

ISSN: 2753-5746

CAMBRIDGE LAW REVIEW

VOLUME 10
ISSUE 1
SPRING 2025

Editor-In-Chief

Wednesday Eden

CAMBRIDGE LAW REVIEW

EDITORIAL BOARD

EDITOR-IN-CHIEF

Wednesday Eden

VICE EDITOR-IN-CHIEF

Christopher Symes

MANAGING EDITORS

Jonathan Rutherford

Samuel Soh

SENIOR EDITORS

Shivam Gandhi

Thomas Loke Zhil Hahn

Ruairi McIntyre

Joshua Metzger

Shalvi Ponwar

Kaden Pradhan

Hannah Reinisch

ASSOCIATE EDITORS

Hannes Bastians

Chow Hey Mann Dawn

Nicholas A Haddad

Eun Seo Jang

Akshat Jha

Jani Knoetze

Jeff Lilley

Sherelle Lim

Liam McCauley

Akola Munyiri

Thomas Musie

Sanat Narayan

Joseph Bingham Opp

Tejasvini Puri

Matthew Ethan Ramli

Roxanne A Schnyder

Kelsey Sena Tamakloe

Kai Zhen Tek

Theo

Awurakua Afaribea Yirrah

INTERNATIONAL EDITORS

Florian Bode

Max Planck Institute for Comparative and International Law, Hamburg

Julia de Moraes Almeida

Universidade de São Paulo

Justin Fei

Monash University

Erik Habich

University of Zurich

Philip Inglis Hudson

University of California Los Angeles (UCLA)

Kirhana Singh Khurana

University of British Columbia, Canada

Rosa Kliesow

University of Passau

Veronica Lynn

University of Zurich

Chris Douglas Malherbe

Stellenbosch University

Paul Mc Burnie

Hugh Wooding Law School, Trinidad & Tobago

Emmanuel Onyebor

Dalhousie University, Canada

Prem Vinod Parwani

National Law School of India University, Bangalore

Bill Puah Ee Jie

Singapore Institute of Legal Education

Divya Sethuraman

National Law School of India University, Bengaluru

Megan Styles

Monash University

Raelee Toh Hsuan Hui

Singapore Management University

Yeeun Uhm

American University

Chenghuai Xu

Indiana University Maurer School of Law

Foreword

This year marks the tenth anniversary of the Cambridge Law Review, which was founded in 2015 by then Editors-in-Chief Ruth McGuinness and Peter Bruland with the goal of providing a contemporary forum for legal scholarship at a time when there were relatively few student-run law journals in the UK. Since then, the journal has remained steadfast to this goal and has consistently provided a forum for students, practitioners, and established academics alike whose articles on diverse and pressing legal issues have been of immense interest to both UK and international readers.

Against this rich history, I am therefore honoured to present the Spring Issue of Volume 10 of the Cambridge Law Review, which includes six articles that each offer stimulating and novel insights on certain challenging legal topics. These articles have been selected for publication because we believe that they represent valuable additions to the academic literature and will likely stimulate further scholarly debate. For this Volume, Serle Court Chambers, a pre-eminent commercial chancery barristers' chambers, are sponsoring a £500 prize for the best submission across both Issues 1 and 2 on a topic in English commercial law and/or equity. The winner will be announced in the foreword of the Autumn Issue.

In preparing this Issue, I am indebted to our student editors (both at the University of Cambridge and through our International Editor programme) who spent several months reviewing the submissions we received and recommending articles for publication. I am also immensely grateful to the members of the Managing Board (Christopher Symes, Samuel Soh, and Jonathan Rutherford) for their expert guidance on specific areas of law. This Issue would not have been possible without their dedicated involvement throughout the review process. Lastly, I would like to thank the authors for their excellent contributions to this Issue and for their attentiveness during the editing stages.

Issue 1 begins with Professor Talia Fisher's article, 'Within the Sound of Silence: Reassessing the Role of Reasoning in Judicial Decision-Making', which challenges the seemingly ingrained assumption within Anglo-American legal systems that it is 'inherently and universally desirable' for judges to give reasons for their decisions. Fisher notes that the US Supreme Court's frequent use of its 'shadow docket', where a number of its decisions are made without outlining the judges' reasoning, has recently been criticised by commentators who regard this as 'circumvent[ing] an essential aspect of the act of judging'. In contrast to these commentators, Fisher presents a more nuanced account of judicial reason-giving, by critically examining five arguments that are typically used to justify judges' provision of reasons. These arguments are the following: first, that judicial reason-giving increases the 'quality' of judicial decision-making by enhancing its transparency and accountability; secondly, that it fosters the development of legal precedent; thirdly, that it fulfils the 'expressive functions' of the law, such as by communicating collective norms; fourthly, that it legitimises judicial authority by, *inter alia*, ensuring that judicial decisions can be understood, and thus 'authored', by the public; and fifthly, that it enables individuals to participate actively in the legal process and upholds their due process rights, such as the right to be heard. Taking each of these arguments in turn, Fisher exposes their fundamental limitations and identifies certain circumstances in which it is arguably preferable for judges *not* to give reasons for their decisions. These limitations include, *inter alia*, her observation with respect to the first argument that requiring judges to provide reasons might undermine the quality of their decisions owing to 'verbal overshadowing' (where judges might privilege reasons that are easier to articulate, but

IV

potentially less relevant, than reasons that are harder to verbalise), and her determination with respect to the fifth argument that reasoned judgments may only have a superficial impact on participatory justice, as they are unable to address deeper, structural imbalances that undermine litigant autonomy.

Turning to international arbitration, Alfie Dove critically examines the impact that inconsistent definitions of due process and varying ethical standards have on the legitimacy and efficiency of the international arbitral system in his article, 'Addressing the Challenges of Due Process and Ethical Disparities in International Arbitration: The Need for Reform'. He begins by considering the lack of a common definition of due process, which he argues is 'problematic' because, inter alia, it triggers 'due process paranoia', where, owing to arbitrators' uncertainty as to whether their decisions will be challenged, they may be excessively cautious when faced with due process objections and some parties may seek to weaponise such objections as a strategic 'sword'. Dove then considers how conflicting ethical standards can 'disrupt' arbitral proceedings, most notably in cross-boarder disputes where counsel are often subject to divergent ethical codes. He argues that the conflict between the International Bar Association Guidelines, as 'soft law', and national ethical codes produces 'double deontology' (where counsel are left confused as to which standard they should adhere to) and may give rise to an 'inequality of arms' (where parties may be significantly disadvantaged during proceedings). To address these challenges, Dove contemplates two possible reforms to international arbitration. The first reform is where the terminology concerning due process is 'harmonised', which he argues would increase the predictability of the 'procedural requirements' that apply within arbitral proceedings. However, he also acknowledges that, given the significant differences that exist between common law and civil law jurisdictions, 'detailed guidance' would need to be provided to supplement this harmonised terminology. The second reform involves the establishment of a single, 'International Code of Ethics', which would supersede the conflicting national codes that counsel abide by. Dove proposes that this International Code could be introduced incrementally, where it starts as 'soft law' and gradually becomes embedded into institutional rules; however, he recognises that it might be overly 'ambitious' to expect this Code to attain the status of a Convention.

In her article, 'Remedying Judicial Intervention into Private Contract: A Case for Abandoning the Creditor Duty Post-*BTI 2014 LLC v Sequana SA* [2022] UKSC 25', Alexandra Kosta-Foti seeks to achieve two aims. The first is to address the 'problematic trend' that commentators have identified since 1974 (namely, that judges have increasingly interfered with parties' contractual allocation of risks in the interests of fairness) and to 'reassert the primacy' of such risk allocations. And secondly, Kosta-Foti seeks to critique the creditor duty in England and Wales and, in particular, what she considers to be the failure of the UK Supreme Court in *Sequana* to establish 'a coherent basis upon which [such] a... duty is recognised'. She begins by examining the two main theories that have shaped the academic debate surrounding the issue of shareholder primacy (specifically, the contractarian and law and economics theories, on the one hand, and the progressivist approach, on the other) and she argues that both theories cannot adequately explain the treatment of corporate purpose and directors' duties under English law. After comparing how the creditor duty has developed in the case law of three jurisdictions (Australia, Delaware, and England and Wales), Kosta-Foti then centres on the Supreme Court's treatment of the creditor duty in *Sequana*, which she argues has 'perpetuated... flaws in the previous line of case law'. Here, she identifies four key flaws in the Justices' interpretations of the 'content' of this duty. According to Kosta-Foti, these flaws include, inter alia, the following: that the creditor duty is effectively a 'fiduciary duty to act in good faith in the interests of the company', which is imposed upon its directors in the

absence of a ‘voluntary undertaking’; and that imposing such a fiduciary duty wrongly assumes that creditors are ‘a homogeneous class’. In her view, the Supreme Court’s position on the creditor duty should therefore be ‘abandoned’. Ultimately, Kosta-Foti argues that the creditor duty conflicts with fundamental principles of fiduciary law and contract law (namely, that fiduciary duties must be ‘voluntarily assumed’ and the principle of sanctity of contract) and she contends that considerations of fairness ought not to ‘trump’ parties’ contractual allocation of risks.

The next article, ‘Repairing the English Civil Law of Bribery: Fixing *Johnson v FirstRand Bank Ltd* [2024] EWCA Civ 1282’, which is co-authored by Rivu Chowdhury and George Beglan, critically examines the reasoning of the Court of Appeal in *FirstRand*. It focuses in particular on the Court’s determination that car dealers who arranged finance for customers were acting as credit brokers, and were therefore under a ‘fiduciary duty’ and a ‘disinterested duty’ to these customers, and that, if the dealers received ‘secret’ or partially disclosed commissions from a lender, the lender would face liability under the tort of bribery or as an accessory to the dealers’ breach of fiduciary duty. This article is timely because the UK Supreme Court heard the appeal of this decision just days ago. Chowdhury and Beglan begin by providing a background overview of the law of bribery and secret commissions, before summarising the Court of Appeal’s decision in *FirstRand*. Although they agree with the Court of Appeal that a fiduciary duty should have been owed in the three cases, they identify key ‘deficiencies’ in the Court’s reasoning. These deficiencies include, inter alia, the Court’s presumption that the car dealers owed a fiduciary duty by virtue of their ‘status’ as credit brokers, without considering that such duties should be consensually agreed to, and the Court’s apparent conflation of ‘dishonesty with knowledge’ when approaching the issue of accessory liability for breach of a fiduciary duty. Subsequently, Chowdhury and Beglan offer suggestions on how to ‘repair’ these deficiencies, such as their recommendation that courts should adopt a ‘fact-specific’ approach when deciding whether to impose a fiduciary duty, which considers whether the dealers made any ‘undertakings’ and whether the customers themselves were particularly ‘vulnerable’. Chowdhury and Beglan also propose ‘fusing’ or ‘absorbing’ the tort of bribery into a breach of fiduciary duty, which they argue is possible because the disinterested duty can be viewed as a ‘modified fiduciary duty’. According to these authors, the result of this fusion would be that the breach of a disinterested duty, being tantamount to the breach of a fiduciary duty, would fall within the jurisdiction of equity, thus removing the current, but unnecessary, existence of ‘two answers to the question of liability in bribery’.

In his article, ‘Severing the Hunt for Tortfeasor-Caregiver Compensation: The Singaporean Rejection of *Hunt v Severs* [1994] 2 AC 350 (HL)’, Jun Xiang Wong compares the contrasting approaches of the Singaporean High Court in *Rajina Sharma v Theyvasigamani* [2024] SGHC 42 (*‘Rajandran’*) and the House of Lords in *Hunt* to the question of whether the victim of a tort can recover compensation from the tortfeasor for the value of voluntary care that the tortfeasor themselves has provided. Wong begins by examining Lord Bridge’s reasoning in *Hunt* and noting the apparent circularity that would ensue if a victim could recover compensation for such care from the tortfeasor owing to the existence of a ‘trust relationship’ between the victim (as trustee) and the tortfeasor-caregiver (as beneficiary). Wong describes such circularity in the following way: the compensation ‘would go from the tortfeasor, to the victim who held it on trust for the tortfeasor, and then back to the tortfeasor’. Next, he examines six criticisms that the Law Commission made of *Hunt* and he proceeds to identify conceptual problems with the ‘damages on trust’ model that Lord Bridge endorsed, such as the ‘lack of clarity’ as to how this model might apply to the recovery of future caregiving

VI

expenses. To address the problems that he identifies with *Hunt*, Wong proposes an alternative model of compensation, which he refers to as a ‘closed loop’ model, under which the victim in *Hunt* could have recovered the costs of voluntary care provided by the tortfeasor-caregiver and without necessitating any consideration of third-party insurance. Finally, after examining the strengths and weaknesses of the approach adopted by Teh J in *Rajandran*, where she determined that the victim could recover compensation for the costs of voluntary care provided by the tortfeasor, Wong argues that this approach represents a ‘better and more principled’ outcome compared to *Hunt*, albeit one that is not ‘doctrinally flawless’, and he considers how his ‘closed loop’ model could be applied to address certain of these flaws.

Lastly, Saniya Mehmood AKC’s article, ‘With Love and Affection: Rethinking the Fairness of Proximity of Relationship in Secondary Victim Claims’, critically examines the fairness of the ‘proximity of relationship’ control mechanism in secondary victim claims for negligently inflicted psychiatric injury. She begins by analysing the House of Lords’ judgment in *Alcock v Chief Constable of South Yorkshire Police* [1992] 1 AC 310 (HL) and highlights three shortcomings in their Lordships’ approach to this control mechanism. According to Mehmood, the first shortcoming is that the establishment of certain relationship categories that are prima facie presumed to have ‘close ties of love and affection’ is ‘unnecessarily restrictive’ because it creates a ‘hierarchy of relationships’ where relevant factors, such as the duration of one’s relationship with the primary victim, are arbitrarily excluded from consideration. Secondly, the requirement that claimants whose relationships do not fall into these categories must prove that they enjoyed a ‘close tie of love and affection’ to the primary victim seemingly ‘trivialises’ their grief. And thirdly, Mehmood argues that their Lordships’ obiter comments on the recoverability of damages by bystanders are contradictory because they indicate that a claimant could bypass the ‘proximity of relationship’ requirement where an accident was ‘particularly horrific’. After identifying these shortcomings, Mehmood analyses 30 cases heard between 1991 and 2024 to test the claim that subsequent courts have applied this control mechanism ‘more liberally’ than was first envisioned in *Alcock* by embracing claimants whose relationships with the primary victim fall outside of the relationship categories that were presumed to be sufficiently proximate in that case. Here, she finds, inter alia, that ‘proximity of relationship’ has consistently ‘played a limited role’ in these cases compared to other control mechanisms and she therefore contends that this control mechanism is in need of reform. Although Mehmood considers ‘more conservative’ proposals for reform, such as the introduction of a ‘closed list’ of relationships that are deemed to be sufficiently proximate, she ultimately argues that the optimal course of action would be to abolish ‘proximity of relationship’ entirely.

Overall, this has been a very promising start to Volume 10 and I look forward to working alongside my fellow Editors on the Autumn Issue later this year.

Wednesday Eden
Editor-in-Chief
Darwin College
6 April 2025

TABLE OF CONTENTS

<i>Within the Sound of Silence: Reassessing the Role of Reasoning in Judicial Decision-Making</i> Talia Fisher	1
<i>Addressing the Challenges of Due Process and Ethical Disparities in International Arbitration: The Need for Reform</i> Allie Dove	25
<i>Remedying Judicial Intervention into Private Contract: A Case for Abandoning the Creditor Duty Post-BTI 2014 LLC v Sequana SA [2022] UKSC 25</i> Alexandra Kosta-Foti	47
<i>Repairing the English Civil Law of Bribery: Fixing Johnson v FirstRand Bank Ltd [2024] EWCA Civ 1282</i> Rivu Chowdhury and George Beglan	75
<i>Severing the Hunt for Tortfeasor-Caregiver Compensation: The Singaporean Rejection of Hunt v Severs [1994] 2 AC 350 (HIL)</i> Jun Xiang Wong	96
<i>With Love and Affection: Rethinking the Fairness of Proximity of Relationship in Secondary Victim Claims</i> Saniya Mehmood AKC	115

Within the Sound of Silence: Reassessing the Role of Reasoning in Judicial Decision-Making

TALIA FISHER*

ABSTRACT

The growing reliance of the US Supreme Court on unexplained orders, commonly known as their ‘shadow docket’, has emerged as a major concern, generating strong criticism for undermining the traditional expectation that judges provide detailed justifications for their rulings. Without presuming to take a stance on the legitimacy of shadow dockets or the motivations that inform them, the increasing reliance on unexplained orders encourages a re-evaluation of the role and function of judicial reasoning within the broader spectrum of case law. This article re-examines the assumption that providing reasons for judicial judgments is inherently and universally desirable. It explores how providing reasons can sometimes conflict with other fundamental values underlying the legal process and argues for considering limits on judicial reasoning that extend beyond pragmatic concerns related to judicial economy (efficient management of judicial resources). The article highlights three key areas where such limits on reason-giving might be warranted. First, witness credibility assessments, in which the verbal articulation of the underlying reasons might distort the decision-making process. Secondly, ‘hard cases’, in which the risk of crafting detrimental precedents suggests a need to separate the resolution of specific cases from the development of broader legal doctrines, as illustrated by the US Supreme Court’s handling of *Bush v Gore*. Thirdly, situations in which judicial silence speaks louder than the articulation of reasons and can serve as a catalyst for democratic deliberation, as manifested in the Supreme Court’s handling of the 2014 same-sex marriage cases. Rather than advocating for Aristotelian intuition-based adjudication (where a judge’s practical wisdom (*phronesis*) and cultivated sense of justice guide decisions, sometimes beyond explicit legal rules), the article emphasises the need to balance the benefits of judicial reason-giving with an awareness of its limitations and proposes a more strategic and deliberate approach to providing reasoned judgments.

Keywords: *reasoned judgments, shadow docket, hard cases, constitutional dialogue, procedural justice*

* Amy and Paul Yanowicz Professor of Human Rights at the Faculty of Law, Tel Aviv University. I am deeply grateful to Eyal Zamir for his wonderful and generous suggestions, and to Miri Gur Arye and the participants in the Criminal Law Workshop at the Hebrew University for their thoughtful comments. My thanks also go to Naama Lipschits, Yael Meskin, and Nadav Zak for their dedicated research assistance. I am profoundly grateful to Wednesday Eden and the editorial team of the *Cambridge Law Review* for their superb editorial work.

I. INTRODUCTION

'Judicial decisions are reasoned decisions.'
*Rita v United States*¹

The US Supreme Court is currently facing mounting criticism for its tendency to issue unsigned and unexplained orders, colloquially referred to as the Court's 'shadow docket'.² A substantial proportion of its judicial decisions are made without providing the reasoning behind them.³ The resort to shadow docket practices, particularly in contentious areas like access to abortion, immigration restrictions, capital sentencing, COVID-19 policies, and election procedures, has recently sparked academic debate.⁴ The issue came into the public spotlight during Justice Amy Coney Barrett's confirmation hearing, where she identified the Supreme Court's shadow docket as a 'hot topic in the last couple of years'.⁵ Critics contend that the Court's decision to forego the detailed reasoning that characterises its formal merits docket, even in highly influential cases, circumvents an essential aspect of the act of judging.⁶ An illustrative example is the January 2022 emergency order blocking the Occupational Safety and Health Administration's COVID-19 vaccination-or-testing mandate for large employers, a *per curiam* decision which directly impacted nearly one-quarter of the country's population.⁷

The COVID-19 pandemic marked the height of the US Supreme Court's reliance on shadow docket practices.⁸ It also provided the context for another case that illuminates the issue of reasoned judicial decisions. In January 2021, during the height of the pandemic,

¹ *Rita v United States* 551 US 338, 356 (2007). See also *Chavez-Meza v United States* 585 US 109, 113 (2018).

² Stephen I Vladeck, 'Putting the "Shadow Docket" in Perspective' (2023) 17 *Harvard Law & Policy Review* 289, 289. The term 'shadow docket' was coined by William Baude; see William Baude, 'Foreword: The Supreme Court's Shadow Docket' (2015) 9 *New York University Journal of Law and Liberty* 1, 1.

³ Richard J Pierce, 'The Supreme Court Should Eliminate Its Lawless Shadow Docket' (2022) 74 *Administrative Law Review* 1, 10 ('We now have a situation in which a high proportion of major judicial decisions are made by the Supreme Court without providing any reasons for the decisions.').

⁴ A symposium dedicated to the Supreme Court's shadow docket was held in 2023; see Leslie C Griffin, 'The Shadow Docket: A Symposium' (2023) 23 *Nevada Law Journal* 669, 669 ('Everyone is talking about the Supreme Court's shadow docket.'). Stephen Vladeck, *The Shadow Docket: How the Supreme Court Uses Stealth Rulings to Amass Power and Undermine the Republic* (Hachette Book Group 2023); Caroline Fredrickson, 'Will American Democracy Last in Light of the Shadow Docket?' (2023) 23 *Nevada Law Journal* 727; Nicholas D Conway and Yana Gagloeva, 'Out of the Shadows: What Social Science Tells Us About the Shadow Docket' (2023) 23 *Nevada Law Journal* 673; Jenny-Brooke Condon, 'The Capital Shadow Docket and the Death of Judicial Restraint' (2023) 23 *Nevada Law Journal* 809; Rachael Houston, 'Does Anybody Really Know What Time It Is?: How the Supreme Court Defines "Time" Using the *Purcell* Principle' (2023) 23 *Nevada Law Journal* 769.

⁵ See Andrew J Wistrich, 'Secret Shoals of the Shadow Docket' (2023) 23 *Nevada Law Journal* 863, 943, fn 4. Another major public encounter with the Court's shadow docket and practice of issuing sparsely explained orders was through the ruling on the Texas Heartbeat Act; see *Whole Woman's Health v Jackson* 141 S Ct 2494 (2021), cited in Vladeck, 'Putting the "Shadow Docket" in Perspective' (n 2) 290.

⁶ Harvard Law Professor Nicholas Stephanopoulos has argued that unexplained, unreasoned court orders contradict fundamental judicial principles, equating them to displays of power rather than legal rulings: 'Missing from Supreme Court's Election Cases: Reasons for Its Rulings' (*National Freedom of Information Coalition*) <<https://www.foic.org/blogs/missing-supreme-courts-election-cases-reasons-its-rulings/>> accessed 24 March 2025. See also Chad M Oldfather, 'Writing, Cognition, and the Nature of the Judicial Function' (2008) 96 *Georgetown Law Journal* 1283, 1285 ('... preparation of a written opinion might be deemed an essential component of a legitimate judicial decision').

⁷ *National Federation of Independent Business v Department of Labor, Occupational Safety and Health Administration* 142 S Ct 661, 662 (2022) (*per curiam*), cited in Vladeck, 'Putting the "Shadow Docket" in Perspective' (n 2) 294.

⁸ See Thomas P Schmidt, 'Orders Without Law' (2024) 122 *Michigan Law Review* 1003.

Armando Sauseda filed a motion for compassionate release under 18 USC § 3582(c)(1).⁹ Sauseda, who was serving a life sentence for murder, had health conditions that placed him at significant risk of severe COVID-19 complications. The provision allows a federal court to reduce a term of imprisonment if it finds that ‘extraordinary and compelling reasons’ warrant such a reduction, after considering the factors listed in 18 USC § 3553(a). Sauseda argued that his heightened health risk met these criteria. The District Court denied Sauseda’s motion in a laconic, four-sentence order: ‘After considering the applicable factors provided in 18 U.S.C. § 3553(a) and the applicable policy statements issued by the Sentencing Commission, the Court denies the Defendant’s Motion[] on its merits’.¹⁰ The Court did not address its consideration of the § 3553(a) factors nor did it indicate whether it found Sauseda’s risk of contracting COVID-19 to be an extraordinary and compelling reason. Sauseda appealed, and the US Court of Appeals for the Fifth Circuit vacated and remanded the District Court’s denial of Sauseda’s motion, citing the District Court’s failure to provide adequate justifications for its decision.¹¹ The appellate court held that the District Court was required to offer specific reasons for its factual ruling.¹²

The Anglo-American legal tradition firmly establishes the expectation that judges provide the reasons behind their decisions. Articulation of the reasons underlying judgment has been deemed ‘always permissible, usually desirable and often obligatory’.¹³ The literature and caselaw supply a variety of compelling justifications for judicial reasoning, assuming its universal value, with exceptions generally confined to pragmatic considerations of judicial economy (efficient management of judicial resources).¹⁴

The object of this article is to challenge the assumption that reasoned judgments are inherently and universally desirable and to explore potential conflicts between providing reasons and other values upheld by the legal process, beyond the efficient use of judicial resources. The article will propose delineating the limits of judicial reasoning not only according to pragmatic considerations, related to judicial economy, but also by identifying categories of cases where the interests underlying judicial reason-giving may warrant exempting—perhaps even prohibiting—courts from offering reasons.

It should be emphasised that the assumption that reason-giving is inherently beneficial is often treated as an axiomatic truth rather than as a proposition requiring justification. Rather than being the product of a rigorous normative balancing exercise, the preference for universal judicial reason-giving (subject to considerations of efficient use of judicial resources) frequently operates as an unexamined background assumption. However, this axiom should be subjected to critical scrutiny. My aim in this article is not to provide an exhaustive assessment of the normative trade-offs of judicial reasoning across all contexts, nor to engage in a mechanistic weighing of advantages and disadvantages. Rather than engaging in a mechanistic cost-benefit analysis, I seek to interrogate the foundational assumptions underlying the expectation of universal judicial reasoning, questioning when and why it advances the integrity of

⁹ *United States v. Sauseda* No 7:09-CR-252-1 (WD Tex 2013).

¹⁰ *United States v. Sauseda* No 21-50210, 2 (5th Cir, 1 April 2022).

¹¹ *ibid* 6.

¹² *ibid* 7.

¹³ Michael Kirby, ‘Reasons for Judgment: “Always Permissible, Usually Desirable and Often Obligatory”’ (1994) 12 Australian Bar Review 121.

¹⁴ Frederick Schauer, ‘The Generality of Law’ (2004) 107 West Virginia Law Review 217, 231–32 (‘reason-giving is a pervasive and frequently praised feature of legal decision-making, and a legal decision-maker who provides reasons for her decisions is considered a better legal decision-maker than one who does not... All sorts of professions and human activities require the giving of reasons, but none obsess about it the way law does, and few – philosophy may be an exception – consider reason-giving a central feature of the enterprise’).

adjudication and when it risks undermining it. The article highlights two key points. First, it challenges the presumption that judicial reasoning is always preferable, urging a more critical evaluation of the forces that shape it. Secondly, it identifies specific contexts in which there is an a priori reason to suspect that requiring reason-giving may be counterproductive—failing to serve, or even undermining, the teleological goals of reason-giving itself. It demonstrates that it does not suffice to assume that judicial reasoning invariably fosters legal soundness, accountability, or legitimacy. One must also consider the countervailing pressures it introduces, such as the risk of incentivising performative reasoning, reinforcing majoritarian biases, or constraining judicial discretion in ways that distort adjudication.

The article proceeds as follows: Section II will offer a detailed taxonomy of the key justifications for the requirement of judicial reason-giving, examining how the provision of reasons serves different purposes for various stakeholders, including litigants, appellate courts, and the public.¹⁵ Section III will undertake a critical assessment of each of these justifications, evaluating their implications and effectiveness. This analysis will be translated and applied to contemporary legal doctrine, highlighting potential shortcomings and areas for reform.

II. A TAXONOMY OF THE JUSTIFICATIONS UNDERLYING REASONED JUDGMENTS

In his seminal article, ‘The Forms and Limits of Adjudication’, Lon Fuller argues that engagement in reasoned argument is a hallmark and unique domain of adjudication.¹⁶ Delivering reasoned judgments functions as ‘both the norm and the ideal’ of adjudication.¹⁷ But, despite judicial reason-giving’s wide-scale endorsement, there is nothing ‘natural’ or ‘pre-political’ about this legal institution. In fact, many pre-modern courts operated without offering reasons for their decisions, viewing their role primarily in terms of rendering judgments.¹⁸ To this day, the requirement to provide reasons does not apply to juries.¹⁹ In certain judicial contexts, as well, judges issue decisions without providing the underlying reasons. This includes decisions regarding objections or juror dismissals, state supreme courts’ refusals to review cases, denials of *certiorari* by the US Supreme Court, and numerous summary decisions by federal courts of appeals.²⁰

The principal justifications for judicial reasoning can be broadly classified into five key arguments. These include enhancing the quality of judicial decision-making, developing legal precedents, fulfilling the expressive functions of trials, reinforcing democratic legitimacy, and protecting litigant autonomy and due process rights. The discussion will commence with the first argument.

¹⁵ Hofit Wasserman-Rozen, Ran Gilad-Bachrach and Niva Elkin-Koren, ‘Lost in Translation: The Limits of Explainability in AI’ (2024) 42 *Cardozo Arts & Entertainment LJ* 391, 392.

¹⁶ Lon L. Fuller, ‘The Forms and Limits of Adjudication’ (1978) 92 *Harvard Law Review* 353, 368.

¹⁷ Frederick Schauer, ‘Giving Reasons’ (1995) 47 *Stanford Law Review* 633, 633.

¹⁸ ‘Reason-giving is a typically modern idea. There have been historical moments when it was deemed valuable *not* to give reasons. For instance, Roman courts, ecclesiastical courts, and a number of aristocratic courts in premodern, continental Europe functioned without giving reasons for their decisions.’: Mathilde Cohen, ‘When Judges Have Reasons Not to Give Reasons: A Comparative Law Approach’ (2015) 72 *Washington & Lee Law Review* 483, 486. See also Michael Akehurst, ‘Statements of Reasons for Judicial and Administrative Decisions’ (1970) 33 *MLR* 154, 154; Doron Menashe, ‘The Requirement of Reasons for Findings of Fact’ (2006) 8 *International Community Law Review* 223, 227.

¹⁹ Bennett Capers, ‘Evidence Without Rules’ (2018) 94 *Notre Dame Law Review* 867, 896.

²⁰ Schauer (n 17) 634.

A. IMPROVING THE QUALITY OF JUDICIAL DECISION-MAKING

One line of argument, invoked to justify the institution of judicial reason-giving, is that providing reasons, especially in writing, polices and refines judicial decision-making.²¹ This is supported by the following considerations: first, judicial reason-giving functions as a check on the exercise of judicial power. It facilitates transparency, allowing judicial decisions to be subjected to scrutiny by relevant stakeholders—including the litigating parties, appellate judges, the media, policymakers, legal scholars, and the broader public.²² The ‘opportunity to evaluate, scrutinize, and possibly assent to the reasons for a decision’²³ allows for ex post correction of judicial error, safeguarding the public from the detrimental effects of judicial rulings that might have been based on erroneous reasoning, personal biases, or judicial partiality.

Reason-giving not only facilitates the rectification of decisions ex post but may also enhance the quality of judicial rulings ex ante.²⁴ It is not merely a communication tool for what has already transpired but can also shape the decision-making process from the very outset. The anticipation of future scrutiny regarding the decision’s underlying rationales may incentivise judges pre-emptively to take only legitimate considerations into account. Put differently, reasoned decision-making promotes accountability, which can have a transformative and de-biasing effect on the decision-making process as it unfolds.²⁵ Research in cognitive psychology indicates that mechanisms of accountability, inherent in reason-giving, can reduce the impact of prejudice on judicial outcomes.²⁶ When decision-makers are required to provide reasons for their judgments, they are prompted to engage System II thinking—a slower, more analytical mode of cognition—rather than relying on System I thinking, which is fast, intuitive, and prone to bias. This shift reduces the influence of implicit biases.²⁷ As a result, reason-giving may serve as a de-biasing mechanism, counteracting prejudiced intuitions.

Lastly, the act of providing written reasons enhances the quality of judicial decisions through reflective introspection. Oscar Wilde’s statement, ‘How do I know what I think until I see what I say?’²⁸ encapsulates this understanding, suggesting that the process of articulating reasons fosters a more rigorous engagement with the factual and legal issues at hand. In the words of Mathilde Cohen, ‘[t]his process is often described as the “it won’t write” phenomenon. In attempting to reason her decision, a judge discovers that she cannot find an appropriate legal justification, leading her to reconsider her initial ruling and make a more accurate

²¹ Oldfather (n 6); Jason Bosland and Jonathan Gill, ‘The Principle of Open Justice and the Judicial Duty to Give Public Reasons’ (2014) 38 Melbourne University Law Review 482, 486; Eyal Zamir, ‘With No Reason: Allowing Courts to Decide Cases without Explaining their Decisions’ (2024) 43 Civil Justice Quarterly 290, 295 (the article proposes that litigants in civil disputes should be allowed mutually to consent to non-reasoned judgments, subject to the Court’s approval).

²² Glen Staszewski, ‘Reason-Giving and Accountability’ (2009) 93 Minnesota Law Review 1253, 1263 (‘reason-giving facilitates transparency’).

²³ Micah Schwartzman, ‘Judicial Sincerity’ (2008) 94 Virginia Law Review 987, 1005.

²⁴ Martha I Morgan, ‘The Constitutional Right to Know Why’ (1982) 17 Harvard Civil Rights-Civil Liberties Law Review 297, 300.

²⁵ Jennifer S Lerner and Philip E Tetlock, ‘Bridging Individual, Interpersonal, and Institutional Approaches to Judgment and Decision Making: The Impact of Accountability on Cognitive Bias’ in Sandra L Schneider and James Shanteau (eds), *Emerging Perspectives on Judgment and Decision Research* (CUP 2003) 431.

²⁶ Christoph Engel, ‘The Psychological Case for Obliging Judges to Write Reasons’ in Christoph Engel and Fritz Strack (eds), *The Impact of Court Procedure on the Psychology of Judicial Decision Making* (Nomos 2007) 71.

²⁷ For further discussion of System I and System II thinking, see Daniel Kahneman, *Thinking, Fast and Slow* (Penguin Books 2011). For further discussion of the debiasing effect of accountability, see Jennifer S Lerner and Philip E Tetlock, ‘Accounting for the Effects of Accountability’ (1999) 125 Psychological Bulletin 255.

²⁸ Oscar Wilde, *De Profundis* (Methuen and Co 1905) 19.

determination'.²⁹ The act of elaborating reasons can, thus, reveal new dimensions of understanding to the judge, refining the arguments and contributing to a more thoroughly considered ruling.

In summary, requiring judges to justify their decisions with reasoned arguments promotes transparency, thereby allowing for greater factual and normative accuracy within the adjudication system. In addition to transparency, the requirement to provide reasons also fosters judicial accountability, which operates as another control on judicial arbitrariness and bias. The potential for future scrutiny motivates judges to rely on relevant facts and sound legal principles from the outset, bearing transformative potential for both the judicial decision-making process and the decisions themselves. The transformative and debiasing capacity of judicial reason-giving is also rooted in the act of reflective introspection. These mechanisms collectively contribute to refining the decision-making process and the judicial outcome.

B. SUPPORTING PRECEDENTS AND THE PRINCIPLE OF *STARE DECISIS*

Another key justification for the institution of judicial reasoning lies in its crucial role within the domain of precedential authority. Reason-giving is instrumental in both the formation and reinforcement of legal precedents. In the context of setting precedents, reason-giving supports generalisation. As Frederick Schauer observes, by articulating reasoned explanations, judges facilitate the extrapolation of broader legal principles from the specific facts of individual cases, thereby elevating the level of abstraction and situating each case within a comprehensive framework of analogous cases.³⁰ This allows judges to extend their role beyond merely resolving the immediate legal dispute before them, converting specific decisions into general directives applicable to future cases. Law develops through such resolution of disputes, grounded on reasons that are both public and broadly applicable.³¹ Even elements of judicial reasoning that do not constitute the formal *ratio decidendi* can impact the evolution of legal doctrine.³²

Reason-giving not only facilitates the formation of new precedents but also demonstrates a court's adherence to existing ones and to the principle of *stare decisis*. By elucidating the rationale behind their decisions, courts enable their commitment to precedential authority to be scrutinised and can forge specific links between their current positions and past judgments. Beyond signalling compliance with prior decisions, reason-giving helps to construct a framework for future proceedings, guiding the judiciary and legal system along a principled and predictable trajectory. In other words, through adherence to the reasoning embedded in precedents and by offering reasons of their own, courts can perpetuate the consistent application of legal rules and facilitate the evolution of the law, both of which are crucial for upholding the rule of law.

²⁹ Cohen (n 18) 511–12.

³⁰ Schauer (n 17). Schauer asserts that the fundamental difference between decisions made without the obligation to provide reasons—such as those by voters and juries—and decisions that require reason-giving, like court judgments and many administrative rulings, is their adherence to the principle of generality. Only the latter are committed to ensuring that like cases in the future are treated alike. This commitment to generality distinguishes reasoned decisions by embedding them within a framework of normative consistency, which is essential for the rule of law. See also Zamir (n 21) 297.

³¹ Robin J Elfron, 'Reason Giving and Rule Making in Procedural Law' (2014) 65 *Alabama Law Review* 683, 713.

³² For further discussion, see Judith M Stinson, 'Why Dicta Becomes Holding and Why It Matters' (2010) 76 *Brooklyn Law Review* 219.

In summary, reason-giving functions as an essential mechanism through which judges engage in an ongoing dialogue with other judges and shape the landscape of legal doctrine. Reason-giving ensures that precedents do not function as static or discrete judgments but, rather, become integrated into a cohesive, dynamic, and continually evolving corpus of law. This fosters a living, adaptive legal system.

C. FULFILLING THE EXPRESSIVE FUNCTIONS OF TRIALS AND FOSTERING PUBLIC TRUST

The role of reason-giving transcends providing interpretative guidance to future courts. It also serves as a means of communicating with the broader public. Another category of justifications for reasoned judgments is that, by articulating reasons, the judiciary fulfils the expressive functions of law, fosters trust in the courts and their judgments, and encourages compliance.

Expressive theories of law centre on the communication of collective attitudes through legal action and on the role of law in ‘making statements’.³³ They emphasise the function of legal institutions in conveying normative messages to the public and shaping social values.³⁴ Trials are a critical component of this expressive endeavour, as they embody the intersection between the individual and the state apparatus and serve as arenas for the clarification of social norms and for the reification of moral commitments. Labelling certain behaviours as ‘criminal’ serves to grant one moral approach precedence over contradictory visions of justice.³⁵ Beyond their role in adjudicating disputes or proclaiming guilt and innocence, trials function as public rituals through which a moral stance on the behaviour in question and on all involved is formed and communicated.³⁶ Criminal convictions, for instance, are a locus of blame and stigma, intended to signal the community’s condemnation of the offender and her actions. The punishment imposed, beyond fulfilling retributive and deterrent functions, serves as a mechanism for expressing the gravity of indignation.³⁷

There is room to claim that a court’s provision of reasons for its judgments is integral to the fulfilment of these expressive tasks. As Cass Sunstein asserts, ‘for law to perform its expressive function well, it is important that law communicate well’.³⁸ Communicating well includes providing reasons, as failing to do so may result in the court’s inability to articulate its moral directive clearly or persuasively. Public articulation of legal reasoning, in other words, is necessary to foster a shared understanding and moral standing within the community.

³³ Cass R Sunstein, ‘On the Expressive Function of Law’ (1996) 144 *University of Pennsylvania Law Review* 2021, 2024.

³⁴ See Elizabeth S Anderson and Richard H Pildes, ‘Expressive Theories of Law: A General Restatement’ (2000) 148 *University of Pennsylvania Law Review* 1503, 1510.

³⁵ Martin Sabelli and Stacey Leyton, ‘Train Wrecks and Freeway Crashes: An Argument for Fairness and Against Self Representation in the Criminal Justice System’ (2000) 91 *Journal of Criminal Law and Criminology* 161, 209.

³⁶ *ibid.*

³⁷ Dan M Kahan, ‘What Do Alternative Sanctions Mean?’ (1996) 63 *University of Chicago Law Review* 591, 593. The expressive argument primarily applies to criminal trials, as civil trials generally address conduct devoid of moral fault, and civil liability does not carry the same connotations of blame, stigma, or condemnation. That said, certain civil adjudications do carry a significant stigmatising effect. The termination of parental rights, for example, imposes a profound stigma on parents, reflecting an implicit judgment of blame underlying such determinations: David D Meyer, ‘Family Ties: Solving the Constitutional Dilemma of the Faultless Father’ (1999) 41 *Arizona Law Review* 753, 782. Likewise, punitive damages function as a key doctrinal tool for expressing moral disapproval. Nevertheless, the expressive dimension of judicial reasoning remains most closely associated with the criminal domain.

³⁸ Sunstein (n 33) 2050.

The process of reasoning also serves to foster public trust in the court's moral³⁹ and legal authority, by demonstrating that judgments are morally and legally grounded, rather than arbitrary or opaque.⁴⁰ Thus, the public trust of judicial rulings hinges not only on their bottom lines and substantive outcomes, but also on the underlying rationales and decision-making processes. Well-articulated reasoning can demonstrate to the public that decisions are anchored in legal principles and are a product of 'judgment' rather than subjective preferences and 'will'. The public is more likely to endorse moral and legal prescriptions they perceive as emanating from legitimate sources and processes. The public is also more inclined to comply with legal norms whose rationales it can comprehend and follow.⁴¹ Research by Daphna Lewinsohn-Zamir, Eyal Zamir and Ori Katz indicates that elucidating the reasons behind legal norms fosters internal motivations to follow them.⁴² This occurs because it enhances individuals' understanding of the norm's purpose, thereby increasing their sense of its legitimacy. Essentially, when people understand 'why' a rule exists, they are more likely to internalise it and adhere to it willingly. Providing reasons can thus complement traditional enforcement-based strategies for generating compliance, proving preferable given the high costs associated with enforcement.⁴³

In conclusion, beyond its role in objectively improving the quality of decision-making, offering reasons for judicial decisions is also justified by its capacity to foster public acceptance and compliance. It is essential for the court's expressive and educative functions, allowing courts to refine their communicative messages.

D. REINFORCING DEMOCRATIC DELIBERATION AND POLITICAL LEGITIMACY

The role of judicial reason-giving extends beyond signalling moral blame to the public or informing future courts. It is also crucial for communication with other branches of government. This brings the discussion to the democratic virtues of reasoned judging.

The justifications provided thus far have predominantly been consequentialist. However, judicial reasoning can also be justified on intrinsic grounds. Accordingly, even if it were to be shown that certain unreasoned decisions are of higher quality than reasoned ones or that obscuring the judicial decision-making process might better foster public trust, this could not undermine the duty to provide reasons. The article will turn to discuss intrinsic arguments

³⁹ This refers to procedural morality rather than substantive morality, which concerns the inherent rightness or wrongness of the outcome.

⁴⁰ See Mathilde Cohen, 'Reasons for Reasons' in Dov M Gabbay and others (eds), *Approaches to Legal Rationality*, vol 20 (Springer 2010) 119.

⁴¹ Tom R Tyler, *Why People Obey the Law* (Princeton University Press 1990) 63.

⁴² Daphna Lewinsohn-Zamir, Eyal Zamir and Ori Katz, 'Giving Reasons as a Means to Enhance Compliance with Legal Norms' (2022) 72 *University of Toronto Law Journal* 316.

⁴³ *ibid* 353. For instance, consider traffic laws. Merely imposing fines (external enforcement) might deter some, but providing clear explanations about the safety benefits of speed limits can foster a sense of responsibility, leading to voluntary compliance. The 'high costs' of enforcement refer to a range of factors. First, there are direct financial costs: maintaining police forces, funding court systems, and administering prisons. Secondly, there are social costs: excessive enforcement can erode public trust, create adversarial relationships between law enforcement and communities, and lead to disproportionate impacts on marginalised groups. For example, overly aggressive policing in certain neighbourhoods can lead to community resentment and a breakdown of cooperation, hindering crime prevention efforts. Additionally, there are opportunity costs: resources spent on enforcement could be allocated to preventative measures, such as education or social programs, which might address the root causes of non-compliance. While it is true that providing reasons also entails costs, such as the time and effort required for judges to write detailed opinions or for policymakers to engage in public discourse, these costs are often outweighed by the long-term benefits of fostering a culture of voluntary compliance and reducing the need for costly enforcement measures.

for judicial reason-giving, emphasising courts' actual, rather than perceived, legitimacy. These arguments, which pertain to the democratic virtues of judicial reasoning and public reason, will be explored in two parts. The first will address the claim that the legitimacy of the judiciary in a democratic system is rooted in its ability to justify its decisions to other branches of government and the constituency. The second will explore the argument that the practice of judicial reason-giving is essential for transforming a judicial decision into one that is made in the name of the 'public'.

