

**Article**

# Negotiating Justice: Examining Restorative Justice Through the Coase Theorem

Andrei Comlosan-Pop<sup>1</sup>

<sup>1</sup> Student (BA Economics, 2021-25), Harvard University, USA; acomlosanpop@college.harvard.edu

Submission Date: 2nd February 2023; Acceptance Date: 11th July 2023; Publication Date: 2<sup>nd</sup> October 2023

**How to cite**

Comlosan-Pop, A. (2023). Negotiating Justice: Examining Restorative Justice Through the Coase Theorem. *UCL Journal of Economics*, vol. 2, pp. 111-119. DOI: <https://doi.org/10.14324/111.444.2755-0877.1595>

**Peer review**

This article has been peer-reviewed through the journal's standard double-blind peer review, where both the reviewers and authors are anonymised during review

**Copyright**

2023, Andrei Comlosan-Pop. This is an open-access article distributed under the terms of the Creative Commons Attribution Licence (CC BY) 4.0 <https://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited • DOI: <https://doi.org/10.14324/111.444.2755-0877.1595>

**Open access**

*UCL Journal of Economics* is a peer-reviewed open-access journal

---

**Abstract**

Restorative justice differs markedly from the traditional law-and-economics model of crime and punishment, as it eschews deterrence-based punishment and grants a far greater role to victims of crime. Existing economic literature on restorative justice is limited, and this increasingly popular paradigm warrants further law-and-economics engagement. This paper adopts a novel approach by using the Coase Theorem to analyze restorative justice negotiations as an alternative that victims and offenders may choose to traditional criminal procedures. I conclude that, with the proper enabling conditions, restorative justice can lead to higher victim and offender welfare compared to traditional criminal procedures by lowering the transaction costs to bargaining and by granting the victim a clear legal entitlement. However, differences in existing programs, behavioral economics principles, and concerns from beyond the economic literature suggest that maximally efficient outcomes may not occur in real life.

**Keywords:** Restorative Justice; Coase Theorem; Welfare; Behavioural Economics

---

## 1. Introduction

For individuals disillusioned with a traditional punitive model of criminal justice, restorative justice represents an enticing alternative. Restorative justice is “an approach to achieving justice that involves” having those implicated in an offense “collectively identify and address harms, needs, and obligations, in order to heal and put things as right as possible” (Zehr 2002, qtd. in Lanni 2021, 641-42). This process is sometimes described in terms of four Rs: ‘repair, restore, reconcile, and reintegrate’ (Menkel-Meadow 2007, 10.2). Pioneered by figures such as John Braithwaite, the ‘modern incarnation’ of restorative justice arose in the 1970s as a reaction to the perceived failures of the justice system to ‘[deter] crime’ and ‘successfully rehabilitate offenders’ (Ibid., 10.3).

In the criminal system, a restorative vision of justice ‘typically’ entails interaction between the victim, offender, and potentially other members of the community (Lanni 2021, 637), with facilitation from professionals or volunteers. For offenders, participating in restorative justice can divert a case away from the traditional trial process, serve as a form of sentencing in and of itself, or help reduce a sentence (Ibid.). Restorative justice also offers victims the opportunity to receive an offender’s acknowledgement of wrongdoing and the process may entail specific steps of restitution for the offender to take (643, 646). Offender participation in a restorative justice process is ‘usually not mandatory’ (654), and victims generally exercise a form of ‘veto power’ over its employment (Fatas and Restrepo-Plaza 2022, 2; see also Lanni 2021, 668). Restorative justice is used around the world as a diversion from or supplement to traditional criminal justice approaches, including in juvenile systems in Australia (Shem-Tov, Raphael, and Skog 2021, 7) and Indonesia (Subekti, Hartiwiningsih, and Handayani 2021, 56), and in both juvenile and non-juvenile systems in New Zealand (Shem-Tov, Raphael, and Skog 2021, 7; New Zealand Sentencing Act 2002 (2022); Lanni 2021, 650), Colombia (Fatas and Restrepo-Plaza 2022, 2), and South Africa (Department of Justice and Constitutional Development 2011, 7-8).

