_{kn}) = \cdots = \frac{\partial^2 EU}{\partial A_j \partial A_i}(r_{kn}), \qquad for\ all\ (legal)\ i \neq j.$$

That is, if criminal acts are the only "means" through the agent can achieve social status, then any increase in the (perceived) social benefits to crime will have no influence on the substitutability of any j contained in Y with any legal activity i. In fact, since this particular reward type is solely addressed by illegal acts, the agent may actually find it *more difficult to forego* some amount of Y in favor of legal alternatives at higher values of r_k. For instance, assume that $e^T A = E$ for all $A \in \mathcal{S}$ (including all prior and posterior solution sets). Since the second derivative of EU with respect to any j in Y can only decrease with r_k, it follows:

$$\frac{\partial^2 EU}{\partial {A_j}^2}(r_{k1}) \geq \frac{\partial^2 EU}{\partial {A_j}^2}(r_{k2}) \geq \cdots \geq \frac{\partial^2 EU}{\partial {A_j}^2}(r_{kn}), \qquad for\ some\ A \in \mathcal{S}',$$

where, in this instance, $A \in \mathcal{S}'$ denotes any fixed selection of A contained in the solution set determined by r_{k1}. Note that $\frac{\partial^2 EU}{\partial {A_j}^2}$ is presumed to decrease with r_{kn} while the cross-partial derivative with respect to i and j remains constant. Because of this, any "proportional" exchange between i and j is likely to have a *stronger impact on the change in expected utility* at higher values of r_{kn}. To illustrate this notion, consider the following:

$$\int_0^1 M(\theta) - EU\left(\theta * A_E^{j=0} + [1-\theta] * A_E^{i=0}\right) d\theta\,, \qquad for\ some\ A \in \mathcal{S}',$$

where $\int_0^1 (*)\, d\theta$ denotes the definite integral with respect to $\theta \in [0,1]$, and $A_E^{j=0}$ denotes a vector equal to some $A \in \mathcal{S}'$ in all components except for j and i, the former of which is equal to zero (i.e., $A_j = 0$), and the latter equals the unique value of A_i which preserves the value of $e^T A$ (that is, if we lower A_j by setting it equal to zero, then we are also lowering the value of $e^T A$, and hence we must also proportionally increase the value of A_i in order to ensure $e^T A = E$). The vector $A_E^{i=0}$ denotes the same for the i-th component (i.e., $A_i = 0$ and A_j is equal to that which preserves the value of $e^T A$), and $M(\theta)$ denotes a function which assigns the "maximum" value of EU to all $\theta \in [0,1]$. That is:

$$M = \sup\left(EU\left(\theta * A_E^{j=0} + [1-\theta] * A_E^{i=0}\right)\right), \qquad for\ all\ \theta \in [0,1].$$

Since the value of M is greater than or equal to all possible values of EU, given θ, it follows that:

$$\int_0^1 M(\theta) - EU\left(\theta * A_E^{j=0} + [1-\theta] * A_E^{i=0}\right) d\theta \geq 0.$$

For the sake of brevity, I will omit the vectors from the $EU(*)$ term from this point onward, and will instead write the above expression as:

$$\int_0^1 M(\theta) - EU(\theta)\, d\theta.$$

Note we have already covered one instance where the value of this integral is exactly equal to zero. Namely, if it is the case that $R^j = R^i$, $C^j = C^i$, and $e_j = e_i$ (i.e., the agent is completely indifferent to either i or j), then, for any $A \in \mathcal{S}$:

$$EU(\theta) = K = M, \qquad for\ all\ \theta \in [0,1],$$

That is, if the value of EU remains constant for all θ, then, by definition, $EU = M$. Hence, in this particular instance:

$$\int_0^1 M(\theta) - EU(\theta)\, d\theta = 0, \qquad given\ R^j = R^i;\ C^j = C^i; and\ e_j = e_i.$$

Now assume we alter the value of any reward or cost vector to either i or j and, in doing so, influence the marginal utility to one or both actions (including their respective second derivatives as well). In this scenario, the value of EU would no longer be constant for all θ, which implies the following:

$$\int_0^1 M(\theta) - EU(\theta)\, d\theta > 0, \qquad given\ A_j^* > 0\ or\ A_i^* > 0,$$

