

University of Chicago Law School

**Chicago Unbound**

---

Coase-Sandor Working Paper Series in Law and Economics Coase-Sandor Institute for Law and Economics

---

2021

## **The Law of Restitution for Mistaken Payments: An Economic Analysis**

Dharmika Dharmapala

Nuno Garoupa

Follow this and additional works at: [https://chicagounbound.uchicago.edu/law\\_and\\_economics](https://chicagounbound.uchicago.edu/law_and_economics)



Part of the **Law Commons**

---

# The Law of Restitution for Mistaken Payments: An Economic Analysis\*

Dhammika Dharmapala  
[dharmap@uchicago.edu](mailto:dharmap@uchicago.edu)  
University of Chicago Law School

Nuno Garoupa  
[ngaroup@gmu.edu](mailto:ngaroup@gmu.edu)  
George Mason University Antonin Scalia Law School

First draft: August 2021  
Revised draft: February 2022

## Abstract

The law of restitution and unjust enrichment has emerged as an important and independent branch of private law globally, but has attracted relatively little economic analysis. The aim of this paper is to develop a general conceptual framework for the economic analysis of the core example of restitution - mistaken payments. We develop a formal model in a parsimonious setting with two buyer-seller pairs, with low (high) transaction costs within (across) pairs. This model generates several novel insights, based on the idea that mistaken payments to strangers impose a “transaction tax” on contracting parties. It sheds new light on distortions generated by the change of position defense, and on the rationale for the discharge for value doctrine (which was implicated in the recent highly controversial *Citibank* case). We show that full restitution is socially optimal when harm is unilateral, and partial restitution is generally optimal when harm is bilateral. We also propose a hypothetical decoupling regime that would (if it were feasible) lead to more efficient outcomes than would partial restitution. The decoupling regime is intended as a thought experiment that can arguably shed light on some of the foundational debates surrounding the law of restitution.

JEL Codes: K10; H20

Keywords: Restitution; Mistaken payments; Transactions tax; Unjust enrichment; Economic analysis of private law

---

\* We are grateful to Elise Bant, Omri Ben Shahaar, Giuseppe Dari-Mattiacci, Andrew Kull, Saul Levmore, Steve Shavell, and Alex Stremitzer for helpful comments and discussions. Dharmapala acknowledges the financial support of the Lee and Brena Freeman Faculty Research Fund at the University of Chicago Law School. The usual disclaimers apply.

## 1) Introduction

The law of restitution (also sometimes referred to as the law of unjust enrichment) has emerged in recent decades as an important and independent branch of private law in many jurisdictions. The common law has a long history of awarding restitution in actions for money had and received (most notably, in Lord Mansfield's famous judgment in *Moses v. Macferlan*<sup>1</sup>). However, it was not until *Lipkin Gorman v. Karpnale Ltd.*<sup>2</sup> that the House of Lords clearly recognized the existence in England and Wales of an independent law of restitution based on the principle of unjust enrichment. A series of important doctrinal developments followed in England and Wales and in other common law jurisdictions outside the United States. These were accompanied by extensive academic debate among private law scholars (almost entirely outside the law-and-economics tradition).<sup>3</sup>

In contrast, the economic analysis of the law of restitution has been quite limited. One strand highlights the symmetry between the law of restitution - conceptualized as addressing nonbargained benefits - and tort law's focus on nonbargained harms (e.g., Levmore 1985; Cooter and Porat, 2014, 2019). Thus, in economic terms, high transaction costs are an essential feature of situations in which the law of restitution applies. Of course, mistaken payments are possible within

---

<sup>1</sup> 2 Burr. 1005 [1760].

<sup>2</sup> 4 All E.R. 512 [1992]. This decision was central to the recognition of the existence of an independent law of restitution - grounded on the unjust enrichment of the defendant at the expense of the plaintiff - in England and Wales (e.g., McKendrick, 1992). Simplifying the facts somewhat and ignoring some extraneous details, Lipkin Gorman was a law firm (of solicitors). One of Lipkin Gorman's partners, Norman Cass, misappropriated approximately £223,000 from the firm and used these funds to place a series of bets at a casino owned by Karpnale Ltd. The casino paid Cass winnings of about £72,000 and retained the remainder of the funds (approximately £151,000). Lipkin Gorman sought to recover the entire £223,000 misappropriated by Cass and spent by him at the casino, in an action for money had and received. The House of Lords for the first time clearly recognized the existence of an independent law of restitution based on the principle of unjust enrichment; it also recognized the change of position defense (discussed below). Thus, their Lordships awarded Lipkin Gorman about £151,000 in restitution, rather than the £223,000 that Cass spent at the casino. Their Lordships determined that the casino had changed its position by the amount of winnings paid out to Cass (about £72,000).

<sup>3</sup> Weinrib (2005, p. 861) describes this situation as follows: "What ever happened to the law of restitution in the United States? Half a century after [the 1937] *Restatement of Restitution*, the subject was more or less dead in American law schools. . . . Meanwhile, the project of developing the law of restitution was taken up by pioneering academics elsewhere in the common-law world. In the past two decades, the leading Commonwealth courts, significantly influenced by academic writing, have recognized or elaborated the basic principle of the law of restitution — that the defendant must restore an unjust enrichment made at the plaintiff's expense. Outside the United States, the law of restitution is now by far the most animated field of private-law scholarship." Note, however, that the American Law Institute's *Restatement (Third) of the Law of Restitution and Unjust Enrichment* has since appeared, in 2011. Saiman (2008) discusses various explanations for the relative lack of interest in the topic in the United States, relative to other common law jurisdictions. Other arguments can be found in Smith (2001) and Priel (2013). For comprehensive accounts of the law of restitution, see Birks (1985) and Burrows (2011). Kull (1995) presents an argument that the principle of unjust enrichment is essential to unifying the disparate scenarios to which the law of restitution is applied.

contractual relationships. However, absent contracting frictions, parties can allocate the risk and costs of mistakes contractually (or accept contract law defaults). Our model, as described below, is thus based on the most parsimonious setting in which mistaken payments to strangers can be analyzed, with two buyer-seller pairs and low (high) transaction costs within (across) pairs.

While the types of scenarios to which it applies are quite diverse, the core example of the law of restitution is traditionally thought to be the legal treatment of mistaken payments. Such payments arise from a spontaneous mistake by the payer, made to a good faith recipient who has done nothing to induce the mistake. For instance, consider the following situations, all of which are encompassed by this paradigm:

- (a) A buyer who is engaged in transactions with multiple sellers mistakenly pays for a particular good or service to the wrong seller.
- (b) A casino with multiple gamblers mistakenly extends credit to the wrong client.
- (c) A bank or other financial institution mistakenly deposits interest or dividends in the account of the wrong investor.
- (d) A debtor mistakenly pays its creditors an amount in excess of that which is owed.

The general principle that courts and scholars have articulated is that (subject to various exceptions and limitations) recipients are required to return mistaken payments to the payer. Only a handful of prior papers (Beatson and Bishop, 1986; Huber, 1988; du Plessis, 2010; Gilboa and Kaplan, 2018) analyze this paradigmatic scenario and its legal treatment from an economic perspective;<sup>4</sup> of these, only the last develops a formal model (focusing on enforcement error, as discussed in Section 3 below).

The formal model that we develop in this paper is based on the idea that the possibility of mistaken payments to strangers imposes what can be viewed as a “transaction tax” on contracting parties. It generates a number of novel insights, and sheds new light on distortions generated by the change of position defense. It also suggests a rationale for the discharge for value doctrine (which was implicated in the recent widely-discussed *Citibank* case, described in Section 2 below). We characterize conditions under which full restitution and partial restitution regimes are socially optimal. Furthermore, our “transaction tax” perspective on mistaken payments focuses attention on the relative social costs and benefits of alternative mechanisms for addressing mistakes. We

---

<sup>4</sup> See also Farnsworth (2014, Ch. 2).

thus propose – primarily as a thought experiment - a hypothetical decoupling regime that can arguably shed light on some of the foundational debates surrounding the law of restitution.

The article proceeds as follows. Section 2 motivates the foundations of our formal model. Section 3 provides further background on the law of restitution and on the prior scholarly literature. Section 4 develops our formal model and derives the paper’s results. Section 5 discusses the implications of these results and offers some additional general observations. Section 6 concludes.

## 2) Motivation

The law of restitution, though largely neglected for many decades in US legal scholarship, has recently attracted attention – and indeed a measure of notoriety – in the United States due to a Federal district court decision denying restitution for a particularly large mistaken payment made by Citibank in August 2020 (de Fontenay, 2021; Kaplan and Gilboa, 2021; Talley, 2021).<sup>5</sup> An error by a Citibank employee resulted in the entire principal amount (about \$900 million) being paid to creditors of an account holder (Revlon); the aim had been to pay only the interest that was due (about \$8 million). Citibank notified the lenders of the mistake about a day later. While many of the lenders returned the mistaken payments, some (representing about \$500 million of the principal amount) did not. Citibank sought restitution from the latter. However, the district court for the Southern District of New York denied restitution, based on the “discharge for value” defense under New York law (which applies when the mistaken payment satisfies a preexisting debt to the recipient). While our framework addresses much more general concerns, it can also shed light on the underlying rationale for the discharge for value defense. However, we do not seek to justify the outcome in the *Citibank* case, where a number of special circumstances applied (for instance, the funds transferred were Citibank’s rather than Revlon’s, and there was an ongoing prior dispute between Revlon and its creditors relating to a proposed debt restructuring).

The aim of this article is to develop a general conceptual framework for the economic analysis of mistaken payments (and thereby of the law of restitution more generally). This framework forms part of a wider conceptualization of the law of restitution as being grounded in an economic notion of expanding parties’ contracting space by unwinding transactions in certain

---

<sup>5</sup> *In re Citibank August 11, 2020 Wire Transfers*, No 20-CV-6539 (JMF), 2021 WL 606167 (SDNY 16 February 2021), currently under appeal at the 2<sup>nd</sup> Circuit. The district court applied the rule articulated by the New York Court of Appeals in *Banque Worms v. Bank America Int’l*, 77 N.Y.2d 362 (N.Y. 1991).

situations (however, because the specific contexts to which restitution applies differ considerably, we only develop this point here in relation to mistaken payments).<sup>6</sup> While the paper discusses some normative issues, its primary aim is to analyze the consequences for economic efficiency of different rules relating to restitution, rather than to seek to explain the doctrines of the law of restitution in terms of economic efficiency. It is motivated by the importance of this body of law within private law doctrine and scholarship, as well as by the significance of the law of restitution in shaping economic activity and the allocation of resources.

We construct a simple taxonomy that encompasses various possible scenarios. The harm from a mistake may be either unilateral (i.e., borne only by the payer, absent full restitution) or bilateral (where there may also be a cost – for instance, of distorted consumption patterns – to the recipient of making full restitution). Precaution against mistakes may be exogenously fixed, unilateral (i.e., available only to the payer), or bilateral (where the recipient can also undertake precaution, for instance by auditing incoming payments). Using this taxonomy, we develop a formal model based on the most parsimonious setting in which mistaken payments can be analyzed. This assumes two buyer-seller pairs, where transactions costs are zero within each pair but prohibitively high across pairs (the payer and recipient are thus strangers who are unable to bargain, an essential feature of the situations in which the law of restitution applies). This model provides a framework that unifies and clarifies a number of existing ideas regarding mistaken

---

<sup>6</sup> Another context that is important in the law of restitution is that of invalid contracts. Indeed, much of the doctrinal development of the English law of restitution occurred as a consequence of the extensive litigation in the 1990s in what are known as the “interest rate swaps” cases (such as *Westdeutsche Landesbank Girozentrale v. Islington London Borough Council* A.C. 669 [1996]). In the statute establishing and regulating local authorities, Parliament imposed restrictions on the ability of local authorities to engage in borrowing. In 1987, Islington Borough Council (a local authority in London) entered into an interest rate swap transaction with a German bank, Westdeutsche Landesbank (referred to more briefly as WestLB). This interest rate swap transaction was economically equivalent to a loan of from WestLB to Islington Borough Council at a floating annual interest rate (a transaction that would have been prohibited by the statute). Nonetheless, at the time the parties entered into the swap arrangement, it was not thought that the statutory restrictions on borrowing applied to swap transactions. In 1992, however, the House of Lords held in *Hazell v. Hammersmith and Fulham London Borough Council*, 2 A.C. 1 (1992), that swap transactions such as that described above were *ultra vires* the local authorities that had entered into them. As Islington Borough Council did not have the legal authority to enter into swap transactions, its contract with WestLB was void. WestLB sought and received restitution of the upfront payment made under the swap arrangement, minus the amount paid by Islington Borough Council up to that point. There were hundreds of similar contracts between financial institutions and UK local authorities, giving rise to extensive litigation and greatly facilitating the doctrinal development of the law of restitution. The payments made pursuant to a contract that is later determined to be invalid may be characterized as being “mistaken”; however, as this setting raises some distinctive issues, we do not directly address it in our model. There is a substantial literature on rescission and restitution as a contract remedy, or as a mechanism used when contracts are void (e.g., Brooks and Stremitzer, 2012).

payments, while also generating several novel insights. These are developed formally in Section 4 below, but we introduce and illustrate them briefly here by means of simple numerical examples.

Prior literature has characterized the primary efficiency cost of mistaken payments (absent restitution) as precaution by the payer that is excessive from a social perspective. Our framework emphasizes instead that – even when precaution is exogenously fixed – there is a social cost from mistaken payments in the form of transactions that buyer-seller pairs do not undertake because their joint surplus is lowered by the expected value of the mistaken payment to an unrelated third party (absent restitution). For example, suppose that the seller’s cost is \$20, while the buyer’s valuation is \$22. The joint surplus is \$2 in the absence of the possibility of mistakes; suppose that they agree on a price of \$21. If we now introduce a 0.1 probability of a mistake (without restitution), then the joint surplus falls by \$2.10 and the transaction is no longer worthwhile for the parties (the price may potentially adjust in response to the possibility of a mistake, but this does not affect the basic point that the joint surplus is lower). However, the transaction remains socially desirable because the loss from a mistaken payment is offset (from society’s perspective) by the gain to the recipient (absent any additional deadweight losses).