As a significant departure from the traditional, rational-choice law-and-economics prescriptions for punishment, restorative justice warrants an economic study. In the former, criminality is a cost-benefit analysis between expected gain and sanction, and the state crafts sanctions to maximize the benefits of deterrence against the costs of law enforcement (see Polinsky and Shavell 1984, among others). In contrast, restorative justice sees crime as a relational problem requiring interpersonal solutions and therefore facilitates reconciliation between the affected parties (see Zehr 2002, qtd. above in Lanni 2021). Another significant disparity between the two frameworks is the role of victims. The rational-choice law-and-economics model does not see the victim as an agent in the criminal justice process. Prosecutors and judges may, by custom or statute, consult victims about plea bargaining or sentences (formalized, in the case of the latter, through victim impact statements), but victims nevertheless lack a clear entitlement to inform the process. Restorative justice affords far greater agency to the victim and sees victim participation as integral to the justice process (Lanni 2021, 642-43). An economic exploration of restorative justice allows us to understand how these differences impact societal welfare.

In this paper, I use the Coase Theorem to analyze whether allowing victims and offenders to choose restorative justice as an alternative to traditional criminal prosecution promotes economic efficiency. I define the Coase Theorem as stating that: under conditions of economic rationality where parties are ready and willing to bargain and do in fact bargain and reach an agreement, entitlements will always move toward their highest-value use, regardless of their initial allocation, provided that there are no (or minimal) transaction costs and property rights are clearly defined (Coase 1960)<sup>1</sup>. Ronald Coase developed the analysis undergirding the Coase Theorem in the context of legal liability, demonstrating the economic irrelevance of such rules in situations of free, rational bargaining over entitlements. The

---

<sup>1</sup> The underlying analysis was first expounded in Coase’s 1960 paper, ‘The Problem of Social Cost,’ although George Stigler was the first to summarize Coase’s reasoning and call it a ‘Theorem’ (Carden 2020, 45).

Coase Theorem is thus a useful framework for conceptualizing restorative justice as a process of negotiation and as one which can partly internalize the externalities of crime. I also define efficiency to mean the maximization of the private welfare of the parties directly involved in the negotiation—victims and offenders. While I briefly comment on the implications of restorative justice for deterrence and recidivism, a thorough analysis of these and other aspects of social welfare exceeds the space available in this paper.

Literature applying an explicit law-and-economics lens to restorative justice is, so far, limited. Bush (2003), who uses this lens to emphasize the shortcomings of the restorative justice paradigm, offers the most thorough treatment. I contribute to the literature by analyzing restorative justice through the Coase Theorem, demonstrating the potential advantages of restorative justice from an economic perspective. To my knowledge, the Coase Theorem has not been used thus far in the English literature to conceptualize restorative justice.<sup>2</sup>

I conclude that, given the proper enabling conditions, the option to resolve a case through restorative justice instead of traditional criminal prosecution leads to more negotiated outcomes between victims and offenders and to outcomes that are better for the parties, thus promoting economic efficiency. Restorative justice achieves this result by lowering the transaction costs to bargaining and by helping internalize the externalities of crime borne by victims. However, transaction costs likely remain even under formal restorative justice programs, and preconditions to negotiation may prevent parties from ever choosing to bargain. Insights from behavioral economics and non-economic literature suggest further reasons why restorative justice negotiations may deviate from Coasean ideals. Consequently, in practice, restorative justice programs may not create maximally efficient outcomes for victims and offenders.

## **2. Literature Review**

The existing law-and-economics literature on restorative justice is limited. Bush's (2003) major treatment gives restorative justice a 'mixed review' (441). In one major critique, he contends that victims should not be satisfied by restorative efforts (459), because 'compensation cannot occur on any significant level in criminal law' (464). Even direct compensation for a crime would not leave the victim as well off as they would have otherwise been, due to the wide variety of costs which crime imposes on its victims (e.g., direct costs, increased expenditure on preventing future crime, and the costs, emotional or otherwise, of participating in a criminal justice process). Thus, Bush interprets the empirical evidence that victims are satisfied with restorative justice as a sign that they are incorrectly comparing their outcomes under restorative justice with outcomes under the criminal justice system, whereas the proper comparison is with a pre-crime state (459).

Here, Bush wrongly conflates satisfaction with indifference between states of utility. While a victim may not be indifferent between their utility before the crime and after restorative justice, their utility may nonetheless rise from a post-crime level through restorative justice. A victim could thus rationally say they are satisfied with the process without being made whole, if the starting point is post-crime, rather than pre-crime as Bush imagines it.

Bush further argues that, despite its claims, restorative justice cannot feasibly compensate a community for the impacts of crime. Crime ripples out from the direct victim in ways which may be too complex for restorative justice to address (461-64). This is an important limitation on the welfare impacts of restorative

---

<sup>2</sup> A Russian-language article may apply a similar idea, but only the title and abstract are available in English. The paper's title is 'Use of Rational Choice Theory in Criminal Law and Criminology (On the Example of the Coase Theorem)

justice: some parties who were harmed by a crime may indeed be left out of a restorative process. But as the number of participants grows, so do the transaction costs. Limiting the participants to those most directly affected by the crime—the victim and the offender—could thus be conducive to a successful agreement, a logical extension of my Coasean analysis below.