Therefore, the value of $\int_0^1 (*)\, d\theta$ can only *increase* as the rewards and costs to i and j *diverge from one another*. Consequently, for the previously established sequence of r_{kn}:

$$I_\theta(r_{k1}) < I_\theta(r_{k2}) < \cdots < I_\theta(r_{kn}), \qquad for\ some\ A \in \mathcal{S}',$$

where $I_\theta(r_{kn})$ denotes the value of $\int_0^1 (*)\, d\theta$ with respect to the n-th iteration of r_{kn}, given any fixed A which belongs to $\mathcal{S}'$. Note that we know this to be the case since any change in r_k does not influence the cross-partial derivative of i and j, and thus any proportional exchange between i and j can only have a *greater impact on the agent's expected utility* as r_k increases. In particular, any reduction in j will come at a greater utility loss at higher values of r_k, and thus the agent will likely find it more difficult to "swap out" any amount of j for some (legal) alternative i. Furthermore, because of this increased "reluctance" to pursue some amount of i in place of j, the agent will also likely be *less responsive* to changes in any other reward component to j. Put differently, we can think of the value of I_θ as a measure of *inelasticity*, such that the higher the value of I_θ for some j and i, the less overall impact changes in the rewards to j will have on the value of A_j^* (Kreps, 2013). Of course, the same also applies to any other activity j contained in Y, as well as other potential legal alternative to Y. Nonetheless, it should be the case that, more often than not, when criminal actions satiate desires not addressed by legal alternatives for the agent, the more he or she will come to view other reward types to crime as "secondary" benefits. Such benefits may, perhaps, justify higher involvement in Y, but should generally tend to have *less impact* on said involvement relative to any situation where the rewards to crime are, on the whole, addressable in some capacity by strictly legal behaviors.

With this in mind, let's now consider a situation where changes in some reward type could potentially *increase* the impact that other types might have on the agent's involvement in Y. Since the value of I_θ is (partially) a function of the degree of "overlap" between R^j and R^i (or the lack thereof), it follows that, given $C^j = C^i$ and $e_j = e_i$:

$$\lim_{R^j \to R^i} \int_0^1 M(\theta) - EU(\theta)\, d\theta = 0.$$

That is, as each reward type to j approaches those offered by i, the value of I_θ approaches zero, assuming all else is equal. Suppose we examine the same sequence of r_{kn} as before, except in this instance we find:

$$I_\theta(r_{k1}) > I_\theta(r_{k2}) > \cdots > I_\theta(r_{kn}), \qquad for\ some\ A \in \mathcal{S}'.$$

Here, the overall change in EU for any proportional exchange between any i and j is actually a *decreasing* function of r_{kn}. The question now becomes: is such a phenomenon possible, given our baseline assumptions established up to this point? Suppose it is the case that:

$$r_k^i > r_{kn}, \qquad for\ all\ r_{kn}.$$

In this scenario, the k-th reward type (i.e., social benefits) addressed by some (legal) i is actually *greater* than that which is associated with any member of the sequence r_{kn} (recall that r_{kn} refers to the overall value of the social rewards offered by each j in Y). Consequently:

$$|r_k^i - r_{k1}| > |r_k^i - r_{k2}| > \cdots > |r_k^i - r_{kn}|,$$

where $|*|$ denotes the absolute value of the difference between the n-th member of r_{kn} and r_k^i. Given all else is equal, the above expression implies:

$$d(R^i, R_1^j) > d(R^i, R_2^j) > \cdots > d(R^i, R_n^j),$$

where d (once again) refers to the Euclidean distance between R^i and R^j at the n-th term of the sequence r_{kn}. This, of course, follows from the distance between r_k^i and r_k^j decreasing with r_{kn} (again, given all else is equal), and therefore:

$$\inf(I_\theta) = I_\theta(\sup(r_{kn})).$$

That is, the smallest value of I_θ is achieved at the highest element of the sequence r_{kn}, wherein the respective reward vectors to i and j are "closest" to one another in $\mathbb{R}_+^k$. In other words, we should anticipate the agent to be *more willing to substitute* i *for any* j (or vice versa) in the event he or she comes to believe criminal actions (i.e., elements of Y) offer similar levels of satiation for a particular reward type to some legal alternative. Note the same can also be applied to the consideration of multiple alternatives (both legal and illegal), meaning it should generally be the case that criminal actions become *more elastic* as r_k^j approaches the value of r_k^i to any number of $i \neq j$. Such an increase in elasticity should result in the agent being *more responsive* to changes in other reward types, as he or she has less to "lose" by foregoing strict conformity in favor of some amount of criminal involvement (i.e., since the value of EU is less "swayed" with any tradeoff between i and j, any reduction in i in favor of j will produce less of an opportunity loss for the agent, and thus any increase in alternative reward types to j should have a *larger impact* on Y).