John Rawls is renowned for his thesis that public reason is a fundamental legitimising feature of the liberal-democratic state.⁴⁴ For Rawls, the legitimacy of state power, including judicial authority, hinges on the ability of public officials to justify their actions to other democratic actors, using reasons that are accessible and comprehensible, even if potentially objectionable.⁴⁵ This applies with special force to the Supreme Court.⁴⁶ In the words of Rawls, 'public reason applies... in a special way to the judiciary and *above all to a supreme court in a constitutional democracy with judicial review...* [T]he court's special role makes it the exemplar of public reason'.⁴⁷

Ronald Dworkin's perspective on judicial reasoning complements that of Rawls by accentuating the judiciary's distinctive role in advancing reasoned argumentation, including outside the courtroom setting.⁴⁸ Dworkin highlights the fact that judicial decision-making is distinct from legislative processes, in that the latter often operate by way of applying sheer political power.⁴⁹ Critical of the scarcity of reasoned discourse in contemporary politics, Dworkin commends the quality of argumentation provided by courts, especially supreme courts.⁵⁰ He contends that constitutional adjudication and judicial reasoning enrich public debate, particularly on issues of political morality, by encouraging sustained public engagement and critical reflection on the courts' decisions and underlying rationales. According to both Rawls and Dworkin, in other words, the institution of judicial reasoning and its fusion with public reason are vital for the legitimacy and efficacy of democratic governance.⁵¹

The democratic rationale for judicial reasoning can also be approached from a different perspective by asserting that reason-giving is essential for transforming judicial decisions into those made in the name of 'the public'. In their book, *Reclaiming the Public*, Avihay Dorfman and Alon Harel argue that the legitimacy of decisions by public officials and institutions hinges on their public characteristics.⁵² These are defined in the following terms: 'public decisions are decisions that can, at least in principle, be shaped by citizens' actions and values, be attributed to the public, and for which the public can ultimately be held responsible'.⁵³ Under this non-instrumental view, institutions and officials are deemed public not merely because they further the interests of, or act on behalf of, the public, but because they speak in the public's name. This distinction differentiates public institutions and actors from other entities that pursue collective interests and goals, like NGOs or businesses. Legitimacy, under this view, is achieved when public institutions genuinely adopt the perspective of those they

⁴⁴ Jeremy Waldron, 'Public Reason and "Justification" in the Courtroom' (2007) 1 *Journal of Law, Philosophy and Culture* 107, 107.

⁴⁵ *ibid.*

⁴⁶ John Rawls, *Political Liberalism* (Columbia University Press 1993) 231–39.

⁴⁷ *ibid.* 215.

⁴⁸ Ronald Dworkin, *Law's Empire* (Harvard University Press 1986) 410.

⁴⁹ *ibid.*

⁵⁰ Cohen (n 18) 502.

⁵¹ *ibid.* 504.

⁵² Avihay Dorfman and Alon Harel, *Reclaiming the Public* (CUP 2024).

⁵³ Alon Harel, Noam Kolt and Gadi Perl, *Automation as Privatization* (unpublished manuscript; on file with author) 15.

serve.⁵⁴ Within this framework, members of the political community can regard the binding decisions of the state and its judiciary as their own, reflecting the principle of self-governance and aligning political and judicial authority with freedom and equality.⁵⁵

Judicial reason-giving is integral to this legitimisation process, as it allows decisions to be understood, scrutinised, and thereby ‘authored’ by the public. For the public to claim authorship and actively participate in the decision-making process, it must grasp the reasoning behind the decisions. In the words of Harel, Gadi Perl, and Noam Kolt, ‘[w]e further establish that for the public to be considered an author of a decision, the public must have an actual opportunity to participate in the decision-making process’.⁵⁶ By providing reasons, courts allow the public such effective participation and authorship over the judicial decisions, transforming them from mere mandates into public acts. Judicial reason-giving plays a fundamental role in turning judicial decisions into expressions of collective, democratic self-governance.⁵⁷ This transcends the instrumental role prescribed to reason-giving and to judicial transparency in the first part of this article, which focused on their de-biasing or quality-enhancing capacities.

In summary, judicial reasoning is essential to legitimising political and judicial authority. One perspective holds that this legitimacy is derived from integrating judicial decisions into the broader democratic framework of public reason, where courts justify their rulings to other political actors and branches of government. Another view posits that the legitimising power of judicial reasoning lies in transforming judicial decisions into acts of collective governance, authored by and made in the name of the public.

E. SECURING LITIGANT AUTONOMY AND PROCEDURAL JUSTICE

From the perspective of another crucial stakeholder, the litigating parties, reason-giving can be justified by its role in safeguarding their agency and autonomy within the judicial process. Autonomy, a core principle in the liberal tradition, literally means ‘self-rule’ or ‘self-government’.⁵⁸ Although the notion of autonomy takes many forms, it is generally understood to include the granting of choice to individuals and the securing of their ability to shape their life stories.⁵⁹ Autonomy assumes particular significance in legal contexts, where decisions often have profound personal bearing.⁶⁰ The adversarial model is particularly occupied with party autonomy and control over the judicial process.⁶¹ Under this model, litigants act as the sovereigns of trial. They are construed not merely as subjects of the legal process but as active agents delineating its borders and controlling legal trajectories. It is not enough that they are given their day in court. They must also be afforded control over how that day unfolds.

When courts provide reasons for their decisions, they facilitate this type and level of engagement. This is particularly potent in the criminal context, where conviction and punishment not only impose suffering but also carry the risk of moral condemnation that can

⁵⁴ Dorfman and Harel (n 52).

⁵⁵ *ibid.*

⁵⁶ Harel, Kolt and Perl (n 53).

⁵⁷ *ibid.*

⁵⁸ The term ‘autonomy’ is derived from the Greek words ‘auto’ (self) and ‘nomos’ (law), collectively meaning ‘self-rule’ or ‘self-government’.

⁵⁹ Robert E Toone, ‘The Incoherence of Defendant Autonomy’ (2005) 83 *North Carolina Law Review* 621.

⁶⁰ Martin H Redish and Nathan D Larsen, ‘Class Actions, Litigant Autonomy, and the Foundations of Procedural Due Process’ (2007) 95 *California Law Review* 1573, 1579, arguing that litigant autonomy refers to the party’s ‘interest in having power to make choices about the protection of her own legally authorized or protected rights’.

⁶¹ Rabecca Assy, *Injustice in Person: The Right to Self-Representation* (OUP 2015) 14.

dehumanise offenders.⁶² By providing a reasoned explanation for punishment, courts affirm the status of offenders as autonomous moral agents and reasoning beings.⁶³ More broadly, courts' provision of reasons transforms the judicial process from a top-down, unilateral imposition of mandates to a participatory dialogue with the litigating parties. As Jack B Weinstein famously claimed, adjudication distinguishes itself from depersonalised bureaucratic processes by providing a platform where litigants' voices are heard and contribute to the integrity of the legal process.⁶⁴ Robin J Effron further noted that reason-giving benefits litigants by grounding decisions in accessible terms that account for their personal perspectives.⁶⁵ Reason-giving, in other words, grants litigants effective voice at trial, ensuring that their factual and legal narratives are meaningfully engaged with. Beyond merely giving litigants a voice, judicial reasoning enhances their ability to make meaningful choices and exercise control over their legal affairs.⁶⁶ When courts provide clear and reasoned justifications for their decisions, litigants can better understand the legal basis of the ruling. This transparency enables them to assess their options and make informed decisions concerning appeal proceedings and other forms of legal action, further solidifying their autonomy.

Reasoned judgments facilitate litigant autonomy not only in managing their personal legal affairs but also in their broader role within society. Autonomy principles emphasise that individuals should be seen as active participants in shaping their society's legal norms, whether through political engagement or legal proceedings. By providing clear reasoning, judicial decisions enable litigants to understand, contest, and contribute to the development of legal principles, reinforcing their role as agents in both the courtroom and the broader legal landscape.⁶⁷ Being subject to well-reasoned judicial authority affirms each litigant's role as an autonomous agent, rather than as merely a passive recipient of legal or political power (as also discussed earlier).⁶⁸ The provision of reasons further secures autonomy by ensuring that trials remain insulated from governmental overreach, thereby shielding the litigating parties from the arbitrary power of the state.⁶⁹ Reasoned justifications for rulings ensure that decisions are rooted in legal principles rather than political or administrative influence. This safeguard preserves judicial independence, preventing courts from becoming mere extensions of governmental authority.

Lastly, judicial reasoning also plays a pivotal role in upholding procedural justice and the parties' due process rights. Procedural justice theories emphasise that the fairness of a legal process is not only about reaching a correct substantive outcome but also about ensuring that the methods by which the decision is reached are just and equitable.⁷⁰ In this sense, the process itself becomes a determinant of the 'correctness' of the judicial outcome.⁷¹ A basic

⁶² RA Duff, *Punishment, Communication, and Community* (OUP 2001).

⁶³ Marah Stith McLeod, 'Communicating Punishment' (2020) 100 Boston University Law Review 2263, 2267.

⁶⁴ Jack B Weinstein, 'Compensation for Mass Private Delicts: Evolving Roles of Administrative, Criminal, and Tort Law' (2001) University of Illinois Law Review 947, 975.

⁶⁵ Effron (n 31).

⁶⁶ For a similar claim as to the general right to explanation, see Bram Vaassen, 'AI, Opacity and Personal Autonomy' (2022) 35 *Philosophy and Technology* 87.

⁶⁷ Cohen (n 18) 505.

⁶⁸ *ibid.*

⁶⁹ William B Rubenstein, 'Divided We Litigate: Addressing Disputes Among Group Members and Lawyers in Civil Rights Campaigns' (1997) 106 *Yale LJ* 1623, 1644. See also Abram Chayes, 'The Role of the Judge in Public Law Litigation' (1976) 89 *Harvard Law Review* 1281.

⁷⁰ Lawrence B Solum, 'Procedural Justice' (2004) 78 *Southern California Law Review* 181, 238; Tom R Tyler and E Allan Lind, 'Procedural Justice' in Joseph Sanders and V Lee Hamilton (eds), *Handbook of Justice Research in Law* (Springer 2001) 65.

⁷¹ Solum (n 70) 191.

tenet of procedural justice and due process is the right to be heard.⁷² By offering reasoned judgments, courts demonstrate that they have effectively heard each of the litigants and seriously weighed their claims. Even when litigants do not prevail, the act of engaging with their arguments and providing a rationale reassures them that their concerns were not arbitrarily dismissed and allows for justice to be both seen and done.

In summary, by providing reasons for their judgments, courts effectively carve out a space for the parties in legal decision-making. This can reaffirm the parties' agency and autonomy, safeguard their due process rights, and reassure them that they have been treated in a manner that corresponds with the principles of procedural justice.

III. REASONED JUDGING AND ITS DISCONTENTS

The preceding discussion has outlined the five principal arguments commonly advanced in support of the requirement for judicial reason-giving. The focus now shifts to a critical evaluation of each of these arguments. This section will delve deeper into these justifications, aiming to uncover their intrinsic limitations and reveal the manners in which they fail adequately to account for prevailing practices.

A. THE QUALITY ENHANCEMENT JUSTIFICATION RECONSIDERED

As previously discussed, the requirement to articulate reasons subjects judicial decisions to both appellate review and public scrutiny, which serves to identify and filter out decisions that are factually or legally flawed. Moreover, the shadow of future review incentivises courts to issue judgments that adhere to sound factual and normative standards from the outset. Thus, providing reasons exerts an *ex ante* debiasing and transformative effect on the decision-making process. This potential for transformation is further amplified by the fact that the requirement to give reasons has the capacity to encourage deeper introspection and more rigorous engagement with the factual and legal issues involved.

Upon closer scrutiny, however, there is room to challenge each of these justifications, starting with the policing effect of public review. While it is often the case that exposing judicial reasoning to public oversight fosters greater accountability and leads to legally and factually sound judgments, there can also be unintended adverse consequences. As evidenced by the growing body of literature on 'judicial populism' and 'popular constitutionalism', the awareness that judicial decisions and their underlying rationales will be subject to public scrutiny can paradoxically push courts to populist decision-making.⁷³ The pressure of anticipated public oversight may compel judges to conform their reasoning to biased public opinion and to prioritise political expediency over correct legal decision-making. This inclination towards populism threatens the court's independence. It risks compromising the precision of the act

⁷² *ibid* 183.

⁷³ Monika Hanych, Hubert Smekal and Jaroslav Benák, 'The Influence of Public Opinion and Media on Judicial Decision-Making: Elite Judges' Perceptions and Strategies' (2023) 14 *International Journal for Court Administration* 1, 10–15 (discussing the intricate relationship between judicial decision-making and public opinion, and demonstrating how the transparency afforded by judicial reasoning may prompt some judges to take popular sentiments into account). For further discussion of 'judicial populism' or 'popular constitutionalism', see Lisa Hilbink, 'Judicial Populism: A Conceptual and Normative Inquiry' (2024) 49 *Law & Social Inquiry* 1.

of judging, resulting in judicial reasoning and rulings that, though aligned with public sentiment, lack robust legal foundations.⁷⁴

In a similar vein, when judges perceive that their reasoning will be scrutinised by elite groups, such as legal academics, this too can skew their decisions in a different—but equally concerning—manner. Elite-driven oversight can compel judges to render decisions and justifications that cater to the preferences of more influential or informed factions of society, potentially undermining the broader public interest. By ‘elite-driven oversight’, I am referring not to any particular ideological or political stance but rather to the influence exerted by highly specialised or institutionally powerful groups, such as legal academics or influential practitioners, whose perspectives may carry disproportionate weight in shaping judicial reasoning. While it is true that legal academics and practitioners often advocate for minority interests, their institutional authority and intellectual influence can nevertheless create an environment in which judicial reasoning becomes oriented towards aligning with academically favoured frameworks or theoretical paradigms, rather than being shaped by broader public considerations. This dynamic does not necessarily produce an undesirable outcome in every instance, but it does introduce a particular kind of external pressure on judicial decision-making. In other words, judicial accountability and reason-giving, intended to secure the integrity of legal proceedings, could inadvertently foster judicial partisanship.

The concerns that reason-giving might paradoxically compromise the court’s impartiality or become a vehicle for populism represent only part of the issue. Additional concerns emerge when considering that, while the articulation of reasons can guide a decision in one direction, it is equally plausible that it could sway the outcome in another direction. The transformative capacity to reduce bias may inherently carry the potential to introduce new forms of bias. Specifically, there is cause for concern that the requirement to provide reasons might distort judicial deliberation by privileging factors that are more readily articulated verbally, potentially at the expense of no less relevant considerations that are challenging to express in words.

As argued by Chad Oldfather, these concerns are also supported by research in cognitive psychology, which has identified the phenomenon of ‘verbal overshadowing’.⁷⁵ Verbal overshadowing refers to a counterintuitive cognitive effect where attempts to verbalise aspects that are not easily captured in words can impair rather than enhance decision-making accuracy.⁷⁶ Introduced by Jonathan W Schooler and Tonya Y Engstler-Schooler, verbal overshadowing reveals that forcing verbal articulation can distort cognitive processing by recoding non-verbal memories into a linguistically biased format, often leading to less accurate judgments.⁷⁷

A series of experiments on facial recognition provided a striking demonstration of this effect. In Schooler and Engstler-Schooler’s study, participants were shown a face and later

⁷⁴ This argument aims to describe the forces and incentives that shape judicial decision-making rather than assert a deterministic outcome. Just as the claim that judges exercise independent reasoning is grounded in an understanding of the institutional and normative structures that guide their role, so too is the argument that public scrutiny can exert pressures that influence judicial reasoning. The emerging literature on judicial populism and popular constitutionalism demonstrates that heightened public oversight does not merely enhance accountability but can also create incentives for judges to align their reasoning with prevailing public sentiment. This is not to suggest that all judges inevitably succumb to such pressures, but rather that the act of judging does not occur in a vacuum—it is responsive to the institutional and sociopolitical context in which it takes place. The concern, then, is not that judicial reasoning will always be compromised by populist forces, but that the very conditions designed to ensure accountability may, in some instances, introduce competing incentives that challenge the ideal of independent adjudication.

⁷⁵ Oldfather (n 6) 1310.

⁷⁶ *ibid.*

⁷⁷ Jonathan W Schooler and Tonya Y Engstler-Schooler, ‘Verbal Overshadowing of Visual Memories: Some Things Are Better Left Unsaid’ (1990) 22 *Cognitive Psychology* 36.

asked to identify it from a lineup. Some participants were required to describe the face verbally before making their selection, while others proceeded directly to the recognition task without verbalisation. Counter to common intuition, those who described the face before attempting recognition performed significantly worse than those who did not. The researchers explained that the verbalisation process altered the way the memory was stored and retrieved—rather than preserving the original, perceptual memory, verbalisation biased participants towards features that were easier to describe in words. This verbal recoding, in turn, impaired their ability to recognise the face based on its actual visual features.⁷⁸

Importantly, this effect was not limited to facial recognition. Additional experiments showed that verbal overshadowing also affected colour memory, reinforcing the conclusion that the phenomenon arises whenever individuals are forced to translate complex nonverbal information into words. The impairment was not temporary; participants continued to show reduced accuracy in recognition tasks even after a two-day delay, suggesting that once a verbally recoded memory replaces the original perceptual representation, the distortion persists. However, the study also revealed that verbal overshadowing does not entirely erase the original memory. When participants were required to make rapid recognition decisions—thus preventing them from relying on their altered, verbally recoded memory—they performed significantly better, suggesting that verbalisation does not destroy the original representation but rather overshadows it, making it less accessible when deliberation is involved.⁷⁹

Similar findings emerged in Sean M Lane and Schooler's study on analogy judgments, which revealed that participants who articulated their reasoning were less successful in identifying deep structural analogies than those who did not.⁸⁰ To understand this phenomenon, it is important to distinguish between surface analogies and deep structural analogies. Surface analogies occur when two situations share obvious, readily observable similarities, such as the same type of characters, objects, or settings, even though they may not be meaningfully related at a deeper level. In contrast, deep structural analogies involve cases that share the same underlying logical structure or relational pattern, even if their superficial details are completely different. Recognising deep analogies requires individuals to abstract away from surface-level details and identify a common principle or reasoning pattern.

Lane and Schooler's study provides direct empirical support for the idea that verbalisation biases individuals towards surface features at the expense of deeper relational insights. In their experiment, participants first read a set of 16 short stories, each with a unique narrative structure. Later, they were given eight test stories, each of which had two possible matches from the original set: one that was a superficial match (i.e. it contained similar characters, objects, or settings but did not share the same deeper relational structure) and another that was a true deep analogy (i.e. it shared the same core reasoning pattern or moral structure but differed in surface details).

The results showed a clear verbal overshadowing effect: participants who were required to think aloud while making their analogy judgments were significantly more likely to select surface-level matches and significantly less likely to retrieve true deep analogies. This suggests that the act of verbalisation disrupts access to deeper relational processing, leading individuals to rely more heavily on information that is easily expressible in words rather than

⁷⁸ *ibid.*

⁷⁹ *ibid.*

⁸⁰ Sean M Lane and Jonathan W Schooler, 'Skimming the Surface: Verbal Overshadowing of Analogical Retrieval' (2004) 15 *Psychological Science* 715.

on abstract, relational reasoning that is more difficult to articulate.⁸¹ These findings have been consistently validated by subsequent research.⁸²

The research on verbal overshadowing offers an important cautionary note regarding the blanket application of a reasoning requirement in judicial decisions. Of course, translating the findings to the judicial context is complex. It requires pinpointing the types of decisions and case categories that correlate with tasks like facial recognition or analogy judgments in a manner which makes them particularly prone to verbal overshadowing. One category that can be hypothesised to share notable similarities is witness credibility assessments. In this context, a gap may be suspected between the court's ability to articulate overt credibility indicators, such as a witness's confidence while testifying, and more nuanced, elusive factors that inform judicial trust, but are difficult to verbalise. The author is currently investigating this further in a separate study aimed at exploring verbal overshadowing specifically in the context of assessing credibility based on impressionistic judgments. Early findings indicate that, in the evidentiary context of witness credibility assessments, verbalisation may introduce unintended biases.⁸³

The verbal overshadowing phenomenon casts new light on Justice Stewart's widely criticised remarks in *Jacobellis v Ohio* regarding the definition of pornography:⁸⁴ '[C]riminal laws in this area are constitutionally limited to hard-core pornography. I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that'.⁸⁵ Justice Stewart's invocation of the subjective standard, 'I know it when I see it', has been disparaged, based on the sentiment that verbal articulation is a necessary condition for legality.⁸⁶ But, this criticism, along with its underlying assumption, may be misplaced when considered through the lens of verbal overshadowing. Verbal overshadowing research highlights the inherent difficulties in articulating certain intuitive judgments and their underlying rationales. Forcing such judgments into precise verbal terms and reasons can obscure rather than illuminate and impair rather than improve them, as the process of verbalisation can shift focus away from diagnostically relevant but hard-to-articulate features and towards aspects that are more easily verbalised, thereby distorting the original judgment. Justice Stewart's reluctance to provide a clear definition might, in fact, reflect a deeper awareness of the limitations of language in capturing subjective legal assessments.

In conclusion, mandating verbal justifications may undermine the quality of judicial decisions. The pressure to conform to public or elite expectations can prompt judges to render decisions that align more with prevailing opinions than with principled legal reasoning. Instead of enhancing the quality of judicial decision-making, the obligation to provide reasons may introduce external pressures that compromise judicial independence and detract from the court's commitment to impartiality. Furthermore, in judicial contexts that parallel facial recognition tasks, including witness credibility assessments based on impressionistic

⁸¹ *ibid.*

⁸² See for example Thomas Hugh Feeley and Melissa J Young, 'Self-Reported Cues about Deceptive and Truthful Communication: The Effects of Cognitive Capacity and Communicator Veracity' (2000) 48 *Communication Quarterly* 101; Christian A Meissner and John C Brigham, 'A Meta-Analysis of the Verbal Overshadowing Effect in Face Identification' (2001) 15 *Applied Cognitive Psychology* 603.

⁸³ While analogical reasoning is central to judicial decision-making, particularly in relation to *stare decisis* and the classification of legal categories, surface-level analogies may be more commonly employed than deeper structural analogies.

⁸⁴ See Oldfather (n 6) 1320.

⁸⁵ *Jacobellis v Ohio* 378 US 184, 197 (1964).

⁸⁶ See Oldfather (n 6) 1320.

judgments, the requirement to articulate reasons can adversely affect both accuracy and decision quality by disproportionately prioritising easily verbalised indicators of credibility.

B. THE PRECEDENT FORMATION JUSTIFICATION RECONSIDERED

The critical examination of the precedent-based justification for reasoned judgments will begin by revisiting its core argument: judicial reasoning is commonly viewed as vital for the development and evolution of legal precedents. By articulating the legal principles that underpin their decisions, judges convert specific rulings into general rules applicable to future cases, thus fostering a coherent legal system. Nonetheless, a more nuanced analysis reveals that, while judicial reasoning is central to the doctrine of precedent, it may also introduce considerable challenges to the broader evolution of the law.

Historical examples illustrate that legal development can occur with minimal reliance on judicial explanations. Roman law, in particular, operated under the assumption that providing reasons may impose constraints on the dynamic progression of legal doctrine. Modern critiques, like those of Michael Stokes Paulsen, argue that if the authority of precedents is based on the idea that judges impart meaning to the law, then contemporary judges should have the same interpretive power as their predecessors.⁸⁷ The notion that judicial authority diminishes over time, or that past decisions should bind current and future interpretations, may undermine the evolution of law and the adaptability of legal reasoning.

The problem is rooted not only in the principle of *stare decisis* or in the institution of legal precedent, as such, but also in the inherent tendency to overextend their reach. Due to structural reasons, courts are driven to overgeneralisation in their reasoning, risking the creation of rigid legal precedents that constrain too many future cases and stifle legal innovation. One reason is that the evolution of legal precedents is fundamentally rooted in the art of comparing and contrasting cases for the purpose of discerning patterns and guiding principles. But the case of highest relevance and significance for such comparative analysis is the ‘next’ case, the one that has yet to be brought before the court, the one that would arise in the future. Given that courts lack the foresight provided by such future cases, there is an inherent risk that the reasoning mechanisms they employ in the precedential case may lead to overgeneralisation. As Richard Posner argued, ‘[w]hat the judge has before him is the facts of the particular case, not the facts of future cases. He can try to imagine what those cases will be like, but the likelihood of error in such imaginative projection is great’.⁸⁸ Since courts operate without the full comparative perspective that future cases might offer, they risk shaping and reinterpreting prior and current rulings in ways that may unduly extend their scope, adversely affecting the development of law.

A compelling illustration is found in Jorge Luis Borges’s article, ‘Kafka and His Precursors’, where he posits that ‘every writer creates his own precursors’.⁸⁹ In this seminal work, Borges illustrates how writers, through their creative endeavours, not only influence contemporary and future discourse but also redefine and reinterpret past literary contributions. Borges supports his argument by demonstrating how earlier figures, such as Svevo Zeno and Søren Kierkegaard, laid the foundation for Franz Kafka’s literary innovations. According to Borges, Kafka’s work does more than merely echo previous ideas; it reconfigures them,

⁸⁷ Michael Stokes Paulsen, ‘The Intrinsically Corrupting Influence of Precedent’ (2005) 22 *Constitutional Comment* 289.

⁸⁸ Richard A Posner, *Law, Pragmatism, and Democracy* (Harvard University Press 2003) 80.

⁸⁹ Jorge Luis Borges, ‘Kafka and His Precursors’ in *Labyrinths* (New Directions Publishing Corporation 1964) 201.

providing new interpretations that transform our understanding of both Kafka and his precursors.⁹⁰

These observations about literary creation are highly pertinent to the field of judicial writing and legal evolution. Meaning, whether in literature or law, is not fixed at the moment of creation but rather shaped and reshaped by future interpretations. Just as literary works acquire new significance through later readings and reinterpretations, judicial rulings too evolve as they are applied in new contexts. In the judicial context, as well, it is future judges, cases, and decisions that provide retrospective context and meaning to their predecessors. Just as Kafka's literary innovations redefine and reinterpret earlier works, future judicial decisions similarly reshape and imbue past rulings with new significance and legal relevance. This dynamic interplay between past and future underscores the necessity of maintaining flexibility in both literary and judicial creation, ensuring that new developments continue to evolve. Attempting to impose a fixed interpretation through judicial reasoning, without the benefit of knowing how future developments will unfold, can hinder the evolution of precedents and the optimal progression of legal doctrine.

Let it be emphasised; the point is not that judicial reasoning should be critiqued simply because future developments are unpredictable. Such unpredictability is, of course, an inherent feature of any legal system. Rather, the argument is that, while certain structural features of judicial reasoning facilitate the evolution of law, others impede it. Judicial reasoning is often presumed to be a wholly progressive force, ensuring coherence and adaptability over time. However, by attempting to impose a fixed interpretation at a given moment, judicial reasoning can also create inertia in legal doctrine, making it more resistant to necessary adaptation.

The purpose of this critique is not to suggest that judicial reasoning should be abandoned. On the contrary, it is to highlight the often-overlooked negative effects of reason-giving that can, in certain contexts, hinder the organic evolution of legal principles. Recognising this duality allows for a more nuanced understanding of how judicial reasoning shapes legal development—not as an unequivocal good, but as a mechanism that operates both as an engine of change and a potential constraint.

The tendency toward overgeneralisation discussed hereto related to institutional or structural features of legal reasoning, rather than judges' personal motivations to amplify the future impact of their normative choices. The overgeneralisation problem is further amplified by judges' personal inclinations and ambitions to extend their normative reach. Schauer attributes judicial overgeneralisation in the act of providing reasons to psychological biases, particularly the 'availability bias'.⁹¹ The availability bias relates to decision-makers' tendency to evaluate the frequency of a case based on how readily information about it can be retrieved from memory or imagination.⁹² Consequently, judges may overestimate the typicality of the precedential case they are considering, which is naturally at the forefront of their attention. As a result, their reasoning might overextend the boundaries of their precedents to a broader range of future scenarios.

Another manifestation of the generalisation problem inherent in judicial reasoning becomes particularly pronounced in 'hard cases', where courts face the challenge of reconciling justice for the immediate parties with the creation of just legal principles for all. Justice

⁹⁰ *ibid.*

⁹¹ Schauer (n 17) 657.

⁹² Amos Tversky and Daniel Kahneman, 'Availability: A Heuristic for Judging Frequency and Probability' (1973) 5 *Cognitive Psychology* 207.

Oliver Wendell Holmes succinctly captured this dilemma with his observation that ‘great cases like hard cases make bad law’,⁹⁰ underscoring a fundamental characteristic of the common law system whereby legal principles are not formulated through hypothetical projections of future case variations, but rather through the adjudication of concrete cases. This introduces a tension between delivering a just resolution for the immediate case and formulating a fair or ideal general rule.⁹¹

The crux of the problem lies in the institution of judicial reasoning, which seeks to align the unique facts of a specific case with overarching legal principles. Legal reasons can inadvertently generate precedents that either impose undue hardship on the parties involved or extend legal doctrines inappropriately. To alleviate the tension, it may be advantageous for courts to decouple the resolution of the hard case from the establishment of broader precedents. This approach appears to have influenced the US Supreme Court’s handling of *Bush v Gore*,⁹² where the Court explicitly limited its decision to ‘present circumstances’, effectively signalling that its ruling was intended to apply only to the unique facts of that case.⁹³ Legal scholars have interpreted this as a deliberate attempt to prevent the decision from setting a lasting precedent, likening it to a ‘one-day only’ ticket.⁹⁴ This perspective suggests that the Court’s decision was crafted to avoid broader implications, with some commentators viewing it as a strategic choice to contain the impact of the ruling and avoid extending its reach beyond the immediate context.

Perhaps a more straightforward method for achieving such a decoupling result may be to render decisions without providing reasons. By separating the resolution of the immediate case from the formulation of general legal principles, courts can address the specific justice concerns of the case without generating potentially problematic precedents. This approach can allow courts to resolve immediate issues effectively while avoiding the imposition of broad doctrines that could skew future legal outcomes. For such ‘hard cases’, in other words, issuing decisions without reasoning may be more advantageous than establishing ‘bad law’.

In addressing the opposition to reason-giving from the lens of precedent-setting, another critical issue emerges that warrants caution: the contemporary judicial landscape is marked by a resurgence of *seriatim*-like tendencies and multiple separate opinions, which make the task of establishing clear precedents more and more difficult.⁹⁵ Fractured and plurality decisions represent instances of extreme dissensus, where no single legal rationale commands majority support. These decisions complicate judicial interpretation, weaken *stare decisis*, and undermine the Court’s institutional authority. They lack a clear majority on every issue, thereby creating complications for the identification of (and therefore formation of) stable legal principles, compelling us to reconsider the boundaries of judicial reasoning in the context of the court’s precedent-setting role.⁹⁶ This is not a call for forced unanimity among judges or for *per curiam* opinions. Rather, it is an acknowledgment of the challenges posed

⁹⁰ *Northern Securities Co v United States* 193 US 197, 364 (1904).

⁹¹ For an innovative empirical study on how reason-giving affects the inclination to depart from formal legal rules in ‘hard cases’, see Ori Katz and Eyal Zanir, ‘Law, Justice and Reason-Giving’ (2025) *Journal of Empirical Legal Studies* (forthcoming).

⁹² *Bush v Gore* 531 US 98, 109 (2000).

⁹³ Chad Flanders, ‘Please Don’t Cite This Case! The Precedential Value of *Bush v. Gore*’ (2006) 116 *Yale LJ Pocket Part* 141.

⁹⁴ Samuel Issacharoff, ‘Political Judgments’ (2001) 68 *University of Chicago Law Review* 637, 650.

⁹⁵ Pamela C Corley and others, ‘Extreme Dissensus: Explaining Plurality Decisions on the United States Supreme Court’ (2010) 31 *The Justice System Journal* 180.

⁹⁶ *ibid.*

by a proliferation of individual reasoning that, while nuanced, may ultimately hinder the court's ability to articulate coherent legal doctrines. The Marshall Court's shift away from *seriatim* opinions was intended to address precisely these concerns by consolidating judicial authority through unified majority opinions, an approach that merits renewed consideration.¹⁰⁰

In conclusion, several challenges confront the precedent-based justification for reasoned judicial decision-making. Courts often face pressures towards overgeneralisation in their reasoning, which can lead to the creation of rigid legal precedents that unnecessarily constrain future cases and legal evolution. This tendency arises from structural limitations inherent in the judicial process, particularly the court's inability to foresee future cases. Psychological factors, such as cognitive biases (like availability bias) and personal motivations to expand one's normative influence further contribute to this overgeneralisation and to the shaping of rulings in ways that may unduly broaden their reach. This generalisation problem permeates judicial reasoning but becomes especially acute in 'hard cases'. Another critical challenge to precedent-setting posed by providing reasons arises from the trend towards multiple, individual opinions within the court. As the frequency of split decisions—especially those without a dominant majority—grows, it becomes increasingly challenging to establish stable legal precedents. These issues prompt a reassessment of the scope and limits of judicial reasoning in relation to its impact on the court's role in setting precedents. Ultimately, a more selective approach to judicial reasoning may better support the law's evolution.

C. THE EXPRESSIVE FUNCTION AND PUBLIC TRUST JUSTIFICATIONS RECONSIDERED

As previously discussed, another justification for judicial reason-giving relates to the court's expressive function. This rationale posits that judicial reasoning engages with the broader public, conveying messages about legality and morality and serving as a critical tool for upholding public trust in the judiciary. This justification, too, requires careful examination. The subsequent analysis will first challenge the role of judicial reasoning in its expressive capacity and then turn to the argument that detailed reasoning is essential for fostering public trust.

Legal reasoning is often marked by complexity, featuring specialised jargon (sometimes termed 'legalese') and intricate legal principles that may be inaccessible to the lay public.¹⁰¹ This complexity may obscure, rather than illuminate, the court's expressive intent. The message might be more effectively communicated by resorting to the simple and straightforward 'language', characteristic of the 'bottom line' of the judicial ruling, rather than through convoluted legal reasoning. For instance, in criminal law, the very act of convicting a defendant conveys moral disapproval of the defendant's conduct. The severity of the punishment reflects the degree of moral condemnation.¹⁰² Criminal verdicts, in and of themselves, can thus articulate and reinforce social and moral values in a manner that aligns more directly with public understanding.

¹⁰⁰ M Todd Henderson, 'From Seriatim to Consensus and Back Again: A Theory of Dissent' (2007) John M Olin and Economics Working Paper No 363 <https://chicagotmbound.uchicago.edu/cgi/viewcontent.cgi?article=1217&context=law_and_economics> accessed 25 March 2025.

¹⁰¹ See also Drury Stevenson, 'To Whom Is the Law Addressed?' (2003) 21 Yale Law & Policy Review 105.

¹⁰² As Dan Kahan insightfully notes, '[w]hat a community chooses to punish and how severely tells us what (or whom) it values and how much.': Dan M Kahan, 'Social Meaning and the Economic Analysis of Crime' (1998) 27 Journal of Legal Studies 609, 615.

Additionally, judicial reasoning is vulnerable to strategic manipulation, whereby courts selectively disclose the rationales behind their rulings to align with broader institutional and political objectives.¹⁰³ The reasoning process can thus be subject to distortion, allowing courts to craft justifications that serve particular interests or mitigate perceived repercussions. In contrast, decisions that focus primarily on the ultimate outcome, without extensive elaboration of the underlying reasoning, are inherently less vulnerable to such strategic manipulations. This is because the essence of ‘bottom line’ decisions lies in their unambiguous nature. Consequently, while detailed judicial reasoning can provide depth, it may also undermine the effectiveness of the trial’s expressive function. Extensive reasoning can sometimes obscure the core message of a decision, as complexity and opportunities for strategic interpretation may weaken the court’s clarity. Prioritising the ultimate outcomes of judicial rulings often serves as a more effective way to communicate the court’s legal, normative, and moral judgments, and to fulfil its expressive function.

The argument that a lack of articulated reasons undermines public trust in the legal system also merits critical evaluation. First, treating public trust as an intrinsic normative goal, in and of itself, is highly debatable. There is room to claim that the primary objective should be to ensure that the judicial system operates in a manner that merits public credence, that it ought to earn the public’s trust through its effective performance.¹⁰⁴ But public trust should not be considered in isolation from the system’s actual performance. Public trust can only be granted when the system is commendable of it. The question of public confidence in the system cannot be separated from the question of the judicial system’s objective performance. Moreover, even if public trust were to be accepted as an intrinsic normative end, in and of itself, there is no empirical evidence to support the notion that providing detailed reasons for decisions is the most effective way to pursue it. The reasoning behind a judgment may inadvertently heighten public dissent, especially in the context of acute public controversies.¹⁰⁵ People may find themselves agreeing with the court’s final decision while objecting to all or part of its underlying rationales, suggesting that, as an empirical matter, withholding detailed reasons can oftentimes prove more conducive to maintaining overall trust in the judiciary.¹⁰⁶ Lastly, even if articulating reasons for judgments were empirically to enhance public trust and even if public trust were deemed an appropriate normative endeavour, the ends do not always justify the means. In situations where providing detailed reasons could lead to verbal overshadowing or adversely affect the quality of judicial decisions, educating the public might be a more appropriate approach than succumbing to the public’s biases. Ensuring that decisions are well-founded and legally sound should take precedence over providing reasons for the mere sake of fostering public trust.

D. THE DEMOCRATIC DELIBERATION JUSTIFICATIONS RECONSIDERED

Another line of justification presented in the article focused on the democratic virtues of judicial reasoning. This class of arguments can also be further addressed. While political and judicial authority should indeed be legitimised, there is room to claim that strategic silence

¹⁰³ See Lynn M LoPucki and Walter O Weyrauch, ‘A Theory of Legal Strategy’ (2000) 49 *Duke LJ* 1405, 1437.

¹⁰⁴ David Enoch and Talia Fisher, ‘Sense and “Sensitivity”’: Epistemic and Instrumental Approaches to Statistical Evidence’ (2015) 67 *Stanford Law Review* 557, 570.

¹⁰⁵ Cohen (n 18) 514.

¹⁰⁶ Moreover, opaque reasoning might also erode public trust if perceived as lacking transparency or credibility.

by the court can function as a form of deliberation and serve as a purposeful tool for democratic engagement. The discussion will now turn to formulate this claim, that judicial silence should not be perceived as merely a passive omission. Rather, it can act as a deliberate strategy designed to promote inter-branch communication and facilitate democratic deliberation.

For as long as there have been courts, there have been unexplained court decisions. And yet, as mentioned at the outset of this article, the Supreme Court's reliance on its shadow docket has recently come under public scrutiny, with its implications for judicial transparency and reason-giving becoming a central point of concern, a concern that was notably raised during Justice Amy Coney Barrett's confirmation hearing. As Justice Elena Kagan recently wrote in a dissenting opinion, 'the majority's decision is emblematic of too much of this Court's shadow-docket decision-making which every day becomes more unreasoned'.¹⁰⁷ And Richard Pierce argues, '[w]e now have a situation in which a high proportion of major judicial decisions are made by the Supreme Court without providing any reasons for decisions'.¹⁰⁸

The phenomenon of unexplained judgments is much wider in scope than the Supreme Court's resort to its shadow docket. Courts at all levels exercise diverse forms of silence in their rulings, including by issuing decisions that deliberately avoid broader questions, by using summary dismissals, or through *per curiam* opinions with minimal explanation. Courts may also issue decisions without immediately providing comprehensive reasons, or withhold reasons for certain aspects of their decisions only to provide detailed explanations at a later time.

In his influential work, *The Supreme Court and the Idea of Progress*, Alexander Bickel highlights that many Supreme Court decisions serve as the start of a conversation between the Court and other branches of government, engaging in an ongoing dialogue rather than delivering a final verdict on complex issues.¹⁰⁹ This perspective underscores the role of judicial silence. The absence of definitive interpretation or guidance may serve to compel legislators and political institutions to deliberate more deeply, propose new legislation, or amend existing laws to address the concerns raised by the case. In this way, judicial silence can stimulate proactive legislative action and ensure that democratic processes remain robust and responsive. Moreover, when courts issue decisions with limited or no reasoning, they invite commentary from scholars and practitioners, thereby enhancing public engagement and contributing to a more informed and active citizenry. Refraining from detailed explanation may also help the judiciary avoid becoming entangled in political controversies, underscoring the judiciary's role in balancing powers and fostering a stable democratic system.

Such strategic use of silence was manifested in the Supreme Court's handling of same-sex marriage cases during its 2014 term.¹¹⁰ When the Court declined to grant *certiorari* in several pivotal same-sex marriage appeals, thereby allowing state rulings permitting same-sex marriage to remain in effect, it did so without offering reasons, signifying a deliberate decision

¹⁰⁷ *Whole Woman's Health* (n 5), cited in Vladeck, 'Putting the "Shadow Docket" in Perspective' (n 2) 290.

¹⁰⁸ Pierce (n 3).

¹⁰⁹ Alexander M Bickel, *The Supreme Court and the Idea of Progress* (Yale University Press 1970).

¹¹⁰ See for example *Bogan v Baskin* 135 S Ct 316 (2014) denying *certiorari* to 766 F3d 648 (7th Cir 2014) (Indiana); *Walker v Wolf* 135 S Ct 316 (2014) denying *certiorari sub nomine* to *Baskin v Bogan* 766 F3d 648 (7th Cir 2014) (Wisconsin); *Schaefer v Bostic* 135 S Ct 308 (2014) denying *certiorari* to 760 F3d 352 (4th Cir 2014) (Virginia); *McQuigg v Bostic* 135 S Ct 314 (2014) denying *certiorari sub nomine* to *Bostic v Schaefer* 760 F3d 352 (4th Cir 2014) (Virginia); *Rainey v Bostic* 135 S Ct 286 (2014) denying *certiorari sub nomine* to *Bostic v Schaefer*, 760 F3d 352 (4th Cir 2014) (Virginia); *Smith v Bishop* 135 S Ct 271 (2014) denying *certiorari* to 760 F3d 1070 (10th Cir 2014) (Oklahoma); *Herbert v Kitchen* 135 S Ct 265 (2014) denying *certiorari* to 755 F3d 1193 (10th Cir 2014) (Utah).

to maintain a degree of judicial silence.¹¹¹ This approach mirrors an historical pattern of using silence to facilitate ongoing deliberation.¹¹² As Chris Schmidt notes, after *Brown v Board of Education*,¹¹³ the Court similarly employed *per curiam* decisions with minimal reasoning to extend its desegregation mandate to additional public institutions. These silent (or minimally reasoned) decisions were intended to advance civil rights discourse in society and allowed legislative and public discourse to evolve without being prematurely closed off.¹¹⁴

Judicial silence, in other words, is far from an empty void. It holds the potential to serve as a constructive element of democratic governance. By choosing not to provide extensive reasoning, courts can stimulate public and institutional dialogue, encourage legislative action, and prevent judicial overreach. Silence, in this context, can indeed speak volumes, facilitating democratic deliberation and strengthening the overall health of democratic institutions.

E. LITIGANT AUTONOMY JUSTIFICATIONS RECONSIDERED

The argument that judicial reason-giving enhances litigant autonomy and procedural justice is compelling, but it, too, merits further scrutiny. It is important to emphasise that, while judicial reasoning provides clarity and justification, it is the substantive outcomes of decisions that most directly shape a litigant's ability to influence their legal situation. Party autonomy and control are primarily determined by the concrete results of legal proceedings, the rulings that dictate real-world consequences, rather than by the explanatory narratives that courts provide. In this sense, the emphasis on judicial reasoning, while significant, should not obscure the fact that it is the decision itself that ultimately defines the litigant's legal reality.