Another important consideration for social welfare is the effect of restorative justice on deterrence. Bush argues that a restorative justice model focused solely on victim compensation would severely weaken, if not eliminate, deterrence, as it results in expected penalties that are too low to dissuade potential offenders (464-65). While Bush is correct that employing restorative justice will result in some offenders receiving lesser sanctions than under a purely retributive system, his analysis assumes that restorative justice would completely replace retributive justice. In a model where restorative justice is an optional alternative requiring victim consent, a potential offender cannot rely on having a victim that is amenable to restorative justice. Thus, the expected penalty will continue to include the potential for traditional criminal sanction, mitigating the drop in deterrence. Adding a restorative element to the justice system may also '[enhance] the community's perception of' the justice system; this makes the public more likely to cooperate with police, increasing the probability of apprehension and mitigating a drop in deterrence (Lanni 2021, 678).

At the same time, Bush sees restorative justice as improving upon traditional criminal processes through lower recidivism rates (464). This conclusion is supported by Shem-Tov, Raphael, and Skog (2021), who conclude that juvenile offenders assigned to a restorative justice program in San Francisco are 32% less likely to be rearrested four years post-intervention compared to non-assigned offenders (2). This result may stem from 'preference shaping devices' in the restorative justice process (Bush 2003, 467). These include 'fostering empathy and dialogue with the victim' (Shem-Tov, Raphael, and Skog 2021, 2-3), and therapy programs that can address how offenders weigh the impacts of their actions and can raise awareness about social supports that eliminate the need to resort to crime (Bush 2003, 468).

Other research characterizes restorative justice as a form of private economic ordering, which is used in lieu of government intervention to achieve an efficient result (Subekti, Hartiwingsih, and Handayani 2021, 59; Lawson and Katz 2004, 179).<sup>3</sup> This is a helpful lens through which to see restorative justice, and naturally suggests the Coase Theorem as an analytical framework. Some theorists add that restorative justice relies on 'social capital,' a concept which Artinopoulou (2016) defines as 'the social bonds, links, networks and connections that bind families, communities and societies' (119). As one example of this exchange of social capital, Bibas and Bierschbach (2004) argue that crime creates a 'moral imbalance' (110) between the offender, and their victim and community (91), as the offender rejects the community's rules and denies that the victim 'deserve[s] respect' (110; see also Mamo 2019, 1447, and Lanni 2021, 646). An apology in this situation 'both teach[es] and reconcile[s] by reaffirming societal norms and vindicating victims' (Bibas and Bierschbach 2004, 113), an opportunity that is largely absent in the current legal system (136). The theories of social capital and apology as moral restitution can explain how Coasean negotiation provides benefits to victims and allows for the trading of valuable psychic goods.

Empirical findings on restorative justice outcomes also indicate benefits to victims and offenders. Strang et al. (2013) conduct a meta-analysis on 15 studies of restorative justice conferencing, concluding that victims who were assigned to restorative justice procedures had 'consistently higher' satisfaction reports than victims who were assigned to traditional criminal procedures (47). They also find a statistically significant decrease in recidivism for participants in restorative justice (25). Wilson, Olaghere, and

---

<sup>3</sup> Subekti, Hartiwingsih, and Handayani (2021) also write that restorative programs with a diversionary element are 'a manifestation of the transaction cost of economy [sic]' (59), although the rest of their article indicates that they may have administrative costs in mind.

Kimbrell (2017) review 84 evaluations in their meta-analysis, focusing on restorative justice with youth offenders. They conclude that victims who went through restorative justice had greater 'perceptions of fairness and greater satisfaction, ... and [were] more likely to feel that the outcome was just' (3). Offenders similarly 'had a greater perception of fairness' (3), were 'more satisfied with the program than youth in the comparison conditions' (38), and had a 'moderate reduction in future delinquent behavior' compared to youth who did not participate in restorative justice (6).