The effect of the (anticipated) mistaken payment is thus analogous to that of a tax (in the example, an *ad valorem* tax at a 10% rate or a specific tax of \$2.10), where the revenue from the tax is received by the mistaken recipient rather than the government.<sup>7</sup> In a unilateral harm setting, a legal regime that provides for full restitution will be optimal – if the administrative costs of obtaining restitution are sufficiently low – because it eliminates this tax at no cost to the recipient. Introducing the buyer’s choice of precaution against mistakes reinforces this conclusion (as full restitution prevents socially excessive precaution), but is not essential to it.

Our framework also generates some new insights about the most important defense to claims of restitution, the change of position defense. Bant (2009, p. 2) provides an “accepted and uncontroversial core example” of this defense:

“Suppose a claimant made a spontaneous mistake (that is, a mistake in no manner generated or induced by the defendant) which caused her to make a payment to the defendant. Suppose further that the defendant, having no grounds to suspect any mistake, purchased a lavish dinner he would not otherwise have bought but for his honest belief that the value of the payment was irrevocably his. If the claimant subsequently brings an action against the defendant for restitution of the value of the payment, the change of position defence

---

<sup>7</sup> Beatson and Bishop (1986, p. 151) mention in passing the idea of mistakes as a tax, but they do not develop it, nor derive any implications.

will apply *pro tanto*, to the extent of the change. Thus, if the defendant received £500 by mistake and the price of the dinner was £200, the defendant will only be required to make restitution to the claimant of the balance of £300.”

The standard account of this “bilateral harm” scenario (when translated into economic terms) posits an excess burden of restitution for the recipient. In the example above, the recipient receives £500 by mistake and commits £200 to the (irreversible) purchase of a luxury good. When she learns that she must unexpectedly make full restitution of £500, she must reduce her consumption of essential goods by £200 to do so. Assume that this suboptimal pattern of consumption entails a utility loss to the recipient of £50. This excess burden of £50 is generally thought to justify departing from full restitution.

However, when restitution is only partial or nonexistent, the payer may also be argued to experience an excess burden. She believes in good faith that her obligation to the intended recipient of the £500 has been discharged. Consequently, she commits to £200 of luxury consumption, only to discover that the payment was misdirected and that she must pay the intended recipient another £500. Absent restitution, she must reduce her consumption of essential goods by £200 to do so. As there is no *a priori* reason that the associated excess burden would be different for the payer than for the recipient, an excess burden of £50 can be assumed here as well. Thus, it is not clear that the mere existence of an excess burden can justify a departure from full restitution. However, our model shows that the convexity of the excess burden implies that partial restitution is optimal. For example, suppose that the excess burden of a £200 loss is £10 and that of a £300 loss is £18. Then, requiring the recipient to pay £300 would generate a combined excess burden of £28 (versus the £50 excess burden for the recipient if she were required to pay the full £500).

When the amount paid in restitution depends on the recipient’s consumption choices, the change of position defense can potentially create an additional distortion (beyond those noted above) to the recipient’s pattern of expenditure. Suppose that after consuming her £200 dinner, the recipient contemplates ordering a bottle of wine that costs £50 and that she values at £47. As a good faith recipient, she does not know that the £500 she received was in error, but believes that in general payments that she receives are mistaken with a probability of 0.1. Then, if she does not order the wine, she expects to pay  $(0.1) \times (300) = £30$  in restitution, leaving her with £270 to spend on other goods. If she orders the wine, she expects to pay  $(0.1) \times (250) = £25$  in restitution, leaving her with £225 to spend on other goods, plus the £47 value of the wine (a total of £272). Thus, it is

(privately) optimal for her to order the wine, although this destroys £3 of value from a social perspective.

Perhaps most fundamentally, the “transaction tax” perspective on mistaken payments focuses attention on the relative social costs of various mechanisms for addressing mistakes. The law of restitution (like other areas of private law) tends to use “balanced-budget” remedies (where what the defendant pays equals what the plaintiff receives, apart from litigation costs). However, we show that in our framework a social planner who has access to lump sum (nondistortionary) taxes, or more generally to taxes that are less distortionary than the implicit tax associated with mistakes, would wish to “decouple” restitution. For example, suppose that the government were to pay £500 to the mistaken payer, while allowing the good faith recipient to retain the mistaken payment of £500. This would simultaneously eliminate the excess burden, the incentive for the payer to take excessive precaution, and the deadweight loss from the transaction tax, without incurring any social cost of raising the required revenue (under the assumption that lump sum taxes are available).

A hypothetical decoupled system would potentially induce collusion among parties to obtain this public subsidy. In some respects, such collusion would be the mirror image of that which may occur in a tort setting under a hypothetical “double strict liability” regime (where the injurer and the victim each pays the full cost of the accident, one to the state and one to the other party). If courts can verify that mistakes are spontaneous and that payments are received in good faith, such collusion may be infeasible or at least not straightforward. Nonetheless, the decoupling idea is intended not as a practical proposal, but rather as a thought experiment that can arguably help us better understand some of the foundational debates surrounding the law of restitution.

Thus, our model generates a number of novel insights. It suggests that a full restitution regime is socially optimal when harm is unilateral – i.e., when a mistaken payment imposes harm only on the payer (absent restitution). When harm is bilateral (i.e., the recipient would suffer some net harm from receiving the payment and making full restitution), our results imply that a regime of partial restitution is optimal among “balanced-budget” remedies. However, a fully decoupled regime would be superior, if it were feasible.

While it is not the paper’s primary focus, we also address some normative issues raised by the law of restitution. From an economic (or more generally any consequentialist) perspective, it is not straightforward to make sense of the central claim that restitution is justified by the unjust

enrichment of the defendant. More broadly, this issue implicates the general tendency of the law to disfavor arbitrary redistributions of wealth (i.e., what are often termed “windfalls”). We discuss in Section 5 below how we might proceed to link the notion of unjust enrichment to one particular form of economic analysis, the Mirrleesian optimal taxation framework. We argue that under certain circumstances, a legal regime providing for the restitution of arbitrary redistributions of wealth may enhance the ability of governments to undertake redistributive taxation.

### 3) Literature Review

While there has been extensive academic discussion and debate concerning the law of restitution among scholars of private law, the economic analysis of the law of restitution has been quite limited. The pioneering contribution by Levmore (1985) conceptualizes restitution as a form of liability distinct from both contract (which presupposes mutual consent to voluntary exchange) and tort (i.e., involuntary exchange without mutual consent).<sup>8</sup> Restitution differs from contracts because there is no mutual consent. At the same time, restitution is about benefits and not harms or wrongdoing, and is therefore also distinct from tort law. Levmore (1985) argues that, from an economic viewpoint, restitution law should generally encourage parties to engage in voluntary exchange, rather than encouraging them to provide benefits unilaterally and thereafter seek an award of restitution. On the whole, Levmore (1985) views the law of restitution as restricting the circumstances in which restitution is awarded in a way that is broadly consistent with this normative prescription. A partially competing principle, however, focuses on the least cost avoider of the mistake; this sometimes leads to restitution being awarded (for instance, when a recipient could have detected the mistake at low cost, but did not).<sup>9</sup>

The intuitive idea that the dissimilarity between tort and restitution simply mirrors negative versus positive externalities is critically analyzed by Dari-Mattiacci (2009). He starts with the observation that while the law aims at internalizing negative externalities by means of general tort liability, no such approach is found for positive externalities. From an economic viewpoint, these situations are symmetric, and so individuals producing a positive externality should be paid

---

<sup>8</sup> Note, however, that there were some earlier contributions that addressed specific aspects of nonbargained benefits, such as the duty to rescue (e.g., Landes and Posner, 1978).

<sup>9</sup> Bouckaert and De Geest (1995) and Elkin-Koren and Salzberger (2000) follow Levmore’s (1985) approach by investigating specific aspects of restitution law. The former provides an economic justification for restitution liability in the context of coerced transfers (such as private takings, private taxes, and private compulsory services). The latter proposes an economic distinction between entitlements and remedies in restitution.

compensation (via “negative liability”). Dari-Mattiacci (2009) argues that the scope and design of restitution rules reflect asymmetries between the tort and restitution contexts with respect to intent, incentives, and evidence. These asymmetries imply that restitution is available in more limited circumstances than are tort damages.<sup>10</sup>

Cooter and Porat (2014, 2019) also highlight the contrast between the symmetry of positive and negative externalities in economic analysis and their generally asymmetric legal treatment. Their explanation for this divergence centers on the differences in the consequences for social welfare of liability in the tort and restitution contexts. In the former setting, failing to allow unilateral activity by injurers would create a holdout problem among potential victims, while an absence of liability would generate excessive harmful activity. In the latter setting, an absence of “negative liability” for benefits creates incentives for free riding by potential beneficiaries, but this is argued to be less socially costly than the potential harms in the tort setting. Cooter and Porat (2019) also make a normative argument that an expanded scope for restitution awards in certain circumstances would increase social welfare.<sup>11</sup>

The contributions discussed above analyze the law of restitution in general terms. Three papers in the law-and-economics literature analyze mistaken payments (the core example of restitution). Beatson and Bishop (1986) develop the fundamental point that a mistaken payment *per se* is a transfer and thus involves no social cost. However, parties may take precautions to avoid making mistakes; these precautions are socially wasteful because the mistaken payment is simply a transfer. Beatson and Bishop (1986) thus justify the award of restitution on the basis that in its absence payers would engage in excessive precaution from a social perspective. They argue that under a legal regime of full restitution, a mistaken payment potentially harms a good faith recipient if the recipient changes her expenditure or consumption patterns as a result of the payment and must then make full restitution. This “detrimental reliance” by the recipient is analogous to the injury suffered by a tort victim. They argue that a change of position defense (which denies full

---

<sup>10</sup> The approach taken by Dari-Mattiacci has generated an important strand of discussion on explaining the asymmetric legal treatment of positive and negative externalities (e.g., Sun and Brigham, 2014).

<sup>11</sup> A recent article by Liu, Avraham and Qiao (2021) considers the interaction between the prohibition of recovery for positive externalities and the incentives to acquire information about possible harms and benefits of activities. The authors note that there are substantially different economic consequences from *ex ante* (fixed) damages versus *ex post* (variable) damages. The reasoning relates to the prohibition of recovery from unrequested benefits truncating the distribution of potential harms (subject to damages) and benefits (outside of legal consequences) from activities. Ben-Shahar and Mikos (2005) analyze how recovery is determined in those circumstances where the law allows recovery for positive externalities, contrasting cost-based and benefit-based measures, and discussing why the law sometimes uses hybrid forms of recovery.

restitution in circumstances of detrimental reliance) can be justified as a means of forcing the payer to internalize the harm to the recipient from mistaken payments.

Huber (1988) focuses on the potential allocative inefficiency of transfers made as a result of mistakes, and argues that the recipient can in some cases take actions to induce the mistaken payment.<sup>12</sup> More recently, Gilboa and Kaplan (2018) develop the first formal model of the economics of the law of restitution. They challenge the tort analogy of Beatson and Bishop (1986) and argue for a regime of full rather than partial restitution. In particular, they argue that because of administrative costs and the possibility that some recipients are never traced, payers have incentives to undertake excessive precaution against mistakes even under a legal regime of full restitution. Allowing defenses that limit restitution would then induce even more excessive precaution. Their pioneering formal model represents an important advance. However, their analysis rests crucially on enforcement error. It does not address the more foundational question of the optimal design of the law of restitution under perfect enforcement, which is the primary focus of this paper. Our formal model is also more general, encompassing a wider set of potential circumstances (as detailed below).

In summary, there is only a very limited body of economic literature on the law of restitution, though it is an important and vibrant topic in the development of contemporary private law and in private law scholarship (especially in common law jurisdictions outside the United States). Perhaps most fundamentally, while identifying social costs of precaution by payers against mistaken payments and of detrimental reliance by the recipient, the existing economic literature does not directly characterize the social benefits of (non-mistaken) payments and hence the social benefits of avoiding mistakes. Thus, our aim is to develop a more general formal framework that can address basic questions about how to characterize these social costs and benefits, and to derive various novel implications for the design of the law of restitution.

Our article is also related – albeit more distantly – to several other strands of the law and economics literature.<sup>13</sup> Models of property rights in situations of sequential exchange (e.g., Arruñada, Zanarone and Garoupa, 2019) are arguably related to our analysis by a common

---

<sup>12</sup> From a comparative perspective, du Plessis (2010) discusses rules of restitution in common law and civil law systems by focusing on risk allocation between the recipient and transferor. The author recognizes that: “In this regard [restitution], it is problematic that use is made of [economic] models developed in the context of tort or delict, which evaluate behavior in the context of a single-loss creating event” (du Plessis, 2010, p. 34).

<sup>13</sup> Some scenarios in which mistaken payments arise are closely analogous to the problem of the good faith purchaser (e.g., Levmore, 1987; Medina, 2003). However, that setting differs from ours in a number of respects.

underlying intuition. In their setting of imperfect information about prior contractual arrangements, there is a tradeoff between the strength of the property rights of the seller (which promotes specialization between asset ownership and control via contracting) and the strength of the property rights of the buyer (which promotes impersonal exchange). Similarly, in our setting where mistaken payments are possible, there is a tradeoff between the certainty and finality of payments received by recipients and the entitlement of mistaken payers to restitution.

Our analysis also points – at least as a thought experiment - towards the decoupling of remedies. It is thus related to the economic literature on decoupling. Polinsky and Che (1991) show in the tort context that, increasing the amount paid by the defendant while lowering the amount received by the plaintiff may increase social welfare. The lower award received will reduce plaintiffs’ incentives to sue, thereby reducing litigation costs. However, the defendant’s expected damages can be held constant by raising the amount paid, so that the defendant’s choice of precaution is unaffected. The difference between what the defendant pays and what the plaintiff receives is paid to the state. Garoupa and Sanchirico (2010) argue that when decoupling is implemented in situations where the parties engage in transactions (such as for contractual breach), the amount paid to the state is analogous to a “transactions tax” that creates an inefficiency by disincentivizing some socially valuable transactions. Although we make a very different point – i.e., that mistaken payments can be analogized to a “transactions tax” – our formal model has some parallels with that of Garoupa and Sanchirico (2010).