Furthermore, while reason-giving is designed to amplify litigants' voices in trial, as mentioned earlier, it is essential to recognise that judicial reasoning is often constrained by the formalistic language and procedural strictures of the law. These constraints can obstruct the effective communication of individual narratives within the framework of legal reasoning. Litigants may find their opportunities to express their perspectives and engage substantively with the court restricted by the rigid structures and formalities that underpin judicial reasoning. This can lead to a perception that their agency is not genuinely affirmed, but rather diminished, by the judicial process. Additionally, the complexity and technical nature of legal reasoning can obscure the clarity necessary for litigants to comprehend fully how their rights were assessed by the court, potentially undermining their ability to protect them. Disadvantaged parties, in particular, may be adversely affected by the technical and legalistic nature of judicial reasoning.

This brings us to the last point: reliance on reason-giving as a mechanism to affirm litigants' agency may create only a superficial semblance of participatory justice, while concealing deeper substantive and systemic threats to litigant autonomy. Although reasoned judgments are intended to promote transparency, they do not inherently address, and may indeed obscure, underlying power imbalances or the inadequate access to legal remedies faced by marginalised groups and individuals. Disparities in legal representation can result in uneven

¹¹¹ Chris Geidner, 'Cert. Denied, Stays Denied, Marriage Equality Advanced: How the Supreme Court Used Nonprecedential Orders to Diminish the Drama of the Marriage Equality Decision' (2015) 76 *Ohio State Law Journal* Furthermore 161.

¹¹² Chris Schmidt, 'Some Thoughts on a "Silent" Supreme Court' (*SCOTUSnow*; 28 October 2014) <<https://blogs.kentlaw.iit.edu/scotus/thoughts-on-a-silent-supreme-court/>> accessed 25 March 2025.

¹¹³ *Brown v Board of Education* 347 US 483 (1954).

¹¹⁴ *ibid.*

articulation within reasoned judgments, with less-resourced litigants frequently unable to advocate as effectively as their more advantaged counterparts.¹¹⁵ Consequently, the reasoning provided may fail to capture fully or address the issues pertinent to less-resourced litigants, thereby perpetuating existing inequities in the legal system.

In closing, while the practice of providing reasons for judicial decisions is often seen as a hallmark of transparency and fairness, it does not necessarily guarantee that litigants can effectively shape their legal trajectory, assert their interests, or fully engage with the judicial process. A reasoned judgment does not automatically enhance an individual's ability to influence outcomes, participate meaningfully, or experience a fair and balanced process.

IV. CONCLUSION

The central aim of this article has been to demonstrate that the considerations against reasoned judgments extend well beyond issues of judicial economy. While the merits of reasoned judgments—such as enhancing transparency, ensuring judicial accountability, fostering a dynamic precedential and legal system, bolstering public trust, legitimising democratic governance, and safeguarding litigant autonomy—are well-founded, the case against reasoned judgments involves deeper, more nuanced critiques.

One illustrative example of the potentially self-defeating nature of universal reasoning is the phenomenon of verbal overshadowing. This cognitive effect introduces the possibility that the verbal articulation of reasons may inadvertently distort the judicial decision-making process. It may be particularly salient in witness credibility assessments, where some of the underlying elements are challenging to convey accurately through verbal expressions. Future research should examine and identify additional areas of judicial decision-making that may be similarly affected.

Another example concerns hard cases at risk of creating bad law. The pressure to provide reasoned judgments in such scenarios may lead to the formulation of precedents that are misaligned with broader legal principles and interests. To mitigate this risk, it may be advisable for courts to decouple the resolution of individual 'hard' cases from the formulation of general legal doctrines. This approach is exemplified by the Supreme Court's handling of *Bush v Gore*.¹¹⁶ A formal mechanism that allows courts systematically to forego reasoning in cases where there is a potential conflict between the adjudication of the immediate matter and the articulation of general legal principles could ensure that the specific complexities of individual cases do not unduly influence the creation of general rules intended to guide future cases.

A third instance where judicial reason-giving may be self-defeating relates to cases where the court's silence speaks louder than its articulation of reasons and contributes in a more profound manner to the democratic deliberation process. Judicial silence can serve as a catalyst for proactive legislative initiatives and bolster the effectiveness of democratic processes by keeping them adaptable and responsive. It can also prompt public debate and engagement. Such strategic use of silence was evident in the Supreme Court's handling of same-sex marriage cases during its 2014 term, where minimally reasoned or silent decisions were intended to allow legislative and public dialogue to evolve organically without being prematurely curtailed.

¹¹⁵ *Assy* (n 61).

¹¹⁶ *Bush* (n 95).

This analysis does not advocate for a return to Aristotelian intuition-based adjudication, an approach rooted in *phronesis* (practical wisdom), where judges rely on experience and contextual judgment rather than rigidly adhering to formal legal rules. Instead, it highlights the imperative to balance the benefits of judicial reasoning with critical awareness of inherent limitations. The article aspires to establish a more nuanced approach to judicial decision-making, one that fosters rigorous debate on the appropriate scope and application of reasoned judgments. It invites further discourse on the appropriate boundaries of the institution of judicial reason-giving. By advocating for a discerning and balanced approach, it aims to refine the use of reasoned judgments, ensuring their effectiveness while mitigating potential drawbacks.

Addressing the Challenges of Due Process and Ethical Disparities in International Arbitration: The Need for Reform

ALFIE DOVE*

ABSTRACT

This article critically examines the challenges facing international arbitration, with a particular focus on the lack of a common understanding of due process and ethical standards. International arbitration has long been plagued by issues such as high costs, unpredictability, and the inability to appeal, but recent concerns about due process and ethics suggest a deeper problem. Due process and ethics are fundamental to the integrity of arbitration and, without proper reform, these issues threaten to undermine the system's credibility. This article begins by defining due process and ethics, followed by a detailed exploration of the problematic concept of 'due process paranoia'. The author's analysis is structured around three key arguments: first, that differing interpretations of due process across jurisdictions lead to its use or misuse as a 'sword' rather than as a protective shield; secondly, that the unequal ethical obligations imposed on counsel by different legal systems create ambiguity and potential ethical conflicts, such as 'double deontology' and 'inequality of arms'; and thirdly, that adopting clearer terminology for due process rights and establishing a universal International Code of Ethics would enhance predictability and fairness in international arbitration. By addressing these gaps, international arbitration could, in the author's view, avoid further erosion of its legitimacy and continue to thrive. Ultimately, this article suggests that, without harmonising these core principles, international arbitration risks being displaced by alternatives, leaving its future uncertain.

Keywords: due process, ethics, arbitration, international arbitration, legal reform

I. INTRODUCTION

International arbitration is growing both in its importance and usage.¹ Proponents argue that several key advantages characterise international arbitration: it promises flexibility and

* LLB, LLM (University of Warwick). I must express my gratitude to Velimir Živković for comments on an earlier draft. All views expressed are the author's own.

¹ See for example Gary B Born, *International Commercial Arbitration* (3rd edn, Kluwer Law International 2021) 7; Derick H Lindquist and Ylli Dautaj, 'AI in International Arbitration: Need for the Human Touch' (2021) 2021 *Journal of Dispute Resolution* 39, 43.

neutrality,² as well as the speedy resolutions of international disputes.³ And as international trade (such as cross-border contracts, maritime disputes, and construction and infrastructure projects) and foreign investment continue to expand, so too does the reliance on international arbitration to ensure a fair and effective resolution between parties from diverse backgrounds. However, the same diversity that gives international arbitration its strength simultaneously makes us question its true effectiveness. This is because, despite some of the more common challenges that threaten to undermine arbitration, such as issues of cost,⁴ issues relating to the unpredictability of the process, and the lack of an ability to appeal,⁵ there are inconsistencies in the interpretation of the concept of due process and disparities in ethical obligations. Such fragmentation not only raises concerns about the coherence of international arbitration but also reflects the shift from a ‘solidaristic’ field, characterised by shared ethical norms and expectations,⁶ to a ‘polar[is]ed’ field in which varying legal traditions, procedural expectations, and ethical standards contribute to the current fragmented landscape.⁷ This polarisation will only grow as we welcome more players into the field. The list of drawbacks seems to be endless and, in the author’s opinion, these are serious challenges that hamper the effectiveness of international arbitration. For if abuse of due process objections and ‘level[ling] the playing field’⁸ of ethical considerations are not given thought, then the credibility of international arbitration as an effective dispute resolution mechanism may be undermined. This article argues that international arbitration must undergo reform to address these concerns. This article’s main argument consists of three components.

First, in Section II, this article will argue that the varied understandings of due process among different jurisdictions result in the principle being malleable and capable of being used as a ‘sword’.⁹ Due process (sometimes referred to as ‘natural justice’¹⁰) typically refers to parties having notice of the proceedings and having the right to be heard and present evidence in front of an independent and impartial tribunal that *ideally* treats all parties involved with

² See for example Tony Dymond and Raeesa Rawal, ‘International Arbitration’ in Anthony Speaight and Matthew Thorne (eds), *Architect’s Legal Handbook: The Law for Architects* (10th edn, Routledge 2021) 276; Mentor Lecaj and Granit Curri, ‘Advantages of International Commercial Arbitration in Resolving the Commercial Contests’ (2021) 10 *Perspectives of Law and Public Administration* 96, 100.

³ Irene Welsler and Christian Klausegger, ‘Fast Track Arbitration: Just Fast or Something Different?’ (2009) *Austrian Arbitration Yearbook* 259, 272.

⁴ See for example Weixia Gu, ‘Security for Costs in International Commercial Arbitration’ (2005) 22 *Journal of International Arbitration* 167, 168; Jean-Claude Najar, ‘Inside Out: A User’s Perspective on Challenges in International Arbitration’ (2009) 25 *Arbitration International* 515, 517–18.

⁵ See for example Rowan Platt, ‘The Appeal of Appeal Mechanisms in International Arbitration: Fairness over Finality?’ (2013) 30 *Journal of International Arbitration* 531, 532; Russell Thirgood, ‘Appeals in Arbitration: “To Be or Not to Be”’ (2021) 87 *Arbitration* 423, 426; Paul Friedland and Loukas Mistelis, ‘2015 International Arbitration Survey: Improvements and Innovations in International Arbitration’ (White & Case 2015) 8 <<https://shop.americanbar.org/PersonifyImages/ProductFiles/262739281/4-And%20Justice%20for%20All.pdf>> accessed 2 January 2024.

⁶ Emmanuel Gaillard, ‘Sociology of International Arbitration’ (2015) 31 *Arbitration International* 1, 14.

⁷ *ibid* 13–14.

⁸ Edna Sussman and Solomon Ebere, ‘All’s Fair in Love and War—Or Is It? Reflections on Ethical Standards for Counsel in International Arbitration’ (2011) 22 *The American Review of International Arbitration* 611, 619; Robert W Wachter, ‘Ethical Standards in International Arbitration: Considering Solutions to Level the Playing Field’ (2011) 24 *Georgetown Journal of Legal Ethics* 1143, 1160; Margaret L Moses, ‘Ethics in International Arbitration: Traps for the Unwary’ (2012) 10 *Loyola University Chicago International Law Review* 73, 73, 80.

⁹ Lucy Reed, ‘Ab(use) of Due Process: Sword vs Shield’ (2017) 33 *Arbitration International* 361, 364.

¹⁰ *ibid* 366.

equality.¹¹ Nothing is regarded as more ‘paramount’¹² to the integrity and credibility of legal dispute resolution mechanisms than due process. Yet, despite national laws,¹³ institutional rules, and conventions covering the principle,¹⁴ due process in the context of international arbitration is neither perfect nor bulletproof. This is because due process does not always function as seamlessly as we may believe and it is not applied in a consistent fashion; rather, its implementation can be flawed or subject to manipulation. In essence, while due process is supposed to protect fairness, in practice it does not function as an absolute safeguard against procedural abuse. This is especially so given the variations in how the concept of due process is interpreted around the world.¹⁵

What is common among the international landscape, though, is that due process has a pivotal role in maintaining both the legitimacy¹⁶ and integrity of the arbitration process.¹⁷ In this context, legitimacy pertains to the degree to which arbitration is recognised as a credible and fair mechanism and integrity encompasses the procedural and ethical fairness of the process. Due process plays a fundamental role in upholding these principles as its absence may lead to perceptions of bias or insufficient procedural safeguards. A lack of confidence in the fairness of arbitral proceedings may prompt parties to challenge arbitral decisions, thereby undermining the effectiveness of the arbitration system. Therefore, if due process rights have been violated, an arbitral award may be set aside or refused recognition.¹⁸ Sometimes parties are fearful of the varied understandings of due process and this gives rise to ‘due process paranoia’,¹⁹ a notion coined by an interviewee in a 2015 survey conducted by Queen Mary University of London (‘QMUL’).²⁰ The survey defines ‘due process paranoia’ as ‘a reluctance by tribunals to act decisively in certain situations for fear of the arbitral award being challenged on the basis of a party not having had the chance to present its case fully’.²¹

¹¹ See for example UNCITRAL Model Law on International Commercial Arbitration (adopted 21 June 1985, amended 7 July 2006) UN Docs A/40/17, annex I and A/61/17, annex I (‘UNCITRAL Model Law’), art 34; Margaret L Moses, *The Principles and Practice of International Commercial Arbitration* (3rd edn, CUP 2017) 219.

¹² Russell Thirgood and Erika Williams, ‘The Non-Responsive Respondent: Taking an Arbitration Forward and How’ (2019) 85 *Arbitration* 65, 70.

¹³ See for example UNCITRAL Model Law (n 11) art 18; Arbitration Act 1996, s 33; International Arbitration Act 1974 (Cth), s 18C; Arbitration and Conciliation Act 1996 (India), s 18; Arbitration and Conciliation Act 1988 (Nigeria), s 15(2).

¹⁴ See for example Convention on the Recognition and Enforcement of Foreign Arbitral Awards (opened for signature 10 June 1958, entered into force 7 June 1959) 330 UNTS 3 (‘New York Convention’), art V(1)(b); UNCITRAL, ‘Arbitration Rules of the United Nations Commission on International Trade Law’ (15 December 1976) UN Doc A/31/98, 31st Sess Supp No 17, as amended in 2010 (A/RES/65/22) and 2013 (A/RES/68/109) (‘UNCITRAL Arbitration Rules’) art 17(1); ICC Arbitration Rules (2021) (‘ICC Rules’) art 22(4); LCIA Arbitration Rules (October 2020) (‘LCIA Rules’) art 14.1(i); Permanent Court of Arbitration (‘PCA’) Arbitration Rules (December 2012) art 17; SCC Arbitration Rules (2023) art 23(2); Swiss Arbitration Centre Swiss Rules of International Arbitration (June 2021) art 32; SIAC Arbitration Rules (2016) r 41.2; Hong Kong Administered Arbitration Rules (2018) art 13; China International Economic and Trade Arbitration Commission Arbitration Rules (2014) art 35(1); WIPO Arbitration Rules (2021) art 37(b); ICSID Rules of Procedure for Arbitration Proceedings (April 2006) (‘ICSID Arbitration Rules’) r 26(3); Japan Commercial Arbitration Association Commercial Arbitration Rules (2019) art 24; American Arbitration Association Commercial Arbitration Rules (2013) r 18(a).

¹⁵ Bernardo M Cremades, ‘The Use and Abuse of “Due Process” in International Arbitration’ (2016) 9 *Arbitraje: Revista de Arbitraje Comercial y de Inversiones* 661, 665.

¹⁶ Enric Picanyol, ‘Due Process and Soft Law in International Arbitration’ (2015) 24 *Spain International Review* 29, 35.

¹⁷ Won L Kidane, *The Culture of International Arbitration* (OUP 2017) 243.

¹⁸ See for example New York Convention (n 14) art V(1)(b); UNCITRAL Model Law (n 11) arts 34(2)(a)(ii)–(iii); *Hui v Esposito Holdings Pty Ltd* [2017] FCA 648, (2017) 345 ALR 287 [183]–[185].

¹⁹ Friedland and Mistelis (n 5) 10.

²⁰ See *ibid.*

²¹ *ibid.*

In this context, the lack of a common understanding and definition of due process, the potential abuse of due process, and unequal ethical obligations collectively undermine the efficiency and legitimacy of international arbitration and contribute to a feeling of due process paranoia. This occurs because tribunals are often hesitant to enforce procedural rules strictly, fearing that doing so may lead to allegations of unfairness and, ultimately, annulment or non-recognition of the award. This is ‘problematic’²² as the resulting unpredictability makes arbitration users sceptical of whether the system can reliably provide timely and fair resolutions.

Secondly, in Section III, this article will contend that the unequal ethical obligations imposed on counsel create a legal landscape of ambiguity where issues of ‘double deontology’²³ and ‘inequality of arms’²⁴ emerge as a result of a conflict in ethical standards. Usually, a code of ethics refers to a set of moral principles and values that guide human behaviour,²⁵ distinguishing right from wrong. These ethical standards regulate parties’ conduct and are therefore crucial for maintaining procedural fairness. However, different jurisdictions (and therefore practitioners) operate under differing levels of ethical obligation, which result in the creation of an uneven playing field at the international level, replicating and amplifying these differences.

Finally, in Section IV, this article will consider how to address these challenges by examining whether it is possible to harmonise the terminology used to describe due process rights. Such harmonisation, in the author’s view, would contribute to predictability and overall efficiency. Further, the author will argue that employing an ‘International Code of Ethics’ for international arbitration will reduce conflicting ethical standards because such a code would provide a single definition of ethics in an attempt to resolve the problem of differing or unequal ethical obligations that currently plague international arbitration. Importantly, as will be discussed, these proposals will maintain party autonomy and flexibility. Thus, this article seeks to contribute to the ongoing discourse on arbitration reform by advocating for clearer procedural definitions and ethical alignment.

II. DUE PROCESS CONCERNS

This article will now examine the concept of due process in more detail, with the aim of revealing broader challenges that result from the differing understandings of due process. There is a well-founded concern that, as the concept of due process is interpreted differently across jurisdictions, this leads to variations in its application for parties to international arbitration.²⁶ In response, the international arbitral system has arguably adopted a ‘hands off’²⁷ approach which has resulted in this lack of uniform understanding being replicated at the international level. For example, it is a core principle of international arbitration that all parties have a right to present their case. For example, article 18 of the United Nations Commission on International Trade Law Model Law on International Commercial Arbitration

²² *ibid.*

²³ Catherine A Rogers, *Ethics in International Arbitration* (OUP 2014) 107–10.

²⁴ *ibid.* 108.

²⁵ Mark S Schwartz, ‘Effective Corporate Codes of Ethics: Perceptions of Code Users’ (2004) 55 *Journal of Business Ethics* 321, 323.

²⁶ See for example Philip Chong and Blake Primrose, ‘Summary Judgment in International Arbitrations Seated in England’ (2017) 33 *Arbitration International* 63, 70; Gabrielle Kaufmann-Kohler, ‘Globalization of Arbitral Procedure’ (2003) 36 *Vanderbilt Journal of Transnational Law* 1313, 1322; Fabrizio Fortese and Lotta Hemmi, ‘Procedural Fairness and Efficiency in International Arbitration’ (2015) 3 *Groningen Journal of International Law* 110, 115–16.

²⁷ Tony Cole and Pietro Ortolani, *Understanding International Arbitration* (Routledge 2020) 128.

(‘UNCITRAL Model Law’) places an emphasis on ‘equality’²⁸ while expressing that parties must have a ‘full’²⁹ opportunity to present their case. In contrast, section 33(1)(a) of the Arbitration Act 1996 emphasises fairness and impartiality and requires that parties must have a ‘reasonable opportunity’³⁰ to present their case. Article 22(4) of the International Chamber of Commerce (‘ICC’) Rules adopts a similar approach.³¹ To complicate matters, the 2014 International Centre for Dispute Resolution (‘ICDR’) Rules provide parties with a ‘fair opportunity’³² to present their case. It is quite clear from a cursory overview of these examples (and there are certainly more³³) that, while there is broad agreement that parties must have a right to present their case, precisely how this right is expressed is often subtly different.

In fact, three different qualifying terms (‘full’, ‘reasonable’, and ‘fair’) are used to describe the scope of rights afforded to parties. These terms function as modifiers because they adjust the meaning of the underlying right, influencing the manner in which it is applied. This subtle variation is significant because the specific choice of wording may shape the obligations imposed on arbitrators and the expectations placed on parties. This is problematic not only due to the subtle and clear differences between the standards of a ‘full’, ‘fair’, or ‘reasonable’ opportunity to present a case but also because these concepts are themselves inherently malleable. For example, what is *fair* is not universal and may differ from person to person (and more importantly from legal system to legal system). In the context of international arbitration, this concern is magnified as, in each case, justice is put in the hands of the arbitrators who act in a manner influenced by their own moral compasses which are invariably grounded in the particularities of their home legal system. And, further, these particularities may (though not always) differ from the standards to which the parties are accustomed.

Some academics have sought to bridge the gaps in the ways that different arbitral standards interpret the concept of due process. For example, when comparing the respective approaches in the UNCITRAL Model Law and the Arbitration Act 1996, Nigel Blackaby and others contend that a *full* opportunity does in fact mean a *reasonable* opportunity.³⁴ Blackaby and others make this suggestion on the basis that the phrase ‘full opportunity’ inherently implies reasonable limitations. They argue that no legal system could uphold an unrestricted right to present one’s case without imposing procedural constraints, such as time limits. These reasonable restrictions, they argue, ensure that arbitration remains efficient and fair.³⁵ However, they assume a widespread acceptance of this interpretation, which is not universally shared among arbitration scholars. For example, Jan Paulsson and Gary Born have explicitly distinguished between these standards, arguing that ‘full opportunity’ establishes a more expansive procedural right than the ‘reasonable opportunity’ formulation.³⁶ This difference in terminology reflects distinct philosophical approaches to arbitral due process, rather than merely semantic variations.

But, in assuming a widespread acceptance, Blackaby and others essentially conflate the two terms and sideline the main issue, namely that there is a difference in how this right

²⁸ UNCITRAL Model Law (n 11) art 18.

²⁹ *ibid.*

³⁰ Arbitration Act 1996, s 33(1)(a).

³¹ ICC Rules (n 14).

³² ICDR Arbitration Rules (2014) art 20(1).

³³ See for example French Code of Civil Procedure, art 1486; Private International Law Act 1989 (Switzerland), art 182(3); Arbitration Act 2019 (Sweden), s 24; Dutch Code of Civil Procedure, art 1039(1).

³⁴ Nigel Blackaby and others, *Redfern and Hunter on International Arbitration* (6th edn, OUP 2015) para 6.14.

³⁵ *ibid.*

³⁶ Jan Paulsson, *The Idea of Arbitration* (OUP 2013) 90; Born (n 1) 455–60.

is expressed between the two systems' rules. A more convincing argument would have acknowledged that a *full* opportunity is referring to an unimpeded chance to present a case, whereas a *reasonable* opportunity varies on a case-by-case basis, depending on the facts of a particular case and connotes sufficient (though not necessarily full) access in disclosure, oral submission, and cross-examination.³⁷ This narrowing of language used over time makes sense:³⁸ the law has clearly put due process at the core of its composition and, by tightening the language used, it has ultimately attempted to reduce due process objections.³⁹ In theory, this seems like an appropriate evolution. But, in practice, this fragmented understanding is problematic for three reasons.

First, the existence of differing definitions of due process across international arbitration gives rise to 'due process paranoia'.⁴⁰ While arbitrators generally apply the standard of the jurisdiction or rules governing the arbitration, the lack of a consistent definition creates uncertainty about how their decisions might be perceived or challenged under different frameworks, leading arbitrators to become overcautious with regard to due process. As a result, they may concede to the parties' requests either out of fear that their award will be challenged or that their impartiality or fairness will be questioned. This is a 'real threat'⁴¹ and was reflected in responses to the updated QMUL survey in 2018,⁴² where a worry was expressed amongst participants that due process issues appeared to restrain arbitrators from taking bold actions in handling cases.⁴³ Unfortunately, this is not rare.⁴⁴ And this raises concerns about a potential imbalance between procedural propriety and the fair and efficient resolution of disputes. For example, an arbitrator who is faced with these pressures may accept continued requests for time extensions,⁴⁵ or allow for continued amendments to be made to written submissions,⁴⁶ or accept the late introduction of evidence,⁴⁷ or grant requests to reschedule hearings at the last minute.⁴⁸

Paul Obo Idornigie, Enuma U Moneke and Ogochukwu Juliet Mgbakogu argue that an arbitrator who is put in the position of having to balance the efficient resolution of arbitral proceedings with upholding the parties' due process rights⁴⁹ becomes confused.⁵⁰ More specifically, they contend that arbitrators are often placed in an impossible position where any decision risks either compromising the timely resolution of disputes or potentially violating a party's fundamental rights. In the author's view, this is a valid contention: balancing efficiency with due process rights is a challenging task. Arbitral tribunals may be inclined to grant applications on due process-related matters out of a concern to maintain the integrity of the

³⁷ Chong and Primrose (n 26).

³⁸ Reed (n 9) 375.

³⁹ *ibid* 368.

⁴⁰ Klaus Peter Berger and J Ole Jensen, 'Due Process Paranoia and the Procedural Judgment Rule: A Safe Harbour for Procedural Management Decisions by International Arbitrators' (2016) 32 *Arbitration International* 415, 421.

⁴¹ *ibid* 420.

⁴² Paul Friedland and Stavros Brekoulakis, '2018 International Arbitration Survey: The Evolution of International Arbitration' (White & Case 2018) 27 <[https://www.qmul.ac.uk/arbitration/media/arbitration/docs/2018-International-Arbitration-Survey--The-Evolution-of-International-Arbitration-\(2\).PDF](https://www.qmul.ac.uk/arbitration/media/arbitration/docs/2018-International-Arbitration-Survey--The-Evolution-of-International-Arbitration-(2).PDF)> accessed 2 January 2024.

⁴³ *ibid*.

⁴⁴ *ibid* 10.

⁴⁵ Berger and Jensen (n 40) 419.

⁴⁶ *ibid*.

⁴⁷ *ibid*.

⁴⁸ *ibid*.

⁴⁹ UNCITRAL Arbitration Rules (n 14) art 17.

⁵⁰ Paul Obo Idornigie, Enuma U Moneke and Ogochukwu Juliet Mgbakogu, 'Due Process Paranoia in Arbitral Proceedings: Myth or Reality?' (2020) 39 *The Arbitrator and Mediator* 9, 12.

proceedings and ensure the enforceability and recognition of their awards.⁵¹ In the author's view, while Idornigie, Moneke and Mgbakogu diagnose a key issue in international arbitration, they omit to explain that this problem is the result of an inherently fragmented understanding of what constitutes due process. This fragmentation creates uncertainty among arbitrators when handling due process requests, which, in turn, contributes to due process paranoia. In this sense, due process paranoia is understandable, as arbitrators may fear that denying such requests could jeopardise the enforceability of their awards.⁵² This makes us question where the line should be drawn between legitimate concerns around due process and obstructionist behaviour. This article will return to this question later.

Secondly, this due process paranoia can manifest itself in some parties choosing to use due process as a sword,⁵³ where they strategically wield due process objections and exploit any perceived vulnerability in the procedural framework to achieve tactical goals.⁵⁴ While due process is primarily intended to safeguard parties' rights and uphold the integrity of proceedings, it is important to acknowledge that there may be circumstances where its use as a sword is justified. For example, a party may legitimately invoke due process objections to exclude improperly obtained evidence or prevent procedural unfairness that could compromise the adjudicative process. The concern arises when such objections are leveraged in bad faith or as a mere excuse for procedural manoeuvring rather than on the basis of genuine concerns about fairness.⁵⁵ But how does this work in practice? How do parties exploit a framework that is designed principally to protect them and the integrity of the arbitral process as a tactical tool to aid them in proceedings? Lucy Reed provides a useful explanation as to how certain parties can abuse due process objections.⁵⁶ Reed argues that certain parties to arbitral proceedings consistently launch due process objections to gain an advantage.⁵⁷ The image painted by Reed makes us question how an alternative dispute resolution mechanism, supposedly built upon integrity and the protection of the parties' due process rights, can be so riddled with 'gamesmanship'.⁵⁸

Reed's concerns are reflected both in case law and in the comments of other authors.⁵⁹ For example, in *China Heilongjiang International Economic & Technical Cooperative Corp v Mongolia*,⁶⁰ Mongolia repeatedly raised due process objections and argued that it had insufficient preparation time. The tribunal decided that these objections were an attempt to disrupt rather than protect procedural fairness. However, objections are not just launched *during* the proceedings. Some parties weaponise due process in an attempt to resist enforcement of an award, as can be seen in *PT First Media v Astro Nusantara International BV*⁶¹ and *CEAC*

⁵¹ Reed (n 9) 372.

⁵² Berger and Jensen (n 40) 428.

⁵³ Reed (n 9).

⁵⁴ Peter A Halprin, 'Resisting Guerrilla Tactics in International Arbitration' (2019) 85 *Arbitration* 87, 88–90.

⁵⁵ Michael Byers, 'Abuse of Rights: An Old Principle, A New Age' (2002) 47 *McGill LJ* 389, 407.

⁵⁶ Reed (n 9) 375.

⁵⁷ *ibid.*

⁵⁸ William W Park, 'Arbitration's Protean Nature: The Value of Rules and the Risk' in Julian DM Lew and Loukas A Mistelis (eds), *Arbitration Insights: Twenty Years of the Annual Lecture of the School of International Arbitration* (Kluwer Law International 2007) 347.

⁵⁹ See for example John P Gaffney, "'Abuse of Process" in Investment Treaty Arbitration' (2010) 11 *Journal of World Investment and Trade* 515, 517–18; Eric De Brabandere, "'Good Faith", "Abuse of Process" and the Initiation of Investment Treaty Claims' (2012) 3 *Journal of International Dispute Settlement* 609, 612.

⁶⁰ *China Heilongjiang International Economic & Technical Cooperative Corp v Mongolia* (PCA, Case No 2010-20, 30 June 2017).

⁶¹ *PT First Media v Astro Nusantara International BV* [2013] SGCA 57.

Holdings Limited v Montenegro.⁶² The facts of each are largely the same: the claimant in *PT First Media* and the respondent in *CEAC Holdings* alleged a due process violation after an unfavourable decision. In the author's view, these cases illustrate the fact that, in some instances, due process can be misused by parties. Instead of being treated as a fundamental protective right, due process instead appears to be used more as a tool for procedural manoeuvring. This aligns with the observations of other academics.

For example, Philippe Fouchard expresses the view that alternative dispute resolution is persistently affected by unethical behaviour.⁶³ In response to the growing concerns about abusive procedural tactics, such as excessive delay tactics or bad faith objections, he argues that these issues are turning what was intended to be a streamlined dispute resolution mechanism into a process resembling litigation. And, in the author's view, this could not be more accurate in this context. There is one explanation. The fragmented understanding of due process that exists internationally is, in the author's view, assisting parties to use these tactical plays to their advantage. But who do we blame in this situation? While the law plays a role in facilitating such tactics, it is ultimately the actions of the parties and their counsel that bring these strategies to life. The law itself is merely a tool, one that can be manipulated to serve strategic ends. The impression left for the world of international arbitration is that the law inadvertently creates opportunities for the use of such strategies.

It is worth noting that some academics, such as William Park, take the view that such tactics only occur in *some* instances, which may hinder an arbitrator's efficiency in decision-making.⁶⁴ In this context, efficiency typically refers to an arbitrator's capacity to conduct proceedings in a timely manner without unnecessary disruptions. Such disruptions can prolong proceedings because, instead of focusing on the substantive merits (such as hearing arguments and considering evidence), arbitrators are forced to allocate time to resolving procedural disputes. In the author's view, while the latter part of Park's argument about hindering efficiency is plausible, given the obstruction to the proceedings, the former argument about the infrequency of such tactics requires some scrutiny. The use of 'some', as a quantifier, is inappropriate when weighed against the evidence. For example, the 2012 QMUL survey found that 68 per cent of respondents reported experiencing guerrilla tactics, such as bad faith objections, intentional withholding of evidence, or frivolous challenges to arbitrators in arbitral proceedings.⁶⁵ However, it is notable that Park's comments pre-date the QMUL 2012 survey by nine years. Perhaps this demonstrates the rise of such tactical foul play. Nevertheless, based on the evidence cited above, it seems more appropriate to replace Park's qualifier of *some* with *many* to better reflect the amplification and frequency of these tactics over time. So widespread has the phenomenon become that Reed cites a case, which she terms the 'Due Process List Case',⁶⁶ where counsel made numerous complaints regarding their due process rights and were urged to construct a list of such violations. This was, in the author's view, an attempt to gain strategic advantage, and such tactical efforts may jeopardise the overall reputation of international arbitration as a reliable method of dispute resolution.⁶⁷

Thirdly, the cumulative effect of these issues has a significant impact on the speed and expense of arbitral proceedings. This cumulative effect can be understood as a chain reaction: when one procedural issue arises within the complex web of arbitration, it often

⁶² *CEAC Holdings Limited v Montenegro*, ICSID Case No ARB/14/8.

⁶³ Philippe Fouchard, 'Où va l'arbitrage international' (1989) 34 McGill Law Journal 435, 436.

⁶⁴ Park, 'Arbitration's Protean Nature' (n 58).

⁶⁵ Sussman and Ebere (n 8) 612.

⁶⁶ Reed (n 9).

⁶⁷ Emmanuel Gaillard, 'Abuse of Process in International Arbitration' (2017) 32 ICSID Review 17, 17.

triggers additional complications. By abusing due process objections, three further problems arise. First, constant objections grounded in due process concerns can slow down the proceedings.⁶⁸ This undermines one of the key advantages of arbitration: efficiency. Arbitration is often preferred over litigation because it is designed to provide a faster resolution. However, when parties exploit procedural objections to create delays, arbitration begins to resemble its fraternal brother of litigation in terms of both length and complexity. Secondly, the inefficiency can make the process more expensive and this can be a deliberate tactic for one party to use (when they have ‘deeper pockets’) to try to make the other party run out of money and so seek settlement.⁶⁹ Thirdly, parties may begin to lose trust in the arbitration system.⁷⁰ When parties perceive arbitration as susceptible to manipulation and excessive delays, they may lose confidence in its effectiveness as an alternative to litigation. This scepticism may lead parties to reconsider their reliance on arbitration. For instance, the 2018 QMUL survey found that due process concerns and procedural inefficiencies were among the top reasons for dissatisfaction with arbitration.⁷¹ A decline in trust may certainly contribute to a decline in use over time. In some contexts, such as investor-state arbitration, some states have withdrawn from treaties containing arbitration provisions. But in international commercial arbitration, there does appear to be a strong reliance on its use. This suggests that, despite frustrations with procedural abuse, parties may still prefer it over litigation. Therefore, the focus should be on strengthening the rules to match the increasing prevalence of arbitration. However, in the author’s view, the persistence of these challenges highlights how international arbitration is ripe for reform.

III. CODE OF ETHICS

This article will now consider how variations in ethical standards can potentially disrupt arbitral proceedings due to the divergence of civil and common law approaches, the plurality of legal systems’ understandings of ethical standards, and overall variations in professional conduct.

Arguably all parties (not just counsel) involved in arbitration are subject to varying ethical obligations. These obligations extend to arbitrators, witnesses, administrative personnel, and other third parties, though the ethical responsibilities of arbitrators have been more widely discussed.⁷² For example, while arbitrators must adhere to strict ethical guidelines concerning impartiality and conflicts of interest, witnesses may not be bound by the same ethical constraints. Catherine Rogers describes this problem as driving international arbitration into an ethical ‘no man’s land’,⁷³ suggesting that there is a perceived lack of well-defined ethical

⁶⁸ Navin G Ahuja, *Taming the Guerrilla in International Commercial Arbitration: Levelling the Playing Field* (Springer 2022) 56.

⁶⁹ Jeff Waincymer, ‘Reconciling Conflicting Rights in International Arbitration: The Right to Choice of Counsel and the Right to an Independent and Impartial Tribunal’ (2010) 26 *Arbitration International* 597, 610.

⁷⁰ Ashish Sharma, ‘Guerrilla Tactics in Arbitration vis-à-vis International Arbitration’ (2021) 4 *International Journal of Law Management & Humanities* 670, 676.

⁷¹ Friedland and Brekoulakis (n 42) 34.

⁷² See for example IBA Rules of Ethics for International Arbitrators (1987); IBA, *IBA Guidelines on Conflicts of Interest in International Arbitration* (10 October 2014); CIArb, ‘The Chartered Institute of Arbitrators Code of Professional and Ethical Conduct for Members’ (October 2009) <<https://www.ciarb.org/media/2x3a0n42/5-ciarb-code-of-professional-and-ethical-conduct-for-members.pdf>> accessed 24 March 2025.

⁷³ Catherine A Rogers, ‘Fit and Function in Legal Ethics: Developing a Code of Conduct for International Arbitration’ (2002) 23 *Michigan Journal of International Law* 341, 341; Catherine A Rogers, ‘The Ethics of Advocacy in International

standards or consensus on ethical principles within the field. Some would go so far as to say that there has been a ‘race to the bottom’,⁷⁴ where ethical standards have eroded over time as different stakeholders seek to establish the lowest common denominator of ethical acceptability. This does not suggest that the importance of ethical standards has diminished; rather, in pursuit of broad international acceptance, arbitration may have compromised on higher standards to create a framework that, while broadly accepted, is ultimately less concerned with maintaining the highest ethical standards.

Essentially, there has been a dilution of ethical standards to accommodate a wider range of perspectives. This makes sense when, in its infancy, international arbitration focused on prioritising efficiency, promoted reaching compromises for parties, and was guided by informal procedures.⁷⁵ As a result, the lack of explicit ethical guidance was not a significant concern and was, in fact, viewed by some as a benefit to distinguish international arbitration from more formal litigation.⁷⁶ But in the present day, as international arbitration has matured, the problem of unequal ethical standards has only grown in importance and can result in some unfair results.⁷⁷ If both parties, including counsel, belong to the same jurisdiction, then that is ideal. Both will know the ethical standards that they are expected to abide by. The problem arises when international arbitration is used to resolve cross-border disputes, where no proper regulation exists for the ethics of counsel or mechanisms for resolving ethical disputes. Aside from the International Bar Association (‘IBA’) Guidelines and the IBA Principles on International Arbitration,⁷⁸ there is no comprehensive framework to address these issues. And while the IBA Guidelines do provide a framework for ethical considerations, they do not intend to supersede mandatory laws or other professional rules.⁷⁹ As such, in the author’s view, they are unsatisfactory for two key reasons.⁸⁰

First, the ‘soft law’ status of the IBA Guidelines gives rise to what Rogers refers to as the issue of ‘double deontology’.⁸¹ This issue arises when counsel are subject to conflicting ethical codes, which can create confusion about which standards should prevail.⁸² Granted, arbitrators are supposed to be knowledgeable, experienced, and have integrity,⁸³ but the law often relies too much on these broad characteristics and essentially advocates that counsel

Arbitration’ (2010) Penn State Legal Studies Research Paper 18/2010, 1 <<https://ssrn.com/abstract=1559012>> accessed 4 January 2024.

⁷⁴ Catherine Rogers, ‘Guerrilla Tactics and Ethical Regulation’ in Günther J Horvath and Stephan Wilske (eds), *Guerrilla Tactics in International Arbitration* (Kluwer Law International 2013) 314.

⁷⁵ Florian Grisel, ‘Arbitration as a Dispute Resolution Process: Historical Developments’ in Stefan Kröll, Andrea K Bjorklund and Franco Ferrari (eds), *Cambridge Compendium of International Commercial and Investment Arbitration*, vol 1 (CUP 2023) 9.

⁷⁶ Rogers, ‘The Ethics of Advocacy in International Arbitration’ (n 73) 5.

⁷⁷ Margie Jaime, ‘Counsel Conduct in International Arbitration: An Ethical “No-Man’s Land”?’ (2019) 85 *Arbitration* 211, 221.

⁷⁸ See for example IBA, *IBA Guidelines on Conflicts of Interest* (n 72); IBA, *IBA Guidelines on Conflicts of Interest in International Arbitration* (25 May 2024); IBA, *IBA Guidelines on Party Representation in International Arbitration* (25 May 2013); IBA Rules of Ethics for International Arbitrators (n 72); IBA, ‘IBA International Principles on Conduct for the Legal Profession’ (28 May 2011) <https://www.icj.org/wp-content/uploads/2014/10/IBA_International_Principles_on_Conduct_for_the_legal_prof.pdf> accessed 24 March 2025.

⁷⁹ See for example Jaime (n 77) 213; Wachter (n 8) 1147.

⁸⁰ See for example Doak Bishop, ‘Ethics in International Arbitration’ (ICCA Congress at the International Arbitration Conference, Rio de Janeiro, May 2010) 3; Jaime (n 77) 228.

⁸¹ Rogers, *Ethics in International Arbitration* (n 23).

⁸² See for example Matthew T Nagel, ‘Double Deontology and the CCBE: Harmonizing the Double Trouble in Europe’ (2007) 6 *Washington University Global Studies Law Review* 455, 456; Nathan M Crystal and Francesca Giannoni-Crystal, ‘“One, No One and One Hundred Thousand”...Which Ethical Rule to Apply? Conflict of Ethical Rules in International Arbitration’ (2013) 32 *Mississippi College Law Review* 283, 291.

⁸³ Park, ‘Arbitration’s Protean Nature’ (n 58) 334.

should intuitively know how to handle themselves in arbitral proceedings. The double deontology issue stems from the tension between national ethical standards, which are deeply ingrained in counsel's practices, and the IBA Guidelines. Rather than providing clear guidance, this conflict can leave counsel uncertain about which standards to prioritise, potentially resulting in contradictory behaviour during proceedings. Moreover, international arbitration inherently involves disputes between parties from different legal traditions, meaning that lawyers are trained under diverse ethical frameworks.⁸⁴ This further reflects the current 'polaris[ed]' landscape, where ethical standards are not only inconsistent *across* jurisdictions but also *within* them.⁸⁵ However, in the author's view, the law has not sufficiently addressed this reality, leaving a gap in how ethical conflicts should be managed in practice. What this means is that counsel are unlikely to know or understand every applicable ethical code that exists in different jurisdictions to any detailed extent.⁸⁶

For example, in the International Centre for Settlement of Investment Disputes ('ICSID') cases of *Hrvatska Elektroprivreda v Republic of Slovenia*⁸⁷ and *The Rompetrol Group NV v Romania*,⁸⁸ the ICSID Arbitration Rules failed to offer any clear unitary ethical standard. In *Hrvatska*, the tribunal had to determine whether newly introduced counsel should be permitted to participate in the proceedings, given potential conflicts of interest. The tribunal excluded the counsel based on fairness considerations rather than a universal ethical standard, highlighting the absence of a consistent ethical framework within the ICSID Arbitration Rules.⁸⁹ Similarly, in *The Rompetrol Group*, the tribunal had to address whether counsel's professional conduct met the appropriate ethical threshold, further illustrating that, in the absence of a unified ethical code, tribunals are left to interpret ethical obligations on an ad hoc basis.⁹⁰ This contributes to greater uncertainty for counsel navigating international arbitration. In the author's view, these cases demonstrate that the ICSID Arbitration Rules fail to impose a unitary ethical standard on counsel and that arbitrators struggle consistently to identify the ethical obligations that are potentially imposed on counsel in international arbitration when grappling with cross border disputes and standards of conduct. In essence, counsel's understanding of these varying ethical codes depends on their experience in handling multiple cases.

This article argues that such an approach amounts to an apprenticeship model, where counsel are left to figure out the applicable ethical rules on the job. Ultimately, this is an unwelcome reality as it increases the likelihood of inconsistent application of ethical rules.⁹¹ This model raises significant concerns regarding power imbalances. It inherently favours established practitioners while placing newcomers at a disadvantage as they must navigate the implicit and 'unwritten' norms of the profession. This dynamic triggers several questions: if counsel can only learn ethical standards through experience, how can they ethically handle their first cases when they lack that experience? Furthermore, how can parties from

⁸⁴ See for example Wachter (n 8) 1144; Cyrus Benson, 'Can Professional Ethics Wait? The Need for Transparency in International Arbitration' (2009) 3 *Dispute Resolution International* 78, 83.