However, study design issues in each meta-analysis limit the generalizability of their results. Strang et al. (2013) only include studies in which victims and offenders agreed to be randomly assigned to restorative justice or an alternative before assignment to treatment or control groups occurred (12). Victims and offenders who agreed to randomization likely differ systematically from the population. Given this, Strang et al. may only indicate that restorative justice is more satisfying among individuals favorably disposed to restoration in the first place. Wilson, Olaghere, and Kimbrell (2017), on the other hand, include restorative justice programs without direct contact between offenders and victims and do not distinguish between different program types when presenting their findings on victim and offender satisfaction (34-35). Thus, their results do not offer a clear evaluation of victim-offender negotiation, although they do indicate that restorative justice generally construed is beneficial for victims and offenders.

An even larger issue arises from underlying studies in both meta-analyses using different control groups. In Strang et al. (2013), the control conditions include traditional prosecution, other diversion programs, or simply not employing restorative justice (in studies where the interventions were post-plea or post-sentence) (20-21). Wilson, Olaghere, and Kimbrell (2017) include studies where the control group was either a 'traditional' criminal justice process or 'an alternative program' (21). Due to these inconsistent comparators, the meta-analyses do not provide a definitive answer as to how restorative justice improves upon alternative options, or upon which alternatives it most improves.

Finally, Fatas and Restrepo-Plaza (2022) conduct an experiment to test the impact of various behavioral interventions on the willingness of participants to forgive offenders. The authors framed the decision to forgive as a risky choice between the benefits of rehabilitation and the costs of reoffending (2-3). For some participants, the risk was presented as the loss of an initial endowment if the offender recidivates, while for others, it was framed as a potential gain if the offender does not recidivate (3). Fatas and Restrepo-Plaza find that participants are more willing to forgive when their choice is framed as a potential loss, rather than a potential gain (8), in accordance with the predictions of prospect theory.<sup>4</sup> This work indicates how behavioral economics concepts may build on a purely rational assessment of restorative justice.

I contribute to the economic study of restorative justice by analyzing this process through the Coase Theorem. I thus extend the law-and-economics work in Bush (2003) and incorporate insights on the nature of restorative justice from non-economic research identified above. I add nuance to my analysis by considering the impact of behavioral economics principles on restorative justice negotiations. My Coasean analysis also provides a testable theory as to the expected impacts of restorative justice for further empirical research.

### **3. Materials and Methods**

This paper focuses on the discussions between victims and offenders that take place in many restorative justice programs and how they change the welfare of both parties. As noted above, the wider question of

---

<sup>4</sup> Prospect theory is a behavioral economic theory of valuation positing that: humans assess changes to wealth (i.e., any entitlement they possess) relative to a reference point; a gain of a given amount brings less utility than a loss of the same amount brings disutility (i.e., loss aversion); and, due to the foregoing, individuals are risk-seeking for losses but risk averse for gains (see Thaler (2015))

efficiency in the justice system, which may be defined as maximizing social welfare, is beyond the scope of this paper. I further limit my focus to forms of restorative justice that involve direct interaction between victim and offender and to crimes with a high degree of personal interaction, including non-fatal violent crime and property crime (e.g., burglary and theft; see Shem-Tov, Raphael, and Skog 2021, 1).

I rely on rational-choice economic theory in this analysis, which assumes that humans rationally seek to maximize their own utility. After presenting my first-stage analysis, I address how the behavioral economic concepts of optimism bias, self-serving bias, bounded self-interest, the endowment effect, and prospect theory could impact my conclusions.

My analysis has six further assumptions. (i) Restorative justice occurs if both the victim and offender choose to participate; they are the only parties in the negotiation. This is often not the case in practice, as participants from the wider community may be involved. (ii) Offenders are guilty of the offenses with which they are charged and are willing to admit the same, given the right incentives. Thus, there are no Type 1 (false positive) error costs. (iii) Restorative justice occurs before a finding of guilt. Restorative justice programs today may also occur as part of sentencing or post-sentencing (Lanni 2021, 649-50). (iv) The offender and victim hold on to their entitlements until they agree on the proper way to divide the cooperative surplus. Some restorative justice programs employed today may require the surrender of some entitlements before the process even begins. (v) If successful, restorative justice replaces a traditional criminal prosecution. However, if either the victim or the offender chooses to forego restorative justice or the process fails to result in an agreement, the case is resolved through ordinary criminal procedures. South Africa, for instance, operates a similar diversionary model, although prosecutors decide whether to refer a case for restorative justice and whether to accept a restorative justice agreement (Department of Justice and Constitutional Development 2011, 7). (vi) Finally, there is no other party (e.g., a presiding judge) who can modify or invalidate a restorative agreement reached by the victim and offender. In practice, this may be true for diversionary restorative justice processes but may not be the case when restorative justice is used as part of sentencing (657-658).