#### **4) Model**

##### ***4.1) Basic Setup***

We assume two buyer-seller pairs. The first consists of B1 and S1, where B1 is the buyer and S1 the seller, while the second consists of B2 and S2 (defined similarly). Transaction costs are zero for transactions within each pair. Thus, B1 and S1 can engage in perfect Coasean bargaining, and decisions taken by each party are assumed to maximize the joint surplus of B1 and S1 (with a similar assumption applying to B2 and S2). However, transaction costs are assumed to be prohibitively high *across* the buyer-seller pairs. For example, B1 and S2 are assumed to be unable to bargain with each other. Therefore, B1 and S2 cannot engage in perfect Coasean bargaining. This latter assumption captures a critical feature of the settings in which the law of restitution applies – as in tort settings, at least some parties are necessarily “strangers” who cannot bargain

with each other and thereby determine via contract how mistaken payments will be allocated or addressed.

The set of assumptions above represents the most parsimonious setting in which mistaken payments involving parties who are not in a contractual relationship can arise. Figure 1 depicts the timing of the sequence of events that occur in the model. Each seller has a fixed, known cost  $c_i$  of producing a good or service ( $i = \{1, 2\}$ ). After being matched with a seller, each buyer draws a valuation  $v_i$  of the seller's good or service from the range  $[0, \bar{v}_i]$ . The probability density function (pdf) over the range of valuations is denoted  $h_i(v_i)$ , where it is assumed that  $h_i(v_i) > 0 \forall v_i$ ;  $H_i(v_i)$  is the corresponding cumulative density function (cdf). Once they are drawn, the valuations are observable to both the buyer and the seller. The buyer's valuation may or may not exceed the seller's cost; a transaction is efficient, of course, only in the former case. Once the valuation is known, each buyer-seller pair determines whether a transaction is mutually beneficial; if so, they contract to trade at a mutually agreed price  $R_i \in [c_i, v_i]$ .

For concreteness, within the perfect Coasean bargaining framework, we impose the assumption in most of our analysis that the buyer makes a take-it-or-leave-it offer, and also make a tie-breaking assumption that the seller accepts the offer when indifferent. These assumptions imply that  $R_i = c_i$ . We discuss below how  $R_i$  may be derived from a more general Nash bargaining framework; this, however, significantly complicates the model without generating any fundamentally different tradeoffs or insights.

In the course of making this payment, the buyer may mistakenly direct it to the wrong seller (e.g., B1 may pay S2 rather than S1). The probability of a mistake is denoted  $q_i$ ; it is initially assumed to be exogenous, but more generally depends on the level of precaution (denoted  $k_i$ ) against mistaken payments chosen by the buyer in the event that a contract is formed. Note that the timing assumed in Figure 1 entails that the buyer chooses  $k_i$  after the parties decide to contract.<sup>14</sup>

If a payment turns out to be mistaken, the consequences depend on the legal regime governing mistaken payments. Under a "no restitution" (NR) regime, the losses from the mistake

---

<sup>14</sup> This is equivalent to assuming that the buyer cannot commit at time 2 to the time-3 choice of precaution. However, the parties fully anticipate the buyer's optimal choice of precaution at time 3 when computing the expected joint surplus (and therefore in determining whether to transact) at time 2. Assuming instead a simultaneous choice of  $k_i$  and of whether to contract at time 2 would significantly complicate the model without generating any fundamentally different tradeoffs or any additional insights.

remain where they fall – e.g. if S2 receives a mistaken payment from B1, she retains the payment (and the assumption of prohibitively high transaction costs with B1 implies that she cannot return it voluntarily even if she wished to do so).<sup>15</sup> Under a “full restitution” (FR) regime, the mistaken payer is entitled to restitution of the full amount paid (thus, in the example above, B1 would be able to obtain restitution from S2). We assume a fixed administrative cost  $A$  of determining that a payment went astray, identifying the mistaken recipient, and obtaining restitution.<sup>16</sup> Note that B1’s mistake is “spontaneous” in that it is not induced in any way by S2 (indeed, our assumption of prohibitively high transaction costs across the buyer-seller pairs rules out any such inducement). S2 is assumed to receive the payment in good faith, being unaware that it was intended for S1.<sup>17</sup>

We do not make any particular assumptions about how B1 and S1 contractually address the possibility of mistaken payments by B1. It is possible that S1 transfers the good to B1 notwithstanding the failure to receive payment (thereby allocating the risk of mistakes to the seller), or that B1 is required to make an additional payment of the full price in order to receive the good (thereby allocating the risk of mistakes to the buyer). We impose only the requirement that a buyer-seller pair maximizes its joint surplus.

In the absence of mistakes ( $q_i \equiv 0$ ), this joint surplus can be expressed as follows, for a given realization of the buyer’s valuation  $v_i$ :

$$W_i(v_i) = \max\{0, v_i - c_i\} \quad (1)$$

Note that a contract is formed, given a realized valuation  $v_i$ , if and only if the joint surplus is positive (i.e. iff  $W_i(v_i) \geq 0$ , with a tie-breaking assumption that the parties transact when the joint surplus is zero). Ex ante social welfare – i.e. social welfare in expectation over all possible valuations – can thus be expressed as:

$$W_s^* = \int_{c_1}^{\bar{v}_1} (v_1 - c_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (2)$$

---

<sup>15</sup> Alternatively, an additional *ad hoc* mechanism (such as reputational losses) would be needed to explain why S2 could have an interest in returning the payment under NR. However, if such a mechanism exists and recipients always return mistaken payments to buyers, the legal regime governing restitution becomes less relevant.

<sup>16</sup> With ongoing technological developments, it is possible to argue that  $A$  is close to zero in the modern economy. We thus focus in the results below on the case of a sufficiently small  $A$ . However, our general expressions do not impose any restriction on the magnitude of  $A$ , and so our framework can accommodate different views with respect to this magnitude.

<sup>17</sup> Payments induced by the recipient in bad faith should, of course, be regulated by a different legal regime (under which S2 should be penalized). This set of circumstances is beyond the scope of our model (and of the basic law of mistaken payments).

$W_S^*$  is an important welfare benchmark in the analysis that follows. Following Figure 1, the model is solved by backward induction, starting from time 4 (the potential restitution stage) back to time 1 (matching of buyers and sellers).

## 4.2) Exogenous Mistakes

### 4.2.1) The Optimal Legal Regime with Exogenous Mistakes

The prior economic literature on mistaken payments (e.g., Beatson and Bishop, 1986; Gilboa and Kaplan, 2018) highlights the primary efficiency cost of mistakes as being a choice of precaution by payers that is excessive from society's perspective. To emphasize the difference in our conceptual framework, we begin with a setting in which the probability of mistakes is exogenously fixed; thus, precaution is unproductive, even from a private perspective. Specifically, we assume that  $q_1 \equiv \bar{q}_1 > 0$ . Without significant loss of generality, we set  $q_2 \equiv 0$ ; it is possible to allow for mistaken payments by B2 to S1 (as well as from B1 to S2), but this generates little additional insight.

Under an NR regime, the joint surplus of B1 and S1, given  $v_1$ , can be expressed as:

$$W_1(v_1) = \max\{0, v_1 - c_1(1 + \bar{q}_1)\} \quad (3)$$

The joint surplus, conditional on a contract being formed, is now lower by  $\bar{q}_1 c_1$  for any given  $v_1$  than it was in the absence of mistakes.<sup>18</sup> Thus, the threshold valuation at which a transaction is jointly worthwhile is correspondingly higher. On the other hand, S2 receives a mistaken payment with probability  $\bar{q}_1$  whenever B1 and S1 form a contract, regardless of whether B2 and S2 decide to transact. Thus, the decision of B2 and S2 as to whether to form a contract is not affected by the possibility of mistaken payments from B1.

Under NR, ex ante social welfare can be expressed as:

$$W_S^{NR} = \int_{c_1(1+\bar{q}_1)}^{\bar{v}_1} (v_1 - c_1(1 + \bar{q}_1))h_1(v_1)dv_1 + \int_{c_1(1+\bar{q}_1)}^{\bar{v}_1} (\bar{q}_1 c_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (4)$$

Here, the second term represents the mistaken payment received by S2. From the social planner's perspective, this simply offsets the cost to B1 and S1 from B1's mistaken payments. Thus:

---

<sup>18</sup> Note that this analytical formulation is derived under the assumption that the price mistakenly paid by B1 to S2 equals the cost borne by S1.

$$W_S^{NR} = \int_{c_1(1+\bar{q}_1)}^{\bar{v}_1} (v_1 - c_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (5)$$

Therefore, mistaken payments are not merely a matter of redistribution in our framework:<sup>19</sup>

**Remark 1:** Under NR, even when the probability of mistakes is exogenous, mistakes reduce social welfare: i.e.,  $W_S^{NR} < W_S^*$ .

The underlying intuition is that mistakes are analogous to a transactions tax that reduces the joint surplus of the buyer-seller pair; this implicit tax raises the threshold valuation required for a transaction to be (privately) jointly optimal. We develop later the idea that the deadweight loss from this implicit tax – i.e. the social value of the transactions that parties forego – can usefully be compared to the deadweight loss from (actual) taxation.

Under FR, B1 is able to obtain full restitution (at an administrative cost  $A$ ) whenever a mistaken payment occurs. Thus, assuming that  $A$  is sufficiently small that it is worthwhile to obtain restitution, the joint surplus of B1 and S1, given  $v_1$ , can be expressed as:

$$W_1(v_1) = \max\{0, v_1 - c_1 - \bar{q}_1 A\} \quad (6)$$

The joint surplus of B2 and S2 is given by Equation (1) – under an FR regime, any mistaken payments received by S2 are returned to B1 without any cost or benefit to S2 (more precisely, this is the case under our current assumptions; these are relaxed below when we consider the case of bilateral harm).

Under FR, ex ante social welfare can be expressed as:

$$W_S^{FR} = \int_{c_1+\bar{q}_1 A}^{\bar{v}_1} (v_1 - c_1 - \bar{q}_1 A)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (7)$$

Note that  $A$  is a social cost, as obtaining restitution entails the use of resources. Thus, the parameter  $A$  enters into the social welfare function.

**Remark 2:** With an exogenous probability of mistakes, FR is socially optimal when  $A$  is sufficiently small:  $\lim_{A \rightarrow 0} W_S^{FR} = W_S^* > W_S^{NR}$

Intuitively, an FR regime mitigates the loss of socially valuable transactions that would fail to occur under an NR regime. Thus, FR – at least under what might be viewed as typical circumstances where  $A$  is sufficiently small – is optimal even when precaution is fixed (so the social planner is not seeking to use restitution to prevent the payer from engaging in socially

---

<sup>19</sup> All proofs are in the Appendix.

excessive precaution). It is also important, however, to highlight the converse of Remark 2, namely that if  $A$  were to be sufficiently large, NR may be optimal (notwithstanding the “unjust enrichment” of S2).<sup>20</sup>

#### 4.2.2) More General Bargaining Assumptions

So far, the price  $R_1$  paid by B1 (and hence the amount of the potential mistaken payment to S2) has been assumed to be determined by a take-it-or-leave-it offer from B1 to S2 (so that  $R_1 = c_1$ ). This has the advantage of simplicity, characterizing the price straightforwardly in terms of one of the primitives of the model (the seller’s cost); it also implies that B1 fully internalizes the cost of mistaken payments. More generally, however, B1 and S1 may engage in other forms of bargaining and divide the joint surplus (characterized for instance in Equations (1) and (3)) more evenly.

For more general bargaining solutions, the parties’ choice of  $R_1$  may depend on the probability of mistaken payments and on the legal regime governing restitution. Suppose that B1 receives a share  $\beta_1 \in [0,1]$  of the joint surplus, and that the outside options are zero for both B1 and S1. Thus, whenever a contract is formed under NR, S1 receives a share  $(1 - \beta_1)(v_1 - c_1 - \bar{q}_1 R_1)$  of the surplus, net of her cost  $c_1$ :

$$R_1 - c_1 = (1 - \beta_1)(v_1 - c_1 - \bar{q}_1 R_1) \quad (8)$$

Rearranging this expression:

$$R_1^{NR} = \frac{(1 - \beta_1)(v_1 - c_1) + c_1}{1 + (1 - \beta_1)\bar{q}_1} \quad (9)$$

The analogous expression under FR is:

$$R_1^{FR} = (1 - \beta_1)(v_1 - c_1 - \bar{q}_1 A) + c_1 \quad (10)$$

Note that the price is higher the lower the probability of mistakes  $\bar{q}_1$ . When we endogenize this probability below, this would imply that greater precaution by B1 against mistakes - while lowering the probability of a mistake - also tends to increase  $R_1$ . This moderates the gain from increased precaution - i.e., there is some leakage of this gain due to the higher loss in the event

---

<sup>20</sup> For NR to be optimal, it would be necessary for  $A$  to be larger than  $c_1$  (or, in more general bargaining environments, to be larger than  $R_1$ ). This seems somewhat unlikely in view of the current state of technology for making and tracing payments. However, it is a possibility that our framework can accommodate. Moreover, even if that scenario is unlikely, it is important to note that restitution is justified here by the magnitude of  $A$  rather than by the “unjust enrichment” of S2.

that there is a mistaken payment.<sup>21</sup> However, this does not fundamentally affect the qualitative conclusions below, or the basic nature of the tradeoff involved. Thus, we continue to use the take-it-or-leave-it assumption (so that  $R_1 = c_1$ ) in the baseline analysis below for simplicity.