⁸⁵ Gaillard 'Sociology of International Arbitration' (n 6) 13-14.

⁸⁶ John M Townsend, 'Clash and Convergence on Ethical Issues in International Arbitration' (2004) 36 *University of Miami Inter-American Law Review* 1, 18.

⁸⁷ *Hrvatska Elektroprivreda v Republic of Slovenia*, ICSID Case No ARB/05/24, Order Concerning the Participation of Counsel (6 May 2008).

⁸⁸ *The Rompetrol Group NV v Romania*, ICSID Case No ARB/06/3, Decision of the Tribunal on the Participation of a Counsel (14 January 2010).

⁸⁹ See *Hrvatska* (n 87) [34].

⁹⁰ *The Rompetrol Group* (n 88) [19].

⁹¹ Gaillard, 'Sociology of International Arbitration' (n 6) 13-14.

developing nations, where arbitration communities may be less established, achieve equal access to justice if their counsel lack the experiential knowledge that comes with repeated engagement in the field? Given that counsel are largely left to discern these norms themselves, it would not be disingenuous to suggest that the prevailing approach is to play the game and find out.

Secondly, this problem is only exacerbated when we compare the differing approaches of lawyers from common law and civil law countries, which feeds into the ‘inequality of arms’⁹² problem. The principle of ‘equality of arms’ is a fundamental aspect of procedural fairness in international arbitration and broader international law, ensuring that no party is placed at a substantial disadvantage when presenting its case.⁹³ But when disparities do arise, such as differing ethical standards for counsel, this ‘equality of arms’ transforms into an ‘inequality of arms’, undermining the procedural balance between the parties. One of the most pressing examples of this imbalance arises in the context of witness preparation, as explained by Tony Cole and Pietro Ortolani.⁹⁴ In England and Wales, counsel can explain to witnesses how the procedure will work and even train them on cross-examination.⁹⁵ In contrast, counsel operating in Austria are encouraged to limit their engagement with the witnesses so as not to influence them inappropriately.⁹⁶ On the other end of the spectrum, in the USA, ‘witness coaching’ is the norm and this includes counsel assisting the witness with their oral articulation of their testimony by encouraging them to do mock cross-examinations.⁹⁷

Monique Sasson notes that these unequal ethical obligations may incentivise clients to seek legal representation in locations with less stringent (or at least more favourable) standards.⁹⁸ This is problematic because it leads to something similar to ‘forum shopping’ where clients cherry-pick counsel whose ethical obligations are perhaps similar to their own moral compass. Unfortunately, this can be damaging to the integrity of the international arbitral process because, if one party is enjoying an ethical advantage due to their counsel operating under more favourable rules, it creates an uneven playing field. For example, in cases involving heavy reliance on witness testimony, parties may find counsel from jurisdictions where ethical rules are more favourable, such as the USA, attractive.⁹⁹ The range of issues impacted by this disparity extends beyond witness preparation, encompassing pre-trial document production and ex parte communications. For instance, in *Fraport AG Frankfurt Airport Services Worldwide v The Republic of the Philippines*, counsel’s ex parte communications became a contentious issue because of differing ethical standards.¹⁰⁰ In the author’s view, the lack of uniform ethical standards enables parties to engage in tactical forum shopping without any repercussions. This

⁹² Rogers, *Ethics in International Arbitration* (n 23) 107–08.

⁹³ Convention for the Protection of Human Rights and Fundamental Freedoms art 3.

⁹⁴ Cole and Ortolani (n 27) 159.

⁹⁵ *ibid.*

⁹⁶ See for example Austrian Code of Professional Ethics, art 8; F von Schlabrendorff, *Interviewing and Preparing Witnesses for Testimony in International Arbitration Proceedings: The Quest for Developing Transnational Standards of Lawyers’ Conduct* (Kluwer Law International 2010) item 2.5.

⁹⁷ Roberta K Flowers, ‘Witness Preparation: Regulating the Profession’s “Dirty Little Secret”’ (2011) 38 *Hastings Constitutional Law Quarterly* 1007, 1009; Bernardo Cartoni, ‘Ethics and International Arbitration: Coaching a Witness under Different Perspectives’ (2023) 15 *Australasian Policing* 28, 29–31; Maia Kvirkashvili, ‘Witness Coaching by a Prosecutor’ (2016) 2016 *Journal of Law* 188, 189–91.

⁹⁸ Monique Sasson, ‘Ethics in International Arbitration’ (*Law360*, 5 February 2016) <<https://www.jamsadr.com/files/uploads/documents/articles/sasson-law360-ethics-in-international-arbitration-2016-02-05.pdf>> accessed 4 January 2024.

⁹⁹ Anna Magdalena Kubalczuk, ‘Evidentiary Rules in International Arbitration – A Comparative Analysis of Approaches and the Need for Regulation’ (2015) 3 *Groningen Journal of International Law* 85, 107.

¹⁰⁰ *Fraport AG Frankfurt Airport Services Worldwide v The Republic of the Philippines*, ICSID Case No ARB/03/25, Decision on Application for Disqualification of Counsel (18 September 2008).

is further exposed through the fact that some counsel leverage the fact that ethical rules are ‘not written down’¹⁰¹ because it enables strategic manoeuvring. This behaviour arguably reveals a tendency among counsel to challenge the boundaries of professional conduct. In the author’s view, their perspective appears to be that, in the absence of explicit guidelines on appropriate behaviour, they are free to act as they deem suitable. But unfortunately, this is the expected result when there is no clarity on professionalism, especially for those non-lawyers acting as legal representatives. This reminds us of how important language construction is: for some, the silence or absence of certain practices equates to acceptability; for others, it signifies prohibition. Again, this illustrates how the law has arguably become complicit in these strategies and highlights the urgent need for clearer ethical frameworks within international arbitration practice.

IV. REFORMS

Building on the previous discussion about the lack of common understanding of what constitutes due process and unequal ethical obligations imposed on counsel, this article now turns to potential reforms that could be implemented to eliminate those issues and allow for arbitration to prosper.

First, before any substantive reform analysis takes place, this article addresses the question posed earlier: where do we draw the line between legitimate concerns and obstruction? Following Reed’s suggestions,¹⁰² arbitrators need not be afraid: most of the time, these sorts of vexatious claims that use due process as an obstruction tool are likely to fail on their merits anyway. The message for arbitrators, then, is to maintain confidence in the process. But they should not indulge in everything that the party asks from them just because they claim that there have been breaches of due process. Of course, if parties genuinely believe that procedural rights are being violated, then it makes sense that they have the right to raise such objections, even if it prolongs the process. Another less dramatic message is for arbitrators to set out clear expectations right at the start of the proceedings.¹⁰³ In the author’s view, Klaus Peter Berger and J Ole Jensen are right to encourage this because it is not too radical a departure from existing processes and will likely be a cost-effective way for arbitrators to lay out their expectations at an early stage.¹⁰⁴ Given these considerations, what potential reforms could help to improve due process and support arbitrators in their role?

A. DUE PROCESS TERMINOLOGY

One possibility is that the terminology of due process could be harmonised. We have already established that the applicable rules include various modes of articulation (e.g. from ‘full’¹⁰⁵ to ‘reasonable’¹⁰⁶ and ‘fair’¹⁰⁷), where each conveys ever so slightly different expectations for how parties present their cases. Perhaps standardising the terminology to an *adequate opportunity* would allow the rules to become clearer and enable arbitrators to have a uniform

¹⁰¹ Malcolm Langford, Daniel Behn and Runar Hilleren Lie, ‘The Revolving Door in International Investment Arbitration’ (2017) 20 *Journal of International Economic Law* 301, 320.

¹⁰² Reed (n 9) 361.

¹⁰³ Berger and Jensen (n 40) 430.

¹⁰⁴ *ibid.*

¹⁰⁵ UNCITRAL Model Law (n 11) art 18.

¹⁰⁶ Arbitration Act 1996, s 33(1)(a); ICC Rules (n 14).

¹⁰⁷ ICDR Arbitration Rules (n 32).

understanding of the expectations regarding due process. In a way, it involves merging the current phrases used to describe how parties can present their cases. However, unlike ‘reasonable’ or ‘fair’, this new terminology carries less jurisprudential weight and is less tied to specific legal traditions, making it more adaptable across diverse systems. We might see that institutions, especially those that are due for rule revisions, will be more willing to adopt the new terminology as it does not favour any one institution’s existing language. To this end, language neutrality plays an important role here.

On the surface, not only would this eliminate the possibility that an institution would perceive that they are abandoning their approach in order to adopt a competitor’s, but it would avoid implying that any one institution’s previous approach is superior. At the same time, this fresh alternative would create a level playing field, allowing all institutions to embrace the new standard without damaging their reputations. At a deeper level, this language would also mitigate bias in arbitrators’ decisions by ensuring that no single system’s expectations take precedence. In other words, this would reduce the likelihood of arbitrators unconsciously favouring familiar standards from their own jurisdictions. Instead, they would approach procedural decisions with a more balanced perspective, recognising the international nature of the proceedings.

However, if we proceed on the basis that institutions are to revise their due process language, would this be so straightforward? Emmanuel Gaillard contends that there would likely be resistance towards terminology uniformity because some institutions want to uphold their distinctive approaches for specific types of disputes.¹⁰⁸ Yet, if we frame this resistance as only an initial hurdle, then what we discover is that terminological uniformity becomes more achievable over time and need not eliminate institutional distinctiveness in practice. And amongst those leading institutions, it seems only that ICSID would be particularly resistant to the proposals given how its primary objective is to maintain state confidence in the system,¹⁰⁹ which encourages a more cautious and incremental approach to procedural reforms. In contrast, institutions like the Singapore International Arbitration Centre (‘SIAC’), Stockholm Chamber of Commerce (‘SCC’), ICC, and potentially the London Court of International Arbitration (‘LCIA’), are more likely to embrace such changes given their openness to procedural innovation. For instance, the SIAC has a strong motivation to adopt the proposed reform because it aligns with Singapore’s strategy of distinguishing itself through progressive dispute resolution.¹¹⁰

The response to Gaillard, then, is that, while initial resistance is expected, broader adoption hinges on proving the success of revised terminology within these more adaptable institutions. Take, for example, the LCIA—it has demonstrated a willingness to modernise language in its rules, particularly in areas like remote hearings and electronic submissions.¹¹¹ This openness suggests that institutions like the LCIA will likely be receptive to revising their due process language. While some institutions may wish to preserve their unique linguistic choices, the benefits of greater clarity and consistency outweigh the challenges of terminological diversity. And because the right to present one’s case is a ‘paramount’¹¹² component of due process, it follows that we must take every effort to ensure that this is interpreted

¹⁰⁸ Emmanuel Gaillard, ‘2018 LALIVE LECTURE: The Myth of Harmony in International Arbitration’ (2019) 34 *ICSID Review* 553, 564–66.

¹⁰⁹ ‘About ICSID’ (*ICSID*) <<https://icsid.worldbank.org/About/ICSID/>> accessed 20 February 2024.

¹¹⁰ Muhammad Farhan Talib and others, ‘Harmonizing Conflict: Exploring Global Applications of Alternative Dispute Resolution Methods’ (2024) 4 *Pakistan Journal of Criminal Justice* 77, 82.

¹¹¹ LCIA Rules (n 14) arts 19.2, 4.1.

¹¹² Thirgood and Williams (n 12).

consistently across international arbitration. Rather than viewing this change as a threat, institutions need to appreciate that clear and consistent language makes arbitration more accessible to parties from different legal traditions.¹¹³

Any reform should include detailed guidance on interpreting this standard because what constitutes *adequate* will invariably involve judgments by the arbitral panel. Won Kidane stresses that substantial differences between Western and Eastern legal traditions necessitate a culturally sensitive definition of adequate opportunity.¹¹⁴ Practically, this means that the definition should offer flexible guidance that accommodates both common law and civil law traditions. So, when making procedural decisions, this article argues that arbitrators should consider specific contextual factors, such as party equality, party expectations, and case complexity. And to illustrate how tribunals might apply just one of these factors, allow us to consider party equality through *Yukos Universal Limited v Russian Federation*.¹¹⁵ In this case, the respondent controlled most key documents related to tax assessments, enforcement, and Yukos Universal's asset auction, giving them a significant advantage.¹¹⁶ However, the claimant, having lost access to the corporate records during Yukos Universal's forced bankruptcy, struggled to present its case. The adequate opportunity standard would require the tribunal explicitly to consider how this imbalance would affect what constituted adequate procedural opportunity. Instead of treating both parties identically, the tribunal might allow the claimant broader document production requests than would normally be allowed under a restrictive interpretation of relevance and materiality. Alternatively, the tribunal might grant the claimant extra time to analyse and respond to documentary evidence solely controlled by the respondent.¹¹⁷ By explicitly addressing party equality, the tribunal would demonstrate that its understanding of adequate opportunity accounted for the fundamental power and information imbalances inherent in the dispute. This contextual framework acknowledges cultural differences in procedural expectations, allowing parties to anticipate how adequacy determinations will be made. Such assessment responds to Park's concern about preventing due process paranoia and Blackaby and others' concern about avoiding vague standards,¹¹⁸ both of which were discussed earlier. In addition, this would help prevent due process being used or misused as a sword and would ultimately improve predictability, which is one of arbitration's greatest strengths.

If we do not seek to reform the rules, divergent interpretations of due process will likely continue. For example, consider a situation where a party is seeking to submit evidence past the relevant deadline.¹¹⁹ When the UNCITRAL Model Law applies, an arbitrator may be more inclined to allow the late evidence submission to ensure that the presenting party has a 'full'¹²⁰ chance to present its case. And this may attract delaying tactics. Yet, if English law or the ICDR Rules applied, the arbitrators would probably weigh the circumstances surrounding the late evidence submission so as to reflect whether this would be 'reasonable'.¹²¹ What we can see from these two conflicting practices is that the terminology is problematic and is

¹¹³ Mary Howard, 'International Arbitration and Cross-Cultural Issues' (2020) 86 *Arbitration* 331, 339–40.

¹¹⁴ Kidane (n 17) 3.

¹¹⁵ *Yukos Universal Limited v Russian Federation* (PCA, Case No 2005-04/AA227).

¹¹⁶ *Ibid.*

¹¹⁷ Ugo Draetta, 'Cooperation Among Arbitrators in International Arbitration' (2016) 5 *Indian Journal of Arbitration Law* 107, 123.

¹¹⁸ Park, 'Arbitration's Protean Nature' (n 58); Blackaby and others (n 34).

¹¹⁹ Reed (n 9) 375.

¹²⁰ UNCITRAL Model Law (n 11) art 18.

¹²¹ Arbitration Act 1996, s 33(1); ICC Rules (n 14).

contributing to the fragmented understanding of due process. An amendment would give parties entering arbitration a more predictable understanding of the procedural requirements related to presenting their case where the likelihood of divergent interpretations is removed.

B. AN INTERNATIONAL CODE OF ETHICS

With regard to altering the ethical landscape, there has long stood a debate as to how this could be achieved. A consensus has now emerged that there needs to be an appetite to deal with the ambiguous nature of ethical conduct in the context of international arbitration.¹²² One of the more recent proposals addressing ethical standards is Cyrus Benson's 'Checklist of Ethical Standards'.¹²³ His proposal entails a thorough examination of nine different types of improper conduct with the goal of assisting the tribunal in evaluating how ethically counsel have behaved and to determine appropriate sanctions based on the nature of any violation and the particular circumstances of the case. However, in the author's view, a primary issue with this proposal is that, as Robert Wachter recognises, this approach exacerbates the ambiguity problem where counsel may selectively apply ethical rules in a way that undermines the coherence and consistency of ethical standards in international arbitration. This selective application risks obscuring a unified approach to ethical integrity in the sense that it soon becomes less predictable and more fragmented.¹²⁴ Thus, simply implementing this reform would not be sufficient to resolve the existing challenges: it would merely reintroduce the same foundational dilemmas. In the author's view, a more comprehensive approach is required.

A better immediate option is to establish an 'International Code of Ethics'. This was first proposed by W Michael Reisman in 1971,¹²⁵ who recognised the need to construct traditional ethical rules that governed counsel conduct in international arbitration. Paulsson extended this in 1992 by suggesting that the community would benefit from a Code that would address the conflicting ethical obligations that are imposed on counsel, particularly regarding *ex parte* communications. Detlev Vagts further refined this suggestion in 1996,¹²⁶ advocating for a single Code that catered to divergent ethical approaches across civil and common law systems. In essence, these scholars envisioned a 'unified' set of ethical standards that would be reflected in a single Code, eliminating the confusion stemming from overlapping (and sometimes contradictory) national codes governing professional conduct and applying to counsel regardless of their base jurisdiction. Edna Sussman and Solomon Ebere assert that this could address the 'double deontology' problem directly.¹²⁷ And this would, arguably, be correct: by having a unified ethical code, a level playing field could be created whereby counsel would no longer be confused as to what ethical standards ought to be applied. International arbitration would no longer be plagued by the issue of potentially abusive tactics being used by participants.

¹²² See for example Jane Wessel and Gordon McAllister, 'Towards a Workable Approach to Ethical Regulation in International Arbitration' (2015) 10 *Canadian International Lawyer* 5, 8; Peter Halprin and Stephen Wah, 'Ethics in International Arbitration' (2018) 87 *Journal of Dispute Resolution* 87.

¹²³ Benson (n 84) 81.

¹²⁴ Wachter (n 8) 1158.

¹²⁵ W Michael Reisman, *Nullity and Revision: The Review and Enforcement of International Judgments and Awards* (Yale University Press 1971) 116-17.

¹²⁶ Detlev F Vagts, 'The International Legal Profession: A Need for More Governance?' (1996) 90 *American Journal of International Law* 250, 250. See generally Jan Paulsson, 'Standards of Conduct for Counsel in International Arbitration' (1992) 3 *American Review of International Arbitration* 214.

¹²⁷ Sussman and Ebere (n 8) 622.

Further, in addition to creating a level playing field, Doak Bishop contends that the introduction of an international code will reduce the ambiguous nature of the rules and offer transparency.¹²⁸ He argues that this is because a unified code would bring certainty over which ethical rules would apply and allows all parties to understand what conduct will be regarded as acceptable.¹²⁹ This reform is not without its critics, though. For example, although this appears attractive on the surface, Rogers argues that the creation of a unitary framework may risk constraining the diversity of legal traditions that is currently represented in the variation in arbitral rules.¹³⁰ The major problem with this interpretation is that variation is, arguably, international arbitration's very issue: not having a unified code of ethics creates situations where ambiguity prevails. What this means is that conflicting ethical obligations can place counsel in impossible positions where complying with one jurisdiction's rules requires violating another's. The impression from Rogers is that conduct should just be left to tribunals to enforce and regulate under potentially divergent standards. Again, this cannot be left to tribunals to enforce for this risks idiosyncratic ideals becoming implanted in the solution.¹³¹ Others, such as Park, have raised the objection that counsel will need to learn something new and not rely on existing frameworks.¹³² But we cannot halt reform just because there is an extra layer of learning. In the author's view, legal counsel are very capable of adapting and learning from a new international code. And if this is going to help level the playing field, it is not an entirely convincing argument to categorise extra learning as a 'risk'.¹³³ In order for international arbitration to continue prospering, an International Code of Ethics is, in the author's view, an attractive and necessary development. This is also the view held by an increasing body of scholars.¹³⁴

Before we proceed to consider what this Code might look like and how it is to be effectuated, it is important to address the obvious objection to this proposal. Opponents are sceptical of one aspect in particular: that is, surely a unified Code contradicts the inherent flexibility of arbitration? Gabrielle Kaufmann-Kohler argues that the success of arbitration is its adaptable quality and that imposing ethical standards that are rigid risks undermining this advantage.¹³⁵ Jane Wessel and Gordon McAllister develop this argument further by arguing that imposing strict ethical rules could make arbitration resemble litigation more closely, where flexibility is eroded.¹³⁶ These concerns are entirely justified, especially the former part of Kaufmann-Kohler's argument commending arbitration for its adaptability. This is not being disputed here; indeed, this adaptability allows parties to structure their proceedings in a way that best suits them. But Kaufmann-Kohler's latter suggestion, along with Wessel and McAllister's argument that rigid ethical standards risk undermining flexibility, should be approached with caution. There are two primary reasons why.

First, the focus of this objection is misplaced as it is preoccupied with the risk of procedural rigidity rather than recognising the greater danger posed by ethical inconsistency. In the author's view, implementing ethical standards does not necessarily compromise the

¹²⁸ Bishop (n 80) 11.

¹²⁹ *ibid.*

¹³⁰ Rogers, 'The Ethics of Advocacy in International Arbitration' (n 73) 10.

¹³¹ Bishop (n 80) 9.

¹³² Park, 'Arbitration's Protean Nature' (n 58) 354.

¹³³ *ibid.*

¹³⁴ See for example Bishop (n 80) 10–12; Rogers, 'The Ethics of Advocacy in International Arbitration' (n 73) 13–14; Crystal and Giannoni-Crystal (n 82) 291; Moses, 'Ethics in International Arbitration' (n 8) 73, 80.

¹³⁵ Gabrielle Kaufmann-Kohler, 'Soft Law in International Arbitration: Codification and Normativity' (2010) 1 *Journal of International Dispute Settlement* 283, 298.

¹³⁶ Wessel and McAllister (n 122) 11.

flexibility that makes arbitration attractive. If the rules are carefully designed (which will be explored later), then they can shape behaviour by promoting openness and honesty without dictating how every procedural step must be followed.¹³⁷ Therefore, such a proposal would not make arbitration as ‘judicial[s]ed’¹³⁸ as one may initially believe. Secondly, ethical certainty does not equate to procedural inflexibility. There is a notable distinction here: while ethical standards govern the conduct of counsel, procedural flexibility governs the process.¹³⁹ In fact, the absence of ethical certainty may actually generate more procedural rigidity because, when arbitrators and counsel face uncertainty about ethical boundaries, they may be especially careful, imposing overly rigid or defensive procedural measures to avoid potential challenges. Therefore, ethical guidelines can coexist with flexibility rather than undermine this advantage. We will see how this can work when we explore witness preparation and ex parte communications later on in the article.

We now return to consider how such code could be effectuated. In the author’s view, the Code should initially take the form of soft law. We do already have the IBA Rules of Ethics for International Arbitrators,¹⁴⁰ the *IBA Guidelines on Conflicts of Interest in International Arbitration*,¹⁴¹ and the Chartered Institution of Arbitrators (‘CIArb’) Code of Professional and Ethical Conduct for Members. However, these soft law instruments primarily focus on the ethical conduct of arbitrators and are therefore insufficient mechanisms for counsel.¹⁴² This is particularly concerning when considering ad hoc arbitration because there is no institution involved to oversee the process.¹⁴³ There is the IBA International Code of Ethics for Lawyers, which lays down general ethical principles, but this is unfortunately not specific to international arbitration and was last revised in 1988. To this end, the Code fails to appreciate the challenges that counsel experience in international arbitration because its ambit is on lawyers more generally. The only instrument that does cover counsel ethical conduct is the IBA Guidelines on Party Representation in International Arbitration, but it is not entirely comprehensive as there are still particularities of counsel conduct that are not precisely defined.¹⁴⁴ For example, while the guidelines rightly discourage strategic tactics aimed at causing unnecessary delay, the definition of unnecessary is not defined. They simply do not provide a precise list of prohibited behaviours, which would be useful for counsel. This approach may appear controversial given its overly prescriptive nature for a mode of dispute resolution that is inherently flexible. However, a more prescriptive ethical regime (in the form of guidelines) may be precisely what international arbitration requires. This is because the unification of ethical standards for counsel does not compromise the broader flexibilities that arbitration offers, including the various freedoms available to parties when constructing their arbitration.¹⁴⁵ This

¹³⁷ Moses, *The Principles and Practice* (n 11) 143.

¹³⁸ Thomas J Stipanowich, ‘Arbitration: The “New Litigation”’ (2010) 1 *University of Illinois Law Review* 1, 8.

¹³⁹ See Catherine A Rogers, ‘The Ethics of Advocacy in International Arbitration’ in Doak Bishop and Edward G Kehoe (eds), *The Art of Advocacy in International Arbitration* (2nd edn, Juris 2010) 49.

¹⁴⁰ IBA Rules of Ethics for International Arbitrators (n 72).

¹⁴¹ IBA, *IBA Guidelines on Conflicts of Interest* (2014) (n 72).

¹⁴² Lord Hacking and Sophia Berry, ‘Ethics in Arbitration: Party and Arbitral Misconduct’ in Julio César Betancourt (ed), *Defining Issues in International Arbitration: Celebrating 100 Years of the Chartered Institute of Arbitrators* (OUP 2016) 139.

¹⁴³ Jaime (n 77) 2.

¹⁴⁴ See Catherine A Rogers, ‘Context and Institutional Structure in Attorney Regulation: Constructing an Enforcement Regime for International Arbitration’ (2003) 39 *Stanford Journal of International Law* 1, 56–57.

¹⁴⁵ Leon Trakman, ‘Legal Traditions and International Commercial Arbitration’ (2007) 17 *American Review of International Arbitration* 1, 29.

further reinforces the necessity of establishing a new International Code of Ethics for international arbitration.

What might this Code look like? Laying out every rule and expectation from this Code would be a mammoth task, and so, rather than becoming a drafting exercise, this article engages only with the most vexing ethical disparities at the moment: witness preparation and ex parte communications. In the context of witness preparation, Jeffrey Waincymer argues that simply labelling witness preparation as either entirely permissible or entirely prohibited is too simplistic.¹⁴⁶ An approach like this risks unfairly restricting legitimate forms of preparation. He therefore suggests that the Code should identify a spectrum of acceptable and unacceptable practices. This is to be encouraged. The Code should clearly define the boundaries between permissible and impermissible conduct in witness preparation. Permissible conduct may include practices, such as presenting contemporaneous documents to refresh a witness's memory or assisting them in understanding the sequence of events—actions that improve clarity without compromising the accuracy of their testimony.¹⁴⁷ Conversely, impermissible conduct should be defined as behaviour such as suggesting specific answers, dictating testimony verbatim, or pressuring the witness to present a misleading narrative.¹⁴⁸ What this means, then, is that the terminology of 'suggesting' and 'pressuring' must be clearly defined so that counsel from all traditions can understand,¹⁴⁹ or else we risk replicating the issues in the IBA Rules of Evidence, which fail to offer a clear distinction between a rehearsal and an interview.¹⁵⁰ To supplement these terminological definitions, Margaret Moses suggests that counsel should be required to document all witness preparation sessions, with these records being made available to the opposing party.¹⁵¹ In the author's view, this would serve as a safeguard, enabling the opposing party to scrutinise the preparation process and ensure that no unethical coaching takes place. But this is a really difficult situation because we may have counsel going above and beyond to find instances of coaching or unethical preparation. In essence, the same problem of strategic play haunts this suggestion. But to remedy this concern, this article advocates for a dual approach: while parties should be encouraged to adhere to the proposed guidelines (especially where consensus on ethical standards may be unattainable for parties), they should also strive to reach a mutual understanding of what constitutes 'suggesting' or 'pressuring' within the context of witness preparation. Ultimately, parties should be able to design the rules of the game that they are playing.

In terms of ex parte communications, whether this practice should be prohibited or accepted reveals yet another disparity regarding ethics because '[o]ne practitioner's guerilla tactic may well be another's legitimate practice'.¹⁵² As Kidane persuasively argues, cultural differences in international arbitration extend far beyond the simplistic common law and civil law divide.¹⁵³ These deeper cultural variations fundamentally affect how parties and counsel perceive appropriate communications and relationships with decision-makers. Consider, for example, the practice in China: counsel are generally permitted to communicate with

¹⁴⁶ Jeffrey Waincymer, *Procedure and Evidence in International Arbitration* (Kluwer Law International 2012) 789–91.

¹⁴⁷ Mark A Cymrot, 'Cross-Examination in International Arbitration' (2007) 62 *Dispute Resolution Journal* 52, 61.

¹⁴⁸ Laurent Hirsch, Rupert Reece and Roxane Reiser, 'Expert Witnesses in International Arbitration' (2019) 3 *International Business Law Journal* 247, 255.

¹⁴⁹ Joshua Karton, *The Culture of International Arbitration and the Evolution of Contract Law* (OUP 2013) 215–18.

¹⁵⁰ Wachter (n 8) 1148.

¹⁵¹ Moses, 'Ethics in International Arbitration' (n 8) 77.

¹⁵² Hacking and Berry (n 142) 141.

¹⁵³ Kidane (n 17) 240.

arbitrators before proceedings; it is simply typical practice in Chinese arbitrations.¹⁵¹ It is not so typical for English arbitration, where *ex parte* communications is a practice that leans towards prohibition rather than acceptance.¹⁵² The most achievable and, indeed, preferred option for parties and counsel is to incorporate flexible options within the guidelines—much like that suggested for witness preparation. Then, during negotiations, both parties and counsel can reach an agreement on the applicable standards. Navin Ahuja suggests that the guidelines should be phrased as questions within this provision.¹⁵⁶ For example, would it be acceptable for any participant (not just counsel of a party to the arbitration) to engage in *ex parte* communications with an arbitrator?¹⁵⁷ In the author's view, Ahuja's suggestion is correct because it recognises that, in drafting the arbitral agreement, parties could explicitly permit or prohibit such conduct or outline specific circumstances under which it may be acceptable. Alternatively, they could delegate the decision to the tribunal. This approach is advantageous for two key reasons. First, it upholds party autonomy, which is the fundamental principle that parties ought to have the freedom to determine the rules governing their arbitration. Secondly, it creates a level ethical playing field for all parties by establishing a common standard that accommodates diverse legal and cultural traditions.

Once this Code has been established, what ought to happen next? This article suggests that, following implementation, the most effective avenue to solidify its existence is through an incremental approach. Initially, the Code should function as soft law, as proposed, before being integrated into institutional rules. The final, more ambitious push is for it to be hardened into a Convention, but this is, indeed, ambitious and unlikely to be accepted.¹⁵⁸ With respect to being incorporated into institutional rules, arbitral institutions such as the LCIA, ICSID, ICC, and SIAC should eventually incorporate identical ethical provisions into their frameworks.¹⁵⁹ Park argues that, just as arbitrators create a distinct reputation for themselves, this also extends to arbitral institutions.¹⁶⁰ In the author's view, this would create an incentive for compliance with these ethical standards: arbitration institutions with strong ethical guidelines may be seen as more reliable and fairer. In turn, this can attract parties to choose those institutions over others that might not have the same meticulous ethical underpinnings. This is a logical contention which positions ethical standards not as burdensome rules, but rather as a way to improve the reputation and appeal of arbitration institutions. The LCIA, for example, has integrated ethical rules (the LCIA 'General Guidance for the Parties' Legal Representatives') into its framework.¹⁶¹ Though there are still weak spots in these guidelines, it represents an enormous step in the right direction, and this ultimately indicates to other institutions the market's acceptance of formalised ethical standards.

The final pressing question that must be answered is one of enforcement. How should the Code be enforced? Should we simply allow national courts to oversee and regulate the conduct of counsel in arbitral proceedings?¹⁶² Should tribunals themselves be entrusted with

¹⁵¹ See for example Jun Ge, 'Mediation, Arbitration and Litigation: Dispute Resolution in the People's Republic of China' (1996) 15 *UCLA Pacific Basin Law Journal* 122, 127; Rogers, 'Fit and Function in Legal Ethics' (n 73) 363.

¹⁵² See Arbitration Act 1996, s 73(1).

¹⁵³ Ahuja (n 68) 267.

¹⁵⁴ *Ibid.*

¹⁵⁵ Isaac J Buckland, 'A Comparative Approach to Consistent Ethical Standards in International Commercial Arbitration' (2019) 85 *Arbitration* 230, 231.

¹⁵⁶ See Bishop (n 80).

¹⁵⁷ William W Park, 'Rectitude in International Arbitration' (2011) 27 *Arbitration International* 473, 496.

¹⁵⁸ LCIA Rules (n 14) arts 18.5, 18.6.

¹⁵⁹ Catherine A Rogers, 'Lawyers Without Borders' (2009) 30 *University of Pennsylvania Journal of International Law* 1035, 1083.

finding solutions on their own?¹⁶³ Or could arbitrators determine the allocation of arbitration costs based on the behaviour of the lawyers involved in the proceedings?¹⁶⁴ We can eliminate the possibility of national regulation by considering the comments made in *Hrvatska*. The tribunal in *Hrvatska* found it problematic for individual national regulatory bodies to be responsible for setting rules for professionals, arguing that a more centralised system should be established.¹⁶⁵ The next two solutions, whilst superficially attractive, are fundamentally flawed. Not only does the option of tribunals finding their own solutions create an unpredictable standard, but counsel would have no notice about how they are expected to behave. Further, cost allocation based on counsel behaviour is equally problematic because, rather than establishing clear professional standards, this route acts as simple incentivisation which impacts clients opposed to counsel. In other words, because clients typically bear the costs, any penalties against their counsel will mean that it is the client who is financially punished. A more troubling consequence of the cost allocation solution is that if counsel are aware that particular behaviours could lead to cost penalties for their clients, they may temper their approach. This tempering results in counsel acting more conservatively so as to protect themselves from penalties. One can see the obvious issue here: if we have practitioners who are softening their approach or are reluctant to engage in justified legal strategies to avoid financial repercussions, then clients will receive a lower standard of representation. Essentially, these solutions are too spread out, too inconsistent, and too indirect.

To this end, given how international trade and foreign investment continue to expand, which increases the reliance on international arbitration,¹⁶⁶ Rogers argues that the only characters capable of such adaptations are institutions themselves because they operate independently.¹⁶⁷ This is a rational approach that distinguishes itself from the three solutions above because it recognises that arbitral institutions have the ability to take direct action against counsel who violate ethical rules; instead of punishing the clients, institutions can impose disciplinary actions directly on counsel responsible for the misconduct. Once ethical uniformity has been achieved, only then can we start discussions of an ethical Convention acting as hard law. This is because such reforms are evolutionary and will require patience.¹⁶⁸ Therefore, an International Code of Ethics should be the immediate starting point to eliminate conflicting ethical conduct.

V. CONCLUSION

Overall, the disjointed understanding of what constitutes due process, coupled with uneven ethical obligations imposed on counsel, is a serious challenge in international arbitration. There were three components to the argument. First, it was argued that a lack of common understanding of what constitutes due process can mean that parties use due process as a

¹⁶³ Mary C Daly, 'Resolving Ethical Conflicts in Multijurisdictional Practice—Is Model Rule 8.5 the Answer, an Answer, or No Answer at All?' (1995) 36 South Texas Law Review 715, 778.

¹⁶⁴ Bishop (n 80).

¹⁶⁵ *Hrvatska* (n 87) [23].

¹⁶⁶ David Salton, 'Recent Trends in International Arbitration and 2021 International Rule Changes' (2021) 17 Construction Law Journal 81, 81–82.

¹⁶⁷ See for example Rogers, 'Fit and Function in Legal Ethics' (n 73) 406; Rogers, 'Context and Institutional Structure in Attorney Regulation' (n 144).

¹⁶⁸ Carrie Menkel-Meadow, 'The Evolving Complexity of Dispute Resolution Ethics' (2017) 30 Georgetown Journal of Legal Ethics 389, 414.

‘sword’¹⁶⁹ by launching unreasonable objections at any stage of the proceedings. The problem generated is that it is extremely difficult to distinguish between legitimate concerns and obstruction, and this leads to delays and increased costs. Secondly, it was asserted that the ‘apprenticeship’ approach adopted by the law with regard to uneven ethical obligations is problematic because it creates confusion, and the ‘double deontology’¹⁷⁰ and ‘inequality of arms’¹⁷¹ between ethical standards place ethics into ‘no man’s land’.¹⁷² Thirdly, it was argued that adapting due process terminology to an ‘adequate opportunity’ offers a clearer, more neutral standard that accounts for diverse legal traditions while enhancing procedural predictability. It was suggested that a contextual approach should be taken in order to ensure fairness without inviting due process paranoia or strategic misuse. In terms of the current fragmented ethical landscape, it was argued that Benson’s ‘checklist’ approach¹⁷³ was far too radical and risks adding to an already ambiguous item. The preferable approach, in this article’s view, is to encourage an ‘International Code of Ethics’ for international arbitration which could harmonise and address the current uneven playing field. To reiterate, this proposal recognises cultural differences and encourages parties to discuss practices that they both deem to be acceptable during negotiations. Though balancing these guidelines with party autonomy was proven to be a difficult task, this can realistically be achieved through the design suggestions explored in this article. Without such reform, the challenges facing international arbitration may continue to undermine its effectiveness and credibility.

¹⁶⁹ Reed (n 9).

¹⁷⁰ Rogers, *Ethics in International Arbitration* (n 23).

¹⁷¹ *ibid* 108.

¹⁷² Rogers, ‘Fit and Function in Legal Ethics’ (n 73).

¹⁷³ Benson (n 84) 78.

Remedying Judicial Intervention into Private Contract: A Case for Abandoning the Creditor Duty Post-*BTI 2014 LLC v Sequana SA* [2022] UKSC 25

ALEXANDRA KOSTA-FOTI*

ABSTRACT

This article identifies a worrisome trend in corporate law: increased judicial intervention in private corporate contracts and the invocation of moralistic conceptions of fairness and effectiveness that depart from the fundamental aspect of bargain. This problematic approach is clear in Australian and English jurisprudence on the debt-equity conflict between shareholders and creditors—precisely, the unprincipled expansion of fiduciary duties to provide enhanced protection to creditors when a company goes insolvent. This article focuses on the theoretical and practical development of the creditor duty since it was introduced by Australian jurisprudence in *Walker v Wimborne*, with two keys aims: first, to reassert the primacy of contractual risk allocation; and secondly, to demonstrate that there exists no principled basis for a creditor duty in England and Wales (specifically focusing on the doctrinal infirmities in the UK Supreme Court case of *BTI 2014 LLC v Sequana SA* [2022] UKSC 25). This article identifies an essential normative conflict: what is worse, a lack of adequate compensation for the level of risk borne by the creditors as the new residual claimants or the imposition of a risk allocation by law that was never bargained for? The position adopted in this article is that fairness grounds should not meddle with debtor-creditor risk allocation—not at the expense of principle, coherence, or contract.

Keywords: creditor duty, fiduciary duties, company law, insolvency, private corporate contracts

I. INTRODUCTION

Corporations: what are they and what is their purpose? A positive ontology of the corporation as a mere private entity risks undermining the complex structural framework that is *facilitated*

* Alexandra Kosta-Foti has a First-Class BA (Hons) from King's College London (where she ranked top of her year) and a Distinction in the Graduate Diploma in Law (GDL) from City St George's, University of London, where she is currently a Visiting Lecturer in Contract Law. Alexandra is a Prince of Wales, Michael Mustill and Dame Juliet Wheldon Scholar at the Honourable Society of Gray's Inn and will be undertaking the LLM (Master of Laws) at Cambridge University in October 2025.

through its legal fiction.¹ This framework consists of a *nexus of contracts* between corporate participants, including creditors, investors, managers, employees, third parties, and others. As explicated by Michael C Jensen and William H Meckling, the corporation ‘serves as a focus for a complex process in which the conflicting objectives of individuals’ are brought into ‘equilibrium within a *framework of contractual relations*’.² Such contractual relations can be express or implied but, most importantly, they are bargains involving a clear allocation of both rewards but also *risks*.³ However, not all corporate participants are exposed to the same level of risks: an important distinction is to be made between debt and equity holders.

The former—creditors—enter into credit contracts with the corporation where they can bargain for specific rights and renegotiate the terms of their contractual undertakings.⁴ As acknowledged by Lord Reed in *BTI 2014 LLC v Sequana SA*,⁵ as a corollary of entering into a contractual relationship with a company, creditors ‘must be the guardians of their own interests’.⁶ That is an inherent consideration present in all credit arrangements—as bargain involves risk, sophisticated economic actors bargaining at arms’ length are presumed to possess the knowledge that they are indeed running their own risk when providing credit. Such risk can, and is, addressed through contractual negotiations and direct protections that exist within debtor-creditor contracts. For example, creditors can negotiate a higher interest rate to reflect a corresponding higher risk undertaken or insist on guarantees or securities for their debt.⁷ While there exist institutional differences in the bargaining power of different creditor groups, indirect protection is nonetheless provided through the fact that it is in the company’s interest to pay its company’s debt in order to carry on its business and to avoid reputational risk that can affect the company’s creditworthiness and access to future credit that will drive its viability.⁸ While creditors are entitled to bargained-for fixed payments, the latter—shareholders—primarily rely on implicit contracts that entitle them to whatever is left after the company has met its express obligations (debts, liabilities, and other claims, such as employee wages or loan

¹ See for example George A Mocsary, ‘Freedom of Corporate Purpose’ (2016) 2016 Brigham Young University Law Review 1319; Edward B Rock, ‘For Whom Is the Corporation Managed in 2020? The Debate over Corporate Purpose’ (2021) 76 *The Business Lawyer* 363; John Armour and Michael J Whincop, ‘The Proprietary Foundations of Corporate Law’ (2007) 27 *OJLS* 429; David G Yosifon, ‘The Law of Corporate Purpose’ (2013) 10 *Berkeley Business Law Journal* 181; Yong-Shik Lee, ‘Shareholder Primacy as an Untenable Corporate Norm’ (2023) 8 *Annals of Corporate Governance* 1; Ann M Lipton, ‘What We Talk about When We Talk about Shareholder Primacy’ (2019) 69 *Case Western Reserve Law Review* 863.

² Michael C Jensen and William H Meckling, ‘Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure’ (1976) 3 *Journal of Financial Economics* 305, 311 (emphasis added).

³ For the creditors, some of the available rewards would be high interest rates, covenants, and collateral to secure the loan. The corresponding risk materialises into credit risk (i.e. that the creditor may not be able to recover the full amount of the loan), interest rate risk if the agreed upon return is lower than market trends, and vice versa, and liquidity risk if the debtor defaults or delays in payments.

⁴ See for example LS Sealy, ‘Directors’ “Wider” Responsibilities – Problems Conceptual, Practical and Procedural’ (1987) 13 *Monash University Law Review* 164; Frederick Tung, ‘The New Death of Contract: Creeping Corporate Fiduciary Duties for Creditors’ (2008) 57 *Emory Law Journal* 809; J William Callison, ‘Why a Fiduciary Duty Shift to Creditors of Insolvent Business Entities Is Incorrect as a Matter of Theory and Practice’ (2006) 1 *Journal of Business and Technology Law* 431.

⁵ [2022] UKSC 25, [2024] AC 211.

⁶ *ibid* [27].

⁷ See for example Simone M Sepe, ‘Directors’ Duty to Creditors and the Debt Contract’ (2006) 1 *Journal of Business and Technology Law* 553, 557; Andrew Keay, ‘A Theoretical Analysis of the Director’s Duty to Consider Creditor Interests: The Progressive School’s Approach’ (2004) 4 *Journal of Corporate Law Studies* 307, 320, 326.

⁸ Alison Grey Anderson, ‘Conflicts of Interest: Efficiency, Fairness and Corporate Structure’ (1978) 25 *UCLA Law Review* 738, 746.

repayments).⁹ Because shareholders bear the most risk, they have the greatest potential for upside—higher risk, higher reward. Such commercial enterprise is actively facilitated by the law; this is precisely the very purpose of *limited liability* whereby the company's separate legal personality protects the shareholders against personal liability arising from the company's failings.¹⁰

Here is where the problem arises. Consider the following propositions: (i) the law facilitates risk-taking; (ii) the courts have legitimised this on the basis of a particular purpose (namely, to encourage investment and profit-making); and (iii) risk is contractually allocated between the company and creditors. If those propositions are accepted as positive statements of corporate governance and law, the subsequent question that arises is the following: on what basis can normative notions of fairness interfere with assumed risk asymmetries that shift as the company's viability changes? Normative concerns emerge in the jurisprudence over the vulnerability of creditors upon insolvency where the lack of equity value means that the costs of management decisions are borne by the creditors (the 'new' residual claimants).¹¹ It is such moralistic concerns over the divergence of interests between shareholders and creditors that have forged a drastic conceptual shift among academics and members of the judiciary as to what corporations are really concerned with—profit-maximisation for the shareholders or distributional fairness among all stakeholders? Determining where to cast the risk of the company's failure is one of the 'principal purposes of limited liability' and of the company's separate legal personality, raising legitimate questions about the level of protection that members of the company deserve.¹² However, while such considerations are meritorious in the light of a growing acceptance of companies' social function and responsibilities, in the author's view, notions of fairness must not *trump* the contractual bargain between creditors and the company and impose *unassumed* obligations upon the directors.