## 4. Results

There are four conditions that must be met for the predictions of the Coase Theorem to hold: (i) rights to entitlements must be clearly defined; (ii) parties must be able and willing to bargain; (iii) there must be no or minimal transaction costs; and (iv) parties must in fact bargain and come to an agreement. Conditions (i) and (iii) are explicitly stated by Ronald Coase in his seminal article (Coase 1960, 15, 19). Conditions (ii) and (iv) are implicit to a successful negotiation process. I analyze each of these conditions in turn, after first identifying the entitlements at play in restorative justice.

The offender has a legal entitlement and two moral entitlements. In the United States context, the offender's legal entitlement is their Sixth Amendment right to contest the charges against them.<sup>5</sup> Their moral entitlements are the ability to grant an apology that a victim may desire and the ability to make restitution for the crime; the two are closely linked, as making amends for the crime validates the sincerity of contrition. The offender's three entitlements exist in both a traditional criminal justice process and a restorative justice process.

---

<sup>5</sup> Prospect theory is a behavioral economic theory of valuation positing that: humans assess changes to wealth (i.e., any entitlement they possess) relative to a reference point; a gain of a given amount brings less utility than a loss of the same amount brings disutility (i.e., loss aversion); and, due to the foregoing, individuals are risk-seeking for losses but risk averse for gains (see Thaler (2015)).

Victims, on the other hand, do not have a true entitlement under the traditional criminal justice system. While prosecutors or judges may solicit their perspectives pursuant to custom or statute, victims lack the ability to determine the trajectory of their case. Under restorative justice, however, victims 'typically' 'hold a de facto veto power' over case resolution (Lanni 2021, 668; Fatas and Restrepo-Plaza 2022, 2). Thus, their legal entitlement is the ability to provide a more favourable outcome to the offender by consenting to a restorative justice agreement. Victims also have the moral entitlement to accept the offender's apology.

The expected value of trial as the next-best alternative for either party determines the value of entitlements. The defendant will not surrender their entitlements if doing so costs them more (in utility terms) than they expect a trial to cost. The expected cost of trial is the highest price an offender would be willing to pay under restorative justice. Similarly, a victim would not accept an agreement that benefits them less than they expect to benefit from a trial. The victim's net utility gain from trial informs the minimum terms they would be willing to accept in a restorative justice agreement.

(i) Rights to these entitlements are, in general, clearly defined. Neither party can be forced to surrender their moral entitlements; the legal entitlement of the offender to contest charges stems from the Sixth Amendment to the United States Constitution; and victims may veto the use of restorative justice in many situations. However, some programs may not guarantee this veto power to the victim. Furthermore, offenders may require statutory protections to access restorative justice without penalty. Massachusetts, for example, prohibits using participation in restorative justice or the offender's statements during such a process as evidence or 'an admission of guilt' (Lanni 2021, 655, quoting Mass. Gen. Laws ch. 276b, § 4 (2018)). Without such protections, defendants must choose between accessing restorative justice and avoiding self-incrimination. Thus, rights to entitlements may not be sufficiently clear in all restorative justice contexts.

(ii) Restorative justice programs provide a formal venue that enables offenders and victims to bargain over their entitlements. This is in contrast to a traditional criminal justice process, where there is likely no established means through which offenders and victims can interact before the trial concludes. However, even with the ability to bargain, not all offenders and victims may be willing to do so. The victim may have a strong desire to see the defendant punished, and the offender may believe that they can avoid a guilty verdict in the first place, or they may simply be unwilling to accept responsibility under any circumstances. In such cases, the expected benefits of restorative justice would not exceed the expected benefits of trial, and restorative justice bargaining would not occur.

Nevertheless, there are potential gains under restorative justice that would lead some parties to bargain. The offender may receive a more favourable disposition through restorative justice than through a traditional criminal process, which strongly incentivizes their participation. The victim could avoid the disutility of a trial, resolve their case more quickly, and receive an apology and restitution. In these situations, the expected value of restorative justice to both parties exceeds the expected value of trial, and the parties would negotiate.

(iii) Compared to traditional criminal procedures, restorative justice programs lower the transaction costs for victims and offenders to negotiate over their entitlements, as the parties no longer bear the logistical expenses required to coordinate such a negotiation. Some transaction

costs would remain, such as the emotional cost of revisiting the crime and the time cost of participating. But since these types of costs also exist under the traditional criminal justice system, restorative justice represents a net decrease in transaction costs for the parties.