### 4.3) Unilateral Harm and Unilateral Precaution

#### 4.3.1) Endogenizing the Probability of Mistakes

In this subsection, we maintain the assumption of unilateral harm – i.e., the assumption that while (absent full restitution) a mistaken payment harms B1, full restitution does not harm S2. However, we now endogenize the probability of mistakes by allowing B1 to choose a level of precaution against mistakes. As shown in Figure 1, this choice is made at time 3, if the parties have decided to form a contract at time 2. The choice of precaution is perfectly anticipated at time 2, and so is reflected in the parties' expectations about the joint surplus. The probability of a mistake is now  $q_1(k_1)$ , where  $k_1$  is B1's cost of precaution (note that, for sake of exposition, although precaution is chosen by B1, it is chosen so as to maximize the joint surplus of B1 and S1).<sup>22</sup> We assume that  $q_1(k_1)$  is decreasing and concave (i.e.  $q_1'(k_1) < 0$  and  $q_1''(k_1) > 0$ ), bounded above (i.e.  $\lim_{k_1 \rightarrow \infty} q_1(k_1) = \bar{q}_1 < 1$ ), and strictly positive for any finite level of precaution (i.e.  $q_1(k_1) > 0 \forall k_1$ ), with  $\lim_{k_1 \rightarrow \infty} q_1(k_1) = 0$ . We also assume that  $\lim_{k_1 \rightarrow 0} q_1'(k_1) = -\infty$ .

Under an NR regime, the joint surplus of B1 and S1, given  $v_1$ , can be expressed as:

$$W_1(v_1) = \max\{0, v_1 - c_1(1 + q_1(k_1)) - k_1\} \quad (11)$$

The jointly optimal choice of  $k_1$ , conditional on forming a contract, is found by maximizing  $v_1 - c_1(1 + q_1(k_1)) - k_1$ , or equivalently by solving (for a given  $v_1$ ):

$$\min_{k_1} \{c_1(1 + q_1(k_1)) + k_1\} \quad (12)$$

The solution is given by the FOC  $-q_1'(k_1)c_1 = 1$ . Recall the timing assumption shown in Figure 1 that  $k_1$  is chosen at time 3, after the parties have chosen whether to contract or not at time 2 (and without any commitment at time 2 to a choice of precaution). This implies that the optimal choice

<sup>21</sup> It might seem that setting  $R_1 = c_1$  would be jointly optimal for B1 and S1 in order to minimize their joint loss in the event of a mistaken payment (while satisfying S2's participation constraint). This point has some validity, but in a more general bargaining setup may require a side-payment from B1 to S1 to make up the difference between  $c_1$  and the payment demanded by S2. This side-payment may potentially be subject to the possibility of being misdirected to S2, thus replicating the original problem.

<sup>22</sup> Under our assumptions, this observation follows directly from the fact that the buyer receives the entire surplus from the contractual transaction.

of  $k_1$  depends only on the price (which under our bargaining assumption is the exogenous seller cost, and which determines how much is lost due to a mistake under NR) and the function  $q_1(k_1)$  that specifies the probability of a mistake. It does not depend on the realized valuation  $v_1$ .<sup>23</sup>

S2 receives a mistaken payment with probability  $q_1(k_1)$  whenever B1 and S1 form a contract, regardless of whether B2 and S2 decide to transact. This probability is not influenced by the choices of B2 and S2, and so the decision of B2 and S2 as to whether to form a contract is not affected by the possibility of mistaken payments from B1. Under NR, ex ante social welfare can be expressed as:

$$W_S^{NR} = \int_{c_1(1+q_1(k_1))+k_1}^{\bar{v}_1} (v_1 - c_1 - k_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (13)$$

As before, the expected mistaken payment cancels out in the social welfare function. However, the cost of precaution entails the use of real resources and so enters into the social welfare function.

Under FR,  $W_1(v_1)$  is analogous to that in Equation (11). The jointly optimal choice of  $k_1$ , conditional on forming a contract, is found by solving

$$\min_{k_1} \{c_1 + q_1(k_1)A + k_1\} \quad (14)$$

Thus, this optimal choice is given by the FOC  $-q_1'(k_1)A = 1$ . Ex ante social welfare can be expressed as:

$$W_S^{FR} = \int_{c_1+q_1(k_1)A+k_1}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1)A - k_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (15)$$

Again, as  $A$  is a social cost, it enters into the social welfare function.

**Remark 3:** With unilateral harm and unilateral precaution, FR is optimal when  $A$  is sufficiently small:  $\lim_{A \rightarrow 0} W_S^{FR} = W_S^* > W_S^{NR}$

In terms of intuition, FR continues to be optimal (for sufficiently small values of  $A$ ) when we endogenize the probability of mistakes. In this setting, there are two distinct social costs that can be mitigated by an FR regime – the loss of transactions between B1 and S1, and the cost of precaution by B1. Thus, the case for full restitution is strengthened under these assumptions, relative to those that generated Remark 2.

---

<sup>23</sup> Allowing for a simultaneous joint choice at time 2 by B1 and S1 of  $k_1$  and of whether to transact would add significantly to the model's mathematical complexity, without generating additional insights.

Introducing the possibility of precaution by S2 – for instance, in the form of auditing incoming payments to verify the intended recipient –, a point we will discuss in more detail later, does not fundamentally affect the results in a setting of unilateral harm. S2 has no reason to choose a strictly positive level of precaution, as she faces no net harm from being the recipient of a mistaken payment (i.e., paying restitution leaves her no worse off than never having received the payment). In addition (as described more fully in subsection 4.5 below) S2’s precaution does not affect the probability that restitution will occur when the law requires it (i.e., it does not create enforcement error). Thus, FR will be optimal under the same conditions specified in Remark 3.

### 4.3.2) The Change of Position Defense

So far, S2 has played a purely passive role as a potential (good faith) recipient of a mistaken payment. However, some doctrines in the law of restitution, notably the change of position defense, hinge on actions taken by S2. To analyze the consequences of the change of position defense, we now allow S2 to choose an allocation of the mistaken payment (if it occurs) between cash ( $y$ ) and luxury consumption ( $(R_1 - y)$ , or more specifically under our bargaining assumption  $(c_1 - y)$ ). Suppose that S2 has a quasi-linear utility function  $y + u(c_1 - y)$ , where  $u(c_1 - y) > 0$  for  $c_1 - y > 0$ ,  $u'(c_1 - y) > 0$  and  $u''(c_1 - y) < 0$ .

Under NR, S2’s valuation of the expected mistaken payment is:

$$q_1(k_1)[y + u(c_1 - y)] \tag{16}$$

This is maximized by setting  $y = y^* \equiv \operatorname{argmax}\{y + u(c_1 - y)\}$ . An interior solution for  $y^*$  is defined by the FOC  $u'(c_1 - y^*) = 1$ . In general, a corner solution where  $y^* = 0$  or  $y^* = c_1$  cannot be ruled out. However, if  $c_1$  is sufficiently large, then the marginal utility from luxury consumption may be expected to fall sufficiently to ensure an interior solution. Thus, we assume that for sufficiently large  $c_1$ ,  $u'(c_1) < 1$  (so that S2 will choose  $y^* > 0$ ).

We introduce the notation  $W_S^{NRu}$  to denote social welfare under this new set of assumptions about S2’s utility from the mistaken payment. We can express  $W_S^{NRu}$  as follows:

$$\begin{aligned} W_S^{NRu} = & \int_{c_1(1+q_1(k_1))+k_1}^{\bar{v}_1} (v_1 - c_1(1 + q_1(k_1)) - k_1)h_1(v_1)dv_1 \\ & + \int_{c_1(1+q_1(k_1))+k_1}^{\bar{v}_1} q_1(k_1)[y^* + u(c_1 - y^*)] h_1(v_1)dv_1 \\ & + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \end{aligned} \tag{17}$$

Here, the first term includes the expected private loss to B1 and S1 from the mistaken payment ( $c_1 q_1(k_1)$ ). This is offset by the expected gain to S2 (in utility terms, using the maximized utility from the mistaken payment), which is represented by the second term in Equation (17). Note that, unlike in our baseline framework, the valuation of a given amount of money may potentially differ between B1 and S2 (i.e.,  $y^* + u(c_1 - y^*)$  may not equal  $c_1$ ) and so there may be a direct social gain or loss from the transfer. The latter is not a general feature of our framework (as, in general, there seems no particular reason to assume that the recipient values the money any more or less than the payer), but is a consequence here of allowing S2 to make consumption choices. Treating B1 symmetrically by positing a similar utility function for B1 would not fundamentally affect the results derived below.

Now, we introduce a legal regime of full restitution with a change of position defense (denoted  $\Delta P$ ). We characterize this regime in very simple terms as follows: suppose that when the change of position defense is available, the amount S2 is required to pay in restitution is  $y$ . Of course, this is a rather extreme formulation of the doctrine, which has many significant nuances. However, scholars of the law of restitution tend to view consumed goods as a particularly appropriate example of the circumstances in which the change of position defense applies (e.g., Bant, 2009:2). Under this assumption, S2's valuation of the mistaken payment will be  $q_1(k_1)[u(c_1 - y)]$ , as any amount held as cash will be subject to restitution. Given the assumption that  $u(c_1 - y) > 0$  for  $c_1 - y > 0$ , S2 will choose the corner solution  $y^* = 0$  under  $\Delta P$ .

There may appear to be some tension between the assumption that S2 is a good faith recipient and the conclusion that S2 will alter her consumption choices to minimize the amount paid in restitution. This apparent tension is largely an artifact of our simple setting with only two buyer-seller pairs. More generally, we might imagine that S2 receives payments (some of which are mistaken) from a large number of payers. S2 would then know that, for instance, 10% of payments she receives are mistaken and subject to restitution, but would not know *which* payments are mistaken. This scenario would create an incentive to choose a higher level of luxury consumption than would be the case absent mistaken payments (albeit not to the same extent as depicted in our corner solution above).

Social welfare under  $\Delta P$  is denoted by  $W_s^{\Delta P}$  and can be expressed as follows:

$$W_S^{\Delta P} = \int_{c_1(1+q_1(k_1))+k_1}^{\bar{v}_1} (v_1 - c_1(1 + q_1(k_1)) - k_1)h_1(v_1)dv_1 + \int_{c_1(1+q_1(k_1))+k_1}^{\bar{v}_1} q_1(k_1)u(c_1) h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (18)$$

Under  $\Delta P$ , B1 fully anticipates S2's optimal choice of  $y^* = 0$  and thus expects not to receive restitution in the event of a mistaken payment. Thus, B1's choice of precaution  $k_1$ , and the conditions under which B1 and S1 choose to transact, are identical to those under NR; this is reflected in the first term of Equation (18) being identical to the first term in Equation (17). The only difference between Equations (17) and (18) is thus in the second term, which represents the expected (maximized) utility that S2 derives from the mistaken payment. This observation leads to the following result:

**Remark 4:** Under the assumptions above, social welfare under a regime of full restitution with a change of position defense is weakly lower than under NR:  $W_S^{\Delta P} \leq W_S^{NRu}$ . If  $c_1$  is sufficiently large that  $u'(c_1) < 1$ , social welfare under a regime of full restitution with a change of position defense is strictly lower than under NR:  $W_S^{\Delta P} < W_S^{NRu}$ .

The underlying intuition is that, under the assumptions made above, the  $\Delta P$  regime distorts S2's choice of consumption patterns without creating any advantage for B1 relative to NR. This intuition is more general, and extends beyond our specific set of assumptions. For any given fraction of the mistaken payment that B1 can recover, the social planner will prefer to set a fixed fractional recovery (independent of S2's actions) rather than achieve the same fractional recovery via a change of position defense. The reason is that B1 – whose choices of precaution and of whether to contract with S1 depend only on the fraction of a mistaken payment that is recovered – is indifferent between these alternatives, while S2's consumption choices are distorted under  $\Delta P$  but not under a fixed fractional recovery. The intuition underlying Remark 4 is thus quite general, and points to an irony. A change of position as a result of a mistaken payment is arguably quite unlikely in most circumstances (except where the payment is very large in relation to the recipient's total wealth).<sup>24</sup> Nonetheless, the existence of the change of position defense may induce recipients to change their position (in order to reduce the amount of restitution paid if the payment

---

<sup>24</sup> An extreme – but quite important – example of a scenario in which we would expect no excess burden is that of tax overpayments, where it is implausible that the (large and well-diversified) government would experience any excess burden from returning overpaid taxes to taxpayers (see e.g., Nussim, 2010, for an economic analysis of one aspect of this issue).

turns out to be mistaken). However, it should be noted that in a more realistic scenario in which S2 receives mistaken payments in good faith along with a large number of other (non-mistaken) payments, the distortion to S2's consumption patterns will be a function of the (typically relatively small) probability that any given payment received is mistaken.

#### ***4.4) Bilateral Harm and Unilateral Precaution***

##### **4.4.1) Basic Scenario**

In the basic framework developed in subsection 4.3.1 above, the harm to S2 from paying restitution is assumed to be equal to the amount paid; therefore, it is exactly offset by the amount received in the form of a mistaken payment. However, a common argument in the literature on the law of restitution is that recipients suffer some net harm from the mistaken payment and its (full) restitution (e.g., Bant, 2009, p. 2). For instance, believing in good faith that she is entitled to the payment, the recipient may alter her pattern of consumption towards luxury goods. Having committed to a higher level of luxury consumption than she otherwise would have chosen, she then finds upon making restitution that her ability to consume essential goods is constrained. This leaves her with a lower level of utility than she would have enjoyed had the mistaken payment never been made.

To formally represent this possibility of bilateral harm, we assume that S2 experiences an “excess burden” – i.e., a cost that exceeds the amount originally received - of making restitution. We also generalize the set of available legal regimes to encompass partial restitution (PR). Let the amount of restitution that is required to be paid be  $\alpha c_1$ , where  $\alpha \in [0,1]$ , and define PR as any regime with  $\alpha \in (0,1)$ . For instance, recall the example of a recipient who receives £500 by mistake and commits £200 to purchasing a luxury good. This leaves £300 for essential goods, but the recipient now learns that she must make restitution of £250 (supposing that  $\alpha = 0.5$ ). This reduces her consumption of essential goods by £250 with an associated utility loss that we denote by the function  $f(250)$  or more generally  $f(\alpha c_1)$ . The harm experienced by S2 under PR is thus:

$$\alpha c_1 - c_1 + f(\alpha c_1) = -(1 - \alpha)c_1 + f(\alpha c_1) \quad (19)$$

The first term on the left-hand side of Equation (19) represents the amount of restitution paid, the second term is the money originally received by mistake, and the third term is the excess burden  $f(\alpha c_1) > 0$  of paying  $\alpha c_1$  in restitution. It is assumed that  $f(\cdot)$  is increasing and convex – i.e.,

$f'(\cdot) > 0$  and  $f''(\cdot) < 0$ . In addition, it is assumed that  $f(0) = 0$  and that  $f(\cdot) > 0$  for any strictly positive argument.