This article identifies a worrisome trend in the jurisprudence. Since Mason J's influential dictum in the Australian case of *Walker v Wimborne*,¹³ there has been a move away from determining risk allocation on the basis of the contractual relations of those voluntarily associated with a company. Instead, this nexus of contracts has been undermined by placing emphasis on distributional fairness—precisely, regulating the implications of risk upon insolvency on the members of the company beyond the shareholders. Most recently, the UK Supreme Court in *Sequana* considered the creditor duty and attempted to shed some light on this nebulous concept.¹⁴ Unsatisfactorily, the Justices were divided over the nature and scope of this duty, failing to provide a clear account of how it manifests in practice and, importantly, *why* it even exists. This lack of a normative explanation of the requirements of such duty led the Supreme Court to adopt justifications and rationales that are devoid of orthodox principles in fiduciary law and contract law: namely, that fiduciary obligations must be voluntarily assumed and that contractual bargains ought to be enforced. The issue identified is the

⁹ Frank H Easterbrook and Daniel R Fischel, *The Economic Structure of Corporate Law* (Harvard University Press 1996) 36.

¹⁰ *Salomon v Salomon & Co Ltd* [1897] AC 22 (HL); *Re City Equitable Fire Insurance Co Ltd* [1925] Ch 407 (CA).

¹¹ See for example Joel Balkan, *The Corporation: The Pathological Pursuit of Profit and Power* (Free Press 2005); Lynn A Stout, 'The Shareholder Value Myth' (2013) Cornell Law Faculty Publications 771 <<http://scholarship.law.cornell.edu/lacpub/771>> accessed 28 March 2025; Collins C Ajibo, 'A Critique of Enlightened Shareholder Value: Revisiting the Shareholder Primacy Theory' (2014) 2 Birkbeck Law Review 37.

¹² *Sequana* (n 5) [52] (Lord Reed). See Judith Freedman, 'Limited Liability: Large Company Theory and Small Firms' (2000) 63 MLR 317.

¹³ (1976) 137 CLR 1 (HCA) 7.

¹⁴ *Sequana* (n 5).

following: what is worse, a lack of adequate compensation for the level of risk borne by the creditors as the new residual claimants or the imposition of a risk allocation by law that was never bargained for? The position adopted is that fairness grounds should not meddle with debtor-creditor risk allocation—not at the expense of principle, coherence, or contract.

Academics and practitioners in the common law jurisdictions have been foreshadowing such a loss of trust in private corporate bargains and their respective contracts.¹⁵ As early as 1974, Grant Gilmore suggested that an expansion of obligations—based on broad notions of duty rather than any contractual basis—pointed towards a problematic trend.¹⁶ This trend was increased judicial intervention in private corporate contracts. The aims of this article are twofold: (i) to revisit this issue and assert the *primacy* of contractual risk allocation; and (ii) to demonstrate that there exists no principled basis for a creditor duty in England and Wales. These arguments are advanced in three sections. Section II tackles the issue of shareholder primacy and the parallel question of whom directors really owe their duties to. Section III analyses the theoretical basis for creditors’ duties, focusing on the key debate between contractarians and law and economics scholars, and progressivist thinkers. Subsequently, Section III critically approaches how this normative debate has emerged in the judicial reasoning in the case law in three jurisdictions: Australia, Delaware, and England and Wales. Section IV focuses on the UK Supreme Court judgment in *Sequana*, identifying its regrettable flaws in the light of the wider academic debate addressed in this article.

II. SHAREHOLDER PRIMACY, DIRECTORS’ DUTIES, AND CORPORATE PURPOSE: IN A STATE OF FLUX?

Company directors are under a fiduciary duty to act in good faith in the interests of the company.¹⁷ The essence of such a relationship is that the fiduciary has *voluntarily* undertaken to act on behalf of somebody else ‘in circumstances which give rise to a relationship of *trust* and *confidence*’.¹⁸ A key principle is that of ‘single-minded loyalty’ to the principal whereby the fiduciary abnegates himself of all self-interest.¹⁹ Such assumption of fiduciary responsibilities to act in the best interest of the company arises when the director is appointed. Directors are automatically subject to statutory duties laid out in part 10 of the Companies Act 2006 (‘CA 2006’), many of which stem from the centuries-long development of the law of equity, fiduciary law, and the law of negligence. According to section 170(3) of the CA 2006, these general duties are based on ‘certain common law rules and equitable principles as they apply in relation to directors and have effect in place of those rules and principles as regards the duties owed to a company by a director’.

Of main importance for the purposes of the analysis of the debt-equity conflict is the ‘success duty’ under section 172(1) of the CA 2006 whereby directors are required to act in the way that they consider, in good faith, to promote the *success* of the company for the

¹⁵ See Easterbrook and Fischel (n 9).

¹⁶ Grant Gilmore, *The Death of Contract* (Ohio State University Press 1974). See also Richard Lindsay Peregrine St John Austen-Baker, ‘Gilmore and the Strange Case of the Failure of Contract to Die after All’ (2002) 18 *Journal of Contract Law* 1; Robert E. Scott, ‘The Death of Contract Law’ (2004) 54 *University of Toronto Law Journal* 369.

¹⁷ *Aberdeen Rail Co v Blaikie Brothers* [1843–60] All ER Rep 249; *Revenue and Customs Commissioners v Holland* [2010] UKSC 51, [2010] 1 WLR 2793.

¹⁸ *Bristol and West Building Society v Mothewe* [1998] Ch 1 (CA) 18 (Millett LJ) (emphasis added).

¹⁹ *Children’s Investment Fund Foundation (UK) v Attorney General* [2020] UKSC 33, [2022] AC 155 [44] (Lady Arden).

benefit of its *members as a whole*.²⁰ Although the notion that directors owe duties to the members of the company is not a neoteric concept, there has been a shift in conceptualising who those members are and the priority in which their interests ought to be considered. This section has three aims. First, to examine the academic theories that have shaped the Anglo-American understanding of the purpose of corporations and demonstrate that they are inadequate to explain the current position under English law. While the contractarian emphasis on contractual risk-allocation is endorsed, the author acknowledges that there exist some doctrinal difficulties in assimilating contractarianism with the English model. Despite this, the progressivist approach is also undesirable due to its disregard for bargain and, instead, its resort to normativism. Secondly, as a result of the critiques of contractarianism and progressivism, to propose an alternative view of how directors' duties operate and precisely who the beneficiaries of those duties are. Such approach attempts to reconcile certain core foundations of contractarianism with English fiduciary law and company law. Thirdly, to contextualise the discussion in the light of the broader debate pertaining to whether shareholder primacy has survived in the CA 2006.

A. THE ANGLO-AMERICAN ACADEMIC DEBATE: INSUFFICIENT?

Since the 1970s, the dominant theory in Anglo-American jurisdictions has been shareholder primacy, narrowing down the purpose of corporations to the maximisation of shareholder profits.²¹ Under this theory, the principal aim of directors' corporate activities is defined by a primary obligation owed to shareholders, with all additional considerations of the interests of other stakeholders taking a secondary status.²² Accounting for the interests of such other groups is often justified on the basis that a lack of focus, at times, on non-shareholder constituencies can negatively impact the maximisation of shareholder wealth by losing specific investments or credit.²³ Stephen M Bainbridge has argued that the shareholder primacy theory is best understood in a twofold manner that correlates economic efficiency-based rationales for prioritising the maximisation of shareholder wealth with control. On this understanding, shareholders should have ultimate control of the corporation because their economic incentives make them the best placed to monitor directors and ensure that corporate resources are allocated efficiently.²⁴ While shareholder primacy continues to take precedence in the US, where it is a foundational principle of corporate law that directors manage corporations primarily for the benefit of their shareholders, the position is not so clear-cut in England and Wales.²⁵ This is because of the introduction of the concept of enlightened shareholder value

²⁰ See for example *Re City Equitable Fire Insurance* (n 10); *Aberdeen Rail* (n 17); *Allen v Gold Reefs of West Africa Ltd* [1900] 1 Ch 656 (CA); *Mutual Life Insurance Co of New York v The Rank Organisation Ltd* [1985] BCLC 11 (Ch); *Item Software (UK) Ltd v Fassihi* [2004] EWCA Civ 1244, [2005] 2 BCLC 91.

²¹ See for example Adolf A Berle and Gardiner C Means, *The Modern Corporation and Private Property* (The Macmillan Co 1933); Steven NS Cheung, 'The Contractual Nature of the Firm' (1983) 26 *Journal of Law and Economics* 1; Stephen M Bainbridge, 'Director Primacy: The Means and Ends of Corporate Governance' (2003) 97 *Northwestern University Law Review* 547; Jonathan R Macey and Geoffrey P Miller, 'Corporate Stakeholders: A Contractual Perspective' (1993) 43 *University of Toronto Law Journal* 401.

²² James J Brummer, *Corporate Responsibility and Legitimacy: An Interdisciplinary Analysis* (Bloomsbury Publishing 1991) 103.

²³ See Oliver Hart and John Moore, 'A Theory of Debt Based on the Inalienability of Human Capital' (1994) 109 *The Quarterly Journal of Economics* 841.

²⁴ Bainbridge, 'Director Primacy: The Means and Ends' (n 21) 573.

²⁵ See for example Leo E Strine, 'The Dangers of Denial: The Need for a Clear-Eyed Understanding of the Power and Accountability Structure Established by the Delaware General Corporation Law' (2015) 50 *Wake Forest Law Review*

(‘ESV’) in section 172 of the CA 2006, which provides that directors owe fiduciary duties to the company and its ‘members as a whole’ (a legitimate basis upon which to question whether that remains the case). This will be explored in the next sub-section. For now, the argument that needs to be tackled concerns the compatibility of prevalent academic theories that have dominated the Anglo-American jurisdictions (primarily the US) with the English law understanding of fiduciary duties and company law. Of specific interest here is the conflict between the dominant US conception of fiduciary duties owed by directors to shareholders and the principle of limited liability, on the one hand, and the statutory provision of section 172 of the CA 2006 that the interests of the company amount to the interests of ‘*its members as a whole*’ (emphasis added) as opposed to merely the shareholders, on the other. The debate has been primarily dominated by two distinct academic viewpoints.

First, the contractarian (otherwise known as the ‘neo-classical’) model and law and economics approaches, which view companies as comprising a ‘*nexus of contracts*’.²⁶ While law and economics theorists largely adhere to the same foundational bases as contractarian theorists, their primary focus is economic *efficiency*. Therefore, rather than viewing shareholder primacy as a normatively desirable model, as contractarians do, they view shareholder primacy as the most economically efficient model.²⁷ Nonetheless, it is not the focus of this article to engage in a detailed examination of the minute theoretical differences between the two theories and so they will be addressed together. The focus of both theories is on the voluntary relationships of individuals associating in a company—viewing the corporation as a complex and consensual amalgamation of contractual bargains.²⁸ Important to this conceptualisation of corporations is the view that contracts can be both explicit *and* implicit.²⁹ For example, when determining the relationship between shareholders and the company, the contractarian approach is twofold. As a first stepping stone, the Articles of Association are perceived as a type of contract where the contractual relationship between the company and its shareholders is created, drawing support from section 33 of the CA 2006 which stipulates that members of the corporation are bound to each other.³⁰ As a latter point, contractarians take the position that there exists an *implicit* contract between the shareholders and the company, which derives from the position of shareholders as residual claimants, and an alleged

761; Lucian Arye Bebchuk, ‘The Case for Increasing Shareholder Power’ (2005) 118 *Harvard Law Review* 833; Richard Williams, ‘Enlightened Shareholder Value in UK Company Law’ (2012) 35 *UNSW Law Journal* 360.

²⁶ For the contractarian approach, see for example Easterbrook and Fischel (n 9); Jensen and Meckling (n 2); Stephen M Bainbridge, ‘Director Primacy and Shareholder Disempowerment’ (2006) 119 *Harvard Law Review* 1735; Henry Hansmann, ‘Ownership of the Firm’ (1998) 4 *Journal of Law, Economics, and Organization* 267. For the law and economics approach, see for example Ronald H Coase, ‘The Nature of the Firm’ (1937) 4 *Economica* 386; Richard A Posner, *Law, Pragmatism, and Democracy* (Harvard University Press 2003).

²⁷ For two distinct conceptions of efficiency, see Laura Lin, ‘Shift of Fiduciary Duty Upon Corporate Insolvency: Proper Scope of Directors’ Duty to Creditors’ (1993) 46 *Vanderbilt Law Review* 1485; Gregory Scott Crespi, ‘Rethinking Corporate Fiduciary Duties: The Inefficiency of the Shareholder Primacy Norm’ (2002) 55 *SMU Law Review* 141.

²⁸ See for example David Millon, ‘New Directions in Corporate Law: Communitarians, Contractarians, and the Crisis in Corporate Law’ (1993) 50 *Washington and Lee Law Review* 1373; Michael J Whincop, ‘Painting the Corporate Cathedral: The Protection of Entitlements in Corporate Law’ (1999) 19 *OJLS* 19; Paddy Ireland, ‘Defending the *Reutier*: Corporate Theory and the Reprivatization of the Public Company’ in John Parkinson, Andrew Gamble and Gavin Kelly (eds), *The Political Economy of the Company* (Hart Publishing 2000).

²⁹ David Millon, ‘Redefining Corporate Law’ (1991) 24 *Indiana Law Review* 223, 230. For a critique of implicit bargains, see Wai Shun Wilson Leung, ‘The Inadequacy of Shareholder Primacy: A Proposed Corporate Regime that Recognizes Non-Shareholder Interests’ (1997) 30 *Columbia Journal of Law and Social Problems* 587.

³⁰ Andrew Keay, ‘Shareholder Primacy in Corporate Law: Can it Survive? Should it Survive?’ (2010) 7 *European Company and Financial Law Review* 369, 397.

agency relationship between the shareholders and company directors whereby the former has *delegated* control to the latter to act in their best interests.³¹

While this theory has been influential due to its emphasis on the existing bargains that formulate the complex net of interactions between the company and its stakeholders, it largely rests on the view that fiduciary relationships are owed *directly* to shareholders and that directors and shareholders are in a principal-agent relationship. The former position, although correct in the US, falls short in England and Wales, where the CA 2006 clearly provides that directors' duties are owed to the company, not the shareholders directly. The latter argument is reliant on shareholder primacy being a foundational principle of company law as such agency relationship would be primarily based on grounds of efficiency: precisely, that agency costs would be eliminated in monitoring the work of managers and ensuring that they do not engage in opportunistic behaviours that would negatively impact shareholders' wealth. However, such position is not tenable under English law, where shareholders do not expressly contract with directors; instead, directors merely contract on behalf of the corporate entity. Shareholders' 'control' power is limited to voting rights, takeovers, and derivative actions but shareholders certainly do not possess any power to instruct directors directly akin to how principals instruct agents in an orthodox agency relationship.³² Even if one attempted to advance an argument that directors owe fiduciary duties to the shareholders directly, such an argument would fail on the basis of the principle of limited liability; for such an argument to succeed, the corporate veil would have to be broken by declaring an agency relationship between the controlling shareholders and 'his' company.³³

Secondly, the progressivist school of thought challenges the shareholder primacy model and advances an inclusive approach which considers the interests of *multiple* stakeholders.³⁴ Such an approach has been influenced by the impact of increasing globalisation and an emphasis on the social obligations of companies that go beyond mere profit maximisation.³⁵ The crux of the problem with the progressivist school of thought, however, is that it is premised upon *normative* notions of what corporate managerial behaviour should be like. The conflict derives from differing ascriptions of importance to how parties themselves choose to form and regulate their relationships—essentially, either adhering to a contractarian view that emphasises the importance of the parties' *contractual* relationship, or adopting *moral* notions that overstep contractual allocations of risk and consider the social effects of corporate activity.³⁶ The genesis of this conflict can be attributed to the presence of two normative considerations as to the role of corporations in the progressivist approach: distribution and fairness.

Distribution is concerned with the allocation of the company's profits or assets among its stakeholders. The distributional conflict arises as a result of the financial distress in which a company finds itself when approaching insolvency—precisely, the concern is risky decision-making by the directors that jeopardises the viability of creditors' claims being paid off. The crux of the issue relates to whether directors' duties to the company—whereby the shareholders are the beneficiaries in the economic sense—warrant a continuation in maximising

³¹ The most well-known account of the agency theory in this regard is Jensen and Meckling (n 2); Eugene F Fama and Michael C Jensen, 'Separation of Ownership and Control' (1998) 26 *Journal of Law and Economics* 301.

³² Robert J Rhee, 'A Legal Theory of Shareholder Primacy' (2018) 102 *Minnesota Law Review* 1951, 1994.

³³ S Ottolenghi, 'From Peeping Behind the Corporate Veil, to Ignoring it Completely' (1990) 53 *MLR* 338, 345.

³⁴ See for example Stout (n 11); Kent Greenfield, *The Failure of Corporate Law* (University of Chicago Press 2006).

³⁵ Cynthia A Williams, 'Corporate Social Responsibility in an Era of Globalization' (2002) 35 *UC Davis Law Review* 705, 720.

³⁶ See for example Callison (n 4); Andrew Keay, 'Directors' Duties to Creditors: Contractarian Concerns Relating to Efficiency and Over-Protection of Creditors' (2003) 66 *MLR* 665.

shareholder return even where the creditors can be financially disadvantaged upon insolvency. *Distribution* being a fundamental concern unsurprisingly derives from the nature of insolvency and the role of contract in allocating security and risk.³⁷ When a company goes insolvent, there will be two classes of creditors to whom the insolvency estate will be distributed. Secured creditors—those holding proprietary security in some of the debtor’s assets—will have priority. Contrastingly, unsecured creditors—those whose financial exposure to the company is usually not commercially significant enough to warrant security—will be subject to the *pari passu* principle, which requires proportionately equal distribution of the remaining assets.³⁸

For progressivist academics, *fairness* cannot be separated from the consideration of distribution for the reason that risk allocation can produce unfair consequences on stakeholders, such as creditors, if not adequately protected.³⁹ Accordingly, they have focused on the fairness of the wealth distribution system when a company goes insolvent, advocating for considerations of social welfare and fair dealing.⁴⁰ This can take two forms: inequality of bargaining power, but also a negative obligation upon directors not to harm stakeholders through risky ventures. Addressing the former, progressivists’ primary focus is the overall fairness of the *transactions* entered into by companies with their creditors—highlighting asymmetries between, for example, banks and trade creditors, when it comes to their monitoring and negotiation abilities.⁴¹ For example, while banks might possess the negotiating power to bargain for covenants that restrict the activities of the company in order to minimise risks, trade creditors will not find themselves in a similar position and will, in fact, often lack the necessary information to price credit accordingly. Addressing the latter, rather than prioritising profit maximisation at times when a company faces financial distress, progressivists advocate for directors’ actions to consider a type of risk-sharing that does not disproportionately affect a group of stakeholders who will ultimately bear the costs of such acts.⁴²

However, such an approach is ultimately grounded in *idealism* in that progressivists attempt to enlarge the considerations of company law beyond the essence of corporate structure which, after all, is how the company regulates its own internal affairs.⁴³ Most importantly, and problematically, they ignore the role that parties’ own contractual agreements play. They thus appear to advance arguments from external disciplines rather than advocate for a *principled* development of the law.⁴⁴

³⁷ Elis Ferran, ‘Creditors’ Interest and “Core” Company Law’ (1999) 20 *The Company Lawyer* 314.

³⁸ See Riz Mokal, ‘The Mysterious *Pari Passu* Principle’ (2024) 39 *Butterworths Journal of International Banking and Financial Law* 443.

³⁹ See Lipton (n 1) 863.

⁴⁰ Keay, ‘A Theoretical Analysis of the Director’s Duty to Consider Creditor Interests’ (n 7) 320. See also AI Ogus, ‘Economics, Liberty and Common Law’ (1980) 15 *Journal of the Society of Public Teachers of Law* 42; Freedman (n 12).

⁴¹ Mark E Van Der Weide, ‘Against Fiduciary Duties to Corporate Stakeholders’ (1996) 21 *Delaware Journal of Corporate Law* 27, 49–50. See also Michael J Whincop, ‘Taking the Corporate Contract More Seriously: The Economic Cases against, and a Transaction Cost Rationale for, the Insolvent Trading Provisions’ (1996) 5 *Griffith Law Review* 1.

⁴² See E Merrick Dodd, ‘For Whom Are Corporate Managers Trustees?’ (1932) 45 *Harvard Law Review* 1145.

⁴³ *Harris v Microfusion 2003-2 LLP* [2016] EWCA Civ 1212, [2017] CP Rep 15.

⁴⁴ Please note that this article employs a distinction between ‘principled’ and ‘normative’ throughout. The former is used to denote a legal basis (upon principle) as opposed to normative considerations devoid of doctrinal coherence. The distinction is in line with the author’s argument which cautions against the oversimplification, and ultimately trumping, of bargains on the basis of notions of fairness.

B. AN ALTERNATIVE VIEW? OWNERSHIP, CONTROL AND RESIDUAL CLAIMANTS

While the above academic theories are inadequate truly to explain the position that English law takes on corporate purpose and directors' duties, this does not mean that the *ipso facto* answer is a simple one, i.e. that directors owe duties to the company and we have moved away from shareholder primacy. For the counter-proposition to be advanced, some time must be devoted to the development of directors' duties in this jurisdiction.

As noted in *Palmer's Company Law*, directors' duties were 'developed by the courts of equity' in 'analogy with' the rules applicable to trustees.⁴⁵ In the former deed of settlement company, a member's share entitled them to some form of beneficial interest in the company's assets, although the precise nature of this interest was hard to identify.⁴⁶ This uncertainty surrounding the nature of the interest was carried on to the position of shareholders who were considered to have a proprietary interest in the assets of the company predicated upon a 'proprietary nexus'.⁴⁷ Accordingly, the courts treated the interests of a company with a share capital as being the same as its members, those being conceived to be the *shareholders*. Any considerations of stakeholder interests were only taken into account if they could affect the interests of the shareholders.

This trust analogy has long been rejected, with it now being established that shareholders do not have 'any right to any item of property owned by the company, for [they have] no legal or equitable interest therein'.⁴⁸ Such development can be attributed to a division of ownership and control facilitated by the acknowledgement of the company's separate legal personality.⁴⁹ Within a property rights framework, the company's nature takes a dual form. On one hand, since *Salomon v Salomon & Co Ltd*, the company is acknowledged as a separate person, legally distinct from its members and subject to its own rights and duties.⁵⁰ It can own property, enter contracts (i.e. the creditor-company contracts addressed in this article), and hold liabilities in its own name. On the other hand, the company remains a *res* in that it can also be the subject of such rights and duties—in simpler terms, it can be the *object* of ownership.⁵¹ The seminal work of Adolf Berle and Gardiner Means accurately captures the revolutionary change that this division between 'ownership' (the *assets* owned by the company versus the *shares* in the entity owned by its members) and 'control' (the appointment of directors to manage the assets) has inflicted on our understanding of property rights.⁵² A company can be an object of ownership while also owning itself because the juridical nature of shares is not associated with equitable proprietary interests anymore. Instead, a share is 'an *interest* measured by a sum of money and made up of various rights contained in the contract [contained in the articles of association], including the right to a sum of money of a more or

⁴⁵ See generally Geoffrey Morse (ed), *Palmer's Company Law* (Release 185, Sweet and Maxwell 2025) para 8.2306.

⁴⁶ Paul L. Davies, Sarah Worthington and Christopher Hare, *Principles of Modern Company Law* (11th edn, Sweet and Maxwell 2021) para 6-001.

⁴⁷ *ibid.*

⁴⁸ *Macaura v Northern Assurance Co Ltd* [1925] AC 619 (HL) 626 (Lord Buckmaster).

⁴⁹ Ernest Lim, 'Of "Landmark" or "Leading" Cases: *Salomon's* Challenge' (2014) 41 *Journal of Law and Society* 523; Paddy Ireland, 'Capitalism without the Capitalist: The Joint Stock Company Share and the Emergence of the Modern Doctrine of Separate Corporate Personality' (1996) 17 *Journal of Legal History* 40.

⁵⁰ *Salomon* (n 10).

⁵¹ Davies, Worthington and Hare (n 46) para 6-002.

⁵² Berle and Means (n 21).

less amount'.³³ From this perspective, shareholders cease being *owners* of the assets of the company, but rather are *investors* entitled to a proportionate share of the company's *profits*. This is precisely why the shareholders are viewed as the residual claimants—their *financial right* to receive dividends if and when declared by the company's board of directors does not provide them with any proprietary interest in specific company assets *prior* to liquidation.

However, just because the shareholders do not have any legal interest in the company's assets does not mean that they do not hold any interest at all. Such interest simply takes a *different form* and is most clearly explicated by contextualising this dual nature of companies in a broader framework that allows us truly to investigate corporate purpose. Such a broader framework requires looking beyond the contractual nexus and, instead, considering the issue within a property rights framework whereby the corporation engages in what Henry Hansmann and Reinier Kraakman refer to as 'affirmative asset partitioning'.³⁴ Andrew Keay has elaborated on the utility of the corporate concept of division of property by emphasising that it is divided into a '*bundle of rights*', with the shares providing shareholders with a '*right to receive some of the fruits of the use of property, a fractional residual right in corporate property, and a very limited right of control*'.³⁵ Such residual essence of rights should not be viewed in the outdated manner of equating residuality with ownership. Instead, it is a residuality that provides a party (such as a shareholder) with an interest in this particular 'bundle of rights'—not the company itself.

This is an important point in the determination of to whom directors' fiduciary duties are owed. As has been argued, the position today is that directors are *not* trustees and that shareholders have no equitable proprietary interest in the company's assets; further, the company remains the legal and beneficial owner of its property.³⁶ But the problem here is that the company is a fictitious person, raising the question of who in fact benefits from directors' duties. Accordingly, adopting an *economic interest* approach, the *real* beneficiaries are the shareholders. However, shareholders' status as 'beneficiaries' is not to be understood in a legal sense; there is a distinction to be made between legal and economic beneficiaries. While directors owe duties to the company, the company's value is ultimately reflected in share prices, providing a measurement of investor sentiment and corporate success. This is supported by Lord Reed's treatment of Nourse LJ's dicta in *Brady v Brady*,³⁷ where his Lordship explicated that, 'where the company is insolvent, or even doubtfully solvent, the interests of the company are in reality the interests of existing creditors alone'.³⁸ Lord Reed disagreed with Nourse LJ, holding that this description 'overstate[s] the position' on the basis that their shares have not yet become 'worthless'.³⁹ Given that the shareholders bear the financial consequences of directors' decisions, it is correct to view them as the *notional* beneficiaries in a financial sense.

Given the above argument concerning the status of shareholders, it would be incorrect to view section 172 of the CA 2006 as a drastic departure from established orthodoxy and the *shareholder primacy* which, it is submitted, has been carried into the Act. In fact, the distinct

³³ *Borland's Trustee v Steel Bros & Co Ltd* [1901] 1 Ch 279 (Ch) 288 (Farwell J) (emphasis added).

³⁴ Henry Hansmann and Reinier Kraakman, 'The Essential Role of Organizational Law' (2000) 110 *Yale Law Journal*, 387, 394. See also Armour and Whincop (n 1).

³⁵ Keay, 'Shareholder Primacy in Corporate Law' (n 30) 395, referring to Dow Votaw, *Modern Corporations* (Prentice-Hall 1965) 96–96.

³⁶ *JJ Harrison (Properties) Ltd v Harrison* [2001] EWCA Civ 1467, [2002] BCC 729.

³⁷ [1988] BCLC 20 (CA).

³⁸ *ibid* 40 (Nourse LJ), cited in *Sequana* (n 5) [50] (Lord Reed).

³⁹ *Sequana* (n 5) [50].

opinions on the creditor duty can be attributed in part to a lack of understanding of this important concept of *shareholder primacy* and what it really entails in practice and theory. Precisely, that shareholder primacy is not antithetical to considerations of other stakeholders' interests. Lady Arden's explanation in *Sequana* is of particular utility as it provides a more nuanced perspective into shareholder primacy than the often unwarrantedly narrow and stringent portrayals of it.⁶⁰ Her Ladyship helpfully outlined that shareholder primacy has 'at least two aspects': (i) a governance aspect and (ii) a right to residual equity.⁶¹

The former explains the fiduciary trust and confidence that exists between the directors and the shareholders who, as the incorporators of the company, have the power to choose, appoint, and remove those who control the assets of the company. However, it is also important to mention that the CA 2006 makes a distinction between a *de jure* and *de facto* director.⁶² A *de facto* director is someone who assumes to act as a director despite not having been validly appointed as such. Section 250 of the CA 2006 defines a *de facto* director as 'any person occupying the position of director, by whatever name called'. Thus, it is not the description of the role that matters but rather its *substance* or *function*. For instance, in *Primlake Ltd v Matthews Associates*⁶³ the defendant was found liable to account for improperly distributed funds on two limbs—either because he was a *de facto* director or because he was a constructive trustee. The point is an important one as it emphasises the *voluntary* aspect of fiduciary duties which, as will be argued later, has been undermined.⁶⁴

The latter explains that the shareholders are the company's residual risk bearers when the company is solvent—a direct result of the company's separate personality, as argued above. Importantly, however, Lady Arden makes the point that shareholder primacy is not *unqualified* in section 172 of the CA 2006. What this suggests is that, instead of witnessing a departure from shareholder primacy, we are merely concerned with a qualification to its application. When viewed in the light of the enlightened shareholder approach enshrined in the Act, it is clear that the consideration of interests has been widened. However, it would be antithetical to interpret this as a removal of shareholder primacy—shareholder gain remains the objective of the success duty when the company is solvent. Where the Supreme Court took the wrong turn, however, is that this qualification of shareholder primacy arises *prior* to insolvency. The sections below provide a critical perspective on the normative reasons for providing such qualification in the form of a creditor duty and provide principled reasons for its untenability under English law.

⁶⁰ See for example Crespi (n 27); Kai Luck and Scott Atkins, 'Shareholder Primacy Already Requires Directors to Actively Consider Non-Shareholder Interests' (*Enterprise Governance eJournal*, 2019) <<https://www6.austlii.edu.au/cgi-bin/viewdoc/au/journals/BondEntGoveJl/2019/5.html>> accessed 29 March 2025; Jonathan Mukwiri, 'Myth of Shareholder Primacy in English Law' (2013) 24 *European Business Law Review* 217.

⁶¹ *Sequana* (n 5) [386].

⁶² Companies Act 2006, s 250.

⁶³ [2006] EWHC 1227 (Ch), [2007] 1 BCLC 666.

⁶⁴ See for example *Bristol and West Building Society* (n 18); *Grimaldi v Chameleon Mining NL (No 2)* [2012] FCAFC 6, (2012) 87 ACSR 260 [177]; *Galambos v Perez*, 2009 SCC 48, [2009] 3 SCR 247; *FHR European Ventures LLP v Cedar Capital Partners LLC* [2014] UKSC 45, [2015] AC 250.

III. TRACING THE DEVELOPMENT OF A CREDITOR DUTY IN THEORY AND PRACTICE

If one accepts the propositions advanced in the previous section, then it remains correct to state that corporate managerial roles are inherently based on the maximisation of stakeholder value. Although the conflict of interest between debt and equity holders over distribution policy or structure implicates an economic question as to the ultimate managerial goal of directors, the role of fiduciary duties in this context must not be overlooked. Stemming from the power of shareholders to choose, appoint, and remove directors lies another corresponding equitable principle: ‘a trustee must use the powers conferred by a trust for the *legitimate purposes* of the trust’.⁶⁵ Directors’ duties are to be exercised bona fide in the manner contemplated by those who gave it. Any modifications of what this power confers and what the ensuing responsibilities and duties of directors are, and to whom they are owed, are matters that derive from this fiduciary relationship and accompanying contract—not normative conceptions of the role of fairness in this schema. This section will consider the development of a creditor duty in the case law of three jurisdictions: Australia, Delaware, and England and Wales. The utility in considering case law from Australia and Delaware arises for two reasons. First, English courts have been largely influenced by Australian courts’ development of the creditor duty, as is evidenced in *West Mercia Safetywear Ltd v Dodd*.⁶⁶ A closer examination of Australian jurisprudence is warranted because this article identifies an incorrect perpetuation and use of proprietary interest to justify a creditor duty. Secondly, a consideration of the Delaware approach provides a useful comparator to how shareholder primacy can be balanced with creditor self-help. It will be demonstrated that a formulation of a creditor duty has been fundamentally reliant upon the invocation of normative theoretical arguments, as opposed to *principle*, to find a legitimate basis for such a duty when one could not be found in the law.

A. AUSTRALIA: THE INTRODUCTION OF THE CREDITOR DUTY

The High Court of Australia in *Walker* was the first Commonwealth court to impose a creditor duty in misfeasance proceedings taken by a liquidator against directors of a corporate group who had moved funds from an insolvent company to other companies in the group.⁶⁷ In Mason J’s influential dictum, it was advanced that, when directors discharge their duty to the company, they must also take account of ‘the interest[s] of its shareholders *and* its creditors’, explaining that any such failure will have ‘adverse consequences for the company as well as for them’.⁶⁸ It is important, however, to contextualise this argument and not merely to construe it as advocating for, let alone setting out, a free-standing creditor duty. The court

⁶⁵ *Sequana* (n 5) [226] (Lord Hodge) (emphasis added); *Howard Smith Ltd v Ampol Petroleum Ltd* [1974] AC 821 (PC); *Re Smith and Fawcett Ltd* [1942] Ch 304; Nathaniel Lindley, *A Treatise on the Law of Companies: Considered as a Brand of the Law of Partnership* (6th edn, Sweet and Maxwell 1902) 510.

⁶⁶ [1988] BCLC 250 (CA).

⁶⁷ *Walker* (n 13); Rosemary Teele Langford and Ian Ramsay, ‘The Contours and Content of the “Creditors’ Interests Duty”’ (2021) 21 *Journal of Corporate Law Studies* 85, 86.

⁶⁸ *Walker* (n 13) (emphasis added). See also *Lonrho Ltd v Shell Petroleum Co Ltd* [1980] 1 WLR 627 (HL) 634, where Lord Diplock stated that the best interests of the company ‘are not exclusively those of its shareholders but may include those of its creditors’.

emphasised the fundamental principle of separate legal personality, noting that, although the companies were part of a wider corporate structure, it was the duty of the directors of Asiatic ‘to consult *its* interests and *its* interests *alone*’.⁶⁹ The subsequent assertion to consider the interests of creditors must therefore be viewed as part of acting in the interests *of the company*. The justification for a creditor duty that emerges from Mason J’s judgment is that prejudicing creditors may have negative consequences for the *company*.

An important endorsement of the view that disregarding creditors’ interests can prejudice the company came in *Kinsela v Russell Kinsela Pty Ltd*.⁷⁰ There, the New South Wales Court of Appeal gave strong judicial recognition of a *fiduciary* duty owed by the directors to the creditors on what can be viewed as a twofold basis. First, competing interests. Street CJ held that when a company is solvent the ‘proprietary interests of the shareholders’ entitle them to be regarded as the company, whereas when the company is insolvent the creditors become ‘prospectively entitled’ through liquidation to the company’s assets.⁷¹ Street CJ described this process as the creditors’ interests *intruding*, resulting in the assets being theirs ‘in a *practical* sense’.⁷² While the conflicting interests of directors and shareholders were acknowledged, the court nonetheless chose to place them under the same fiduciary umbrella. This presents difficulties as fiduciary duties are premised upon a *voluntary assumption* of a duty of loyalty to the principal. As argued earlier, such voluntary assumption occurs the moment a director is appointed and consequently *accepts* such appointment. It is unclear how a director can be held to owe a fiduciary duty to creditors when they never made such an undertaking or purported to act as if they had undertaken such duty. Further incoherence is created by the fact that to hold that a director owed a fiduciary duty to both the shareholders and the creditors would be a breach of the obligation of undivided loyalty as the two groups of principals may have conflicting interests.

Secondly, proprietary interest. Despite the above addressed rejection that shareholders acquire equitable proprietary interest in the company, notions of proprietary interest shifts continued to permeate the case law that developed the conception of a creditor duty. For example, consider Cooke J’s judgment in *Nicholson v Permakraft (NZ) Ltd*,⁷³ where he argued that, when a company is solvent, the shareholders are entitled to be regarded as the company when questions of directors’ duties arise on the basis of their *proprietary interest*.⁷⁴ This misstep was regrettably followed in the dictum of Street CJ in *Kinsela* (later endorsed in *West Mercia* by Dillon LJ).⁷⁵ Despite this theory having also been rejected by the UK Supreme Court in *Sequana*, the point remains that the development of a creditor duty was based on a *fundamentally* flawed conception of proprietary interest in the company’s assets that would shift to the creditors upon insolvency.⁷⁶ Influenced by Cooke J’s approach in *Nicholson*, where he emphasised the shifting beneficial interest in times of ‘marginal commercial solvency’, Street CJ held that directors cease to be able to ratify breaches upon insolvency.⁷⁷ This is because, in his view, the shareholder approval stems from them being, in substance, the company. However, once they cease to have any proprietary interest in the company’s assets, they

⁶⁹ *Walker* (n 13) (emphasis added).

⁷⁰ (1986) 10 ACLR 395 (NSWCA).

⁷¹ *ibid* 401.

⁷² *ibid* (emphasis added).

⁷³ [1985] 1 NZLR 242 (NZCA).

⁷⁴ *ibid* 249.

⁷⁵ *Kinsela* (n 70) 401; *West Mercia* (n 66) 252–53.

⁷⁶ *Sequana* (n 5) [44], [102] (Lord Reed), [132], [147] (Lord Briggs), [246] (Lord Hodge), [256], [346] (Lady Arden).

⁷⁷ *Nicholson* (n 74).

also lose the power or authority to absolve directors of any breaches. Contrastingly to *Walker*, where the duty was expressed as one owed to the company, *Kinsela* appears to be based upon a notion of risk shifting and a corresponding transfer of rights. It thus appears that Street CJ was actually insinuating that the fiduciary duty would *solely* be owed to the creditors upon insolvency due to their acquisition of proprietary interest in the assets. Nonetheless, this argument is flawed as shareholders do not have proprietary interest in the company's assets and so there can be no such actual transfer of rights.⁷⁸ The point at which such transfer is meant to happen was also left unclear as Street CJ refused to 'formulate a general test of the degree of financial instability which would impose upon directors an obligation to consider the interests of creditors'.⁷⁹

What does emerge from this analysis of *Walker* and *Kinsela* is that policy considerations surrounding the fairness of distribution of assets upon insolvency prompted the courts to sidestep *principle*. Due to a lack of principled basis, there exist clear inconsistencies as to whether the duty is an independent one owed *directly* to creditors or whether it is an *indirect* duty owed to the company to take into account the interests of creditors. Uncertainty remains as to whether this duty requires exclusive consideration of creditors' interests or whether other groups are also to be considered and to what extent.

B. DELAWARE: CREDITOR SHIFT AND ECONOMIC EFFICIENCY

While Australian jurisprudence appears to be more aligned with the progressivist school of thought, Delaware law embodies the emphasis on economic efficiency from the law and economics approach in its evolution of a creditor duty. The starting point is the case of *Credit Lyonnais Bank Nederland NV v Pathe Communications Corp.*⁸⁰ where Allen J held that, when a corporation is 'operating in the vicinity of insolvency', directors are not merely the agent of the residual risk bearers—the shareholders—but rather owe their duty to the 'corporate enterprise'.⁸¹ The case emerged in connection with a failed leveraged buyout where, within months of Pathe Communications acquiring a controlling interest in MGM-Pathe Communications Co, its trade creditors forced it into bankruptcy court. The bank attempted to finance MGM's escape from bankruptcy by installing its own management team, which was subsequently challenged by the controlling shareholder on the basis that the management's refusal to sell MGM's interest was against the interest of shareholders and thus constituted a breach of management's fiduciary duty. Ultimately, the court held that the MGM board or its executive committee had an *obligation* towards the 'community of interest that sustained the corporation' to maximise the corporation's 'wealth creating capacity' in good faith.⁸² Allen J was alert to the 'differing interests' at stake and explicated that, '[a]t least where a corporation is operating in the vicinity of insolvency, a board of directors is not merely the agent of the [residual] risk bearers, but owes its duty to the corporate enterprise'.⁸³ *Credit Lyonnais* offers an interesting and nuanced perspective into the justifications for a creditor duty—these can be divided into two aspects.

⁷⁸ See n 14.

⁷⁹ *Kinsela* (n 70) 404. See also John Quinn and Philip Gavin, 'The Creditor Duty Post *Sequana*: Lessons for Legislative Reform' (2023) 23 *Journal of Corporate Law Studies* 271.

⁸⁰ Civ A No 12150 (Del Ch 30 December 1991).

⁸¹ *ibid.*

⁸² *ibid.*

⁸³ *ibid.*

First, the corporation was depicted as representing a ‘community of interests’ which is to be considered on both efficiency and fairness grounds when a company finds itself in the vicinity of insolvency. Importantly, the court recognised that this ‘may diverge from the choice that the stakeholders... would make if given the opportunity to act’.⁸⁴ What is interesting is that Allen J described a corporation as both a *legal* and ‘*economic entity*’,⁸⁵ alluding to the importance of risk shifting in the enforcement of duties. What this means is that, if a duty is owed to the corporate enterprise as a whole, and thus includes creditors, a breach of such a duty would entitle creditors to sue directors directly. This is precisely how *Credit Lyonnais* was subsequently interpreted in *Production Resources Group LLC v NCT Group Inc*,⁸⁶ where it was confirmed that such suits are derivative.

Secondly, the duty has been presented as one of *deterrent* rather than *prospective* effect. Allen J suggested that a board of directors is generally faced with three types of course that differ on the basis of risk: (i) a very risky course that could produce the highest shareholder equity return but has a low success probability; (ii) a moderately risky course that has a higher chance of success but lower return; and (iii) a least risky course that would leave no shareholder equity but would pay creditors in full.⁸⁷ Logically, depending on how the duty is formulated (i.e. whether it is owed to the creditors directly or whether it is an indirect duty to act in the interests of the company), the appropriate course of action would differ. If we, for example, adopt the *Walker* approach, directors would most likely have to engage in a balancing act of interests to determine which course of action is most appropriate, this most likely being the moderate course of action. Contrastingly, adopting the *Kinsela* approach, where the duty is owed directly to creditors, the least risky approach is most suited. Allen J elucidated that the most *efficient* approach would be the moderately risky one.⁸⁸ Thus, the approach is that of deterrence in that it dissuades directors from adopting either the riskiest or most risk-averse courses of action.

Despite this deterrent effect, the Supreme Court of Delaware in *North American Catholic Educational Programming Foundation v Gheewalla* rejected a creditor duty owed directly to creditors.⁸⁹ The Supreme Court emphasised orthodoxy and adopted the contractarian approach espoused by Bainbridge—that, in essence, creditors are limited to whatever rights that the contract they bargained for provides.⁹⁰ The ‘vicinity of insolvency’ espoused in *Credit Lyonnais* was clearly rejected, holding that, regardless of whether the company is insolvent or in the zone of insolvency, the creditors ‘have no right, as a matter of law, to assert direct claims for breach of fiduciary duty against its directors’.⁹¹

C. ENGLAND AND WALES: A MOVE AWAY FROM CONTRACT

English jurisprudence on creditor duty was largely influenced by Australian jurisprudence. In the same year that the *Kinsela* judgment was given, Lord Templeman in *Winkworth*

⁸⁴ *ibid* fn 55 (emphasis added).

⁸⁵ *ibid* (emphasis added).