(iv) Provided that the victim and offender do engage in negotiation and their reservation prices (i.e., the 'price' each is willing to accept or pay to give up their entitlements) overlap for some mutually acceptable outcomes, they should reach an agreement. The Coase Theorem does not predict the precise division in cooperative surplus between victim and offender, but under conditions of rationality, any agreement within the settlement range would be acceptable to both parties. Once a restorative agreement is reached, the offender accepts responsibility and commits to certain actions to make amends, surrendering their entitlements, and the victim consents to the resolution of the case through such actions and surrenders their entitlements. The surrender of entitlements also signals that the welfare of each party is higher than it would have been under a traditional criminal process.

This negotiation demonstrates how restorative justice helps internalize the externalities of crime and criminal justice. An external cost, or externality, is a negative consequence of some main economic transaction. In Coase's (1960) classic example of a cattle rancher and farmer, the victim of the cattle ranching is the farmer—but the rancher's goal is to grow cattle, not to harm the farmer. Similarly, non-fatal violent crime and property crime functions as an economic transaction where the offender unlawfully takes some tangible or intangible good, such as stealing a valuable object. Even acts that physically harm the victim may simply use the victim's body to attain a wider goal, like instilling fear. The emotional, monetary, and other impacts on the victim are the externalities. The traditional criminal justice process also imposes externalities on victims. Choices made by prosecutors and defendants result in costs to the victim. For instance, while the victim may prefer to avoid testifying at trial, they may have to incur emotional and time costs to appear on the stand if the defendant contests the charges. An adversarial process fails to internalize these two sets of externalities and thus leads to economic inefficiency.

In a restorative justice process, the victim's ability to negotiate with the offender and determine the outcome of their case allows them to fully bring their externalities into the discussion. The negotiating parties directly address the costs to the victim when forming the restorative justice agreement, thus helping to mitigate the externalities of crime. The victim can also avoid the costs of trial by entering into a restorative justice agreement, resolving an additional externality problem that would be present in the traditional system.

## **5. Discussion**

If the four conditions above hold, the Coase Theorem predicts that the victim's and offender's entitlements have moved to their highest-value use. Thus, adding the option to pursue restorative justice instead of traditional criminal prosecution promotes economic efficiency by raising the welfare of offenders and victims. Importantly, no offender or victim would be worse off in my model than in the criminal justice status quo since a failed negotiation would simply return a case to the traditional process.

My conclusion addresses part of the critique of restorative justice in Bush (2003), showing that victims can be rationally satisfied under restorative justice because of how the process internalizes the crime-

and trial-related externalities that they face. It also provides an explanation for the findings in Strang et al. (2013) and Wilson, Olaghere, and Kimbrell (2017) that restorative justice benefits victims and offenders.

A Coasean framework shows that the efficiency benefits of an optional restorative justice process arise in large part from two key features. The first is that restorative justice programs reduce the transaction costs that victims and offenders face in negotiating: they remove the logistical burden of coordinating these types of discussions and shift the expense onto the program provider. Furthermore, restorative justice grants the victim a clear legal entitlement, allowing them to then direct the outcome of their case and internalize their crime-related externalities. Victims who expect greater benefits from avoiding, rather than participating in, a trial can negotiate with the offender to obtain these benefits. Such an entitlement does not exist in the current system.

However, the conditions for Coasean bargaining may not always be present in real-world settings. There may be high transaction costs due to the requirements of a restorative justice process. Indeed, in restorative justice programs, rates of victim participation 'at or below fifty percent are not uncommon, and ... rarely exceed eight percent,' often due to 'the time and effort involved' in participating (Lanni 2021, 662). A further challenge could be that a victim's legal entitlement is unclear because their control over the process is poorly defined. Either of these possibilities threatens an ideal Coasean resolution.

Another complication is that restorative justice programs may require the offender to surrender one or more of their entitlements as a precondition to participation. Programs 'typically require the offender to accept responsibility' (Lanni 2021, 654) at the start of the restorative process—a surrender of their moral entitlement. In such cases, the defendant retains their legal entitlement (i.e., the right to contest the charges), and the parties may still bargain over the resolution of the trial. But some programs go one step further and require the defendant to formally plead guilty before restorative justice negotiations occur. One jurisdiction with this model is New Zealand (Lanni 2021, 650, 656-57). In this case, the defendant must give up both a moral entitlement and a legal entitlement to participate. This may decrease the defendant's expected value from engaging in restorative justice, making it less likely that the defendant will enter the process in the first place and thus less likely that the efficiency benefits of restorative justice will be realized.