Note that an alternative interpretation of the excess burden is as a cost of risk-bearing. In a more realistic scenario, S2 would receive mistaken payments in good faith along with other (non-mistaken) payments and would be unable to distinguish between them. Then, the possibility that some incoming payments are mistaken and will be subject to restitution creates uncertainty that imposes a risk-bearing cost on (a risk-averse) S2.

The formulation above formalizes how the harm to the recipient is widely conceptualized among scholars of restitution. Note, however, that it is asymmetric, in the sense that an excess burden is assumed to exist only for the recipient. However, under an NR or PR regime, B1 may also be argued to experience an excess burden, believing in good faith that her obligation to S1 has been paid but then having to adjust her consumption patterns to make another payment to S1 that is not fully offset by the amount received in restitution. For example, suppose that B1 pays £500 intended for S1 to S2. Believing that her obligation to S1 has been discharged, she commits to £200 of luxury consumption. Then, she learns that the payment was misdirected to S2, and (while, with  $\alpha = 0.5$ , she receives £250 in restitution from S2) she still owes S1 £500. She must thus reduce her consumption of essential goods by £250 in order to pay S1. The overall harm experienced by B1 is:

$$-\alpha c_1 + c_1 + f((1 - \alpha)c_1) = (1 - \alpha)c_1 + f((1 - \alpha)c_1) \quad (20)$$

The first term on the left-hand side of Equation (20) represents the amount of restitution received by B1, the second term is the money originally paid by mistake, and the third term is the excess burden  $f((1 - \alpha)c_1) \geq 0$  of failing to receive  $(1 - \alpha)c_1$  in restitution.

Thus, a symmetric formulation entails that B1 experiences an excess burden of  $f((1 - \alpha)c_1) \geq 0$ . As there is no *a priori* reason to expect that this excess burden will be higher or lower for B1 relative to S2, we assume the same function  $f(\cdot)$  as characterized above. The excess burden could alternatively be interpreted as a risk-bearing cost, as the possibility of mistaken payments combined with the lack of full insurance under a PR regime creates uncertainty for B1. Again, however, there is no reason to assume that risk preferences differ across parties.

Social welfare under PR is:

$$W_s^{PR} = \int_{c_1^{PR}}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1)[A + f((1 - \alpha)c_1) + f(\alpha c_1)] - k_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (21)$$

where the lower limit of integration in the first term is:

$$c_1^{PR} = c_1 + q_1(k_1)(A + (1 - \alpha)c_1 + f((1 - \alpha)c_1)) + k_1 \quad (22)$$

It is assumed that, since the administrative cost  $A$  is a fixed cost, it is incurred in full even when restitution is partial. That is, B1 faces the same cost of identifying a mistaken payment and its recipient and of obtaining restitution as under FR (as reflected in the formulation in Equation (21)). However, the results below would not be significantly affected if the administrative cost were instead to be assumed to be some fraction of  $A$  under PR.

Note that as  $\alpha$  approaches 1, the legal regime converges to FR. We introduce the notation  $W_s^{FR} = \lim_{\alpha \rightarrow 1} W_s^{PR}$  to denote social welfare under FR in the bilateral harm setting (i.e., where an excess burden exists). We also denote by  $k_1^{FR}$  B1's (privately) optimal choice of  $k_1$  under an FR regime. This choice is unaffected by the excess burden, as under FR that is borne entirely by S2. Thus, this privately optimal choice  $k_1^{FR}$  given by the FOC  $-q_1'(k_1^{FR})A = 1$  (as in the unilateral harm setting). Analogously, we can define  $W_s^{NRe} = \lim_{\alpha \rightarrow 0} W_s^{PR}$  to denote social welfare under NR in the bilateral harm setting, and denote by  $k_1^{NRe}$  B1's (privately) optimal choice of  $k_1$  under an NR regime. This choice is affected by the excess burden, as under NR that is borne entirely by B1. Thus, this privately optimal choice  $k_1^{NRe}$  given by the FOC  $-q_1'(k_1^{NRe})[c_1 + f(c_1)] = 1$ .

The symmetric formulation in Equations (19) and (20) entails that the existence of an excess burden *per se* cannot justify PR. If we start with an FR regime and lower  $\alpha$ , S2's excess burden will fall, but B1's excess burden will increase. Rather, it is the convexity of the excess burden that implies the optimality of PR under certain conditions. In particular, the convexity of  $f(\cdot)$  entails that (all else equal) the aggregate excess burden will fall when it is divided between B1 and S2 rather than placed entirely on S2. Proposition 1 makes this point more precise for a sufficiently small change in  $\alpha$ , starting from an FR regime with  $\alpha = 1$ :

**Proposition 1:** Under the assumptions of bilateral harm and unilateral precaution, as specified above:

- i) There exists some  $\alpha$  sufficiently close to 1 such that social welfare is higher under PR than under FR for any convex excess burden  $f(\cdot)$ . In particular, there exists some  $\alpha < 1$  such that  $W_s^{PR} > W_s^{FR}$ ,
- ii) There exists some  $\alpha$  sufficiently close to 0 such that social welfare is higher under PR than under NR for any convex excess burden  $f(\cdot)$ . In particular, there exists some  $\alpha > 0$  such that  $W_s^{PR} > W_s^{NR}$ .

Importantly, this argument holds only for a sufficiently small variation in  $\alpha$ . For instance, part (i) applies only to an infinitesimal reduction in  $\alpha$  below 1.<sup>25</sup> For larger changes, it is not possible to hold  $k_1^{FR}$  fixed. Instead, B1 will re-optimize her choice of precaution (and the set of transactions she enters into with S1). There would then be a tradeoff between the gains from sharing the excess burden (or sharing risk, in an alternative but equivalent interpretation) and the social losses from excessive precaution and foregone transactions. It is thus unclear how a large change in  $\alpha$  would affect social welfare. However, our framework identifies the factors that play a role in determining the efficiency of the alternative legal regimes – the social gains from sharing the excess burden, and the social costs of precaution and foregone transactions.

#### 4.4.2) Decoupling Restitution

Under a PR regime, there exist three types of social losses relative to social welfare in “no-mistake” setting (as characterized in Equation (2)). First, there are some foregone transactions between B1 and S1, even as  $A$  tends towards zero, because a component  $(1 - \alpha)c_1$  of the mistaken payment is not paid in restitution. For the same reason, there is a higher level of precaution by B1 (relative to the choice under FR or when mistakes are impossible). Finally, there are losses from the excess burdens  $f(\cdot)$  borne by B1 and S2. Nonetheless, a PR regime may be optimal among the class of restitution regimes that satisfy the property that may be termed “budget balance” (e.g. Dharmapala and Hoffmann, 2005) - i.e. that what the defendant pays in restitution is what the plaintiff receives (apart from litigation costs). Budget balance is a widespread – indeed, virtually universal - feature of private law remedies, although there has been some theoretical discussion of “decoupling” what defendants pay and what plaintiffs receive (e.g., Polinsky and Che, 1991; Garoupa and Sanchirico, 2010).

---

<sup>25</sup> Note that the case of bilateral harm with an exogenous probability of mistakes is simply a special case of the scenario analyzed above, with  $q_1$  being fixed. Thus, the result in Proposition 1 applies.

If we relax the “balanced-budget” assumption, then it is possible to imagine legal regimes that are more efficient than PR. As discussed above, (uncompensated) mistakes create a transactions tax and generate other sources of deadweight loss. Suppose, then, that lump sum (nondistortionary) taxes exist. With access to nondistortionary financing, the social planner will wish to eliminate the distortions due to mistakes by using the revenue from lump sum taxes to provide subsidies to B1 and S1 that offset the transaction tax that they face due to mistaken payments, and to S2 to offset the excess burden she faces from full restitution. Of course, lump sum taxes do not exist in reality. However, this intuition extends to situations in which taxes are distortionary, but less so than mistakes. If the distortions from taxes are sufficiently large, then the social planner will prefer to allow the distortions due to mistakes to occur.

Let the amount of restitution that is paid to B1 be  $\alpha_B c_1$ , and the amount paid in restitution by S2 be  $\alpha_S c_1$  (with the difference between these being made up by lump sum taxes or subsidies). Then, social welfare under a decoupled regime (denoted  $W_S^D$ ) can be characterized as follows:

$$W_S^D = \int_{c_1^D}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1)[A + f((1 - \alpha_B)c_1) + f(\alpha_S c_1)] - k_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (23)$$

where the lower limit of integration in the first term is:

$$c_1^D = c_1 + q_1(k_1)(A + (1 - \alpha_B)c_1 + f((1 - \alpha_B)c_1)) + k_1 \quad (24)$$

A fully decoupled (FD) regime sets  $\alpha_B = 1$  and  $\alpha_S = 0$ . Social welfare under full decoupling and lump sum taxes (denoted  $W_S^{FD}$ ) can be characterized as:

$$W_S^{FD} = \int_{c_1 + q_1(k_1)A + k_1}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1)A) - k_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (25)$$

An immediate and obvious challenge for such a decoupled regime is that there would exist an incentive for parties to collude and fabricate mistaken payments in order to receive what amounts to a joint subsidy from the government. We discuss this issue in more detail in Section 5. For now, it suffices to note that the underlying assumption of prohibitively high transaction costs between B1 and S2 essentially rules out this type of collusion, as do the implicit assumptions in the prior analysis that courts can verify that the mistake is spontaneous, and that the payment is

received in good faith. The possibility of collusion when these assumptions do not hold is discussed in Section 5.

Proposition 2 establishes the social optimality of a fully decoupled regime under the conditions described above:

**Proposition 2:** Assume a setting of bilateral harm and unilateral precaution, in which lump sum taxes exist and prohibitively high transactions costs between B1 and S2 prevent collusive behavior. Then, a decoupled regime that sets  $\alpha_B = 1$  and  $\alpha_S = 0$  is socially optimal, and superior to both PR and FR, when  $A$  is sufficiently small:  $\lim_{A \rightarrow 0} W_S^{FD} = W_S^* > \lim_{A \rightarrow 0} W_S^{PR}$  and  $\lim_{A \rightarrow 0} W_S^{FD} = W_S^* > \lim_{A \rightarrow 0} W_S^{FR}$ .

The intuition underlying Proposition 2 is the following. The FD rule is equivalent to a FR regime from the perspective of B1 and S1. Thus, B1 has no incentive to undertake socially excessive precaution (in particular, B1 will only take account of the social cost  $A$  of obtaining restitution, and precaution will approach zero as  $A$  becomes very small). Moreover, B1 and S1 jointly have no reason to forgo socially valuable transactions (for which  $v_1 > c_1 + q_1(k_1)A$ , or  $v_1 > c_1$  as  $A$  becomes very small). At the same time, the FD rule is equivalent to an NR regime from the perspective of S2. Thus, S2 does not experience any excess burden from making restitution. Decoupling thus allows us to reach the social optimum along multiple dimensions. It can be viewed as an application of a very basic and general economic principle (e.g., Tinbergen, 1952) – namely, that achieving multiple objectives (in this instance, neutralizing the transaction tax faced by B1 and S1, and mitigating the excess burden faced by S2) requires the use of multiple instruments (in this case, FR for the payer and NR for the recipient).

Note that the retention by S2 of a mistaken payment does not create a social cost, notwithstanding the “unjust enrichment” of S2, under the assumption that S2 is a good faith recipient who has no available means to influence the probability of receiving a mistaken payment. The use of general tax revenue to compensate B1 for the loss from the mistake likewise does not create a social cost under the assumption of lump sum taxation.

#### ***4.5) Bilateral Harm and Bilateral Precaution***

Previously, under unilateral harm and unilateral precaution, we derived Remark 3 – when  $A$  is sufficiently small, FR is optimal. In the following subsection, under bilateral harm and unilateral precaution, we proved Proposition 1 – in general, PR is better from a social welfare

perspective than either FR or NR. In this subsection, we consider the case of bilateral harm and bilateral precaution. We present a verbal discussion rather than a full formal proof since the main conclusions can be easily derived by extending previous results.

Each party seeks to choose individual precaution to limit exposure to mistaken payments. B1 invests in precaution to minimize the expected costs from engaging in a mistaken payment to S2, as modeled in subsection 4.3. At the same time, S2 invests in precaution to minimize the expected costs from the “excess burden” induced by a potential mistaken payment from B1. For instance, S2 may audit incoming payments to determine if they are truly intended for S2. These audits would have some probability of correctly identifying mistaken payments, which S2 would then immediately return to B1. These returned payments would have no impact on S2’s consumption or expenditure patterns, and so would avoid the excess burden.

More explicitly, we can envisage that the probability of a mistaken payment from B1 to S2 taking place is generally captured by  $q_1(k_1, k_2)$ . We can reasonably assume that this function is decreasing in each argument - i.e., the probability of a mistaken payment (or of incurring a cost from a mistake) is lowered by either party’s precaution. However, in general, the productivity and cost of each party’s precaution may differ, and precaution by the two parties are not necessarily perfect substitutes.

Both parties, B1 and S2, choose individual precautions against mistakes at time 3 (in the timeline illustrated in Figure 1). Therefore, they can be modeled as playing a Nash-Cournot game. In other words, when minimizing expected losses from mistaken payments, each individual derives a reaction function, that is,  $k_1^*(k_2)$  and  $k_2^*(k_1)$ . Since these individual precautions reduce the probability of mistaken payments, the intuitive assumption is that  $k_1$  and  $k_2$  are strategic substitutes, i.e., the reaction functions both have a negative slope (so that B1 engages in less precaution when she anticipates that S2 will engage in more precaution). Technically, this requires imposing the assumption that the second-order cross-derivative is positive. Under these conditions, the standard conditions for the existence and uniqueness of a Nash-Cournot equilibrium apply. Thus, a unique Nash-Cournot equilibrium exists, with the parties’ precaution choices defined by  $k_1^*(k_2^*)$  and  $k_2^*(k_1^*)$  being mutual best responses.