⁸⁶ 863 A 2d 772 (Del Ch 2004).

⁸⁷ *Credit Lyonnais Bank Nederland* (n 80) fn 55; Callison (n 4) 435.

⁸⁸ *Credit Lyonnais Bank Nederland* (n 80) fn 55.

⁸⁹ *North American Catholic Educational Programming Foundation v Gheewalla* 930 A 2d 92 (Del 2007).

⁹⁰ See Stephen M Bainbridge, ‘Much Ado About Little? Directors’ Fiduciary Duties in the Vicinity of Insolvency’ (2007) 1 *Journal of Business and Technology Law* 335.

⁹¹ *Gheewalla* (n 89) 103.

*v Edward Baron Development Co Ltd*⁹² advocated for an independent creditor duty for the purpose of ensuring that ‘the affairs of the company are properly administered and that its property is not dissipated or exploited for the benefit of the directors themselves to the prejudice of the creditors’.⁹³ The emphasis on *prejudice* by the House of Lords demonstrates a clear influence from *Walker*—expanding the reasoning by holding that the directors are the custodians of the ‘conscience of the company’.⁹⁴ The use of the word *conscience* is an interesting choice—it denotes that the company’s interests encompass moral considerations of right and wrong. Such approach leaves the discretion of course of action to the directors; however, it unhelpfully does not provide any guidance, especially in the light of the difficulty and lack of clarity between determining the risk of insolvency and the extent of permissible risk.⁹⁵

The subsequent case of *West Mercia* is of immense importance—precisely, the question of whether it sets out a common law rule that, in certain circumstances, the interests of the company are to be understood as including the interests of its creditors as a whole was one of the key issues in the Supreme Court case of *Sequana*. In the *Modern Company Law for a Competitive Economy: Final Report* produced by the Company Law Review Steering Group (‘CLRSG’), it was recognised that a key question when considering the codification of the CA 2006 is whether the orthodox rule that a company is to be run in the interests of its shareholders is to be maintained or modified by an obligation to have regard to the interests of creditors.⁹⁶ The report’s justifications for the consideration of departing from orthodoxy surrounded risk shifting—paragraph 3.15 explicitly refers to limited liability exposing creditors to a risk of loss that warrants consideration of protecting the creditors’ interests. It is interesting that *protection* underpinned the reasoning of the CLRSG—‘protection’ denoting both a positive or affirmative act to protect but also an implied negative obligation not to harm prospectively. Lord Hodge in *Sequana* identified two such protections: (i) section 214 of the Insolvency Act 1986 (‘IA 1986’), which imposes personal liability on directors who fail to minimise the potential loss to the company’s creditors when the director knew or ought to have known that there was no reasonable prospect that the company would avoid going into insolvent liquidation; and (ii) an obligation to consider the interests of creditors at an earlier stage in the onset of insolvency, as recognised in Australian case law and by the Court of Appeal in *West Mercia*.

The Justices of the Supreme Court all agreed that a creditor duty is supported in the case law—notably *West Mercia*—and that the duty has been preserved by section 172(3) of the CA 2006.⁹⁷ However, the reasoning in *West Mercia* itself is dubious. Dillon LJ endorsed the reasoning of Street CJ in *Kinsela* where he based the absence of shareholders’ power to ratify directors’ breaches on the basis that a proprietary interest in the company’s assets transfers to the creditors upon insolvent liquidation.⁹⁸ However, as has already been argued, the proprietary interest argument is flawed and was, in fact, rejected by the Supreme Court.⁹⁹ Lord Hodge acknowledged that if the court were to overrule the *West Mercia* judgment it would be going

⁹² [1986] 1 WLR 1512 (HL).

⁹³ *ibid* 1516.

⁹⁴ *ibid*.

⁹⁵ See Keay, ‘A Theoretical Analysis of the Director’s Duty to Consider Creditor Interests’ (n 7).

⁹⁶ CLRSG, *Modern Company Law for a Competitive Economy: Final Report* (Department of Trade and Industry 2001).

⁹⁷ *Sequana* (n 5) [29]–[36], [43] (Lord Reed), [152] (Lord Briggs), [207], [209], [244] (Lord Hodge), [344], [387]–[416] (Lady Arden).

⁹⁸ *Kinsela* (n 70) 401.

⁹⁹ *Sequana* (n 5) [44] (Lord Reed); *Marex Financial Ltd v Sevilleja* [2020] UKSC 31, [2021] AC 39 [31] (Lord Reed), [105] (Lord Hodge).

against the recognition by Parliament of the existence of such a common law duty to creditors.¹⁰⁰ While that is correct—in that section 172(3) of the CA 2006 effectively *authorised* the courts to develop the common law duty of directors in relation to the interests of the company's creditors as a company nears insolvency—the Supreme Court had an opportunity at least to rationalise the reasoning in *West Mercia* and to provide a coherent basis upon which a creditor duty is recognised. Unfortunately, the Supreme Court did not do this. Instead, although agreeing that a common law duty to consider the interests of creditors exists, the Justices of the Supreme Court disagreed on its content.¹⁰¹ Most problematically, they adopted reasoning that not only departed from orthodox principles in fiduciary law but also often perpetuated the flaws in the previous line of case law. These flaws in *Sequana* are explored in the following section.

IV. *BTI 2014 LLC v SEQUANA SA*—A LACK OF COHERENT FOUNDATION

Sequana concerned a dividend declared to Sequana by the directors of AWA, a wholly owned subsidiary of Sequana. This dividend virtually extinguished the debt that Sequana owed to AWA which, at the time, was solvent on both a cash and balance sheet basis. Despite this, there was a real risk that AWA might become insolvent in the future as, at the time that the dividend was declared, there existed significant long-term contingent liabilities. After 10 years, AWA became insolvent. BTI was assigned by AWA to recover the amount of the dividend from the company directors on the basis that the declaration of the dividend was a breach of the creditor duty. While the Supreme Court unanimously held that AWA's directors were not at the relevant time under a duty to consider, or to act, in accordance with the interests of the creditors, the Justices of the Supreme Court differed in their conceptions of the creditor duty. This section identifies and analyses four key principled flaws in the Supreme Court's position on the *content* of the duty.

A. THE CREDITOR DUTY AS A FIDUCIARY DUTY

The first problem with the position adopted in the *Sequana* case is that the creditor duty is essentially a modified *fiduciary duty* to act in good faith in the interests of the company.¹⁰² The essence of a fiduciary relationship has been described by Lady Arden in *Children's Investment Fund Foundation (UK) v Attorney General* as the fiduciary owing 'the duty of single-minded loyalty to his beneficiary'.¹⁰³ Although there is no legal definition of when someone can be considered a fiduciary, the distinguishing feature is that 'B is entitled to expect that A will act either in B's best interests or in their joint interests, to the exclusion of A's own interest'.¹⁰⁴ The nature of fiduciary duties is inherently prospective in effect which indicates that the concern is not about what fiduciaries must do but rather what they should not do—in that sense, they are *prophylactic*. The strict and onerous nature of these duties can be

¹⁰⁰ *Sequana* (n 5) [232] (Lodge Hodge)

¹⁰¹ See *ibid* [11] (Lord Reed), [152], [153] (Lord Briggs), [207], [209], [224] (Lord Hodge), [250] (Lady Arden).

¹⁰² *ibid* [1], [74], [96] (Lord Reed), [252], [265] (Lady Arden).

¹⁰³ *Children's Investment Fund Foundation* (n 19).

¹⁰⁴ Graham Virgo, *The Principles of Equity and Trusts* (5th edn, OUP 2023) 468.

justified only upon the voluntary undertaking by the fiduciary.¹⁰⁵ As explained by Millet LJ, ‘every fiduciary relationship is a *voluntary* relationship. No one can be compelled to enter into a fiduciary relationship or to accept fiduciary obligations, any more than he can be compelled to enter into a contract or to accept contractual obligations’.¹⁰⁶ Thus, as a matter of principle, it is hard to conceive the basis upon which the director assumed a duty of loyalty to the company’s creditors before or upon insolvency. While a counterargument to this might be that the fiduciary duty is owed to the company, as opposed to the shareholders or creditors directly, the manifestation of the company’s interests is most accurately calculated on the basis of the worthiness of the shareholders’ shares. Predicated upon this lies the parallel between shareholders’ economic interest and the directors’ fiduciary duty to the company. The two cannot be logically separated when the latter is measured by the former. From this perspective, it cannot be said that a director has voluntarily undertaken to act for the benefit of the creditors, whether that is directly or indirectly or through the success duty under section 172 of the CA 2006. The success duty remains measured by shareholders’ wealth maximisation. If this were not to be considered by directors, then their duty of loyalty would be devoid of utility and value as there would be no coherent way to ascertain what the company’s success can amount to. Further, from that perspective, it is clear that shareholders purchase shares with an expectation that directors will not harm their interests but will, in fact, maximise their value. Otherwise, companies would simply not attract investment. Accordingly, while creditors deal with the company as a matter of contractual bargain, for shareholders it is *trust*.

This is even harder to justify when one considers that the duty is *ex post* in nature—directors are not subject to it upon entering into credit arrangements on behalf of the company but rather they become subject to it at some point after the contract has been concluded. Proponents of the duty, such as Keay, argue that, although it might be unfair for directors to be held liable for actions that are not covered by the credit contract, the law does impose *ex post* duties where appropriate to remedy *unfairness*.¹⁰⁷ An example provided is section 239 of the IA 1986, which attempts to prevent a situation where *pari passu* distribution is circumvented by giving a creditor priority in relation to past indebtedness at the expense of other creditors. The *pari passu* principle provides an interesting analogy on normativism within this area of corporate and insolvency law. Riz Mokal notes that the *pari passu* principle is insensitive to considerations of fairness surrounding a ‘myriad [of] normatively significant differences amongst claimants’, such as the ability to monitor and influence the terms on which one becomes a creditor; instead, the rule remains centred on the *formal status* of a creditor.¹⁰⁸ As argued by Jan H Dalhuisen, ‘[i]t is... the ranking..., and not the equality, that is the essence of bankruptcy and of [creditors’] relationships more generally’.¹⁰⁹ Viewing the *pari passu* principle as one of equality or fairness would thus be incorrect—the principle only treats equals equally, taking creditors ‘exactly as it [the IA 1986] finds them’ from their pre-insolvency arrangements.¹¹⁰ Both Mokal and Dalhuisen support the argument that this indicates that the *pari passu* rule is ultimately reflective of a wider basis upon which the system of creditors’

¹⁰⁵ *Bristol and West Building Society* (n 18).

¹⁰⁶ PJ Millet, ‘Restitution and Constructive Trusts’ (1998) 114 LQR 399, 404–05 (emphasis added).

¹⁰⁷ Keay, ‘A Theoretical Analysis of the Director’s Duty to Consider Creditor Interests’ (n 7) 313.

¹⁰⁸ Mokal (n 38) 445 (emphasis removed); however, note that the position varies when employee rights and claims are concerned. See also Look Chan Ho, ‘On *Pari Passu*, Equality and Hotchpot in Cross-Border Insolvency’ [2003] LMCLQ 95.

¹⁰⁹ Jan H Dalhuisen, *Dalhuisen on International Commercial, Financial and Trade Law* (Hart Publishing 2000) 661 (emphasis removed).

¹¹⁰ *Re Smith, Knight & Co, ex p Ashbury* (1868) LR 5 Eq 223, 226 (Lord Romilly MR).

protection and insolvency law is based: *equity*, not equality.¹¹¹ What ensues from a normative analogy with the *pari passu* principle is confirmation that at the core of corporate arrangement lies precisely this: *contract*. Parties bargain over distribution, not fairness, and it certainly is not the role of the courts to question their bargains by imposing *ex post* duties. While some level of regulation is accepted even amongst contractarians and law and economics academics, a creditor duty reverses contractual risk allocation and thus imposes a new distributional mechanism that parties never agreed upon.

A potential counterargument as to whether a director can be held to owe fiduciary duties to creditors is that a fiduciary relationship can also arise when one voluntarily places oneself ‘in a position where he is obliged by equity to act in the interests of another’.¹¹² But, even then, it is difficult to see how a director would have placed himself in such a position. This is because the duty of loyalty is owed by the director to the company and so to its *notional* beneficiaries—the *shareholders*. Oppositely, creditors are in an arm’s length relationship with the company and so a consideration of their interests prior to insolvency would logically be a breach of the duty of loyalty. Even where risk shifting would result in the displacement of the shareholders’ interests by the creditors as the new residual claimants, the same problem of conflicting interests, and so a breach of the duty of loyalty, would remain if the duty was triggered at any moment *prior* to insolvent liquidation. If the duty of loyalty is owed to the company, but the interests of the company are measured in a financial sense by the maximisation of shareholders’ wealth, then the manifestation of that duty would involve the shareholders being its beneficiaries, not in a legal sense but in a financial one. From that perspective, any prioritisation of creditors’ interests would inherently go against such wealth maximisation and would involve a conflict of interest. A different view would suggest that the success duty is measured by a company’s viability even when its profits measured through shareholder wealth are underperforming. Wealth maximisation remains the core corporate purpose. This is clearly acknowledged by Lord Reed who adopts a primary duty versus secondary obligation analogy.¹¹³ The relevant primary duty expressed in section 172 of the CA 2006, although expressed to the company, involves the beneficiaries as the ‘intended beneficiaries of that duty’. The subsequent secondary obligation that arises to consider the interests of other stakeholders is employed as part of the wider multi-faceted consideration of what amounts to the ‘interests of the company’. As argued earlier, actions that can harm stakeholders might have reputational impact on a company (amongst other things) which would subsequently negatively affect the company’s profitability and, thus, its interests. It would be an elliptical approach to ignore how primary duty is measured (i.e. that its success or failure are only measured through its effect on shareholders’ wealth). Peter Watts KC has eloquently made the point: ‘it is one thing to remind ourselves of the intrusion of limits on the vires of the operators of companies after insolvency, it is quite another to assume that creditors become the focus of the exercise of all powers by directors’.¹¹⁴

From this perspective, Lady Arden’s judgment is preferable to the view adopted by the majority which resorts to notions of shifting economic interests.¹¹⁵ Lady Arden takes the

¹¹¹ Mokal (n 38); Dalhuisen (n 109). See also Keay, ‘Directors’ Duties to Creditors’ (n 36) 665, where Professor Keay notes that permitting creditors to recover under a direct duty could damage the *pari passu* principle.

¹¹² Millett (n 106) 405.

¹¹³ *Sequana* (n 5) [66].

¹¹⁴ Peter Watts, ‘Why as a Matter of English-Law Principle Directors Do Not Owe a Duty of Loyalty to Creditors upon Insolvency’ [2021] *Journal of Business Law* 103.

¹¹⁵ See *Sequana* (n 5) [42], [83], [86] (Lord Reed), [246] (Lord Hodge).

view that the creditor duty is, in essence, a modification of the success duty by section 172(3) of the CA 2006, which requires directors to consider and act in the interests of creditors at a specific moment in time. She rejects any explanations on the basis of shifting interests; instead, she advances that the true effect of section 172(3) introduces a *qualification* arising from the rule of law that *restricts* the acts of directors in how they perform their duty. Further, this rule concerns *protecting* directors from harm rather than requiring directors to run the business for the benefit of creditors. The requirement is that the directors should consider creditors' interests and act on those interests in *certain* circumstances. The directors may make a decision that benefits shareholders, but this would not be a breach of the requirement that they act in the interests of creditors if the creditors would not be worse off in a liquidation. This is a slight, yet interesting, distinction, encapsulating a *negative* obligation not to ignore the interests of creditors at insolvency rather than a *positive* obligation to take acts to benefit the position of the creditors. It is rather a battle of interests rather than a displacement of shareholder primacy. Shareholder primacy does not mean that shareholders' interests *exclude* those of others with legitimate interests.

Lady Arden employs the concept of a 'sliding scale' to assist with her modification argument, which allows her to overcome the previously mentioned issue of potential breach of the duty of loyalty prior to insolvent liquidation. A sliding scale analogy is mere guidance rather than a literal rule.¹¹⁶ It guides directors in managing the interests of a company until the point is reached where the creditors' interests are predominant (which, in the author's view, would logically be the point of insolvent liquidation given that, prior to liquidation, shareholders still have an economic interest in the company). Where, however, Lady Arden's judgement can be criticised is that she agrees with Lord Reed that '[i]t is only where an insolvent liquidation or administration is unavoidable that the shareholders cease to have any interest in the company, and their interests can therefore be left out of account'.¹¹⁷ This is problematic because her Ladyship bases this argument on the view that the key moment when the interests of shareholders are displaced as a matter of law is when the company is 'irretrievably insolvent'.¹¹⁸ However, this ignores the earlier point made by Lord Briggs that insolvency is not actually the moment upon which the economic stakeholding is truly transferred; it is the moment of *liquidation*.¹¹⁹ From this perspective, any arguments upon reliance of transfer of residual risk to creditors are untenable as such risk is (i) addressed in creditors' credit arrangements; and, (ii) if it is not, it is the moment of liquidation that will provide them with the right to deal with the company's assets under the insolvency regime.

Lord Briggs rejected the argument that the duty requires treating creditors' interests as paramount by adopting a contractarian view—precisely, that it is inherent in the law's encouragement of risk-taking and commercial enterprise under limited liability that voluntary creditors are able to make judgments about risk and accordingly take sufficient precautions as they see fit.¹²⁰ His Lordship therefore rejects the notion that mere risk can elevate the position of creditors to the status of 'paramount stakeholders'.¹²¹ As a consequence, he adopted a more *practical* notion of economic interest, holding that it is the 'onset of liquidation' rather

¹¹⁶ *ibid* [303], [419].

¹¹⁷ *ibid* [80], referred to by Lady Arden: [291].

¹¹⁸ *ibid* [231], [247] (Lord Hodge), [302], [323] (Lady Arden).

¹¹⁹ *ibid* [165].

¹²⁰ *ibid* [132], [164], [175].

¹²¹ *ibid* [164].

than insolvency that converts creditors into the main economic stakeholders in the company.¹²² He compares the position to that under section 214 of the IA 1986, which imposes liability upon directors for failure to act in the interests of creditors once liquidation becomes inevitable.¹²³ It is on this basis that he rejects up to date justifications for the creditor duty as rendering creditors' interests paramount upon insolvency. Prior to liquidation, the position in his view is that the creditor duty is merely a duty to consider creditors' interests, to give them appropriate weight, and to balance them against shareholders' interests where they may conflict.¹²⁴ From this perspective, it appears that the duty is essentially a *prospective statutory priority* that kicks in when the company goes into insolvent liquidation.¹²⁵ However, the problem with this approach is that it is devoid of any justificatory arguments that advocate for the existence of creditor duty when insolvency is imminent; if the duty only arises at the moment of liquidation, then its *prophylactic* nature will never actually reap its benefits.

Nonetheless, what this analysis demonstrates is that the approaches adopted by the Justices of the Supreme Court do not speak in one voice. It is regrettable that, despite the above flaws, the majority made the same error as that in *West Mercia*—holding that a director can owe a fiduciary duty towards the creditors without there ever having been the requisite voluntary undertaking.

B. TREATMENT OF CREDITORS AS A CLASS AND THE DIMINISHMENT OF CONTRACT

The idea that directors owe a fiduciary duty of loyalty to creditors after insolvency, assumes—and accordingly treats—creditors as a *class*. Lord Reed explicitly made this point when he said that creditors must be treated as ‘a class’ rather than ‘as a *fixed group* of individuals’.¹²⁶ This is the second problem with the position adopted in *Sequana*, as creditors are not a *homogeneous* class.¹²⁷ Treating creditors as a class assumes that they are a homogeneous group which, as a corollary, deserves the same treatment reflecting such homogeneity. Contrastingly, fixed groups are fixed because they have heterogenous rights and obligations that arise from distinct contractual relationships. Indeed, some are voluntary creditors; some are involuntary. Some negotiate detailed contracts; some do not. Some are fully secured with the firm's assets; some have either no security or lesser security. This indicates that the treatment of creditors as a class and the expansion of fiduciary duty to them ignores the conflicting interests that creditors can have between themselves. Frederick Tung emphasises the importance of *intercreditor conflict*, noting that, when the company approaches financial distress, creditors become competitors rather than an allied class with the same interests.¹²⁸ On this basis, he argues that a creditor duty reduces a complex multiparty conflict into a seemingly bilateral bargain which treats debt as undifferentiated and unitary.¹²⁹ The problem with such an approach is that it is a highly stylised and idealistic model of debt-equity conflict which

¹²² *ibid* [165].

¹²³ *ibid* [148], [165], [176].

¹²⁴ *ibid* [176].

¹²⁵ A ‘prospective statutory priority’ alludes to the prospective entitlement that creditors acquire to the company's assets upon its winding up. Such notions of ‘prospective entitlement’ appear in *Sequana*; see for example *ibid* [92] (Lord Reed).

¹²⁶ *ibid* [48] (emphasis added).

¹²⁷ See Callison (n 4) 450.

¹²⁸ Tung (n 4) 828.

¹²⁹ *ibid* 828, 841.

ignores the real commercial considerations and conflicts that exist between creditors. It reduces complex and variable interests into simplistic and unitary bargains which are unrepresentative of the reality of creditors' credit contracts. What ensues from this is that any arguments for expanding a fiduciary duty of loyalty to creditors *imputes* such intention to creditors who would not—and, importantly, *have not*—agreed to any generalised, judicially imposed creditor-shareholder bargain. For instance, it is hard to imagine that a creditor with a detailed negotiated contract that offers him superior protection to another creditor would agree to inferior rights. While Lord Reed acknowledges that limited liability reflects the fact that the relationship between creditors and the company is purely contractual, and so creditors can *negotiate* the terms on which credit is given to reflect the risk undertaken, his Lordship concludes that this does not mean that creditors are 'bereft of legal protection'.¹³⁰ This posits yet another *normative* conflict: what is worse, a lack of adequate compensation for the level of risk borne by the creditors as the new residual claimants or the imposition of a risk allocation by law that was never bargained for?

The Supreme Court clearly decided that imposing an unbargained-for risk allocation is justified. This is most evident in the dicta of Lord Hodge.¹³¹ His Lordship examined in detail the report produced by the CLRSG, which stated that 'these interests [the creditors' interests] are covered by contract while the company is solvent; but if insolvency threatens they *override* any considerations of the success of the company for members'.¹³² Lord Hodge endorsed this view and stated that, if there was no such override on insolvency, 'our company law in relation to directors' duties would lack both clarity and coherence'.¹³³ This is because directors would remain under an apparently unqualified duty in section 172(1) of the CA 2006 to promote the success of the company for the benefit of its shareholders even when the company was insolvent. This view is problematic, not only because it trumps contract, but also because its justification for doing so (i.e. the notion that duty in section 172(1) would be unqualified) is flawed. This is precisely why Lady Arden's position is preferred—her Ladyship's emphasis that the duty remains owed to the company but merely incorporates a wider group of interests (not just the creditors') in the form of a sliding scale when insolvency is imminent tackles Lord Hodge's concern. The duty is not a shift but rather a *qualification* arising from the rule of law.¹³⁴ Nonetheless, the issue of lack of voluntary undertaking remains the most fatal and, unfortunately, unsatisfactorily addressed flaw of the creditor duty.

Another issue with the majority's position is that the displacement of contract is justified upon a moralistic view that creditors deserve *enhanced* protection, adopting rationales that are premised on *fairness* rather than legal principle.¹³⁵ While it is correct that credit contracts vary and that not all creditors have the same ability to negotiate their position, such variation does not provide any justificatory basis upon which the courts should remedy this imbalance of power. Such imbalance is merely the reality of commercial bargaining. Nevertheless, it would be incorrect to adopt the view that creditors are not already protected. Two ways have been identified in which this happens.

¹³⁰ *Sequana* (n 5) [53].

¹³¹ See *ibid* [210]–[222].

¹³² *ibid* [214], [229], [230].

¹³³ *ibid* [230].

¹³⁴ *ibid* [265].

¹³⁵ In this case, the relevant legal principle addressed is sanctity of contracts. Normative considerations of fairness cannot, and should not, displace contractual bargains which already reflect the desired level of protection negotiated by creditors.

First, creditors are protected *directly* by the ‘panoply’ of remedies available in insolvency law.¹³⁶ This point was made by Lady Arden when she referred to Lord Hodge’s example at paragraph 238, where he expressed that, without a creditor duty, there would be no other remedy available under the general law for creditors. This position omits the fact that directors can actually bring a claim in *misfeasance* against a company director for having breached their duty to act with reasonable skill and care when making decisions.¹³⁷ Further, adopting the interpretation of section 170(1) of the CA 2006 as requiring the considerations of the company’s members as a whole would mean that, when making decisions for the benefit of the company, disregarding the interests of creditors can have negative consequences for the company. An example of this is *Colin Gwyer & Associates Ltd v London Wharf (Limehouse) Ltd*,¹³⁸ where the director did not show single-minded loyalty to the company by having no regard to the interests of the creditors. Lastly, section 214 of the IA 1986 already imposes a duty on directors to take reasonable care to minimise the potential loss to the company’s creditors.

Secondly, creditors are protected *indirectly* by delegated monitoring. The theory of delegated monitoring provides that credit contracts with minimal credit protection indirectly benefit from secured banks’ monitoring.¹³⁹ The starting point of this theory is that creditors—regardless of their variations—have a collective interest in minimising transaction costs and efficient delegation of monitoring responsibilities.¹⁴⁰ A borrower is more likely to grant covenant protections to banks as the low-cost monitor in order to minimise monitoring costs; the same does not apply to other creditors who are not equipped to monitor, such as public bondholders. A rich finance literature supports this argument, pointing out that banks *typically* play the role of delegated monitor. As noted by Saul Levmore, other creditors, aware of the bank’s extended credit and monitoring abilities, avoid incurring monitoring costs by simply free-riding.¹⁴¹ Accordingly, the reduction of risk-credit and overinvestment by banks benefit the creditors with less secure positions and explain why an imbalance in protection in credit contracts does not warrant judicial involvement. Given those direct and indirect protections, the subsequent question is whether there exist any grounds upon which creditors should receive *enhanced* protection. The answer is no: fairness grounds should not displace contractual bargains and meddle with parties’ risk allocation.

C. ECONOMIC INTEREST MASKED AS PROPRIETARY INTEREST

Problematically, the argument that creditors are to be treated as a class served the basis for the notion of shifting *economic interest* adopted by Lord Reed, with whom Lords Briggs and Hodge agreed.¹⁴² As has already been pointed out, the notion that there exists a transfer of proprietary rights was rejected by the Supreme Court. In order to rationalise the case law where such a proprietary interest explanation was adopted—such as *Kinsela* and its endorsement in *West Mercia*—Lord Reed transposed the analysis from an emphasis on

¹³⁶ *Sequana* (n 5) [328] (Lady Arden).

¹³⁷ *ibid* [330].

¹³⁸ [2002] EWHC 2748 (Ch), [2003] BCC 885.

¹³⁹ Douglas W Diamond, ‘Financial Intermediation and Delegated Monitoring’ (1984) 51 *Review of Economic Studies* 393; Alan Schwartz and Robert E Scott, ‘Contract Theory and the Limits of Contract Law’ (2003) 113 *Yale Law Journal* 541; Tung (n 4).

¹⁴⁰ Tung (n 4) 837.

¹⁴¹ Saul Levmore, ‘Monitors and Freeriders in Commercial and Corporate Settings’ (1982) 92 *Yale Law Journal* 49, 53–54. See also *Sequana* (n 5) [12] (Lord Reed).

¹⁴² *Sequana* (n 5) [56], [83] (Lord Reed), [246] (Lord Hodge), [256], [263], [311] (Lady Arden).

proprietary interest to economic interest. His Lordship drew a distinction between economic interest, on the one hand, and legal interest, on the other, holding that ‘economic interest in the company’ by the company’s creditors provides them with the ‘entitlement to be paid the debts owed to them’, which is ‘ultimately enforceable against the proceeds of realisation of the company’s assets’.¹⁴³ An attempt can be made at rationalising Lord Reed’s stance by examining in more detail his position on Nourse LJ’s dicta in *Brady*, where Nourse LJ was adamant that, at times when the company is insolvent or even doubtfully insolvent, the interests of the company are the creditors’ alone.¹⁴⁴ Interestingly, Lord Reed rejected this contention, holding that it overstates the position. Contrastingly, his Lordship held that it is only in the eventuality of an insolvency that shareholders have no remaining interest in the company as ‘their shares become worthless’.¹⁴⁵ This provides more insight into what he meant by ‘economic interest’. Specifically, the allusion to the financial worth of shares suggests that reference to economic interest is really the *monetary conversion* of legal interest.

Despite this argument being eloquently and attractively put forward, on a *superficial* level, it is not dissimilar to the line of thought advanced by Dillon LJ in *West Mercia*. Precisely, the notion that creditors become ‘prospectively entitled, through the mechanism of liquidation, to displace the power of the shareholders and directors to deal with the company’s assets’ resulting in the assets being ‘in a practical sense’ the creditors’.¹⁴⁶ The flaws of the reasoning in *West Mercia*—already addressed above—have been reiterated in Lord Reed’s judgement, albeit through a different lens. It is, however, unclear upon what basis a non-legal interest can give rise to legal obligations which in turn have the potential to make directors personally liable directly to creditors. While it is not incorrect to acknowledge that a corporation—as in *Credit Lyonnais*—is a legal and economic entity, and thus has both legal and economic interests, it is incorrect to conflate the two. Economic interest in the case of creditors is merely the *result* of acquiring legal interests, such as through debt-equity contracts; it cannot be used to entitle parties to new legal interests that they did not possess upon contracting. From this perspective, Lord Reed’s reasoning is *conclusory*, or explanatory, focusing on explaining the *effect* or outcome without delving into its *causes*.

D. ‘BORDERING ON INSOLVENCY’ AND UNCERTAINTY PERTAINING TO THE RATIFICATION PRINCIPLE AND SECTION 214 OF THE IA 1986

The case law on the trigger of the duty is full of linguistically variable determinations. Lord Toulson’s formulation in *Bilta (UK) Ltd v Nazir (No 2)* was that of ‘bordering on’¹⁴⁷ insolvency; Nourse LJ described it as ‘doubtfully solvent’ in *Brady*,¹⁴⁸ and other cases used concepts, such as ‘verge of insolvency’¹⁴⁹ or ‘potentially insolvent’.¹⁵⁰ Rather than remedying this uncertainty, the Supreme Court was, in fact, divided over the preferred linguistic choice—the division between ‘bordering on’ insolvency and ‘imminent insolvency’.¹⁵¹ On a superficial

¹⁴³ *ibid* [12] (Lord Reed).

¹⁴⁴ *Brady* (n 57) 40.

¹⁴⁵ *Sequana* (n 5) [50] (Lord Reed).

¹⁴⁶ *Kinsela* (n 70) 401.

¹⁴⁷ *Bilta (UK) Ltd v Nazir (No 2)* [2015] UKSC 23, [2016] AC 1 [123].

¹⁴⁸ *Brady* (n 57) 40.

¹⁴⁹ *Colin Gwyer & Associates* (n 138) [74] (Leslie Kosmin QC).

¹⁵⁰ *Re Loquinn Ltd* [2003] EWHC 999 (Ch), [2003] 2 BCLC 442 [240] (Etherton J).

¹⁵¹ Lord Reed, Lord Hodge, and Lady Arden preferred the former, while Lord Briggs and Lord Kitchin preferred the latter: see *Sequana* (n 5).

level, these might appear to be synonymous descriptions, but the two choices of words indicate different *states of time*. ‘Bordering on’ insolvency indicates oscillation between solvency and insolvency which, when viewed in the light of Lord Briggs’s ‘light at the end of the tunnel’, is a state of fluctuation and impermanence.¹⁵² This grey area between insolvency and solvency leaves room for uncertainty as to the end result. Contrastingly, ‘imminence’ carries with it more certainty in that it denotes that insolvency is about to happen, or is on the verge of happening, conveying immediacy and thus more certainty than the meaning of ‘bordering on’. Of course, subjective value judgments by judges mean that these two standards can be interpreted in other ways too, demonstrating just how *uncertain* the Supreme Court’s trigger test is.

Nevertheless, it is not the purpose of this article to consider the arguments pertaining to the trigger of the duty. In the author’s view, the question of ‘trigger’ is a secondary concern once a coherent conception of the content of the duty is acquired. However, as this article has argued, the content of a creditor duty is *de facto* incoherent as a matter of principle. The above analysis of the trigger test in *Sequana* is thus only engaged with to demonstrate the incoherent basis of the duty in the light of the ratification principle and section 214 of the IA 1986.

(i) Ratification Principle

Although fiduciary duties are strict, they are not *absolute*. The fiduciary can immunise themselves against liability by obtaining fully informed consent for their conduct.¹⁵³ Similarly, the shareholders can ratify the director’s breach, after full disclosure, by making the act the company’s *own* act through the law on agency.¹⁵⁴ It was common ground between counsel (and the Justices of the Supreme Court also adhered to this view) that the creditor duty cannot apply at the same time as the ratification principle does. Lord Briggs noted that the essence of the difference between the applicability of these two rules is centred around insolvency—the creditor duty is *engaged* by insolvency, whereas the ratification principle is *disapplied* by insolvency.¹⁵⁵ The significance of determining when the two rules are engaged is because, as per *Ciban Management Corpn v Citco (BVI) Ltd*, the shareholders cannot authorise or ratify a breach that would jeopardise the company’s solvency and cause loss to its creditors. Therefore, the application of the former will displace the latter.¹⁵⁶ However, there exists uncertainty as to the exact moment that this will occur. Peter Walton notes that whether a breach is ratifiable is a ‘binary decision’—this means, in simple terms, that a breach will either be ratifiable, when a company is solvent, or will not, when a company is insolvent.¹⁵⁷ Problematically, the concept of ‘bordering on’ or ‘imminent’ insolvency is that insolvency has not yet occurred, which means that the ratification principle is not technically disapplied yet. Therefore, there exists a moment when both the ratification principle and the creditor duty can apply at the

¹⁵² *ibid* [120], [164], [174]–[176].

¹⁵³ *Boardman v Phipps* [1967] 2 AC 46 (HL); *Regal (Hastings) Ltd v Gulliver* [1967] 2 AC 134 (HL).

¹⁵⁴ *Sequana* (n 5) [125] (Lord Briggs).

¹⁵⁵ *ibid*.

¹⁵⁶ See also *ibid* [91], where Lord Reed expressed that, ‘where the directors are under a duty to act in good faith in the interests of the creditors, the shareholders cannot authorise or ratify a transaction which is in breach of that duty’.

¹⁵⁷ Peter Walton, ‘Triggers, Content, and Enforcement: Directors’ Duties to Creditors – Where Are We After *Sequana*?’ (2024) 9 *Wolverhampton Law Journal* 5, 13.

same time. This, however, goes against previous authority such as *Ciban* whose correctness was not challenged in the Supreme Court.

Further, the idea that the duty is owed to the creditors through the medium of the company complicates the position even more. At the core of the ratification principle lies the equitable rule derived from trust law that ‘those to whom duties are owed may release those who owe the duties from their legal obligations... prospectively or retrospectively’.¹⁵⁸ This means that *if* the creditor duty applies when insolvency is bordering on or is imminent, then it is the creditors to whom the duty is really owed and who have the right to ratify any breaches. However, given that both the rule and the duty cannot apply at the same time, one has to give way to the other, and so creditors might be deprived of this right, going against the equitable principle.

(ii) Section 214 of the IA 1986

The Justices of the Supreme Court took the position that the creditor duty—or, as Lady Arden called it, the rule in *West Mercia*—works ‘harmoniously’ with section 214 of the IA 1986 on the basis that they ‘do not cover the same legal space’.¹⁵⁹ Lord Reed elaborated in depth on the differences between the two, noting that the points in time at which the relevant duties arise differ. There are three key distinct timing differences: first, the modified fiduciary duty is not subject to any limitation period; secondly, it applies *prior* to insolvency and is therefore not limited to commencement of insolvency proceedings; and thirdly, while the fiduciary duty and its subsequent modification applies at all times, the duty under section 214 only applies when a reasonably diligent and competent director would know that there was no reasonable prospect of avoiding insolvency proceedings. However, this position is not as clear as Lord Reed made it out to be. There can be a situation where liability arises for *both* the creditor duty and section 214 of the IA 1986, giving rise to potential double recovery. It is unfortunate that the Supreme Court did not consider this issue.

The effects can be seen in the recent *Wright v Chappell*¹⁶⁰ cases brought by the liquidators of the BHS group of companies against its former directors. The claims were brought on two grounds under the IA 1986: wrongful trading under section 214; and misfeasant trading under section 212. The latter includes ‘breach of any fiduciary or other duty in relation to the company’ and thus encompasses a failure to comply with the modified duty under section 172 of the CA 2006. Accordingly, the creditor duty principles from *Sequana* were applied. The timings for the application of these sections differ—however, only *marginally*. Section 214 of the IA 1986 is engaged when it is borne in mind by the directors that *liquidation* is inevitable, imposing personal liability if they do not minimise loss to creditors. The *onset of liquidation* is thus the key timing for a wrongful trading claim. Contrastingly, the modified creditor duty arises where the company is ‘insolvent’ or ‘bordering insolvency’—the timing thus being a much earlier one. On the facts, the directors were found to be personally liable at a time when the BHS companies were not yet cash flow insolvent, but insolvent administration *could* have been avoided—about six months earlier than the date on which the claim for wrongful trading was, in fact, established.

¹⁵⁸ Paul L. Davies, *Gower, Principles of Modern Company Law* (6th edn, Sweet & Maxwell 1997) 644; Jennifer Payne, ‘A Re-Examination of Ratification’ [1999] CLJ 604, 605–06.

¹⁵⁹ *Sequana* (n 5) [326].

¹⁶⁰ [2024] EWHC 1417 (Ch); [2024] EWHC 2166 (Ch), [2024] BCC 1343.

While the two duties clearly apply at different times, there is indeed overlap. The moment at which a company approaches insolvency—to the point that this insolvency is ‘bordering’ or is ‘imminent’—could also be the moment when the directors know or should have known that liquidation will be inevitable. Indeed, the approach for calculating the quantum of the claims set out by Leech J points towards significant overlap between the two claims—providing an easier alternative or back-up claim to the higher bar required for a wrongful trading claim. The approach adopted focused on the net deficiency caused by continued trading after the point of insolvency for wrongful trading, and the losses directly arising from breaches of fiduciary duties for misfeasance. The potential for double recovery arises because the nature of the relevant conduct amounting to wrongful trading can also constitute a breach of fiduciary duty, with there being no accompanying clear-cut approach to how losses are to be distinguished. Further, liability for breach of the creditor duty is arguably very similar to wrongful trading, raising questions over the principled and normative justifications for the Supreme Court’s acknowledgement of such a creditor duty, especially when a remedy under the general law already exists.¹⁶¹ The position, ultimately, remains unsatisfactory.

V. CONCLUSION: REFLECTION UPON THE UTILITY OF A CREDITORY DUTY AND ITS CONTENT

As a matter of principle, a creditor duty is contradictory to sanctity of contract and fiduciary law. Fairness considerations do not warrant *ex post* judicial interference with parties’ private bargains and the further expansion of fiduciary duty for creditors should be rejected with the *Sequana* approach abandoned. Professor Sealy accurately captured the position when he argued that ‘creditors deal with a company as a matter of *bargain*, not of trust, and bargain involves *risk*’.¹⁶² The Supreme Court sought too readily to gloss out the provisions in the CA 2006 to reverse parties’ risk allocation on normative conceptions of equality and fairness that should play no role in the distribution of the assets of an insolvent company. Problematically, this extension was done on the basis of a confused line of case law with clear flaws in judicial reasoning. The unfortunate result from *Sequana* is a vague obligation with a contradictory basis and a troubled interaction with other fundamental aspects of company law, such as shareholder primacy.

The subsequent question, however, is *what next?* Can the incoherence created be remedied the next time the Supreme Court is called to decide on the issue of creditor duty or should the matter be left for Parliament to reconsider? Section 172(3) of the CA 2006 does, after all, refer to a ‘rule of law’ and effect must be given to Parliament’s intent. Watts has argued that this ‘rule of law’ refers to the rules of capital maintenance invoked.¹⁶³ While this is an alternative option, this article takes the position that Parliament should reconsider section 172(3). Another possibility, arguably a better suited alternative to address directors’ risky business decisions with detrimental effect upon creditors, is to consider the development of an implied term in law. Such term would function as a negative covenant that deters directors from undertaking risky decisions that could prejudice the interests of creditors and other members of the company. The matter would thus largely turn on any counter evidence by directors that they did, in fact, consider those interests and that the economic decision

¹⁶¹ *Sequana* (n 5) [328].

¹⁶² Sealy (n 4) 176 (emphasis added).

¹⁶³ Watts (n 114).

undertaken was justifiable on a proportionality basis. While developing this idea is not within the scope of this article to explore, it was the aim of this article to re-trigger interest in the basis of the creditor duty post-*Sequana* and to raise legitimate questions about its principled basis—or rather, its lack thereof.

Repairing the English Civil Law of Bribery: Fixing *Johnson v FirstRand Bank Ltd* [2024] EWCA Civ 1282

RIVU CHOWDHURY* AND GEORGE BEGLAN**

ABSTRACT

The recent Court of Appeal decision of *Johnson v FirstRand Bank Ltd* (*FirstRand*) held that a car dealer acting as a broker between a consumer and a lender for car finance will be under a ‘disinterested duty’ and a fiduciary duty. If the lender has paid a ‘secret’ or partially disclosed commission to the dealer, the dealer will be liable under the tort of bribery or for a breach of fiduciary duty. The lender will also be liable primarily under the tort of bribery or as an accessory to a breach of fiduciary duty. This result has shocked the car finance industry, lenders, and banks. It is submitted that there are issues in the reasoning of *FirstRand* and that this is because of deficiencies in the law of bribery itself. Conflicting authorities and incoherent differences between the common law and equitable action have resulted in a widening of civil liability to an unjustified degree. We examine the history of the law of bribery and the reasoning in *FirstRand*, and we suggest more careful analysis of fiduciary duties and more careful application of accessory liability. We also propose subsuming the common law action into the equitable one. This reform, alongside improved legal analysis of the duties and accessory liability, is more in line with the economic realities of brokerage and is better for consumers as it reduces the likelihood of price increases and facilitates the availability of finance. It is also more desirable from the perspectives of coherence between common law and equity, as the law should not have two answers to one question.

Keywords: law of bribery, equity, tort, fiduciary duties, lending

I. INTRODUCTION

The English law of bribery is complex and wide. The payment and receipt of a bribe will attract criminal liability at common law and under statute.¹ It will also attract civil liability under

* LLM Candidate, University College London, BA (Oxon). We are grateful to Ruby Sanders-English, Eben MacDonald, and an anonymous reviewer for their comments on our earlier drafts. Any errors that remain are our own.

** LLM (Durham), BA (Oxon).

¹ See Bribery Act 2010. For a detailed treatment, see also Richard Lissack and Fiona Horlick, *Lissack and Horlick on Bribery and Corruption* (3rd edn, LexisNexis 2020). The criminal law of bribery is complex and beyond the scope of this article.

both the tort of bribery² and as a breach of fiduciary duty.³ Actions are available against both bribe-givers⁴ and bribe-takers,⁵ and bribes include secret commissions.⁶ Secret commissions are unauthorised payments that an agent or fiduciary receives from a third party. The issue of civil liability arose recently when the Court of Appeal handed down its decision in three combined appeals in the case of *Johnson v FirstRand Bank Ltd* ('*FirstRand*').⁷ The case concerned secret and half-secret commissions given by lenders to car dealers who acted as credit brokers for consumers. It was held that the brokers had breached their 'disinterested' duty and fiduciary duty owed to the consumers and that, essentially, customers were mis-sold finance. The case has sent shockwaves through the motor finance industry, impacting brokerages,⁸ car manufacturers,⁹ and banks, and encouraging consumer claims.¹⁰ The cause of concern is that, in the short term, both car dealers and lenders will be exposed to significant liability (Lloyd's Banking Group have set aside £1.1 billion so far to cover claims¹¹) and that, in the long term, car finance will become more expensive. The impact of the case is not limited to the motor finance industry and applies to all credit brokers as most brokers operate on a commission model which is usually secret or half-secret.¹² Given the wide and significant impact of the case, the UK Supreme Court has granted permission to appeal,¹³ and the case and the law of bribery are ripe for review.