Of course, this is merely one example of a potential "positive" interactive effect between multiple reward types to some nonlegal behavior (or possibly set of illegal actions). In general, we can expect such an effect to occur any time some increase in a particular reward type to criminal activities results in those activities becoming "easier to substitute" in place of legal alternatives. It may even be the case that, for some measure of reward type k (e.g., $r_k \in [L_k, U_k]$), the overall influence of some measure of reward type t (e.g., $r_t \in [L_t, U_t]$) on Y will either increase *or decrease* at different values of r_k. That is, some marginal increase in r_k could lead to a stronger influence of r_t on Y for "low enough" values of r_k, while marginal increases at higher values of r_k may elicit a weaker overall influence of r_t on Y (or vice versa). Such a scenario may occur in the event:

$$r_{k1} < r_{k2} < \cdots < r_{km} < r_k^i < r_{km+1} < \cdots < r_{kn},$$

where $r_{kn} \in [L_k, U_k]$ for all members of r_{kn}. More specifically, the "elasticity" of criminal acts may be higher for all r_k closest to r_k^i (lower otherwise). There, of course, may be other reasons for such a "variable" influence of r_t on Y with respect to r_k, and vice versa (e.g., an increase in r_k may lead to an optimized value of Y greater than zero, and thus marginal changes in r_t are more likely to have an observable impact on a person's overall involvement in criminal activities). As with the previous two sections, we can summarize this notion via a restatement of Hypotheses 3a through 3c:

Proposition 3*: For every $r_k \in \mathbb{I}_k$ at which β_t achieves its global maximum value (for some $k \neq t$), it follows that $\beta_t\left(r_k^i\right) \geq \beta_t\left(r_k^j\right)$ for all $r_k^i, r_k^j \in \mathbb{I}_k$ for which:*

$$d\left(r_k, r_k^j\right) = d\left(r_k, r_k^i\right) + d\left(r_k^i, r_k^j\right).$$

PROOF: Note that r_k denotes the value of a particular reward type (e.g., social rewards) to some set of criminal activities Y, $\mathbb{I}_k$ denotes the range of r_k (for some empirical measure), and β_t denotes the overall criminogenic influence of a different reward type (e.g., intrinsic rewards) to crime for some $r_k \in \mathbb{I}_k$. The proof of this proposition can be carried out in a similar fashion to the approach taken in the proof of Proposition 1.■

This concludes the discussion of the more technical aspects of this dissertation's conceptual foundation, as well as overall predictions related to the interdependent influences of perceived incentives on criminal behavior.

APPENDIX B: SUPPLEMENTARY MODELS

The contents of this appendix provide an overview of a series of models which aim to achieve a similar set of goals to that of each of the generalized additive models (as provided in the main body of results) but do so in a considerably "simpler" way. Specifically, I divided the entire sample of participants within the Pathways to Desistance Study into several "sub-conditions" for each moderation variable, as established in each primary set of hypotheses presented in Chapter 2, and then estimated the value of β for each primary variable of interest using the model outlined in Equation 4 for each respective sub-group. To examine Hypotheses 1a through 1c, I restricted sample participants to one of 9 different conditions determined by the following pair-wise combinations of:

$$Low\ Certainty = \ P_{it} < 4.29$$

$$Mid\ Certainty = \ 4.29 \leq P_{it} \leq 7.00$$

$$High\ Certainty = \ P_{it} > 7.00$$

with respect to:

$$Low\ Social\ Costs = \ C_{it} < 2.80$$

$$Mid\ Social\ Costs = \ 2.80 \leq C_{it} \leq 3.50$$

$$High\ Social\ Costs = \ C_{it} > 3.50$$

where P_{it} denotes the value of perceived arrest risk for participant i at time t, and C_{it} similarly denotes the value of the perceived social losses. Each "cut-off" point defined within each of the above expressions corresponds to the 33rd and 66th percentiles for both the arrest risk and social losses variables (e.g., 4.29 is the 33rd percentile for the arrest risk variable and 7.00 is its 66th percentile). The results of this analysis mostly seem to mimic those seen in Figures 2.1 through 2.6, although a handful of odd "jumps" can be seen (e.g., the social rewards exhibit a fairly weak influence for those with middling perceptions of both the risks and costs to crime, yet also a considerably *stronger* influence for those in every other group aside from that of persons who reported both a high probability of arrest and high social losses). The details of this analysis are provided in Table B1.