Bilateral precaution generally reinforces the insights from Proposition 1 regarding the optimality of PR. However, there is now an additional reason to favor PR (beyond the social gain from sharing the excess burden), based on the precaution technology of each party. Consider, for

instance, starting with an FR regime. Moving away infinitesimally from this is optimal, by Proposition 1, because of the convexity of the excess burden. Suppose, in addition, that B1's precaution technology is significantly more effective than S2's precaution technology.<sup>26</sup> Then, reducing  $\alpha$  below 1 has the added benefit of inducing more precaution by B1 and less by S2 (thereby lowering the overall social cost of reducing  $q_1(k_1, k_2)$  by a given amount). A "large" reduction in  $\alpha$  – moving towards an NR regime - may be optimal, to the extent that this social gain (together with the benefits of sharing the excess burden) outweighs the social cost of foregone transactions.

Consider the opposite scenario, i.e., B1's precaution technology is significantly less effective than S2's precaution technology. In this case, incentivizing S2 to take more precaution than B1 would be optimal. Approximating a FR regime would then be desirable, since it imposes relatively more costs on S2 and thereby substitutes S2's precaution for that of B1. Therefore, a PR regime should be responsive to the effectiveness of the precaution technology, not simply to the convexity of the excessive burden as in Proposition 1. In particular, the optimal restitution regime should take into account which precaution technology is more "wasteful" from a social welfare viewpoint. Importantly (and unlike in the unilateral harm and unilateral precaution scenario where Remark 3 was derived), it is not possible to argue that FR is always preferable to NR. If PR is unavailable, for instance because it is too complex to administer, the choice between FR and NR would ultimately depend on the effectiveness and cost of each party's precaution technology.

The discussion above presumes the unavailability of an FD regime. Note, however, that even when the PR regime induces the optimal allocation of precaution across parties and optimally shares the excess burden, some excess burden will nonetheless be incurred in equilibrium. In addition, there will be some foregone socially valuable transactions under the PR regime. If lump sum taxes and an FD regime are available, then these social costs can be eliminated. The argument here is essentially identical to that in Proposition 2 – an FD regime functions as an FR regime from the perspective of B1 and so eliminates excessive precaution and ensures that B1 and S1 engage in all socially valuable transactions. At the same time, an FD regime functions as an NR regime from the perspective of S2, and so eliminates S2's auditing of incoming payments. Note that S2's

---

<sup>26</sup> For example, in the *Lipkin Gorman* scenario, it might be the case that it is very costly for the casino to investigate the provenance of the funds used by gamblers, whereas it is less costly for the law firm to monitor or deter embezzlement by its partners.

excess burden is a social cost, and so S2's precaution is socially desirable, conditional on B1's precaution. However, the FD regime eliminates S2's excess burden, and so there is no role for precaution by S2.

## 5) Discussion

### *5.1) A Review of the Main Results*

Table 1 represents the various possible scenarios and the corresponding optimal legal regime. In general, it is clear from Remark 3 that FR is socially optimal whenever harm is unilateral – i.e., when a mistaken payment imposes only on the payer (absent restitution). This is true regardless of whether precaution is fixed, unilateral or bilateral. The intuition is that an FR regime mitigates both sources of inefficiency arising from mistaken payments in this setting (absent restitution) – the incentive for the payer to take socially excessive precaution, and the loss of transactions that the payer foregoes because of the reduced joint surplus. These inefficiencies can be corrected by FR without imposing any cost on the recipient (who is no worse off than if she never received the mistaken payment). If the administrative cost  $A$  is not close to zero, there will remain some distortions along both dimensions relative to a world without mistakes; however, these are inescapable because implementing FR is not costless. As  $A$  approaches zero, these distortions disappear. In the unilateral harm setting, there is thus no advantage to an FD regime (even assuming this to be feasible) because FR already achieves the social optimum (for sufficiently small  $A$ ).

When harm is bilateral (i.e., the recipient would suffer some net harm from receiving the payment and making full restitution), our results imply that a regime of PR is socially optimal (as shown in Table 1). Again, this is independent of whether precaution is fixed, unilateral, or bilateral, though the factors at play differ across these settings. When precaution is fixed or unilateral, the driving force is the convexity of the excess burden. For instance, a small departure from FR will result in the parties sharing the excess burden (or, in an alternative interpretation, sharing risk) and will increase social welfare (as shown in Proposition 1). When precaution is bilateral, there is an added impetus towards PR in order to encourage greater precaution by the party with access to a superior precaution technology (and this can justify substantial departures from FR or NR).

While PR is generally optimal in these conditions, it should be noted that our results apply to PR regimes with fixed fractional recovery (i.e., a share  $\alpha$  of the payment that can be recovered,

where the share is unaffected by the parties' behavior). Regimes that make the fraction recoverable conditional on the recipient's behavior (as in at least some formulations of the change of position defense) potentially create an additional distortion to the recipient's consumption or expenditure patterns (as suggested by Remark 4).

If PR is thought to be too administratively complex, the remaining choice is between FR and NR. In the unilateral harm case, this is straightforward as FR is always socially optimal. In the bilateral harm case, however, our framework does not yield an unambiguous answer. The relative merits of FR and NR depend on the precaution technology (i.e., which party's precaution is more effective and which party faces lower costs of precaution). Thus, whether FR or NR is preferable cannot be determined without making additional assumptions about the parties' precaution technology.

### ***5.2) Conceptual Challenges of Decoupling***

The general optimality of PR for the bilateral precaution case holds within the class of traditional ("balanced-budget") private law remedies. However, a PR regime entails inefficiencies with respect to the excess burden borne by the parties and the loss of socially valuable transactions (due to the "transaction tax" associated with mistaken payments). As shown in Proposition 2, when lump sum taxes exist, it is possible to do better, and indeed to eliminate these inefficiencies altogether for a sufficiently small  $A$  through a (hypothetical) FD regime. This involves allowing the recipient to retain the mistaken payment, while fully compensating the payer from government funds raised from (hypothetical) nondistortionary lump sum taxes.

As mentioned earlier, an obvious conceptual challenge for an FD regime is that there would exist an incentive for parties to collude and fabricate mistaken payments in order to receive subsidies from the government. Our assumption of prohibitively high transaction costs between B1 and S2 rules out this type of collusion. In addition, underlying the analysis is a set of implicit assumptions that courts can verify that putative mistaken payments are spontaneous (i.e., not induced by the recipient) and that the payment is received in good faith. These assumptions would entail that fabrications of mistaken payments would be detected (and potentially penalized). In more general settings, however, collusion may be possible. To be clear, though, such collusion would require the putative payer to make a payment to a stranger, claim the amount in compensation from the government, and then seek the (surreptitious) return of part of the initial

payment from the recipient (without, of course, any enforceable contract between them). Reputational mechanisms may in principle sustain such cooperation, but repetition of this practice would increase the likelihood of its detection.

To assess how likely such collusion might be, it is helpful to consider the analogy to a (hypothetical) “double strict liability” regime in tort law. In bilateral-precaution accidents, imposing strict liability on injurers leaves victims with inadequate incentives to take precaution and restrict activity levels, while a “no liability” regime (where victims bear their own losses) leaves injurers with inadequate incentives to take precaution and restrict activity levels. In principle, a solution is to impose strict liability on *both* injurers and victims (e.g., by requiring injurers to pay the government an amount equal to the victim’s loss and requiring the victim to bear her own loss).<sup>27</sup> This aligns incentives correctly, but is subject to a serious potential problem of collusion. After an accident, the injurer and victim have a joint incentive to conceal its occurrence to avoid the joint “tax” imposed on the parties. In relation to this scenario (where the injurer and victim face low transaction costs following an accident), it would seem that collusion to fabricate mistaken payments would be quite challenging.

A second serious difficulty for an FD regime is that lump sum taxes do not exist in reality. However, it should be noted that the advantages of an FD regime generalize to a scenario in which taxes are distortionary, but where the implicit “transaction tax” from mistakes is even more distortionary. A partial decoupling regime (where the government compensates the payer for part of the loss) may then be optimal, even in the presence of some – though not too much – collusion. On the other hand, if taxes are more distortionary than the implicit “transaction tax” from mistakes, then even partial decoupling would not be optimal.

---

<sup>27</sup> This “double strict liability” regime has been discussed for decades in the economic literature – see e.g. Shavell (2020). Note that the problem of excessive activity levels also occurs under a negligence rule (a well-known result in the economic analysis of tort law establishes that all forms of the negligence rule induce socially optimal precaution by injurers and victims in bilateral-precaution accidents; however, the activity levels of injurers are excessive). In the restitution context, it is possible to analogize an FR regime to one of strict liability (for S2) and an NR regime to one of no liability (again, from S2’s perspective). The latter induces excessive precaution by B1, while the former imposes an excess burden on S2 and (in the bilateral precaution setting) may induce excessive precaution by S2. An FD regime (if it were feasible) aligns incentives appropriately, somewhat like the double strict liability regime.

While the difficulties noted above may or may not be insuperable, the FD regime is presented here as a thought experiment rather than as a practical proposal.<sup>28,29</sup> In particular, it can arguably help us better understand some of the foundational debates in the law of restitution. For instance, it confronts the unjust enrichment perspective on the law of restitution with a hypothetical situation in which the payer suffers no harm whatsoever from the absence of restitution from the recipient (as the government has already restored the mistaken payer to the *status quo ante*). This raises the question of whether restitution is nonetheless desirable in order to confiscate the recipient's unjust gain.<sup>30</sup>

This discussion also highlights a point noted earlier, that private insurance or hedging by B1 does not solve the efficiency problems that we highlight. Actuarially fair insurance against mistaken payments would give rise to the same distortions as in our model, although it would spread the risk of mistakes and thereby reduce risk-bearing costs (and thus one possible source of an excess burden). Even when B1 is perfectly insured by a third party against mistaken payments,

---

<sup>28</sup> Nonetheless, there do exist some analogous real-world mechanisms. For instance, the Torrens system of land registration – in which landowners register their title in a public register (e.g. Miceli and Sirmans, 1995) – establishes a public fund to pay monetary compensation to claimants whose title has been canceled by good faith registrants (through mistakes or third-party fraud). While typically funded by registration fees, the public fund mechanism is not necessarily budget-balanced, because ex post realizations of errors in the past conveyance of title are unpredictable.

<sup>29</sup> The conceptual value of the hypothetical FD regime can also be illustrated by revisiting the cases introduced above. The *Westdeutsche Landesbank* scenario may be viewed as one of unilateral precaution (recall that the local authority was not legally competent to enter into the transaction, and hence presumably unable to undertake meaningful precaution). WestLB (the payer) can undertake costly precaution to investigate whether financial transactions into which it seeks to enter are *ultra vires* its counterparties. An NR regime would induce excessive precaution by WestLB from a social perspective: WestLB would fail to take into account that if the contract is void, its upfront payment will finance public services for the residents of Islington and therefore be socially valuable. Under an FR regime, WestLB will take optimal precaution. However, the cash flows (an upfront payment by WestLB and restitution by Islington) would closely resemble those of the forbidden transaction (a loan from WestLB to Islington). The latter was prohibited by Parliament, presumably on the grounds that it would create social harms (some of which would perhaps be replicated by the analogous cash flows under the FR regime). An FD regime, however, would lead to optimal precaution by WestLB, while avoiding the replication of the forbidden loan.

The *Lipkin Gorman* scenario might be viewed as involving bilateral precaution: the casino can investigate the provenance of funds used by gamblers, while the law firm can monitor its partners to prevent the misappropriation of its funds. It is likely that the former is more costly and less effective. An FR regime would induce excessive levels of (highly costly) precaution by the casino. An NR regime (or a PR regime that approximates NR) would reduce or eliminate the casino's costly precautions and substitute the less costly precaution of the law firm. While that generates a social gain, the law firm's level of precaution will be excessive because it fails to take into account the increase in the casino's profits from bets placed using the misappropriated funds. An FD regime, on the other hand, would eliminate the casino's costly precaution while also discouraging excessive precaution by the law firm.

<sup>30</sup> In *Kleinwort Benson Ltd. v. Birmingham City Council*, 4 All ER 733 (1996), a bank that engaged in an interest rate swap transaction (similar to that of WestLB) hedged its exposure to the local authority and thereby divested itself of the risk associated with the transaction. The Court of Appeal of England and Wales rejected the existence of a "passing on" defense, arguing that Birmingham City Council's unjust enrichment "at the expense of" the bank did not require that the bank experience an actual loss. However, the hypothetical FD regime poses this question more sharply, as hedging transactions are costly (entailing transaction costs, fees, and forgoing potential upside gains).

the expected value of her transaction with S1 falls (by the insurance premium) under NR. In contrast, in the FD regime as *A* approaches zero, there is no reduction in the expected value of the transaction between B1 and S1, and no loss of socially valuable transactions.

### ***5.3) Revisiting the Symmetry of Positive and Negative Externalities***

While it is focused on mistaken payments, our framework also has more general implications for the economic analysis of restitution. The most prevalent unifying theme in the economic analysis of the law of restitution is the contrast between the symmetry of positive and negative externalities in economic theory and their asymmetric legal treatment (e.g., Levmore, 1985; Ben-Shahar and Mikos, 2005; Dari-Mattiacci, 2009; Cooter and Porat, 2014, 2019; Liu, Avraham and Qiao, 2021). Levmore (1985) rationalizes this asymmetry by arguing that restitution should be limited in order to encourage voluntary market transactions. Our framework suggests a contrary idea, at least within the setting that we study. In particular, our model implies that it is the *availability* of full restitution – rather than its denial – that tends to encourage market transactions (in particular, between B1 and S1).