While the result of the case is defensible, the reasoning is, in the authors' view, deficient. We argue that the issues in the case arise from confusion and difficulties between the tort of bribery and a breach of fiduciary duty, conflicting authorities, and the lack of explanation as to why a fiduciary duty and disinterested duty are imposed. We propose that the common law action should be subsumed into the equitable one and that the disinterested duty can and should be viewed as a fiduciary duty. We also put forward a better explanation of why a fiduciary can be imposed in the case itself. This is a better result as a matter of principle,

² *T Mahesan v Malaysia Government Officers' Co-Operative Housing Society Ltd* [1979] AC 374 (PC) 383. See Section II(B)(ii) below.

³ *FHR European Ventures LLP v Cedar Capital Partners LLC* [2014] UKSC 45, [2015] AC 520 [7]. See Section II(A) below.

⁴ *Mahesan* (n 2) (common law); *FHR European Ventures* (n 3) (equity).

⁵ See for example *Johnson v FirstRand Bank Ltd* [2024] EWCA Civ 1282 [124] (equity); *Group Seven Ltd v Nasir* [2019] EWCA Civ 614, [2020] Ch 129 [110] (equity); *Mahesan* (n 2) (common law). See Section II below for a detailed account.

⁶ *Hurstanger Ltd v Wilson* [2007] EWCA Civ 299, [2007] 1 WLR 2351 [39]; *Industries and General Mortgage Co Ltd v Lewis* [1949] 2 All ER 573 (KB) 575 (Slade J) ('bribe means the payment of a secret commission').

⁷ *FirstRand* (n 5).

⁸ See Jack Williams, 'Close Brothers Stops Underwriting New Dealer Finance after Landmark Court Ruling' (*Car Dealer*, 25 October 2024) <www.cardealermagazine.co.uk/publish/close-brothers-stops-underwriting-new-dealer-finance-after-landmark-court-ruling/309297> accessed 15 December 2024.

⁹ Michael Bow and Matt Oliver, 'Car Deliveries Halted amid Fears Motor Finance Scandal Is "Bigger than PPI"' *The Telegraph* (London, 31 October 2024) <www.telegraph.co.uk/business/2024/10/31/car-deliveries-halted-amid-fears-motor-finance-scandal/> accessed 15 December 2024.

¹⁰ See for example Stephen Fairclough, 'Motorists Who Bought Cars on Finance Could Share in Billions' (*BBC News*, 11 November 2024) <www.bbc.co.uk/news/articles/cg7dy50p6vo> accessed 15 December 2024; Kalyeena Makortoff, 'Claims Management Companies Circle after UK Court Ruling on Mis-sold Car Finance' *The Guardian* (London, 9 December 2024) <www.theguardian.com/business/2024/dec/09/car-finance-scandal-uk-claims-management-consumer-compensation> accessed 15 December 2024.

¹¹ Ben Martin, 'Lloyds Banking's Bill for Possible Car Loan Misselling Rises to £1.1bn' *The Times* (London, 20 February 2025) <www.thetimes.com/business-money/companies/article/lloyds-bankings-bill-for-possible-car-loan-misselling-rises-to-11bn-wblygsv6h> accessed 11 February 2025.

¹² Larry Harris, *Trading and Exchanges: Market Microstructure for Practitioners* (OUP 2003) ch 5.

¹³ 'Announcement from the UK Supreme Court' (*UK Supreme Court*, 11 December 2024) <<https://www.supremecourt.uk/news/uksc-announcement-1/>> accessed 15 December 2024. See *Wrench v FirstRand Bank* (UKSC 2024/0159); *Johnson v FirstRand Bank* (UKSC 2024/0158); *Hopcraft v Close Brothers* (UKSC 2024/0157).

as the law should not have two answers to the same question on liability for bribes. Equity also has stricter requirements for accessory liability, which limits exposure to lenders. Both the result and the ultimate reasoning of the case can, therefore, be rescued. A separate issue is whether it should be rescued at all. We argue that it should, as an ad hoc fiduciary duty will only arise where parties have made consensual undertakings to that effect.

This article will, first, give a background on the law of secret commissions and bribery; secondly, it will detail and explain the reasoning in *FirstRand*; thirdly, it will highlight the flaws in the reasoning of the case; fourthly, it will offer suggestions on how to reform the law of bribery and provide a justification for the imposition of a fiduciary duty; and finally, it will discuss whether a fiduciary duty should be imposed as a matter of policy.

II. THE LAW OF SECRET COMMISSIONS AND BRIBERY

There are actions for bribes and secret commissions at common law¹⁴ and in equity.¹⁵ A bribe consists of any ‘commission or other inducement, which is given by a third party to an agent [or other fiduciary] and which is secret from his principal’.¹⁶ A secret commission is best viewed as part of a fiduciary’s ‘no conflict’ and ‘no profit’ rule.¹⁷ By accepting the secret commission, the fiduciary has placed themselves in a position that conflicts with their fiduciary duty or has made unauthorised profit as a result of their fiduciary position which is why they are liable to account to their principal. Secret commissions thus fall under the definition of bribes, and the terms are often used interchangeably.¹⁸ We will use the term ‘bribe’ as including, and as a shorthand for, secret commissions. We will use the following terminology: the payer is the person paying the bribe; the payee is the recipient; and the principal is the person to whom the payee owes a duty.

A. IN EQUITY: BREACH OF FIDUCIARY DUTY

The origin of civil actions for bribery lies in equity.¹⁹ When a secret commission is paid contrary to a fiduciary duty, the payee is liable for breach of fiduciary duty,²⁰ and the payer is liable in dishonest assistance.²¹ This is because the payment of a commission generates a conflict of interest between the payee and the principal.²² This conflict of interest can be overcome by consent, so the secrecy renders the commission a breach of duty. *Shipway v Broadwood* sets out this justification most clearly: ‘the real evil [of bribery] is not the payment

¹⁴ *Mahesan* (n 2) 382–83; *Hovenden & Sons v Millhof* (1900) 83 LT 41.

¹⁵ *Attorney General for Hong Kong v Reid* [1994] 1 AC 324 (PC) 330; *Reading v Attorney-General* [1951] AC 507 (HL).

¹⁶ *Anangel Atlas Compania Naviera SA v Ishikawajima-Harima Heavy Industries Co Ltd* [1990] 1 Lloyd’s Rep 167 (QB) 171 (Leggatt J); *Industries and General Mortgage* (n 6).

¹⁷ *FHR European Ventures* (n 3) [5] (Lord Neuberger); *Regal (Hastings) Ltd v Gulliver* [1967] 2 AC 134 (HL) 144G–145A (Lord Russell). See also *Boardman v Phipps* [1967] 2 AC 46 (HL); *Murad v Al-Saraj* [2005] EWCA Civ 959.

¹⁸ See for example *Industries and General Mortgage* (n 6); *FHR European Ventures* (n 3).

¹⁹ *Mahesan* (n 2) 380; *Wood v Commercial First Business Ltd* [2021] EWCA Civ 471, [2022] Ch 123 [96]. The earlier actions in equity arose as a breach of fiduciary duty. See Derek Whayman, ‘Liability for Bribes and Secret Commissions at Common Law: Obsolete, Unnecessary and Probably a Fusion Fallacy’ (2022) 86 *Conveyancer and Property Lawyer* 184, 190–01 for a fuller account and examples of earlier cases in equity.

²⁰ See for example *Regal* (n 17); *Boardman* (n 17); *Murad* (n 17).

²¹ *Hurstanger* (n 6) [35]; *Novoship (UK) Ltd v Mikhaylyuk* [2014] EWCA Civ 908, [2015] QB 499 [67], [93]; *Group Seven* (n 5); *FirstRand* (n 5) [124]–[142].

²² See for example *FHR European Ventures* (n 3) [5]; *Fiona Trust & Holding Corporation v Privalov* [2010] EWHC 3199 (Comm) [73]; *Novoship* (n 21) [106].

of money, but the secrecy attending it.²² A bribe is thus a helpful shorthand for a secret commission. The remedies available against a payee are the same for any breach of fiduciary duty, including an account of profits,²³ equitable compensation,²⁴ a proprietary claim,²⁵ and rescission.²⁷ The remedies available against a payer include an account of profits, but as a dishonest assister, payers are only liable for their own profits.²⁶ This will usually not include the bribe, as they have paid this away and have not gained it.

Moreover, half-secret commissions can attract liability, as established by the Court of Appeal in *Hurstanger Ltd v Wilson*.²⁹ The case involved a lender's payment of a commission to a broker, and the document signed by the borrowers included a statement that the lender paid a commission to its brokers. However, the quantum of the commission was not disclosed, and this amounted to a breach of fiduciary duty. This disclosure was a 'half-way house',³⁰ which was enough to negate secrecy but insufficient to obtain informed consent. For half-secret commissions, the 'full armoury of remedies' is not available.³¹ A principal could seek rescission, but not as of right. Rescission would be discretionary and could not be elected for.³² The remaining remedies, including an account of profits, equitable compensation, and a proprietary remedy, are still available.³³ While secret commissions will have concurrent liability for breach of fiduciary duty,³⁴ only an equitable action will be available for half-secret ones.³⁵

So, in equity, the payee must owe a fiduciary duty to the principal and disclosure of the commission must be insufficient to obtain informed consent to the potential conflict of interest. The payer can be liable as an accessory to the breach of fiduciary duty. While primary liability will arise for secret commissions at common law,³⁶ the payee must have dishonestly assisted the breach to be liable as an accessory for either a secret or half-secret commission in equity.³⁷

B. AT COMMON LAW: TORT OF BRIBERY

(i) History and Development

The first instances of common law actions occurred in 1808 and 1811 with *Thompson v Havelock* ('*Havelock*')³⁸ and *Diplock v Blackburn*.³⁹ Derek Whayman has convincingly

²² *Shipway v Broadwood* [1899] 1 QB 369 (CA) 373 (Chitty LJ), quoted in *FirstRand* (n 5) [59].

²³ See for example *Novoship* (n 21) [119]. An account of profits will be available in principle but will be discretionary.

²⁴ See for example *Mahesan* (n 2); *Swindle v Harrison* [1997] 4 All ER 705 (CA).

²⁵ The payee holds the bribe or its traceable proceeds on constructive trust: *FHR European Ventures* (n 3).

²⁶ See for example *Reading* (n 15); *Mahesan* (n 2) 380; *Wood* (n 19) [98].

²⁷ *Ultraframe (UK) Ltd v Fielding (No 2)* [2005] EWHC 1638 (Ch), [2006] FSR 17 [1589]–[1594]; *Novoship* (n 21) [84].

²⁸ *Hurstanger* (n 6).

²⁹ *ibid* [45] (Tuckey LJ).

³⁰ *ibid* [46]; *Panama and South Pacific Telegraph Company v India Rubber, Gutta Percha, and Telegraph Works Company* (1874–75) LR 10 Ch App 515 (DC) 526.

³¹ *Hurstanger* (n 6) [48]; *Spence v Crawford* [1939] 3 All ER 271 (HL) 288; *Johnson v EBS Pensioner Trustees Ltd* [2002] EWCA Civ 164 (CA) [78]–[79].

³² *Hurstanger* (n 6) [49].

³³ *Wood* (n 19) [100], [125]; *FirstRand* (n 5) [18].

³⁴ *Hurstanger* (n 6) [46].

³⁵ *FirstRand* (n 5) [18].

³⁶ *ibid*.

³⁷ (1808) 170 ER 1045.

³⁸ (1811) 170 ER 1300.

argued that these cases were really an application of equitable principles of fiduciary duties in the common law courts using a plea replication mechanism.⁴⁰ The plea replication mechanism permitted equitable defences, but not claims, in common law courts.⁴¹ *Havelock* uses the language of fiduciary duties ('[n]o man should... have an interest against his duty'⁴²) and *Diplock* states that the commission 'belonged to the owner',⁴³ which was a result of imposing a constructive trust. The plea replication mechanism was needed due to the deficiencies in the Court of Chancery, including delays arising from a backlog, a reluctance at the time to award an account of profits, and the inability to hear live evidence.⁴⁴ This is similar to the case of *Taylor v Plumer*,⁴⁵ a case on tracing heard in the common law courts. Unlike the tracing cases, the bribery cases were later integrated into the common law action or tort of bribery, instead of viewing them as cases dealing with equitable principles. For example, in the 1874 case of *Morison v Thompson*,⁴⁶ the plea replication mechanism could not be used, yet the plaintiffs still argued a breach of fiduciary duty in the common law courts. The case also treated *Diplock* and *Havelock* as common law authorities, as well as master and servant cases. Cases in equity were cited separately and were seen as a different category. Since then, there has been a common law action for bribery. This amounts to a 'fusion fallacy',⁴⁷ where the courts have taken substantive equitable rules and applied them to a common law claim.⁴⁸

The issue is not that a secret commission or bribe can result in an action under the common law but rather that an action can arise in both equity and common law. This can, and has, led to differences in the claims and incoherence as a result of diverging rules.

(ii) The Modern Law

The term 'tort of bribery' first appeared in the case of *Fiona Trust & Holding Corporation v Privalov* but is not universally used.⁴⁹ The common law action has been described as a 'tort of fraud'⁵⁰ and as 'sui generis and [defying] classification'.⁵¹ We will use the term 'tort of bribery' as a convenient, albeit contested, term. *Industries and General Mortgage Co Ltd v Lewis* sets out the classic test of bribery:

⁴⁰ Whayman (n 19) 186-91.

⁴¹ Common Law Procedure Act 1854, s 85: 'The Plaintiff may reply, in answer to any Plea of the Defendant, Facts which avoid such Plea upon equitable Grounds; provided that such Replication shall begin with the Words "For Replication on equitable Grounds," or Words to the like Effect.'

⁴² *Havelock* (n 38) 1046.

⁴³ *Diplock* (n 39) 1300.

⁴⁴ Whayman (n 19) 187.

⁴⁵ (1815) 105 ER 721. See Lionel D Smith, 'Tracing in *Taylor v. Plumer*: Equity in the Court of King's Bench' [1995] LMCLQ 240.

⁴⁶ (1874) 9 LR 9 QB 480.

⁴⁷ The term 'fusion fallacy' refers to the argument that some substantive rules have been changed as a result of the Supreme Judicature Act 1873. It is not to say that rules cannot be changed by developing case law, but is a result of the confusion caused by assuming that the Act fused substantive rules rather than fusing the administration of the courts. The term was coined by Australian authors but applies equally to English law: see JD Heydon, MJ Leeming and PG Turner, *Meagher, Gummow and Lehane's Equity: Doctrine and Remedies* (5th edn, LexisNexis 2014) paras 2-135, 2-140.

⁴⁸ See Whayman (n 19) 190 for the full argument.

⁴⁹ The term first appeared in *Fiona Trust* (n 22) [177] (Andrew Smith J). See also *Otkritie International Investment Management Ltd v Urumov* [2014] EWHC 191 [66]-[73]; *Mototrak Ltd v FCA Australia Pty Ltd* [2018] EWHC 1464 (Comm) [15]; *FM Capital Partners Ltd v Marino* [2018] EWHC 1768 (Comm) [550]. The term is not used in *FirstRand* (n 5), *Wood* (n 19), or *Hurstanger* (n 6).

⁵⁰ *Mahesan* (n 2) 382 (Lord Diplock). However, it is distinct from the tort of deceit and does not depend on any form of representation: *Petrotrade Inc v Smith* [2000] 1 Lloyd's Rep 486 ff.

⁵¹ *Mahesan* (n 2) (Lord Diplock).

For the purposes of the civil law a bribe means the payment of a secret commission, which only means (i) that the person making the payment makes it to the agent of the other person with whom he is dealing; (ii) that he makes it to that person knowing that that person is acting as the agent of the other person with whom he is dealing; and (iii) that he fails to disclose to the other person with whom he is dealing that he has made that payment to the person whom he knows to be the other person's agent.⁵²

The tort of bribery is similar to the equitable claim in that the payer pays a bribe. The payee owes some form of duty to the principal and the principal does not have knowledge of the bribe.

The key difference is the duty. The payee must be 'under a duty to provide information, advice or recommendation on an impartial or disinterested basis'.⁵³ The authority for this 'disinterested duty' is *Wood v Commercial First Business Ltd*.⁵⁴ There is a long line of cases describing the duty as being 'disinterested'.⁵⁵ *Wood* formulates this into a disinterested duty which is distinct from a fiduciary duty.⁵⁶ To be liable at common law for bribery, all that is needed is a disinterested duty, not a fiduciary one. There will still usually be a fiduciary duty as well, as there was in *Wood*.⁵⁷ Both a breach of a disinterested duty and a breach of fiduciary duty are sufficient to attract liability for a payee for a secret commission.⁵⁸ They will arise concurrently. Half-secret commissions can be a breach of fiduciary duty and will not attract liability at common law. Whether it is coherent to categorise a disinterested duty as not being fiduciary will be challenged in Section IV, as we argue that a disinterested duty is just a fiduciary duty of limited scope.

Another difference is that the payer is also liable for the tort of bribery.⁵⁹ As the action is available against both the payer and payee, it is perhaps inaccurate to call it a single tort of bribery, but it is well-established that both are primarily liable at common law.⁶⁰ What is required is knowledge that the payee is an agent, fiduciary, or under a disinterested duty.⁶¹

The remedies available will be compensation for loss suffered by the principal, liability for which is joint and several.⁶² This will typically be the value of the bribe⁶³ or losses due to entry into a transaction that was caused by the bribe. There will also be a restitutionary

⁵² *Industries and General Mortgage* (n 6).

⁵³ *Wood* (n 19) [48] (David Richards LJ).

⁵⁴ *ibid*.

⁵⁵ See for example *Panama and South Pacific Telegraph* (n 31) 528–29 (Mellish LJ); *Logicrose Ltd v Southend United Football Club Ltd* [1988] 1 WLR 1256 (Ch) 1260–61 (Millett J).

⁵⁶ *Wood* (n 19) [48]–[49].

⁵⁷ *ibid* [110].

⁵⁸ *ibid* [128].

⁵⁹ *Mahesan* (n 2).

⁶⁰ *FirstRand* (n 5) [18], [77].

⁶¹ *Industries and General Mortgage* (n 6).

⁶² *Mahesan* (n 2).

⁶³ This is often assumed: see for example *ibid* 381; *The Mayor, Aldermen, and Burgesses of the Borough of Salford v Lever* [1891] 1 QB 168 (CA) 176–77; *Grant v The Gold Exploration and Development Syndicate Ltd* [1900] 1 QB 233 (CA) 244.

remedy for the value of the bribe.⁶⁴ This is referred to as an action for ‘money had and received’ in *Mahesan* and other common law cases.⁶⁵ This can be either as a restitution for wrong or an account as against the payee, and a change of position defence will not be available.⁶⁶ Bribery itself is a ground for rescission at common law,⁶⁷ and rescission will also be available as an absolute right subject to the bars.⁶⁸

C. DIFFERENCES BETWEEN THE ACTIONS

At common law, the payer is subject to primary liability under the same cause of action as the payee.⁶⁹ This is distinct from equity, where accessory liability for breach of fiduciary duty must be proved, which means establishing a claim for dishonest assistance which is more onerous. There is also no need to prove a fiduciary duty and, as a tort, remedies (in particular rescission) are available as a right.

However, half-secret commissions do not fall within the tort, as discussed above, and only attract liability as a breach of fiduciary duty.⁷⁰ A proprietary remedy is likely unavailable as their availability arises from the imposition of a constructive trust owing to a breach of fiduciary duty.⁷¹ *FHR European Ventures LLP v Cedar Capital Partners LLC*, the leading authority on proprietary remedies for secret commissions, explicitly reasons using fiduciary duties and does not mention any common law cases.⁷²

III. THE DECISION OF *JOHNSON V FIRSTRAND BANK LTD*

When a consumer goes to buy a second-hand car, the dealer will often offer to act as a credit broker and arrange financing to help them buy the car. The dealer will usually receive a commission from the lender for this. The issue is where the commission has been kept secret from the consumer or is hidden away in the terms and conditions. While the dealer’s profit on the sale itself would be obvious, a commission from the lender would likely be a surprise. The customer may not have agreed to the finance, having this knowledge.

The facts in the three appeals follow the general fact pattern set out above. In all three, the claimants were ‘financially unsophisticated consumers on relatively low incomes’,⁷³ who bought cars on finance brokered by the car dealers. In *Hopcraft v Close Brothers Ltd*, there was no disclosure whatsoever and it was secret.⁷⁴ In *Wrench v Firstrand Bank Ltd*, it was partially disclosed in the lender’s standard terms and conditions that a commission ‘may be payable by us to the broker who introduced this transaction to us. The amount is available from the Broker on request’.⁷⁵ However, the statement was so ‘buried’ in the lender’s standard

⁶⁴ *Mahesan* (n 2).

⁶⁵ *ibid*; *Boston Deep Sea Fishing and Ice Co v Ansell* [1886–90] All ER Rep 65 (CA) 75 (Bowen LJ); *Salford* (n 63) 176 (Lord Esher MR).

⁶⁶ *FM Capital Partners Ltd v Marino* [2020] EWCA Civ 245, [2021] QB 1.

⁶⁷ *Smith v Sorby* (1875) 3 QBD 552 (QB); *Wood* (n 19) [99].

⁶⁸ *Conway v Eze* [2018] EWHC 29 (Ch) [143]–[156].

⁶⁹ *Wood* (n 19) [94].

⁷⁰ *Hurstanger* (n 6).

⁷¹ *FHR European Ventures* (n 3) [46], [47]–[48].

⁷² *ibid* [47]–[48].

⁷³ *FirstRand* (n 5) [6] (Andrews, Birss and Edis LJ).

⁷⁴ *ibid* [27].

⁷⁵ *ibid* [34], referring to cl 12.6.

terms and conditions that it was insufficient to negate secrecy and was a secret commission.⁷⁶ In *FirstRand*, it was the same as *Wrench*, but the dealer also provided a ‘Suitability Document’ which stated that ‘we may receive a commission from the product provider’.⁷⁷ It was therefore treated as a half-secret commission.⁷⁸

Andrews LJ, Birss LJ and Edis LJ held in a joint judgment that the dealers acted as both sellers of the cars and credit brokers on behalf of the claimants. In all three cases, the dealers owed both a ‘disinterested duty’, described in *Wood*,⁷⁹ and a fiduciary duty to the principals in their capacity as credit brokers.⁸⁰ In *Hopcraft* and *Wrench*, the commissions were fully secret, so the payees were liable for breach of a disinterested duty which was sufficient to give rise to a primary liability of the lenders.⁸¹ The fiduciary duty owed was parallel.⁸² In *FirstRand*, the commission was half-secret,⁸³ so only a breach of fiduciary duty would suffice to make the payee liable,⁸⁴ and a claim for accessory liability would have to be sought against the lenders.⁸⁵ Both claims were successful. The claimant in *FirstRand* did not need to bring specific evidence of dishonesty of the payer; knowledge of the agency relationship and thus a fiduciary duty was enough.⁸⁶

Therefore, in fully secret commissions where there has been no disclosure, the payee must owe either a duty to provide information, advice, and recommendation on an impartial and disinterested basis (disinterested duty), or a fiduciary duty.⁸⁷ The payer of the commission will be primarily liable for bribery where there is a breach of disinterested duty.⁸⁸ Disgorgement of the secret commission and rescission will be available as of right.⁸⁹ Accessory liability of the payer must be established for breach of fiduciary duty. In half-secret or partially disclosed (using the language of *FirstRand*) commissions, the payee must owe a fiduciary duty to the principal and disclosure must be insufficient to obtain informed consent to the potential conflict of interest.⁹⁰ The payee must owe a fiduciary duty for the payer to attract accessory liability for dishonest assistance.⁹¹ The payer must have knowledge that the broker was a fiduciary.⁹²

There was a further ground of appeal for a claim under sections 140A–C of the Consumer Credit Act 1974. The court found that the relationship between a lender and consumer was unfair under the Act. This claim is beyond the scope of this article.

⁷⁶ *ibid* [119].

⁷⁷ *ibid* [8].

⁷⁸ *ibid* [120].

⁷⁹ *ibid* [87]–[88].

⁸⁰ *ibid* [91]–[92], [173].

⁸¹ *ibid* [173].

⁸² *ibid*.

⁸³ *ibid* [112].

⁸⁴ *ibid* [77].

⁸⁵ *ibid* [124], [137].

⁸⁶ *ibid* [133].

⁸⁷ *ibid* [77].

⁸⁸ *ibid* [81].

⁸⁹ *ibid* [77].

⁹⁰ *ibid* [60].

⁹¹ *ibid* [80].

⁹² *ibid* [133].

IV. DEFICIENCIES IN THE REASONING

A. WHY IS THERE A DISINTERESTED DUTY AND FIDUCIARY DUTY?

The first key issue is why the disinterested duty and fiduciary duty are imposed upon the dealer. The issue is not necessarily the imposition of these duties, but rather the lack of convincing reasoning and the inadequate normative explanations that are provided. We will begin by dealing with when a fiduciary duty does arise.

Fiduciary duties arise in settled categories of relationships without further inquiry, such as for trustees and beneficiaries,⁹³ agents and principals,⁹⁴ solicitor and clients,⁹⁵ partners in a partnership, and directors and their company.⁹⁶ English law also recognises ad hoc fiduciary relationships, which arise by undertaking, as opposed to status. As Millett LJ states in *Bristol and West Building Society v Mothew*, a person ‘is not subject to fiduciary obligations because he is a fiduciary; it is because he is subject to them that he is a fiduciary’.⁹⁷ This position is the same in Australia⁹⁸ and Canada.⁹⁹

Millet LJ defines a fiduciary as ‘someone who has undertaken to act for or on behalf of another in a particular matter in circumstances which give rise to a relationship of trust and confidence’.¹⁰⁰ This has been cited with approval in the Supreme Court in *FHR European Ventures*¹⁰¹ and is acknowledged in *Sheikh Tahnoon Bin Saeed Bin Shakhboot Al Nehayan v Kent* (*‘Al Nehayan’*)¹⁰² by Leggatt J who, approving Mason J’s approach in *Hospital Products Ltd v United States Surgical Corporation*,¹⁰³ states that fiduciary duties ‘arise where one person undertakes and is entrusted with authority to manage the property or affairs of another and to make discretionary decisions on behalf of that person’.¹⁰⁴ When considering if a fiduciary duty is imposed, the duty we are looking at is the duty of undivided loyalty, meaning that you are guided solely by the interests of the principal and not the fiduciary’s own interests. This is because the duty of undivided loyalty gives rise to the rules that justify liability for bribes.

The Court in *FirstRand* states that, like the case of *McWilliam v Norton Finance (UK) Ltd*,¹⁰⁵ ‘[t]he very nature of the duties which the credit broker undertook gave rise to a “disinterested duty”’.¹⁰⁶ It is not explained why undertaking duties as a credit broker gives rise to this duty. Moreover, the Court states that there is an ‘ad hoc fiduciary duty... arising from the nature of the relationship, the tasks with which the brokers were entrusted, and the obligation of loyalty which is inherent in the disinterested duty’.¹⁰⁷

⁹³ *Keech v Sandford* (1726) 25 ER 223.

⁹⁴ *Kelly v Cooper* [1993] AC 205 (PC).

⁹⁵ *Bristol and West Building Society v Mothew* [1997] 2 WLR 436 (CA).

⁹⁶ *Guinness Plc v Saunders* [1990] 2 AC 663 (HL).

⁹⁷ *Bristol and West Building Society* (n 95) 18 (Millett LJ), quoting the classic formulation in PD Finn, *Fiduciary Obligations* (Law Book Co 1977) 2.

⁹⁸ *Hospital Products Ltd v United States Surgical Corporation* [1984] HCA 64, (1984) 156 CLR 41 (HCA) 96–97 (Mason J).

⁹⁹ *Galambos v Perez*, 2009 SCC 48, [2009] 3 SCR 247 [37].

¹⁰⁰ *Bristol and West Building Society* (n 95) 18.

¹⁰¹ *FHR European Ventures* (n 3) [5].

¹⁰² [2018] EWHC 333 (Comm), [2018] 1 CLC 216.

¹⁰³ *Hospital Products* (n 98).

¹⁰⁴ *Al Nehayan* (n 102) [159].

¹⁰⁵ [2015] EWCA Civ 186, [2015] 1 All ER (Comm) 1026.

¹⁰⁶ *FirstRand* (n 5) [87].

¹⁰⁷ *ibid* [91] (Andrews, Birss and Edis LJ).

The underlying transaction in these cases is that of a sale, and the primary relationship is that of a buyer and seller. This is true even if the car is sold immediately or on credit, and it is still true if a third party lender is introduced to the transaction. The dealer is not necessarily subordinating their interest to that of the customer. The idea that the duties of a credit broker can justify imposing a disinterested interest are simply not true. There must be something more to justify imposing a duty.

Compare a *Quistclose* trust, for example. It is often said that it arises when money is transferred for a specific purpose,¹⁰⁸ most commonly in a contract for a loan. But most loans are under a contractual relationship and you would expect that all the rights and relationships of parties are in the contract. If one of them has bargained for a security, it must be in an express term in a contract; there must be something more. Yet a trust arises in a contract with no mention of a security, property interest, or trust. That something more is necessarily objective intention that the money be used for that purpose and *no other purpose*, that it is not at the free disposal of the parties,¹⁰⁹ and is of a fiduciary nature.¹¹⁰ Further, if the money is segregated, that is an indicative, but not necessary, factor.¹¹¹ There is a presumption to overcome.

The same is true of a contract for sale or hire-purchase agreement. It is a consensual, contractual relationship and you would expect the rights, relationships, and duties of each party to the contract to be contained within the contract. A fiduciary duty is consensual. When a fiduciary duty falls into one of the established categories, such as that of a trustee, agreeing to take on that role means that you have agreed to be a fiduciary. Likewise, an ad hoc fiduciary duty must be consensually agreed to. If the car dealers had intended to act as agents or fiduciaries when brokering finance, we would expect them to enter into an agency agreement. Any undertaking to become a fiduciary must amount to subordinating their interest to that of the customer and must be substantial enough to overturn the presumption of sale. Equity's intervention in this commercial transaction must be justified by strong indications of consent. Equity's role here is not to intervene to protect the vulnerable, a role being fulfilled by consumer law, including the unfair relationship claim under sections 140A-C of the Consumer Credit Act 1974, which is being argued in this very case.

FirstRand seems to look at this backwards. It states that, 'unless the broker made it clear to the consumer that they could not act impartially because they had a financial incentive [from the lenders]',¹¹² this gave rise to a disinterested duty. The Court treats credit brokerage as having a presumption of fiduciary duties when they should be doing the opposite and trying to overcome the presumption that it is a sale. One party's role as a credit broker cannot and should not without more impose fiduciary or disinterested duties; the evidence must demonstrate that such duties have been undertaken or agreed to. We must look elsewhere for a justification in the case.

The Court in *FirstRand* also looked at the relationship between the brokers and the buyers. The claimants were 'more vulnerable than someone who might have had the choice to pay in cash' and they all 'relied on the dealer to find them an offer which met their needs'.¹¹³ The issue with this reasoning is that an undertaking justifies a fiduciary duty, not the nature of the relationship itself. The vulnerability of the claimants means that they expect a relationship

¹⁰⁸ *Carreras Rothmans Ltd v Freeman Mathews Treasure Ltd* [1985] Ch 207 (Ch) 222B.

¹⁰⁹ *Twinssectra Ltd v Yardley* [2002] UKHL 12, [2002] 2 AC 164 [74].

¹¹⁰ *ibid* [76].

¹¹¹ *Re Kayford Ltd* [1975] 1 WLR 279 (Ch) 282.

¹¹² *FirstRand* (n 5) [87] (Andrews, Birss and Edis LJ).

¹¹³ *ibid* [91].

of ‘trust and confidence’, but the ‘trust and confidence’ that Millett LJ speaks of is a consequence of an undertaking, not a prerequisite. Leggatt LJ in *Al Nehayan* reminds us that there are many commercial transactions that create trust and confidence, but which are not fiduciary duties. In *Re Goldcorp Exchange Ltd*,¹¹⁴ the customers trusted a company to do what it promised and keep gold bullion safe for them, but the Privy Council did not find them to be fiduciaries. Any undertaking must create a legitimate expectation of a relationship, not just of trust and confidence, but one where the dealers would put aside their own self-interest. This is not to deny the importance of how vulnerable the claimants are, but it is to say that the Court’s arguments are insufficient.

The issue in *FirstRand* was that the Court operated from a presumption that acting as a credit broker would impose a disinterested duty. They are not treating credit brokers as status-based fiduciaries and they are not arguing to that effect. But they discuss the brokers being ‘in a position to take advantage of their vulnerable customers’ and mention that there was a ‘reasonable and understandable expectation that they would act in their best interests’,¹¹⁵ without discussing whether the brokers had undertaken that expectation or whether it had simply arisen from the buyer’s perspective. The only thing that the brokers undertook, as the court discusses, was to act as credit brokers.¹¹⁶ The court is therefore presuming that the undertaking to act as a credit broker is sufficient to be an undertaking to act as a fiduciary. This presumption essentially conflates status-based and ad hoc fiduciaries. Once we reverse the presumption, given that the transaction is not one that is typically fiduciary, we can see that the Court has insufficiently justified the imposition of fiduciary duties. The vulnerability of the claimants is also insufficient to justify the duties.

B. ACCESSORY LIABILITY

The Court also takes a broad-brush approach to accessory liability for breach of fiduciary duties. The Court is bound by the well-established decision, *Twinspectra Ltd v Yardley*,¹¹⁷ which requires that, for accessory liability in any breach of fiduciary duty, the accessory must act dishonestly. The Court said that dishonesty in this context meant ‘knowing about’ or deliberately turning a blind eye to the breach of fiduciary duty.¹¹⁸ First, this conflates dishonesty with knowledge. For knowledge to amount to dishonesty, the knowledge must have been against the standard of an ordinary honest person in the circumstances of the defendant.¹¹⁹ There is no indication of this whatsoever. Applying *Twinspectra* as it stands as the test for dishonesty (as the Court in *FirstRand* did¹²⁰) ignores the change in the definition of dishonesty as introduced by *Ivey v Genting Casinos (UK) Ltd*¹²¹ and as applied to equity in *Group Seven Ltd v Nasir*.¹²² Even in tort, the lender must know that they are bribing an agent. Here this is impossible as no one thought that the brokers were an agent prior to the decision. While in some cases, brokers had been made fiduciaries, such as in *Wood*, it was seen as ad hoc and not applying to all brokers as a class. Knowledge of the commission is not knowledge of a

¹¹⁴ [1995] 1 AC 74 (PC).

¹¹⁵ *FirstRand* (n 5) [100] (Andrews, Birss and Edis LJ).

¹¹⁶ *ibid* [86]–[87], [89], [91]–[93].

¹¹⁷ *Twinspectra* (n 109).

¹¹⁸ *FirstRand* (n 5) [128] (Andrews, Birss and Edis LJ).

¹¹⁹ *Group Seven* (n 5) [52]–[58], applying *Ivey v Genting Casinos (UK) Ltd* [2017] UKSC 67, [2018] AC 391 [74].

¹²⁰ *FirstRand* (n 5) [127].

¹²¹ *Ivey* (n 119).

¹²² *Group Seven* (n 5).

fiduciary duty. It also requires that the lender should assume that the dealer will not disclose the commission.

The source of this issue is a ‘fusion fallacy’. When the tort of bribery split off from the breach of fiduciary duty, accessory liability in equity was knowing assistance,¹²³ not dishonest assistance,¹²⁴ and thus only knowledge was required to make the payer of a bribe liable. By paying the bribe, knowledge is almost automatic in the case of secret commissions; the payer is always liable, and thus accessory liability is transmuted into primary liability by virtue of this inevitability. As tortious liability arises concurrently with a breach of fiduciary duty in the case of secret commissions, there would be no need to address dishonesty for accessory liability. It then seems that dishonesty is no longer a requirement for accessory liability for breach of fiduciary duty in the context of bribes, and some cases such as *Wood* go as far as saying that the payer is not an accessory but a primary wrongdoer in equity as well as common law.¹²⁵

When such beliefs arise for breaches of fiduciary duty for secret commissions, they remain for half-secret commissions. For example, there is not a single mention of dishonesty in *Hurstanger*. The presence of two causes of actions has resulted in divergence on this issue, not just by virtue of the cause of action being in tort or equity, but because the tort of bribery applies outdated equitable rules on accessory liability.

V. REPAIRING THE REASONING

A. CAN A FIDUCIARY DUTY BE IMPOSED?

A fiduciary is someone who has ‘undertaken to act for or on behalf of another in a particular matter in circumstances which give rise to a relationship of trust and confidence’.¹²⁶ This is not a test, but an indicative factor as there is no single authoritative test for ad hoc fiduciary duties. The common factor in *Bristol and West Building Society*, *Al Nehayan*, and *Hospital Products* was the idea of an undertaking. Since, in *FirstRand*, there were no express undertakings, the test is essentially the same as the test of implication of terms in *Attorney General of Belize v Belize Telecom Ltd*:¹²⁷ did the party, by their words or conduct, give rise to an understanding or expectation in a reasonable person that they would act in a particular way?¹²⁸ Other indicative factors include authority and discretion,¹²⁹ which were present given the inexperience of the claimants who essentially allowed the brokers to choose for them. Another indicative factor is whether the claimants have a ‘legitimate expectation’ that the brokers will not act in their own self-interest.¹³⁰

The questions we are asking are the following: did the dealers in *FirstRand* make undertakings that gave rise to a relationship of trust and confidence (not was there trust and confidence) and could the claimants have legitimate expectations of undivided loyalty? In

¹²³ *Barnes v Addy* (1874) LR 9 Ch App 244. This is the most well-known authority for knowing assistance.

¹²⁴ *Royal Brunei Airlines Sdn Bhd v Tan* [1995] UKPC 4. This was the first instance of dishonest assistance and all cases of accessory liability for breach of fiduciary duty prior to it were for knowing assistance.

¹²⁵ *Wood* (n 19) [94].

¹²⁶ *Bristol and West Building Society* (n 95) 18 (Millet LJ). See James Edelman ‘When Do Fiduciary Duties Arise?’ (2010) 126 LQR 302 for a detailed exploration.

¹²⁷ [2009] UKPC 10, [2009] 1 WLR 1988 [21].

¹²⁸ Edelman (n 126) 314–17.

¹²⁹ *Al Nehayan* (n 102) [159].

¹³⁰ *Arklow Investments Limited v Maclean* [2000] 1 WLR 594 at 598; *Farrar v Miller* [2018] EWCA Civ 172 [75]; *ibid* [161], [166], [167] (Leggatt LJ); *Kelly & Anor v Baker & Anor* [2022] EWHC 1879 (Comm).

Wrench, a sales representative made an express assurance twice that the dealership would get him the best deal from their panel of lenders and one that was most suitable for his needs.¹³¹ The language of ‘most suitable’ and ‘best’ gives rise to a legitimate expectation that it is the claimant’s interest that is being prioritised and that the broker is putting their own interests aside (apart from their interest in the sale, which the claimants have consented to). In *FirstRand*, the dealers brokered not only a hire-purchase agreement but a personal loan, and they made misleading and false statements which created an impression that the dealer was exercising its judgment in selecting a finance provider that ‘may be most appropriate’ for the customer’s needs from a panel of 22 lenders, when in fact there was an undisclosed obligation to give first refusal to FirstRand Bank.¹³² These assurances are what give rise to a legitimate expectation. The Court considered these additional factors in *FirstRand* in deciding whether it gave rise to an unfair relationship for the Consumer Credit Act 1974 claim, but not in relation to whether it gave rise to fiduciary duties.

Moreover, the higher the degree of trust, vulnerability, and confidence placed in the fiduciary, the more likely that a reasonable person would view it as creating legitimate expectations or undertakings that would give rise to a relationship of trust and confidence, not whether the pre-existing relationship was one of trust and confidence (which is how the Court in *FirstRand* approached it). Thus, Miss Hopcraft being ‘naïve’ and ‘vulnerable’¹³³ and her assumption that the dealer would give her the best deal are strong indicators that there was such a relationship.¹³⁴ Therefore, the imposition of a fiduciary duty in that case must be considered a fact-specific one and not one that applies to all credit brokering.

This also explains why the dealer is treated separately as a credit broker and as a seller. The division of the relationship is justified as the undertaking to obtain finance is only related to getting finance. While a consumer might not separate the transaction in their mind, the question is what the dealer has undertaken, and they have only undertaken as a credit broker. It also explains the narrowness of the fiduciary duty, as it is not an all-encompassing fiduciary duty but only in relation to the transaction.

This is not how the court viewed it. They applied a broad-brush approach, and this result should apply to all car finance cases involving partially disclosed or undisclosed commissions.¹³⁵ Instead, deciding whether there is a fiduciary duty should be a fact-specific exercise. The court should have looked at whether undertakings were made by the dealers and if there was a particular vulnerability of the buyers.

B. ACCESSORY LIABILITY

For the lender to be liable as an accessory in these cases, dishonesty must be proved and evidenced. Knowledge is insufficient. Otherwise, the application of accessory liability will continue to conflict with the rest of the law on dishonest assistance.

To be liable for assisting in a breach of fiduciary duty, the lender must act dishonestly. As Lord Hutton stated in *Twinsidestra*, ‘dishonesty requires knowledge by the defendant that what he was doing would be regarded as dishonest by honest people’.¹³⁶ This is the common

¹³¹ *FirstRand* (n 5) [95].

¹³² *ibid* [48] (Andrews, Birrs and Edis LJ).

¹³³ *ibid* [23].

¹³⁴ *ibid*.

¹³⁵ *ibid* [91], [105].

¹³⁶ *Twinsidestra* (n 109) [36].

standard of dishonesty for accessory liability in breach of fiduciary duties. However, this language of dishonesty appears nowhere in *Hurstanger*, requiring only knowledge.¹³⁷ *FirstRand* states that dishonesty means ‘knowing about, or deliberately turning a blind eye to, the breach of the broker’s fiduciary duty to their principal’.¹³⁸ The Court in *FirstRand* cites *Twinsectra* and recognises the tension between the two cases.¹³⁹ However, instead of trying to reconcile these two cases, the law should adopt the dishonesty standard. A single standard has the benefit of consistency and, importantly, it makes it harder for lenders to be liable, which, as we will argue in Section IV, is more desirable.

Group Seven makes it clear that the knowledge and belief of the defendant are subjective and that the standard of dishonesty is objective.¹⁴⁰ Blind-eye knowledge is dishonest when you avoid confirming what you suspect to be true, but not when you suspect something and you ‘negligently refrain[] from making further inquiries’;¹⁴¹ this is the correct interpretation of blind-eye knowledge as discussed in *Twinsectra*. *FirstRand* expands ‘turning a blind eye’ well beyond its natural meaning and beyond how it applies in law. *FirstRand* states that knowing that a dealer acts as a credit broker for a fiduciary means that they cannot pay an undisclosed commission to them.¹⁴² It is reasonable to expect someone to know that a trustee is a fiduciary, as all trustees are. Thus, knowingly paying a secret commission to a trustee will always be a breach of fiduciary duty and thus dishonest. As fiduciary duties for brokers are ad hoc, one cannot expect a reasonable lender to assume a fiduciary duty in every case, even if it was present in some cases, such as in *Wood*. To do so would transmute brokerage into a new category of status-based fiduciary. Therefore, knowledge of a commission is not necessarily knowledge that it is secret (as a lender does not know that there is a duty of disclosure), that it is to a fiduciary, or that it is in breach of fiduciary duty from the point of view of the reasonable person. Knowing that a shadow moves does not necessarily mean knowing what casts it. Further, *FirstRand* posits that failing to ensure that the disclosure is made means that the lender deliberately takes the risk that it will not be disclosed. But a failure to prevent an action is not automatically a deliberate omission. Negligent failures do not fall under turning a blind eye, as noted in *Group Seven* above.