Furthermore, behavioral economics principles of optimism bias, self-serving bias, bounded self-interest, and the endowment effect suggest that optimal Coasean exchange is less likely in the real world compared to a perfectly rational world as the reservation prices and bargaining attitudes of the parties may not reflect what is in their rational self-interest. Prospect theory also provides insights into how the framing of restorative justice impacts the likelihood of success.

Optimism bias, as defined by Jolls (2004), is a cognitive bias whereby individuals 'underestimate the probability that negative events will happen to them as opposed to others' (4). Optimism bias could induce offenders to overestimate their chances of a favourable outcome at trial; they would then demand a more lenient restorative justice agreement. Such an outcome is particularly likely since the prosecution's standard of proof (i.e., beyond a reasonable doubt) is very high. Self-serving bias, as defined by Babcock and Loewenstein (1997), leads parties 'to conflate what is fair with what benefits oneself' (110). A related concept is bounded self-interest, defined as when one's utility function depends in part on the utility of others (Jolls, Sunstein, and Thaler 1998, 1479). One implication of this idea is that individuals may reject outcomes that are in their self-interest but nonetheless do not appear fair to themselves or others (Ibid.). Applied to restorative justice, self-serving bias and bounded self-interest would lead parties to reject offers that would be in their rational self-interest but are less than what is perceived to be fair. This narrows the range of mutually acceptable outcomes, and in extreme situations may eliminate the settlement range altogether (Babcock and Loewenstein 1997, 110).

The endowment effect predicts that economic actors value a particular good or entitlement more when they possess it than when they need to obtain it (Jolls, Sunstein, and Thaler 1998, 1484). In other words, the amount they are willing to accept to surrender an entitlement is higher than the amount they would be willing to pay to gain the same entitlement (Ibid.). As a result, the initial endowment of entitlements affects their ultimate allocation (1483). In the context of restorative justice, the endowment effect would make optimal Coasean outcomes less likely. As victims and offenders demand higher prices for giving up their own entitlements than they would pay to gain the entitlement, the settlement range grows narrower or disappears.

Finally, as demonstrated by Fatas and Restrepo-Plaza (2022), framing the choice to engage in restorative justice—specifically, to forgive an offender—as a potential loss rather than a potential gain improves the likelihood that a victim will participate. This is consistent with prospect theory (see note 4 for a definition). While participants in the study occupied the role of a victim, the same results may hold for offenders. Thus, Fatas and Restrepo-Plaza’s findings suggest that the Coasean outcome of a restorative justice process can be made more likely by emphasizing the benefits that the parties may lose if they fail to participate.

In addition to these concerns from behavioral economics, non-economic literature provides additional reasons why restorative justice in practice may deviate from an ideal Coasean outcome. Restorative justice programs include the potential for victims to feel ‘coerced’ to accept a restorative agreement and eschew justice through traditional criminal procedures, while offenders may also feel ‘coerced’ to surrender their entitlements (Menkel-Meadow 2007, 10.5). In economic terms, such coercion may mean that victims and/or offenders accept restorative outcomes that do not maximize their welfare. There are additional concerns about the social implications of a model which allows for private resolution of offenses, such as whether disparate outcomes for ‘similarly situated’ victims and offenders challenge legal ‘equality,’ and whether the normative and ‘precedent-setting’ role of the legal system is lost when crime resolution becomes privatized (Ibid., 10.6). This set of concerns bears on the social welfare effects of restorative justice through its impact on deterrence.

Further research may test my theoretical conclusions with real-world data. Another line of future inquiry would be how the option to pursue restorative justice impacts the welfare costs of Type I errors, as innocent offenders and their ‘victims’ may also find themselves in this process. There is also a need to model the wider social welfare effects of restorative justice through its impacts on deterrence and recidivism, which merits further study but was not the primary focus of this paper.

## **Acknowledgements**

I would like to thank Ken Felter for his extensive feedback on the drafts of this paper and for his continued support of my work. I would also like to thank the editorial team at the UCL Journal of Economics and an anonymous peer reviewer for their thoughtful comments on a draft of this paper.