Cooter and Porat (2019) develop a normative argument for expanding the scope of restitution awards in order to encourage the production of positive externalities. However, in our setting, the benefits conferred by the creators of putative positive externalities – whether they be mistaken payers or officious meddlers – may not be an unmitigated blessing for recipients. More generally, our framework suggests that the application of the economic logic of symmetrical externalities to the law of restitution is not as clear as it might appear. For instance, a mistaken payer under an FR regime may impose a negative – rather than positive - externality on the recipient.

### ***5.4) The Discharge for Value Defense***

A defense to claims of restitution that has recently attracted considerable attention (as a result of the *Citibank* case discussed in Section 2 above) is the “discharge for value” defense. This denies restitution in situations where the mistaken payment is made to a creditor of the payer in satisfaction of an existing debt, where the recipient has not induced the error and has no notice of

the mistake.<sup>31</sup> This defense applies even though the payment is mistaken (i.e. although the payer is a debtor of the recipient, the payer did not intend to repay the debt at that particular time), and regardless of whether the recipient has experienced a change of position.

In our simple setting, suppose that B1 makes a mistaken payment to S2 (as before), but now imagine that S2 is a creditor of B1. The mistaken payment merely accelerates the repayment of B1's debt. That is, the gain to S2 and the loss to B1 represent the time value of the mistaken payment, and will typically be much smaller than the amount that was mistakenly transferred. More fundamentally, the preexisting debt between B1 and S2 indicates that transaction costs between them are, by definition, low. This implies that B1 and S2 can contractually specify how a mistaken payment (or a mistaken acceleration of the repayment of the debt) is to be addressed. Moreover, even when the gain and loss in terms of time value are nontrivial, the joint surplus across B1, B2, S1 and S2 is unaffected by the mistaken payment. The low transaction costs among B1, S1, S2 (and therefore implicitly also B2, who contracts with S2) imply that the contractual solution to mistakes that the parties adopt will maximize this joint surplus. Of course, in practice, the payment would be made not directly by B1 but rather via B1's bank. Nonetheless, B1 and B1's bank are necessarily in a contractual relationship, and (absent contracting frictions) will choose terms (including precautions taken by the bank against mistakes) to maximize joint surplus.

This perspective suggests that the discharge for value defense might best be understood as carving out from the domain of the law of restitution a scenario in which the mistaken payer and the recipient face low transaction costs.<sup>32</sup> Awarding restitution in these circumstances may possibly restrict the contracting space available to B1 and S2 with respect to mistakes. For instance, suppose that – in view of the modest gain and loss associated with the time value of the payment – B1 and S2 choose to specify in their debt contract that any mistaken acceleration of the repayment of the debt may be retained by S2. An FR rule under the law of restitution would be harmless if it were clear that B1 could waive her right to restitution in all circumstances. However, if there is some possibility that ex post B1 could claim the contract is invalid and seek restitution

---

<sup>31</sup> The discharge for value defense can be viewed as forming part of a more general principle in the law of restitution that denies restitution when payments are made to “bona fide payees” or “bona fide creditors” (e.g. Kull, 2001).

<sup>32</sup> Kaplan and Gilboa (2021) develop a theory of the discharge for value defense that views the recipient's status as a preexisting creditor as indicating a greater likelihood of detrimental reliance by the recipient – i.e. the recipient is more likely to believe the payment is intended rather than mistaken when there is a preexisting debt, and therefore more likely to experience a change of position. This view, however, does not account for the differences in transaction costs when the recipient is a preexisting creditor.

for a mistaken payment, the parties may be better served by an NR rule (or, equivalently, a refusal to apply the law of restitution). Parties who prefer that mistaken payments be returned could instead contract accordingly.

Talley (2021) presents evidence that many publicly-traded corporations issuing debt after the *Citibank* decision chose (“Revlon blocker”) contractual provisions to prevent the outcome that resulted from Citibank’s mistake. This might be viewed as an example of the use of the contractual freedom provided by the discharge for value doctrine in a setting with low transactions costs. Talley’s (2021) interpretation of this evidence is that it indicates that the parties’ preferred default rule is for restitution to be paid, with the *Citibank* decision imposing costs on the parties of contracting around the default rule. Within our perspective, however, the economic role of the law of restitution is not to establish contract law defaults for contracting parties, but rather to establish rules for interactions among strangers who face high transaction costs.

Contracting frictions may, of course, prevent parties from contracting optimally regarding the treatment of mistakes. Applying the law of restitution in such a setting is analogous to the application of tort law in certain contexts – such as products liability or medical malpractice – in which transaction costs are not necessarily high but where contracting is nonetheless thought to be imperfect. It is possible that (essentially mandatory) rules of restitution may improve on imperfect contracting outcomes. However, the content of those restitution rules would depend crucially on the reasons for contracting failure, rather than on the framework developed here.

It is also worth emphasizing that the unappealing outcomes in situations such as *Citibank* are, to a significant degree, attributable to the insolvency of the debtor, rather than to the discharge for value defense as such. Consider a stylized example based on the *Banque Worms* facts.<sup>33</sup> Suppose that a debtor (D) owes \$100 each to financial institutions F1 and F2. D directs its bank F3 to pay its \$100 debt to F2, but F3 mistakenly pays \$100 to each of F1 and F2. The discharge for value defense would deny F3 restitution of the \$100 mistakenly paid to F1. However, as long as D is solvent (with assets  $\geq$  \$200), the stakes involved are quite modest. In particular, F3 can recover the \$100 from D, or (in effect) become D’s creditor, just as Citibank became Revlon’s creditor. This anticipated recovery, moreover, limits potential over-precaution by F3 in making the transfer.

---

<sup>33</sup> *Banque Worms v. Bank America Int'l*, 77 N.Y.2d 362 (N.Y. 1991).

However, if D is insolvent (e.g., with assets = \$100), then the discharge for value doctrine allocates the loss from D's insolvency among the various financial institutions in a seemingly arbitrary manner. The contractual freedom (from a mandatory or sticky restitution rule) to address mistakes is less relevant in the case of D's insolvency; moreover, the financial institutions may not be in direct contractual relationships with each other. Note, however, that the problem of the arbitrary allocation of losses from D's insolvency is fundamentally symmetric. In particular, absent a discharge for value defense, the loss from D's insolvency would instead be allocated to F1, which (like F2) is completely blameless with respect to the mistake. Thus, the problem is by no means straightforward. Solving the issues raised by insolvency, moreover, requires first developing a framework that can illuminate the economic role of the law of restitution in more standard settings; such a framework is precisely what this paper seeks to construct.

### ***5.5) Unjust Enrichment, Restitution, and Redistribution***

Our model is intended to elucidate the efficiency consequences of different legal regimes governing restitution for mistaken payments, rather than to explain the doctrines of the law of restitution using economic analysis. Nonetheless, there is a disjuncture between the dominant theory of the law of restitution among legal scholars – that restitution is justified by the unjust (or, in some accounts, the “unjustified”) enrichment of the defendant – and the consequentialist normative commitments typically associated with economic analysis. More generally, it has been widely noted that there is a general tendency for the law to disfavor arbitrary redistributions of wealth (“windfalls”),<sup>34</sup> without necessarily identifying any social cost or inefficiency of the windfall. A satisfactory economic account of the law of restitution must grapple with this issue. As a tentative step towards bridging economic analysis with the notion that the recipient's enrichment may be “unjust”, we offer the following thoughts, based on the Mirrleesian optimal taxation framework. In a simplified version of this framework (e.g., Stiglitz, 1982), there are two types: high-ability and low-ability. The government wishes to redistribute income from the former to the latter, but can observe only incomes and not ability levels. The problem it faces is to redistribute to the low-ability type, while satisfying the high-ability type's incentive-compatibility constraint (i.e., inducing the high type not to mimic the low type by choosing a reduced level of labor supply or effort).

---

<sup>34</sup> See Hurt (2010) for a description and taxonomy of this phenomenon.

Stiglitz (1982) argues that random taxation of the low-ability type can potentially enhance the scope for redistribution. Random taxation makes the low-ability type's bundle of labor supply and after-tax consumption less attractive to the (risk-averse) high type. This makes it possible for the government to redistribute more from the high type to the low type while still satisfying the high type's incentive-compatibility constraint (i.e., inducing the high type to choose the bundle of labor supply and after-tax consumption that the government wishes it to choose).

This idea may potentially be applied to the context of restitution, albeit in reverse. Suppose that mistaken payments occur mostly among high-ability types (who choose the bundle intended for them). A legal regime that denies restitution would create random variation in after-tax consumption among individuals of the high-ability type (when they choose the high-ability bundle). Given risk-aversion, this makes the high-ability bundle less attractive to high-ability types. This tightens the incentive-compatibility constraint (i.e., it makes the high type more likely to mimic the low type), and thereby reduces the amount of redistribution from high to low types that the government can accomplish. A legal regime that provides for the restitution of mistaken payments (even at some, though not too high, administrative cost) would eliminate this risk and make the high-ability bundle more attractive to the high type. This would enable the government to undertake more redistribution.

## **6) Conclusion**

The law of restitution has emerged as a central focus for private law scholarship, particularly in common law jurisdictions outside the United States. This article develops a general conceptual framework grounded in economic analysis for understanding the structure and consequences of the core example of the law of restitution: mistaken payments. Our model both unifies existing insights and provides a new perspective based on the idea that mistaken payments impose a "transaction tax" on contracting parties and thereby lead to foregone socially valuable transactions. The results imply that full restitution is socially optimal when harm is unilateral and partial restitution when harm is bilateral. Importantly, we propose a novel hypothetical decoupled regime that would (if it were feasible) lead to more efficient outcomes than would partial restitution. Our conceptual framework helps us better understand the social costs and benefits of the legal regimes that govern the restitution of mistaken payments, and illuminate some of the foundational debates regarding the law of restitution.

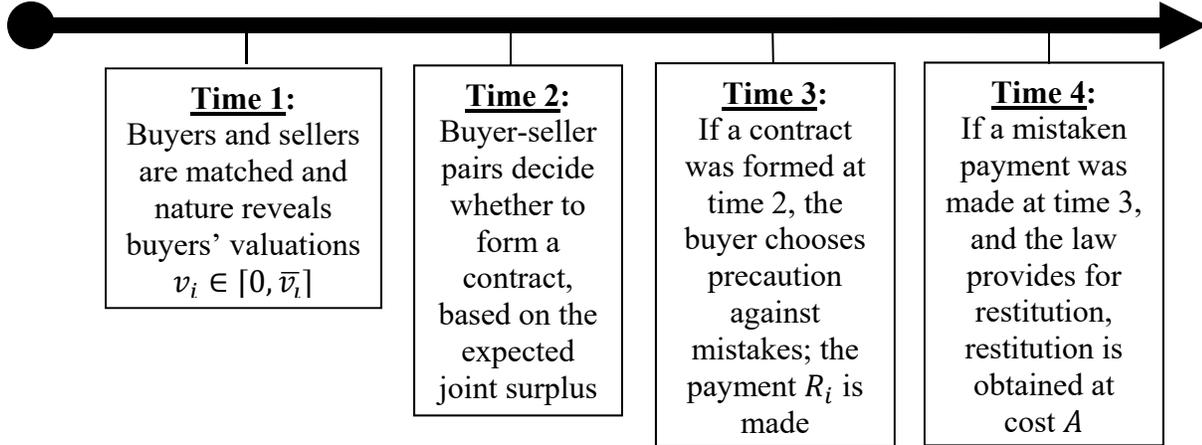
## References

- Arruñada, B., G. Zanarone, and N. Garoupa (2019) “Property rights in sequential exchange” *Journal of Law, Economics, and Organization*, 35(1), pp.127-153.
- Bant, E., 2009. *The Change of Position Defence*. Bloomsbury Publishing.
- Beatson, J. and W. Bishop (1986) “Mistaken payments in the law of restitution” *University of Toronto Law Journal*, 36, pp.149-185.
- Ben-Shahar, O. and R. A. Mikos (2005) “The (legal) value of chance: Distorted measures of recovery in private law” *American Law and Economics Review*, 7 (2), pp. 484-522.
- Birks, P., 1985. *An Introduction to the Law of Restitution*. Oxford University Press.
- Burrows, A. 2011. *The Law of Restitution*. 3<sup>rd</sup> edition. Oxford University Press.
- Bouckaert, B. and G. De Geest (1995) “Private takings, private taxes, private compulsory services: The economic doctrine of quasi-contracts” *International Review of Law and Economics*, 15, 463-487.
- Brooks, R. R. and A. Stremitzer (2012) “On and off contract remedies inducing cooperative investments” *American Law and Economics Review*, 14(2), pp. 488-516.
- Cooter, R. D. and A. Porat (2014) *Getting Incentives Right*. Princeton University Press.
- Cooter, R. D. and A. Porat (2019) “Torts and restitution: Legal divergence and economic convergence” *Southern California Law Review*, 92, pp. 897-919.
- Dari-Mattiacci, G. (2009) “Negative liability” *Journal of Legal Studies*, 38 (1), pp. 21-59.
- De Fontenay, E. (2021) “The \$900 Million Mistake: *In re Citibank August 11, 2020 Wire Transfers* (SDNY 16 February 2021)” *Capital Markets Law Journal*, 16(3), pp. 307-313.
- Dharmapala, D. and S. A. Hoffmann (2005) “Bilateral accidents with intrinsically interdependent costs of precaution” *Journal of Legal Studies*, 34(1), pp. 239-272.
- du Plessis, J. (2010) “The restitution of undue transfers from an economic perspective” *Restitution Law Review*, 18, pp. 34-50.
- Elkin-Koren, N, and E. M. Salzberger (2000) “Towards an economic theory of unjust enrichment law” *International Review of Law and Economics*, 20, pp. 551-573.
- Farnsworth, W. (2014) *Restitution*. University of Chicago Press.
- Garoupa, N. and C. W. Sanchirico (2010) “Decoupling as transactions tax” *Journal of Legal*

- Studies*, 39(2), pp. 469-496.
- Gilboa, M. and Y. Kaplan (2018) “The mistake about mistakes: Rethinking partial and full restitution” *George Mason Law Review*, 26, pp. 427-451.
- Huber, P.K. (1988) “Mistaken transfers and profitable infringement on property rights: An economic analysis” *Louisiana Law Review*, 49, pp. 71-109.
- Hurt, C. (2010) “The windfall myth” *Georgetown Journal of Law & Public Policy*, 8, pp. 339-393.
- Kaplan, Y. and M. Gilboa (2021) “The Costs of Mistakes” Working paper. Available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3951197](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3951197)
- Kull, A. (1995) “Rationalizing restitution” *California Law Review*, 83, pp. 1191-1242.
- Kull, A. (2001) “Defenses to Restitution: The Bona Fide Creditor” *Boston University Law Review*, 81, pp. 919-955.
- Landes, W.M. and R. A. Posner (1978) “Salvors, finders, good Samaritans, and other rescuers: An economic study of law and altruism” *Journal of Legal Studies*, 7(1), pp. 83-128.
- Levmore, S. (1985) “Explaining restitution” *Virginia Law Review*, 71, 65–124.
- Levmore, S. (1987) “Variety and uniformity in the treatment of the good-faith purchaser” *Journal of Legal Studies*, 16(1), pp. 43-65.
- Liu, Z. (J.), R. Avraham, and Y. Qiao (2021) “Unrequested benefits, damages assessment, and information acquisition” *American Law and Economics Review*, 23 (1), pp. 207-253.
- McKendrick, E. (1992) “Restitution, misdirected funds and change of position” *Modern Law Review*, 55(3), pp. 377-386.
- Medina, B. (2003) “Augmenting the value of ownership by protecting it only partially: The “market-overt” rule revisited” *Journal of Law, Economics, and Organization*, 19(2), pp. 343-372.
- Miceli, T.J. and Sirmans, C.F. (1995) “The economics of land transfer and title insurance” *Journal of Real Estate Finance and Economics*, 10(1), pp. 81-88.
- Nussim, J. (2010) “On A Passé defense: Unjust enrichment and the recovery of overpaid taxes” *Supreme Court Economic Review*, 18(1), pp. 233-257.
- Polinsky, A.M. and Y. K. Che (1991) “Decoupling liability: Optimal incentives for care and litigation” *RAND Journal of Economics*, 22(4), pp. 562-570.