A similar set of tests were carried out for Hypotheses 2a through 2c using the following:

$$Low\ Social\ Benefits = S_{it} < 1.93$$

$$Mid\ Social\ Benefits = 1.93 \leq S_{it} \leq 2.07$$

$$High\ Social\ Benefits = S_{it} > 2.07$$

as well as:

$$Low\ Personal\ Benefits = T_{it} = 0$$

$$Mid\ Personal\ Benefits = 0 < T_{it} \leq 2$$

$$High\ Personal\ Benefits = T_{it} > 2$$

Once again, each cutoff point was determined by the 33rd and 66th percentiles for each of the perceived reward variables. The model outlined in Equation 4 was then estimated with respect to each pairwise combination of the above conditions (9 groups in total), the results of which are provided in Table B2. On the whole, relatively minor differences are observed with respect to each of the strengths of the coefficients, although a generally "decreasing" pattern is observed at higher values of reward perceptions (i.e., the overall influence of the perceived risk and cost variables seemed to strengthen at higher social and personal reward values, although the difference is fairly minimal on the whole). Finally, Hypotheses 3a through 3b were re-examined using the above "grouping" criteria.

The results of this analysis are presented in Table B3, and overall seem to mimic that which was observed in Figures 4.1 and 4.2.

Table B1: Supplementary Tests of Hypotheses 1a through 1c.

	Low Certainty	*Mid Certainty*	*High Certainty*
Low Social Costs	$\beta_1 = .392^{***}$ $\beta_2 = .114^{***}$ $N = 1{,}808$	$\beta_1 = .226^{*}$ $\beta_2 = .112^{***}$ $N = 1{,}147$	$\beta_1 = .298^{**}$ $\beta_2 = .060^{**}$ $N = 1{,}092$
Mid Social Costs	$\beta_1 = .255^{**}$ $\beta_2 = .080^{***}$ $N = 1{,}356$	$\beta_1 = .040$ $\beta_2 = .129^{***}$ $N = 1{,}255$	$\beta_1 = .351^{***}$ $\beta_2 = .104^{**}$ $N = 1{,}096$
High Social Costs	$\beta_1 = .282^{**}$ $\beta_2 = .121^{***}$ $N = 901$	$\beta_1 = .221^{*}$ $\beta_2 = .034$ $N = 1{,}170$	$\beta_1 = .030$ $\beta_2 = .078^{***}$ $N = 1{,}704$

NOTE: β_1 denotes the estimated coefficient for the perceived social benefits, while β_2 denotes the same for the perceived personal benefits.
*p < .05; **p < .01; ***p < .001

Table B2: Supplementary Tests of Hypotheses 2a through 2c.

	Low Social	*Mid Social*	*High Social*
Low Personal	$\beta_3 = -.049^{***}$ $\beta_4 = -.031$ $N = 1{,}973$	$\beta_3 = -.029^{\dagger}$ $\beta_4 = -.078$ $N = 1{,}547$	$\beta_3 = -.036^{\dagger}$ $\beta_4 = -.070$ $N = 776$
Mid Personal	$\beta_3 = -.051^{***}$ $\beta_4 = -.031$ $N = 3{,}068$	$\beta_3 = -.049^{***}$ $\beta_4 = -.065^{\dagger}$ $N = 2{,}670$	$\beta_3 = -.042^{***}$ $\beta_4 = -.063^{\dagger}$ $N = 1{,}860$
High Personal	$\beta_3 = -.050^{**}$ $\beta_4 = -.111^{*}$ $N = 841$	$\beta_3 = -.068^{***}$ $\beta_4 = -.118^{*}$ $N = 1{,}150$	$\beta_3 = -.058^{***}$ $\beta_4 = -.089^{**}$ $N = 1{,}940$

NOTE: β_3 denotes the estimated coefficient for perceived arrest risk, while β_4 denotes the same for the perceived social costs.
$^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Table B3: Supplementary Tests of Hypotheses 3a through 3c.

	Social Rewards Effect	*Personal Rewards Effect*
Low Alternative	$\beta_1 = .303^{***}$ $N = 4{,}296$	$\beta_2 = .089^{***}$ $N = 3{,}909$
Mid Alternative	$\beta_1 = .210^{***}$ $N = 7{,}598$	$\beta_2 = .108^{***}$ $N = 3{,}820$
High Alternative	$\beta_1 = .239^{***}$ $N = 3{,}943$	$\beta_2 = .079^{***}$ $N = 3{,}800$

NOTE: β_1 denotes the estimated coefficient for the perceived social benefits, while β_2 denotes the same for the perceived personal benefits.

$^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$