## References

- Artinopoulou, Vasso. 2016. 'Restorative Justice: A Value for Money Justice.' *Regional Science Inquiry* 8, no. 3: 107-123.
- Babcock, Linda, and George Loewenstein. 1997. 'Explaining Bargaining Impasse: The Role of Self-Serving Biases.' *Journal of Economic Perspectives* 11, no. 1 (Winter): 109-126. DOI: <https://doi.org/10.1257/jep.11.1.109>.
- Bibas, Stephanos and Richard A Bierschbach. 2004. 'Integrating Remorse and Apology into Criminal Procedure.' *The Yale Law Journal* 114, no. 1 (October): 85-148. DOI: <https://doi.org/10.2307/4135717>.
- Bush, Darren. 2003. 'Law and Economics of Restorative Justice: Why Restorative Justice Cannot and Should Not Be Solely About Restoration.' *Utah Law Review* 2003, no. 1: 439-468.
- Carden, Art. 2020. 'A Useful Introduction.' *Regulation* (Fall): 44-46.
- Coase, Ronald. 1960. 'The Problem of Social Cost.' *The Journal of Law & Economics* 3, (October): 1-44. <https://www.jstor.org/stable/724810>.
- Department of Justice and Constitutional Development, Republic of South Africa. 2011. 'Restorative Justice: The Road to Healing.' <https://www.justice.gov.za/rj/2011rj-booklet-a5-eng.pdf>.
- Fatas, Enrique and Lina Restrepo-Plaza. 2022. 'When losses can be a gain. A large lab-in-the-field experiment on reference dependent forgiveness in Colombia.' *Journal of Economic Psychology* 88 (2022): 1-19. DOI: <https://doi.org/10.1016/j.joep.2021.102463>.
- Jolls, Christine, Cass R. Sunstein, and Richard Thaler. 1998. 'A Behavioral Approach to Law and Economics.' *Stanford Law Review* 50: 1471-1504.
- Jolls, Christine. 2004. 'On Law Enforcement with Boundedly Rational Actors.' *Harvard Law School John M. Olin Center for Law, Economics and Business Discussion Paper Series, Paper 494*.
- Lanni, Adriaan. 2021. 'Taking Restorative Justice Seriously.' *Buffalo Law Review* 69, no. 3 (May): 635-681. <https://digitalcommons.law.buffalo.edu/buffalolawreview/vol69/iss3/2>.
- Lawson, Catherine L., and JoAnne Katz. 2004. 'Restorative justice: an alternative approach to juvenile crime.' *Journal of Socio-Economic* 33, (2004): 175-188. DOI: <https://doi.org/10.1016/j.socec.2003.12.018>.
- Mamo, Andrew B. 2019. 'Three Ways of Looking at Dispute Resolution.' *Wake Forest Law Review* 54 (2019): 1399-1457. <https://ssrn.com/abstract=3386840>.
- Menkel-Meadow, Carrie. 2007. 'Restorative Justice: What Is It and Does It Work?' *Annual Review of Law and Social Science* 3 (2007): 10.1-10.27. <https://doi.org/10.1146/annurev.lawsocsci.2.081805.110005>.
- New Zealand. *Sentencing Act 2002*. 2022.
- Polinsky, A. Mitchell, and Steven Shavell. 1984. 'The Optimal Use of Fines and Imprisonment.' *Journal of Public Economics* 24, no. 1 (June): 89-99.
- Shem-Tov, Yotam, Steven Raphael and Alissa Skog. 2021. 'Can Restorative Justice Conferencing Reduce Recidivism? Evidence From the Make-It-Right Program.' *National Bureau of Economic Research, Working Paper 29150*, August 2021. DOI: <https://doi.org/10.3386/w29150>.
- Strang, Heather, et al. 2013. 'Restorative Justice Conferencing (RJC) Using Face-to-Face Meetings of Offenders and Victims: Effects on Offender Recidivism and Victim Satisfaction. A Systematic Review.' *Campbell Systematic Reviews*, 2013:12, 1 November 2013. DOI: <https://doi.org/10.4073/csr.2013.12>.
- Subekti, Hartiwingsih, I Gusti Ayu Ketut Rachmi Handayani. 2021. 'Diversion in the Settlement of Child Cases to Realize Restorative Justice in Richard A Posner's Theory View.' *Advances in Social Science, Education and Humanities Research* 583: 56-61. DOI: <https://doi.org/10.2991/assehr.k.211014.014>.
- Thaler, Richard H. 2015. 'Value Theory.' In *Misbehaving: The Making of Behavioral Economics*, 25-34. New York: W. W. Norton & Company.

Wilson, David B., Ajima Olaghere and Catherine S. Kimbrell. 2017. 'Effectiveness of Restorative Justice Principles in Juvenile Justice: A Meta-Analysis.' U.S. Department of Justice Office of Justice Programs, Grant No. 2015-JF-FX-0063, 12 May 2017. <https://www.ncjrs.gov/pdffiles1/ojdp/grants/250872.pdf>.