- Priel, D. (2013) "The law and politics of unjust enrichment" *University of Toronto Law Journal*, 63, pp. 533-569.
- Saiman, C. (2008) "Restitution in America: Why the US refuses to join the global restitution party" *Oxford Journal of Legal Studies*, 28(1), pp. 99-126.
- Shavell, S. (2020) "On the redesign of accident liability for the world of autonomous vehicles" *Journal of Legal Studies*, 49(2), pp. 243-285.
- Smith, L. (2001) "Restitution: The heart of corrective justice" *Texas Law Review*, 79, pp. 2115-2176.
- Sun, L. G. and B. Daniels (2014) "Mirrored externalities" *Notre Dame Law Review*, 90, pp. 135-185.
- Stiglitz, J. (1982) "Utilitarianism and horizontal equity: The case for random taxation" *Journal of Public Economics*, 18(1), pp. 1-33.
- Talley, E. (2021) "Discharging the Discharge for Value Defense" *NYU Journal of Law & Business*, 18, pp. 147-220.
- Tinbergen, J. (1952) *On the Theory of Economic Policy*. North-Holland Pub. Co.
- Weinrib, E. (2005) "Restoring restitution" *Virginia Law Review*, 91, pp. 861-878.

**Figure 1: Timing**



**Table 1: Optimal Legal Regimes**

		<b>Precaution</b>		
		None/Exogenous	Unilateral	Bilateral
<b>Harm</b>	Unilateral	FR	FR	FR
	Bilateral	FD (if available) PR (if FD is not available)	FD (if available) PR (if FD is not available)	FD (if available) PR (if FD is not available)

FR= full restitution; FD= full decoupling; PR = partial restitution (the fraction varies with precaution)

## Appendix: Proofs of Results

**Proof of Remark 1:** Remark 1 follows straightforwardly from comparing Equations (2) and (5). In particular, note that  $c_1(1 + \bar{q}_1) > c_1$  (as  $\bar{q}_1 > 0$  by assumption) and recall the assumption that  $h_1(v_1) > 0 \forall v_1$ . It follows from these assumptions that:

$$\int_{c_1(1+\bar{q}_1)}^{\bar{v}_1} (v_1 - c_1)h_1(v_1)dv_1 < \int_{c_1}^{\bar{v}_1} (v_1 - c_1)h_1(v_1)dv_1 \Rightarrow W_s^{NR} < W_s^* \quad (\text{A1})$$

**Proof of Remark 2:** Follows straightforwardly from taking the limit of the expression in Equation (7) as  $A \rightarrow 0$ ; i.e.

$$\lim_{A \rightarrow 0} \int_{c_1 + \bar{q}_1 A}^{\bar{v}_1} (v_1 - c_1 - \bar{q}_1 A)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 = W_s^* \quad (\text{A2})$$

where  $W_s^*$  is defined by Equation (2). Applying Remark 1,  $W_s^* > W_s^{NR}$ .

**Proof of Remark 3:** The result follows from taking the limit of the expression in Equation (15) as  $A \rightarrow 0$ . When taking that limit, it is necessary to determine B1's optimal choice of  $k_1$  as  $A \rightarrow 0$ . From Equation (14), B1's goal is:

$$\min_{k_1} \left\{ \lim_{A \rightarrow 0} q_1(k_1)A + k_1 \right\} \quad (\text{A3})$$

Since by assumption  $q(k_1)$  is bounded above by  $\bar{q}_1 < 1$  and is a well-behaved convex function with  $\lim_{k_1 \rightarrow 0} q'_1(k_1) = -\infty$ , the FOC  $-q'_1(k_1)A = 1$  and SOC  $q''_1(k_1)A > 0$ , the optimization problem is equivalent to:

$$\min_{k_1} \left\{ q_1(k_1) \lim_{A \rightarrow 0} A + k_1 \right\} \equiv \min_{k_1} \{k_1\} \quad (\text{A4})$$

Therefore, it follows that as  $A \rightarrow 0$ , the optimal choice of  $k_1$  by B1 becomes 0. Thus,

$$\lim_{A \rightarrow 0} W_s^{FR} = \int_{c_1}^{\bar{v}_1} (v_1 - c_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 = W_s^* \quad (\text{A5})$$

Finally, we apply Remark 1 to establish that  $W_s^* > W_s^{NR}$ .

**Proof of Remark 4:** As noted in the text, Equations (17) and (18) are identical apart from the expected (maximized) utility that S2 derives from the mistaken payment. In Equation (17), this is given by  $y^* + u(c_1 - y^*)$  and in Equation (18) by  $u(c_1)$ . Consider the following cases:

i) Suppose that under NR, S2 chooses the corner solution  $y^* = 0$ . Then,  $u(c_1) = y^* + u(c_1 - y^*)$ , implying that Equations (17) and (18) are identical. Thus,  $W_s^{\Delta P} = W_s^{NRu}$ .

ii) Suppose that  $c_1$  is sufficiently large that  $u'(c_1) < 1$ . Then, starting from  $y^* = 0$ , it will be optimal for S2 to increase  $y^*$  until the marginal utility of luxury consumption rises sufficiently that the FOC  $u'(c_1 - y^*) = 1$  is satisfied. This yields an interior solution  $y^* > 0$ . It follows that  $y = 0$  (the corner solution under the  $\Delta P$  regime) is distinct from  $y^*$  and does not maximize the expression  $y + u(c_1 - y)$ . Therefore:

$$u(c_1) < y^* + u(c_1 - y^*) \quad (\text{A6})$$

It follows that  $W_s^{\Delta P} < W_s^{NRu}$ .

iii) Suppose that under NR, S2 chooses the corner solution  $y^* = c_1$ . It again follows that  $y = 0$  (the corner solution under the  $\Delta P$  regime) is distinct from  $y^*$  and does not maximize the expression  $y + u(c_1 - y)$ . Therefore, using the same argument as in case (ii) above,  $W_s^{\Delta P} < W_s^{NRu}$ .

**Proof of Proposition 1:** (i) When  $\alpha = 1$ , social welfare can be characterized as:

$$\begin{aligned} W_s^{FRE} &= \lim_{\alpha \rightarrow 1} W_s^{PR} \\ &= \int_{c_1 + q_1(k_1^{FR})A + k_1^{FR}}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1^{FR})[A + f(c_1)] - k_1^{FR})h_1(v_1)dv_1 \\ &\quad + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \\ &= W_s^{FR} - \int_{c_1 + q_1(k_1^{FR})A + k_1^{FR}}^{\bar{v}_1} q_1(k_1^{FR})f(c_1)h_1(v_1)dv_1 \\ &= W_s^{FR} - q_1(k_1^{FR})f(c_1)(1 - H_1(c_1 + q_1(k_1^{FR})A + k_1^{FR})) \end{aligned} \quad (\text{A7})$$

where  $W_s^{FR}$  is defined in Equation (15) and  $H_1(v_1)$  is the cdf for the distribution of B1's valuation. Now, consider reducing  $\alpha$  at the margin by an  $\varepsilon$ -infinitesimal variation. For a sufficiently small change in  $\alpha$ , an envelope theorem argument can be invoked – because B1 has optimized her choice of precaution  $k_1^{FR}$  (and hence the threshold  $c_1 + q_1(k_1^{FR})A + k_1^{FR}$  at which contracts are formed), a small change in  $\alpha$  does not have a first-order effect on B1's payoff. Thus, for a sufficiently small change in  $\alpha$ , we can hold  $k_1^{FR}$  fixed when calculating social welfare under the  $\varepsilon$ -infinitesimal variation:

$$\begin{aligned}
W_s^{PR} &= \int_{c_1 + q_1(k_1^{FR})A + k_1^{FR}}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1^{FR})[A + (f(\alpha c_1) + f((1 - \alpha)c_1))]) \\
&\quad - k_1^{FR}h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \\
&= W_s^{FR} - \int_{c_1 + q_1(k_1^{FR})A + k_1^{FR}}^{\bar{v}_1} q_1(k_1^{FR})(f(\alpha c_1) + f((1 - \alpha)c_1))h_1(v_1)dv_1 \\
&= W_s^{FR} - q_1(k_1^{FR})(f(\alpha c_1) + f((1 - \alpha)c_1))(1 \\
&\quad - H_1(c_1 + q_1(k_1^{FR})A + k_1^{FR}))
\end{aligned} \tag{A8}$$

For any well-behaved convex function,

$$f(\alpha c_1) + f((1 - \alpha)c_1) < \alpha f(c_1) + (1 - \alpha)f(c_1) = f(c_1) \tag{A9}$$

Therefore, it follows that the expression in Equation (A8) is larger than that in Equation (A7) for any convex excess burden  $f(\cdot)$ . That is,  $W_s^{PR} > W_s^{FRe}$ .

(ii) When  $\alpha = 0$ , social welfare can be characterized as:

$$\begin{aligned}
W_s^{NRe} &= \lim_{\alpha \rightarrow 0} W_s^{PR} \\
&= \int_{c_1 + q_1(k_1^{NRe})[c_1 + f(c_1)] + k_1^{FR}}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1^{NRe})[f(c_1)]) \\
&\quad - k_1^{NRe}h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2
\end{aligned} \tag{A10}$$

Using a similar argument to that for part (i), for a sufficiently small change in  $\alpha$  away from 0, we can hold  $k_1^{NRe}$  fixed when calculating social welfare under an  $\varepsilon$ -infinitesimal variation.  $W_s^{PR}$  under this variation will be identical to the expression in Equation (A8), except that  $f(c_1)$  is replaced by  $f(\alpha c_1) + f((1 - \alpha)c_1)$ . From Equation (A9), it follows that for any convex excess burden  $f(\cdot)$ ,  $W_s^{PR} > W_s^{NRe}$ .

**Proof of Proposition 2:** We begin by taking the limit of the expression in Equation (25) as  $A \rightarrow 0$ , noting that we can apply the same argument as in the proof of Remark 3 to establish that as  $A \rightarrow 0$ , the optimal choice of  $k_1$  by B1 becomes 0:

$$\lim_{A \rightarrow 0} W_s^{FD} = \int_{c_1}^{\bar{v}_1} (v_1 - c_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 = W_s^* \tag{A11}$$

where  $W_s^*$  is defined in Equation (2). Then, we take the limit of the expression in Equation (21) as  $A \rightarrow 0$ :

$$\lim_{A \rightarrow 0} W_s^{PR} = \int_{\lim_{A \rightarrow 0} c_1^{PR}}^{\bar{v}_1} (v_1 - c_1 - q_1(k_1)[f((1 - \alpha)c_1) + f(\alpha c_1)] - k_1)h_1(v_1)dv_1 + \int_{c_2}^{\bar{v}_2} (v_2 - c_2)h_2(v_2)dv_2 \quad (\text{A12})$$

where the lower limit of integration in the first term is:

$$\lim_{A \rightarrow 0} c_1^{PR} = c_1 + q_1(k_1)((1 - \alpha)c_1 + f((1 - \alpha)c_1)) + k_1 \quad (\text{A13})$$

Clearly,  $W_s^* > \lim_{A \rightarrow 0} W_s^{PR}$

We now take the limit of the expression in Equation (A7) as  $A \rightarrow 0$ , noting again that we can apply the same argument as in the proof of Remark 3 to establish that as  $A \rightarrow 0$ , the optimal choice of  $k_1$  by B1 becomes 0:

$$\begin{aligned} \lim_{A \rightarrow 0} W_s^{FRe} &= \lim_{A \rightarrow 0} W_s^{FR} - \lim_{A \rightarrow 0} [f(c_1)(1 - H_1(c_1 + q_1(k_1)A + k_1))] \\ &= W_s^* - f(c_1)(1 - H_1(c_1)) \end{aligned} \quad (\text{A14})$$

Thus,  $W_s^* > \lim_{A \rightarrow 0} W_s^{FRe}$  (given the assumptions that  $f(\cdot) > 0$  for any strictly positive argument, that  $h_i(v_i) > 0 \forall v_i \in [0, \bar{v}_i]$ , and that  $c_1 > 0$ ).