Further, if the financier contracted that the broker must make a disclosure, as a matter of contract law they can assume that that disclosure will be made.¹⁴³ There is no reason why this should not also be the case in the context of brokerage, so that there is no reason to suspect a lack of disclosure, let alone require inquiry.

C. FUSING THE TORT OF BRIBERY AND BREACH OF FIDUCIARY DUTY

We propose absorbing the tort of bribery into a breach of fiduciary duty. In this section, we aim to show that it is possible as the disinterested duty can be considered the same type of duty as the fiduciary duty of loyalty, just with a more limited scope.

¹³⁷ *Hurstanger* (n 6) [35].

¹³⁸ *FirstRand* (n 5) [128] (Andrews, Birss and Edis LJJ).

¹³⁹ *ibid* [177].

¹⁴⁰ *Group Seven* (n 5) [57].

¹⁴¹ *ibid* [59] (Henderson, Peter Jackson and Asplin LJJ).

¹⁴² *FirstRand* (n 5) [128].

¹⁴³ *Cehave NV v Bremer Handelsgesellschaft MBH* (*The Hansa Nord*) [1976] QB 44 (CA) 71 (Lord Mustill) (‘contracts are made to be performed and not to be avoided’); *Ruxley Electronics and Construction Ltd v Forsyth* [1996] AC 344 (HL) 360; Daniel Friedmann, ‘The Performance Interest in Contract Damages’ (1995) 111 LQR 628, 629.

While a fiduciary duty was found on the facts of *Wood*, it was desirable for the claimant to try to obtain two key advantages found in the tort of bribery. These advantages are (i) that the payer is primarily liable as long as they have knowledge that the payee was subject to a duty¹⁴⁴ and (ii) that rescission is available as of right (subject to the bars).¹⁴⁵ If the language of fiduciary duties is used to describe the duty, it makes it seem close to the equitable action. This was particularly risky as *Hurstanger* imposed a very wide discretion in awarding damages. Citing *Johnson v EBS Pensioner Trustees Ltd*,¹⁴⁶ Tuckey LJ considered that awarding rescission would be disproportionate and unfair and so he awarded damages to the borrowers instead. Such a power exists by virtue of section 2(2) of the Misrepresentation Act 1967, which allows upholding contracts and awarding damages in the case of *innocent misrepresentation*, but it is not a general power. It is constitutionally inappropriate to expand statutory power in this manner. More importantly to claimants, the exercise of this discretion is far more unpredictable. Equitable rescission was, like other equitable remedies, subject to a ‘weak’ discretion.¹⁴⁷ This means that it was subject to the bars to rescission, such as *restitutio in integrum*, clean hands, third-party impact, or affirmation. In addition, there is now a wider discretion to refuse rescission. There was thus a significant incentive to avoid falling within equity’s jurisdiction that was not present in earlier cases. It is submitted that this is a key reason for trying to argue the case as a disinterested duty, as opposed to a fiduciary one.

David Richards LJ in *Wood* considered the fiduciary duty required for bribery to be the ‘core’ fiduciary duty of undivided loyalty.¹⁴⁸ But the disinterested duty is not a different duty to that of the duty of undivided loyalty. The duty of undivided loyalty means that you must prioritise the principal’s interests over your own interests and the interests of other people.¹⁴⁹ Ultimately, it is about conflict of interest. The disinterested duty in *Wood* is derived from cases that discussed equitable principles.¹⁵⁰ None of the cases cited in *Wood* held that a disinterested duty is different from a fiduciary one; instead, in each case, the payee was a fiduciary and so a fiduciary duty was owed. The language of ‘disinterest’ is used as an alternative formulation of ignoring others’ interests and prioritising the principal’s interests. The duty of undivided loyalty does not have a set definition, and all that any definition can do is capture the issue of conflicts of interest; ‘disinterest’ does just that. The duty of undivided loyalty is wider than the disinterested duty, but it is best to view the disinterested duty as a fiduciary duty in relation to the advice given. The scope is limited by the undertaking made by brokers. The fiduciary duty of loyalty is not absolute and can be modified. Contractual trusts can exclude or limit the duty of loyalty for certain purposes, such as in *Citibank NA v MBLA Assurance SA*,¹⁵¹ and conflicts can be agreed in the trust deed, such as where a trustee is also a beneficiary

¹⁴⁴ *Mahesan* (n 2).

¹⁴⁵ *Wood* (n 19) [101].

¹⁴⁶ [2002] Lloyds Rep PN 309.

¹⁴⁷ The distinction between ‘weak’ and ‘strong’ discretion is based on one that is drawn in Ronald Dworkin, *Taking Rights Seriously* (rev edn, Duckworth 1978) 31–39. There is support for strong discretion in choosing remedies where judges can choose the most appropriate remedy on the facts, but see for example Peter Birks, ‘Rights, Wrongs, and Remedies’ (2000) 20 OJLS 1; Peter Birks, ‘Three Kinds of Objection to Discretionary Remedialism’ (2000) 29 University of Western Australia Law Review 1.

¹⁴⁸ *Wood* (n 19) [27]. The idea of the duty of loyalty being the ‘core’ duty comes from *Bristol and West Building Society* (n 95) 18 (Millett LJ).

¹⁴⁹ *Bristol and West Building Society* (n 95) 18.

¹⁵⁰ *Panama and South Pacific Telegraph* (n 31) 528–29 (Mellish LJ) (‘honest and disinterested advice’); *Shipway* (n 23) (Chitty LJ) (‘his duty conflicted with his interest’); *Logicrose* (n 55) (Millett J) (‘the agent has put himself in a position where his interest and duty may conflict. A principal is entitled to the disinterested advice’); *Anangel Atlas Compania Naviera* (n 16) 169 (Leggatt J) (‘his duty and his interest conflict’).

¹⁵¹ [2006] EWHC 3215 (Ch).

(the trust deed has agreed that a trustee can act in their own interest as well) or for discretionary trusts where conflict between the beneficiaries is built into the trust. Similarly, any written or oral undertakings made by the broker can modify the scope of the fiduciary duty that is adopted. Therefore, we can see the disinterested duty as a case of a modified fiduciary duty, which relates only to the information, advice, or recommendations given. Thus, a breach of a disinterested duty is a breach of fiduciary duty, which falls within the equitable jurisdiction.

What David Richards LJ ends up doing is assessing whether there is a ‘requirement for a fiduciary relationship’,¹⁵² instead of asking whether a fiduciary duty is imposed. As the tort of bribery is not a breach of fiduciary duty, it is not a necessary element of the action. ‘The law of bribery is, in its fundamentals, a manifestation of the law of agency and fiduciary duties’,¹⁵³ so the common law simply acts as ‘tortious interference’,¹⁵⁴ in getting to the heart of the action. He states that, while it may be ‘accurate’ that a disinterested duty is characterised as a fiduciary duty of loyalty, we should not engage in this analysis as it is ‘complex’ and you should instead ask the ‘straightforward question’ of whether a disinterested duty is owed.¹⁵⁵ But, as a new development, there is little case law to answer when a disinterested duty is owed, while there is a large body of fiduciary law to answer when a fiduciary duty is owed. The fact that fiduciary law is complex does not mean that we should ignore it. Owing a fiduciary duty is onerous, and ad hoc fiduciary duties are and should be relatively rare and difficult to find. Being ad hoc means that the relationship falls outside the norm of fiduciary relationships. Pretending that the rules of fiduciary law do not exist does not make the rules disappear and our considerations of when they should be owed are not properly argued. Moreover, the Court of Appeal in *FirstRand* states that the ‘obligation of loyalty... is inherent in the disinterested duty’.¹⁵⁶ The Court recognises that a disinterested duty and fiduciary duty (as a duty of undivided loyalty) are the same thing or, at least, that both respond to ‘loyalty’. However, as the Court of Appeal, they are bound by the errors in *Wood* as a matter of precedent or they refused to consider *Wood* to be incompatible with *Hurstanger* and the rest of the law.¹⁵⁷

If we view the disinterested duty as a modified fiduciary duty, we can consider both duties to be equitable. The differences are also explained by the undertakings made and not by status or the action from which they arose. As the disinterested duty is a modified fiduciary duty, the tort of bribery should be absorbed into the breach of fiduciary duties. Why is this desirable? We take the position of Andrew Burrows that different rules at common law and in equity cannot be justified by reference to their historical and jurisdictional origins,¹⁵⁸ although we recognise the contrary position that the Judicature Acts of 1873 and 1875 brought

¹⁵² *Wood* (n 19) [39], [52], [92].

¹⁵³ Thomas Grant and David Mumford (eds), *Civil Fraud: Law, Practice and Procedure* (1st edn, Sweet & Maxwell 2018) para 7-006.

¹⁵⁴ The term is used non-technically.

¹⁵⁵ *Wood* (n 19) [102] (David Richards LJ).

¹⁵⁶ *FirstRand* (n 5) [91] (Andrews, Birss and Edis LJ).

¹⁵⁷ The Court of Appeal are bound by their own previous decisions, unless (i) there is conflicting House of Lords or Supreme Court authority, (ii) there are two earlier conflicting Court of Appeal decisions, or (iii) a decision was made *per incuriam*: *Young v Bristol Aeroplane Co Ltd* [1944] KB 718 (CA).

¹⁵⁸ Andrew Burrows, ‘We Do This at Common Law But That in Equity’ (2002) 22 OJLS 1, 3–5. This position is controversial, but now widely adopted in English private law to varying degrees of success. See for example *Target Holdings Ltd v Redfern* [1996] AC 421 (HL) and *AIB Group (UK) Plc v Mark Redler & Co Solicitors* [2014] UKSC 58, [2015] AC 1503 [71], in the context of compensation in equity, and *Times Travel (UK) Ltd v Pakistan International Airlines Corporation* [2021] UKSC 40, [2023] AC 101 [5]–[9], [89]–[90], for cases of undue influence now being treated as duress.

only administrative and not substantive fusion of common law and equity.¹⁵⁹ The ‘fusion fallacy’, proposed in *Meagher, Gummow and Lehane’s Equity: Doctrine and Remedies*¹⁶⁰ and taken up by later authors of the book, just means that one must recognise that they are separate doctrines and that we can develop and fuse them using the basic common law techniques of analysing the principles and progressing the law slowly by analogy. We cannot automatically fuse rules by reference to history. The tort of bribery is particularly egregious in that it created a new concurrent jurisdiction instead of fusing two separate doctrines. As Whayman has argued, the doctrines have a common origin. They both deal with the same question: how should the law respond to bribes and secret commissions paid to agents and fiduciaries? The law should not have two answers to the same question without good reason. We can think of none.

We are then left with the choice of constructing a new action or choosing between the two. We propose choosing the equitable action of breach of fiduciary duty. This has the benefit of having a wider body of law around it with the entire body of fiduciary law and equity, while the common law action is ‘sui generis’.¹⁶¹ The disinterested duty is also best viewed as a modified fiduciary duty, as we have argued above, which responds to the same issue as a fiduciary duty of loyalty. Dealing with it under equity means that we deal with the purpose of a disinterested duty instead of pretending it does not have anything to do with a fiduciary. Moreover, all the major cases, including *Wood* and *FirstRand*, find a fiduciary duty as well as a disinterested one. The fiduciary duty is wider, but only because the disinterested duty is a fiduciary duty for specific purposes. Equity can therefore accommodate the common law.

One key benefit of the tort of bribery for claimants is that it is easier to prove as it only requires knowledge that the payee is an agent or fiduciary or owes the disinterested duty. But this primary liability tort is a result of using outdated principles of knowing rather than dishonest assistance, and ease of use for claimants is no justification for any law. For coherence and clarity within secret commissions and with equitable principles generally regarding accessory liability, fusion would be a better result. Dishonest assistance also has the benefit of disgorging any secondary profits made as a result of a bribe which is not available at common law.¹⁶² A further benefit of it being a breach of fiduciary duty is that it permits proprietary claims as a constructive trustee. It would also create coherence with regard to half-secret commissions, as there would only be one claim. Given the current incoherence, it is illogical to have two separate causes of action and different remedies for the same wrong. Not only are equitable remedies already available for the tort, but equity has the flexibility to accommodate the common law, as opposed to the other way round.

Another justification for requiring dishonesty in accessory liability is that fewer lenders will be liable in any brokerage case. We argue in Section VI that this is desirable as a matter of policy. As a summary, we propose that the tort of bribery should be subsumed into a breach of fiduciary duty. The result is that both secret and half-secret commissions are included within the scope of a breach of fiduciary duty if the disclosure is insufficient to obtain the consent of the principal. A payee is primarily liable for breach of fiduciary duty, and a payer is liable as an accessory to the breach. The standard for accessory liability is dishonesty and requires evidence of dishonesty.

¹⁵⁹ Heydon, Leeming and Turner (n 47).

¹⁶⁰ *ibid.*

¹⁶¹ *Mahesan* (n 2) (Lord Diplock).

¹⁶² *Novoship* (n 21) [84].

VI. SHOULD A FIDUCIARY DUTY BE IMPOSED?

We agree with the Court of Appeal in *FirstRand* that a fiduciary duty should be imposed in the three cases therein. However, we disagree with the way in which they determine that a duty is imposed. They determine that the duty arises from the parties' status as credit brokers. This is a presumption that the credit brokers must rebut.¹⁶³ With respect, we suggest that these duties, whether under common law or equity, should be applied more carefully. Ad hoc fiduciary duties only arise consensually, based on the undertaking made. Therefore, they do not apply to situations where no such undertaking has been made and to brokers generally. The ad hoc approach is better than a presumption as it has a stronger basis in law; it makes brokers liable when they have actually made the undertaking, which respects the autonomy of individuals and when customers are particularly vulnerable, and it means that most brokerages will reflect the commercial reality which we will detail further below.

As argued, a fiduciary duty is defensible as a matter of law, but whether a fiduciary duty is appropriate as a matter of policy can be discussed separately. Three policy concerns we have identified include the protection of consumers, the facilitation of transactions, and the incentives governing both parties. These will inform our conclusions on the desirability of the decision in *FirstRand* and our proposal that breaches of fiduciary duties should be the sole action in relation to bribes. We begin by discussing three financial effects of the decision.

A. THREE FINANCIAL EFFECTS

There are three financial effects which are, in our opinion, issues caused by the decision in *FirstRand* as presently constituted. These are unintentional effects of the judgment. Having fiduciary duties be ad hoc rather than based on a presumption reduces the number of brokers that are liable. This reduces the impact of the following financial effects.

First, it requires lenders and other financiers to seek disclosure of whether the commission was disclosed to the principal. In order to avoid liability on their own part, they must incur search costs.¹⁶⁴ Brokerages make their cases to financiers by saving them these very search costs. This disincentivises using commission as an automated searching process by forcing disclosure. The brokerage uses commission to chase perpetually what the financiers' search role would otherwise be. Duplicating the very costs that the brokerage aims to save will damage motor finance brokerage. Consumer protection is based on advancing consumers' interests; this case may thus fall into the trend predicted by Briana Chang and Martin Szydłowski,¹⁶⁵ that, whilst the quality of information available to a consumer may improve, consumer results may not. If the decision is at least in part justified by consumer expectations or protections, as one of the policy concerns that we have identified, then it must be re-examined.

Secondly, the decision tolerates disclosed commissions. Its effect, assuming all commissions are fully disclosed, is to shift commission costs onto brokerages and then onto financiers. This would be sustainable if financiers had no alternative, but they do. Consumers prefer cheaper brokers, all other things being equal. As long as the customer is paying some element of this commission at some point, they have very little reason to opt into a commis-

¹⁶³ *FirstRand* (n 5) [87], [100].

¹⁶⁴ *ibid* [128], [159], [168].

¹⁶⁵ See Briana Chang and Martin Szydłowski, 'The Market for Conflicted Advice' (2020) 75 *The Journal of Finance* 867.

sion where a cheaper option exists. This cost, previously paid by consumers, shifts onto financiers where commission remains at all; brokerages absorb it otherwise. A race to the bottom likely emerges, where no individual consumer wants to pay a commission. This is not a problem in and of itself. If this were a simple two-party relationship of payer versus payee, there would be no issue. However, given the presence of financiers, one should consider what is likely to occur next. As Roman Inderst and Marco Ottaviani observe, '[d]isclosure unambiguously reduces the equilibrium level of commissions',¹⁶⁶ meaning that this kind of market activity lessens.

Brokerages depend on commissions, at least, in contrast to financiers. Commissions incentivise brokerages to pair consumer and financier accurately in a deal that benefits both, plus brokerage costs. This adds information, which other tools, like bid-ask spreads, order flow data, and historical trading patterns, may not completely capture. Commissions capture data hidden beyond these more precise tools, with tighter portfolios, into price. Both financier sentiment and the urgency of consumer demand are difficult to pin down with such niche tools; both are more neatly but implicitly demonstrated by commission. Other models, like fixed brokerage fees or subscription models (the second is difficult to imagine in the motor finance market in any case), do not provide the same real-time updates to all parties involved. Furthermore, commissions seem to behave differently when quality differs across markets. Alexandre de Cornière and Greg Taylor's research suggests that, whilst the removal of biased advice will help consumers in lower quality markets, it may injure consumers in higher quality markets.¹⁶⁷ In higher quality markets, biased advice given by a brokerage incentivises their counterparties to match the consumer's lofty expectations, though not so in lower quality markets. A commission for selling a Ferrari is not the same as a commission for selling a Volkswagen. The law is an inappropriate forum in which to draw this distinction; however, so far as this ruling is premised on consumer protection, this issue must be noted. Similarly, Shubhranshu Singh's research suggests the possibility of a competitive market in bias.¹⁶⁸

Finally, lenders can respond more easily and effectively to changes in commission than consumers, meaning that they may choose to stop providing motor finance. A commission in one place is not a commission in all places. Commissions have different effects depending on the product that they concern. George A Akerlof's seminal paper hints at this difference, though from the point of view of discerning between qualities.¹⁶⁹ Cars are expensive. Consumers often want a specific model or to choose from a limited set of models. The relevant set of consumers needs a car or at least strongly desires one. However, the same is not true from the financier's perspective. There is little reason to presume, at least from the outside looking in, that motor finance loans are necessary to the operation of a given financier or that they radically outperform any other kind of finance. Motor finance can, as an asset on a balance sheet, be bundled, traded, and speculated upon like any other, regulations permitting. Financiers can trade loans far more easily than consumers can cars. The two effects described above disincentivise the financier from engaging in motor finance brokerage. It is both easier for the financier to withdraw from the market than the consumer, and he has

¹⁶⁶ Roman Inderst and Marco Ottaviani, 'Competition through Commissions and Kickbacks' (2012) 102 *American Economic Review* 780, 781. Equilibrium level is the level of commission where the market clears and where supply and demand meet.

¹⁶⁷ See Alexandre de Cornière and Greg Taylor, 'A Model of Biased Intermediation' (2019) 50 *RAND Journal of Economics* 854.

¹⁶⁸ See Shubhranshu Singh, 'Competition in Corruptible Markets' (2017) 36 *Marketing Science* 361.

¹⁶⁹ See George A Akerlof, 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism' (1970) 84 *The Quarterly Journal of Economics* 488.

more reason to do so. Santosh Anagol, Shawn Cole and Shayak Sarkar demonstrated that sellers prefer products where they do not have to disclose commissions.¹⁷⁰ We extrapolate this into market exits where the product is just another loan for lenders.

This imbalance matters because it determines who ultimately bears the cost of a legal intervention which has policy implications. Consumers cannot exit motor purchases as easily as financiers can exit motor finance. Warnings against this sort of mandatory disclosure have been present in the literature for 13 years, sometimes explicitly in the UK context.¹⁷¹ By failing to account for this, the ruling risks making motor finance less competitive, more expensive, or less available altogether. If we want motor finance to be available to consumers, lenders to lend, and to incentivise transactions, then the effect of *FirstRand* is undesirable. Following our approach to ad hoc fiduciary duties, the duty should be imposed in fewer circumstances, requiring disclosure in fewer circumstances.

B. THE LIABILITY OF LENDERS

A complete implosion of the brokerage industry or model is unlikely. Our personal concerns about aftereffects on brokerages are ultimately limited. The concerns that motivate us, the response of the industry, and the Treasury's intervention in this case, are the implications for widespread primary or accessory liability of lenders as a result of this ruling. If the tort of bribery remains, lenders paying the commission will be liable if they knew that the payee owed a disinterested or fiduciary duty. Similarly, if accessory liability for breach of fiduciary duty is applied as it is in *FirstRand*, this knowledge suffices to be dishonest.¹⁷² Our proposals, as discussed above, deal with the issue in two ways: first, by removing the tort of bribery, lenders will not be primarily liable without more; and secondly, we require that the test of dishonesty for dishonest assistance be applied more stringently in line with other case law on the issue. If knowledge does not suffice for primary liability or dishonesty, then far fewer lenders will be liable.

Finally, we wish to comment on the Financial Conduct Authority ('FCA') rules in relation to discretionary finance agreements from 2021. It would be easy to misunderstand the breadth of those rules. They only banned agreements that specifically aligned the brokerage commission with the interest rate payable by the consumer. This is not a comprehensive ban of secret and half-secret commissions. Updated guidance on that topic was published with the FCA's Policy Statement (PS20/8).¹⁷³ Consumer Credit Sourcebook ('CONC') 4.5.3R requires disclosure of some elements of the commission framework, but seems to require no more than the law as presently constituted.¹⁷⁴ Therefore, the FCA rules leave our proposal unaffected.

¹⁷⁰ See Santosh Anagol, Shawn Cole and Shayak Sarkar, 'Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market' (2015) Harvard Business School Working Paper 12-055 <https://www.hbs.edu/ris/Publication%20Files/12-055_13c23c02-c57f-4aca-9630-316aa4b772ce.pdf> accessed 30 March 2025.

¹⁷¹ Inderst and Ottaviani (n 166) 782.

¹⁷² *FirstRand* (n 5) [128].

¹⁷³ Financial Conduct Authority, 'Motor Finance Discretionary Commission Models and Consumer Credit Commission Disclosure - Feedback on CP19/28 and Final Rules' (Policy Statement PS20/8, July 2020) <www.fca.org.uk/publication/policy/ps20-8.pdf> accessed 15 December 2024.

¹⁷⁴ Financial Conduct Authority, *FCA Handbook* (2024) <www.handbook.fca.org.uk/handbook/CONC/4/5.html#DES18> accessed 15 December 2024.

VII. CONCLUSION

FirstRand should have applied the law differently in two ways. First, we argue that fiduciary duties should be owed by credit brokers in fewer circumstances. We agree with the Court of Appeal in *FirstRand* that a fiduciary duty should be owed in the three cases. However, we disagree that this should arise as a presumption out of their status as credit brokers. Ad hoc fiduciary duties are consensual and should be applied carefully based on the undertakings in a relationship. The relationship itself only affects how undertakings are understood; it does not create undertakings.

Secondly, accessory liability for breach of fiduciary duty should strictly apply the test for dishonesty. *FirstRand* essentially applies a test of knowledge or, on an alternative interpretation, it applies the wrong test of dishonesty or applies it incorrectly. The source of this inconsistency is *Hurstanger* which dispenses individual justice in the case of a vulnerable consumer but is a creature of its facts. Dishonesty is consistent with the remainder of the law, whereas knowledge is not.

We also propose fusing the tort of bribery into a breach of fiduciary duty. The law should not have two answers to the question of liability in bribery without good reason. The reason is not a difference in duty. The disinterested duty in *Wood* is just a fiduciary duty modified by the undertaking, so the disinterested duty can be accommodated in equity. Consistency and coherency should prevail; the desirability of these qualities in a legal system is uncontroversial.

However, coherency and consistency only mean that we must choose one of the two; it does not dictate choosing equity. To begin with, there is a wider range of remedies available in equity, including a proprietary remedy. More importantly, equity provides a result that is better from a policy perspective based on the policy factors of consumer protection, facilitating finance, and incentivising transactions. Treating the disinterested duty as a fiduciary duty means that more careful analysis is applied due to the entire body of fiduciary law being relevant. This careful application of the facts when looking at the undertakings made and the nature of the relationship means that fewer credit brokers will be treated as fiduciaries. This better reflects the underlying transaction of sale in car cases.

An equitable action also reduces the number of lenders that are liable. This is because the tort of bribery only requires knowledge for a lender to be primarily liable and equity requires dishonesty for a lender to be liable as an accessory. This, combined with a more careful application of dishonesty, means that fewer lenders will be liable. Lenders and financiers will abide by a situation where the risk of accessory liability outweighs the utility of brokerages. There is no prima facie reason why financiers cannot operate and offer these products themselves without the involvement of brokerages. The only reasons to do so, from their point of view, are to save time, extrapolate demand, and simplify problems. If financiers veto the use of brokerages, then consumers lose out too.

We hope that our observations are timely and that our reform to *FirstRand* will be ordained by the Supreme Court to establish a more coherent and less punitive position. Subsequent debate on the nature of brokerage and its interaction with agency is for Parliament.

Severing the Hunt for Tortfeasor-Caregiver Compensation: The Singaporean Rejection of *Hunt v Severs* [1994] 2 AC 350 (HL)

JUN XIANG WONG*

ABSTRACT

In 2024, in the case of *Rajina Sharma v Theyvasigamani* [2024] SGHC 42 (*'Rajandran'*), the Singaporean High Court was faced with a highly similar fact pattern to that which arose in the English case of *Hunt v Severs* [1994] 2 AC 350 (HL) 30 years earlier. A defendant tortfeasor had caused a motorcycle accident to be suffered by a plaintiff victim, and the former subsequently provided voluntary care to the latter at his own expense. However, the Singaporean court departed from *Hunt* and allowed the plaintiff to claim compensation from the tortfeasor for the voluntary care that he had provided himself. This article will compare the approaches taken by the two courts, as well as the problems that arose from the original *Hunt* judgment. It is submitted that the *Rajandran* approach represents a more principled and just solution than its English predecessor, but that this could be further improved by adopting a 'closed loop' model of compensation proposed by the author.

Keywords: *Hunt v Severs, law of tort, damages, personal injury, Singapore*

I. INTRODUCTION

On 28 April 1994, the House of Lords handed down judgment in the case of *Hunt v Severs*,¹ holding that a victim of a tort who had received voluntary care from the tortfeasor could not sue the tortfeasor for damages for the value of the voluntary care that the tortfeasor himself had rendered. The case confirmed that, in English law, the underlying rationale for the recovery of care rendered voluntarily to a victim in damages was to compensate the voluntary carer for their services.²

Though this outcome may look simple and sensible, it ran up against several problems. Arguably, the problem that left the most room for wider unfairness when applying *Hunt* to future cases was the lack of consideration of the tortfeasor's insurance coverage in the House of Lords. The existence of insurance coverage was not regarded as a separate party

* BA Law student, Hughes Hall, University of Cambridge. I am grateful to the anonymous reviewers for their comments on earlier drafts of this article. All errors that remain are my own. E-mail for correspondence: jxw22@cam.ac.uk.

¹ [1994] 2 AC 350 (HL).

² *ibid* 363.

from the defendant himself in the finding of liability or calculation of damages, yet it was materially important to how damages were to be paid.

30 years later, as misfortune would have it, a highly similar fact pattern arose in Singapore in the case of *Rajina Sharma v Theyvasigamani* (*Rajandran*).³ The High Court of Singapore would take advantage of this opportunity to clarify the principles behind the recovery of damages for gratuitous voluntary care and the position in Singaporean law when a tortfeasor provided the voluntary care. In doing so, the High Court departed from *Hunt* and reached the opposite solution, namely that the victim could be entitled to recover damages from the tortfeasor without the risk of circularity of compensation arising. Essentially, this meant that damages would not proceed in a circular fashion from the defendant to the plaintiff, and then back to the defendant, aligning with the reality (where an insurer was present). Interestingly, this was the case even though the victim's litigation representative was the tortfeasor himself.

This article will seek to analyse the criticisms raised against *Hunt*, and how *Rajandran* departed from *Hunt*, as well as the strengths and weaknesses of the *Rajandran* approach. In Section II, the article will provide background on the case of *Hunt*, and Section III will examine several criticisms that have been raised in response to that case. The particular problems with the trust approach adopted in *Hunt*, and the author's proposed model for resolving the conundrum posed by that case, will be considered in Section IV. Section V will then consider the background and holding of *Rajandran*, with Section VI examining the strengths and weaknesses of that approach.

It is submitted that the outcome of *Rajandran* is a better and more principled one compared to *Hunt* in the light of the criticisms of the latter, while any weaknesses in the *Rajandran* approach are unlikely to bear heavy consequences in practice. Although the original case of *Hunt* is dated and the rule that it provides has been established for decades, the recent example provided by the Singaporean model has provided a sound rejoinder to the *Hunt* rule, prompting a fresh re-examination of the *Hunt* rule.

II. BACKGROUND: *HUNT V SEVERS*

A. THE FACTS AND JUDGMENT

On 14 September 1985, 22-year-old Ms Hunt was riding pillion with Mr Severs on his motorcycle when she got into a serious accident. Her injuries left her paraplegic, and she subsequently spent a lot of time in various hospitals. However, Severs dutifully remained by her side and provided her with care, including visiting her in hospital numerous times. The two were married in 1990.

Therefore, at first blush, it might seem odd that Hunt would seek to sue Severs for damages inclusive of the value of the services he had rendered gratuitously to her. However, this becomes comprehensible when one considers the position of the insurer behind Severs, providing him with coverage and a source of damages. The one who would be paying damages for a successful claim in negligence, therefore, would not be Severs but his insurer. Severs, presumably, was hopeful that, should his wife's claim fully succeed, not only would she be compensated out of his insurance plan for the costs of future care and her loss of earnings,

³ [2024] SGHC 42.

but he would also be able to recoup the costs of his voluntary care from his insurance provider. (This was not stated in the case itself, but given his loyal care and support and their eventual marriage, it is an inference that can reasonably be drawn from the wider context.)

However, with an insurance payout at stake, *Severs's* insurance provider appears to have asserted strong control over the way his arguments were presented during the proceedings, and the House of Lords ruled in favour of *Severs* (though potentially against his personal interests). A single judgment delivered by Lord Bridge, with whom the other four Law Lords agreed, determined that 'there can be no ground in public policy or otherwise for requiring the tortfeasor to pay to the plaintiff, in respect of the services which he himself has rendered, a sum of money which the plaintiff must then repay to him'.⁴

B. THE REASONING

Lord Bridge began by recognising that English law could allow an injured plaintiff⁵ to recover the value of voluntary services rendered to them, but that a 'consistent juridical principle' had been elusive.⁶ Tracing previous case law on the issue, he arrived at a conflict between two Court of Appeal cases: *Cunningham v Harrison*⁷ and *Donnelly v Joyce*.⁸

In *Cunningham*, Lord Denning MR held that:

when a husband is grievously injured—and is entitled to damages—then it is only right and just that, if his wife renders services to him, instead of a nurse, he should recover compensation for the value of the services that his wife has rendered. It should not be necessary to draw up a legal agreement for them. On recovering such an amount, the husband should hold it on trust for her and pay it over to her. She cannot herself sue the wrongdoer... but she has rendered services necessitated by the wrong-doing, and should be compensated for it.⁹

Lord Denning therefore justified the recovery of damages for voluntary care on the basis of the existence of a trust relationship between the victim (as trustee) and the voluntary carer (as beneficiary).¹⁰ The rationale for this head of recovery was the compensation of the voluntary carer.¹¹

In contrast, the Court of Appeal in *Donnelly* instead sought to place the focus for this head of recovery on the plaintiff's loss. In this case, a mother gave up her part-time employment to look after her son, who had to wear a surgical boot and suffered recurrent skin breakdown following a road accident caused by the defendant. Her son sought to recover damages for his mother's loss of wages. For Megaw LJ, delivering the judgment of the court:

⁴ *Hunt* (n 1) 363.

⁵ *Hunt v Severs* uses the term 'plaintiff', but ever since the Civil Procedure Rules came into force in 1999, the term 'claimant' rather than 'plaintiff' has been used in England and Wales. The term 'plaintiff', however, is still used in Singapore. For reasons of consistency, the term 'plaintiff' will be used throughout this article, other than in direct quotations that use the term 'claimant'.

⁶ *Hunt* (n 1) 358.

⁷ [1973] QB 942 (CA).

⁸ [1974] QB 454 (CA).

⁹ *Cunningham* (n 7) 952.

¹⁰ *ibid.*

¹¹ *ibid.*

The question from what source the plaintiff's needs have been met, the question who has paid the money or given the services, the question whether or not the plaintiff is or is not under a legal or moral liability to repay, are, so far as the defendant and his liability are concerned, all irrelevant.¹²

Megaw LJ did not see the loss suffered by the plaintiff as being his mother's expenditure on medical care and equipment, but the 'existence of the need' for said medical care and equipment.¹³ Her loss of wages had been the 'proper and reasonable cost of supplying those needs'.¹⁴ Therefore, it was recoverable against the defendant. The rationale for this head of recovery was that it had been a loss suffered by the plaintiff himself.¹⁵

Faced with the conflict in rationales, in *Hunt*, Lord Bridge held that he found the *Donnelly* justification unconvincing. He reasoned that the source of treatment was relevant, because a plaintiff could not recover from the defendant the costs of being treated by the National Health Service, but could recover the costs of being treated by a private hospital.¹⁶ Furthermore, adopting the *Cunningham* rationale better harmonised with the position taken by the law of Scotland on the same issue.¹⁷ Lord Bridge therefore approved the *Cunningham* view: 'the injured plaintiff who recovers damages under this head should hold them on trust for the voluntary carer'.¹⁸

With the focus on compensating the voluntary carer established, it became a natural conclusion that allowing a tortfeasor to pay a victim damages for voluntary care that he himself had rendered was circular—it would go from the tortfeasor, to the victim who held it on trust for the tortfeasor, and then back to the tortfeasor. Lord Bridge held that there was 'no ground in public policy or otherwise' for allowing recovery for the value of voluntary care in such a scenario.¹⁹

One additional factor, raised before the House of Lords, but not the Court of Appeal, was the existence of liability insurance. Counsel for Severs argued that his insurance coverage provided a policy reason to allow *Hunt* to claim the value of Severs's voluntary care from him (or rather, from his insurer). However, Lord Bridge dismissed this argument as 'a novel and radical departure in the law of a kind which only the legislature may properly effect'.²⁰ Lord Bridge then held that '[a]t common law the circumstance that a defendant is contractually indemnified by a third party against a particular legal liability can have no relevance whatever to the measure of that liability'.²¹ One further point had been raised, and accepted without dispute, in the Court of Appeal without being considered in the House of Lords. This was that *Hunt* could have recovered the reasonable cost of services rendered by Severs 'pursuant to a bona fide contract between the two had there been such'.²²

¹² *Donnelly* (n 8) 462.

¹³ *ibid.*

¹⁴ *ibid.*

¹⁵ *ibid.*

¹⁶ *Hunt* (n 1) 361.

¹⁷ *ibid.* 363.

¹⁸ *ibid.*

¹⁹ *ibid.*

²⁰ *ibid.*

²¹ *ibid.*

²² *Hunt v Severs* [1993] QB 815 (CA) (*Hunt* (CA)) 824 (Sir Thomas Bingham MR).

III. PROBLEMS WITH THE *HUNT V SEVERS* APPROACH

The Law Commission identified six recurring criticisms of the way that the House of Lords had applied the relevant principles to cases where the tortfeasor provided voluntary care, in their Consultation Paper on *Damages for Personal Injury: Medical, Nursing and Other Expenses*.²³ These were:

1. It remained possible, and in fact became encouraged by the decision in *Hunt*, that a plaintiff and tortfeasor in the position of Hunt and Severs could enter into an agreement where the tortfeasor would provide the plaintiff with the requisite care for money.²⁴
2. *Hunt* might encourage plaintiffs to accept nursing services from relatives and friends other than the tortfeasor, who might be less well positioned to provide such aid.²⁵
3. *Hunt* might discourage plaintiffs from accepting gratuitous care entirely, pushing them to opt for professional care at full commercial rates instead.²⁶
4. Where there are two tortfeasors responsible for the injury to the plaintiff, but only one of them provides voluntary care to the plaintiff, the tortfeasor who provided voluntary care would not be able to claim contribution from the other tortfeasor.²⁷
5. The tortfeasor's insurance coverage was considered irrelevant. The plaintiff and tortfeasor were unable to rely on the proceeds of insurance to cover the costs of the tortfeasor's voluntary care.²⁸
6. In *Hunt v Severs*-type cases, the plaintiff and tortfeasor generally form one household and one economic unit. It is meaningless to talk of money being held by one of them for another.²⁹

The Law Commission rejected the fifth and sixth criticisms.³⁰ On the fifth criticism, they responded that the existence of insurance should not be permitted to affect liability and that the question of the tortfeasor's liability should necessarily be answered before their insurer's liability ever became relevant.³¹ On the sixth criticism, they responded that litigation *ex hypothesi* treated family members as separate economic units and that framing the case as being between the plaintiff and tortfeasor, on one side, and the third-party insurer, on the other, would turn the third-party insurance into first-party insurance.³² However, they saw the other four criticisms as persuasive.³³

²³ Law Commission, *Damages for Personal Injury: Medical, Nursing and Other Expenses* (Law Com CP No 144, 1996) ('Law Commission (CP No 144)').

²⁴ *ibid* paras 3.60–3.61.

²⁵ *ibid* para 3.62.

²⁶ *ibid* para 3.63.

²⁷ *ibid* para 3.64.

²⁸ *ibid* para 3.65.

²⁹ *ibid* para 3.66.

³⁰ *ibid* paras 3.65–3.66.

³¹ *ibid* para 3.63.

³² *ibid* para 3.65.

³³ *ibid* para 3.67.

It is agreed that the sixth criticism identified above is a weak one: there is no case in law for a strict rule against treating individual members of the same household as separate economic units. That money, or other property, cannot be held on trust by one family member for another is an absurd suggestion. For instance, land can be held on trust by an individual for themselves, their spouse, and their children.³⁴ However, it is suggested that the question of insurance is still worthy of deeper examination given the questions that it raises in *Hunt v Severs*-type scenarios, despite the Law Commission's dismissal of that point.

The Law Commission agreed that a plaintiff should be under a personal legal obligation to account for damages for past care to a voluntary caregiver who provided it.³⁵ However, they responded to the application of that rule to tortfeasor-caregiver cases with much less enthusiasm, proposing that the case be reversed by legislation, such that the tortfeasor's liability to pay damages to the plaintiff should be unaffected by any duty on the plaintiff to repay said damages back to the tortfeasor.³⁶ Furthermore, they raised a potential concern regarding the application of the 'damages held on trust' approach to cases where future caregiving costs were being claimed for, namely that it was difficult for this model to take into account the uncertainty of whether care could continue to be provided in the future.³⁷

Criticisms 1 to 5 above will be examined in further depth below. Meanwhile, Section IV of this article will be devoted to conceptual difficulties regarding the trust model adopted by the English courts and will propose an alternative working model that addresses all the concerns raised.

A. PLAINTIFF-TORTEFEASOR CARE AGREEMENTS

As mentioned above, the undisputed contention that the plaintiff would have been able to recover the reasonable cost of gratuitously rendered services from the tortfeasor, had there been a bona fide contract between them, was not analysed in the House of Lords. Yet, as the Law Commission identified, Megaw LJ, in *Donnelly*; and Sir Thomas Bingham MR, in the Court of Appeal hearing of *Hunt*, had already expressed discontent about the undesirability of encouraging contracts for services with relatives.³⁸ Megaw LJ had found it 'repulsive' that a tortfeasor's liability should depend upon whether the victim had made a prior legally binding agreement with her family members that she would repay them for voluntary care.³⁹ Sir Thomas Bingham MR, citing Megaw LJ's judgment in *Donnelly*, stated that it would be 'regrettable' if the law encouraged plaintiffs to make contracts with those in the position of the defendant.⁴⁰ Roderick Doggett, approaching the problem from an insurance standpoint, wrote that it would bring the courts and litigants into a procedural muddle. Insurers wishing to argue against the victim-tortfeasor contract would require separate legal representation from their insured tortfeasor, who would want the victim to obtain the highest amount of compensation possible.⁴¹

³⁴ Law of Property Act 1925 (UK), ss 34, 36.

³⁵ Law Commission, *Damages for Personal Injury: Medical, Nursing and Other Expenses; Collateral Benefits* (Law Com No 262, 1999) ('Law Commission (No 262)') paras 3.62, 3.66.

³⁶ *ibid* para 3.76.

³⁷ *ibid* paras 3.56–3.59.

³⁸ Law Commission (CP No 144) (n 3) para 3.60; *Hunt* (CA) (n 22) 831; *Donnelly* (n 8) 463–64.

³⁹ *Donnelly* (n 8) 463–64.

⁴⁰ *Hunt* (CA) (n 22) 831.

⁴¹ Roderick Doggett, 'Hunt v Severs - A Pyrrhic Victory for Insurers?' (Quantum, 6 May 1995).

These criticisms may be viewed in the light of the fact that most victims, in the domestic sphere, would not think to form a legally binding contract with their relatives in order to make enforceable damages claimed against the tortfeasor for the costs of voluntary care. Often, leaving behind one's career to become a full-time caregiver is an act done out of a sense of moral and not legal obligation. Research has shown that voluntary caregiving may be motivated by feelings of duty, guilt, love and affection, a desire to reciprocate past help, social norms, altruism, and even egotism, all of which would not intuitively seem to accord with the idea of the parties setting out their mutual rights and obligations in a formal agreement.⁴² It is therefore arguable that the law should avoid requiring the added formality of a legally binding agreement between the parties where voluntary caregiving has been offered. This would avoid creating unnecessary complexity between the parties.

B. DISCOURAGING VICTIMS FROM ACCEPTING THE TORTFEASOR'S SERVICES

The second and third criticisms identified above by the Law Commission will be grouped under this shared heading. The victim may be encouraged to rely on relatives or friends other than the tortfeasor to provide voluntary care, or to eschew voluntary care entirely and seek commercial nursing care at a financial cost to herself.⁴³ Both of these outcomes may be considered undesirable. Their negative social impacts may run counter to the wishes or best interests of the victim and, in addition, they may create thorny evidential issues. Both of these drawbacks will be examined below.

(i) *The Victim's Best Interests*

The Law Commission stated that '[t]he plaintiff should receive gratuitous care from the most appropriate person free from tactical constraints regarding the award of damages'.⁴⁴ Similarly, Alan Reed noted, in *Hunt*, that the husband-tortfeasor *Severs* was best placed to provide care to *Hunt*, rather than a third party or another family member.⁴⁵

Admittedly, the case law and commentary in this area do not expand much on what makes the tortfeasor-caregiver the most tactically placed to provide voluntary care, but some relevant factors can be inferred. The tortfeasor-caregiver may be well-placed to provide voluntary care to the victim because of their physical proximity to the victim, the emotional closeness of their relationship with the victim (enabling them to understand best the victim's needs, preferences, and desires), their long history with the victim (giving them the best knowledge of the victim's medical history), and, most importantly, the fact that they are the only one who has voluntarily stepped up to provide care at all (making a search for other voluntary carers totally futile and potentially intrusive). The facts of *Hunt* show that the defendant lived together with the plaintiff both before and after their marriage and would certainly have been very well-acquainted with her medical needs, in addition to being willing and able to provide voluntary care.⁴⁶

⁴² Mikołaj Zarzycki and Val Morrison, 'Getting Back or Giving Back: Understanding Caregiver Motivations and Willingness to Provide Informal Care' (2021) 9 *Health Psychology and Behavioral Medicine* 636, 644.

⁴³ Law Commission (CP No 144) (n 3) paras 3.62–3.63.

⁴⁴ *ibid* para 3.62.

⁴⁵ Alan Reed, 'A Commentary on *Hunt v Severs*' (1995) 15 *OJLS* 133, 138.

⁴⁶ *Hunt* (CA) (n 22) 822–23.

In contrast, by denying a claim for the tortfeasor-caregiver's own voluntary expenditure, the law might have the unfortunate effect of driving the victim to seek care from other family and friends who might not have the willingness or means that the tortfeasor-caregiver does have.⁴⁷ It is questionable whether the law should leave the victim in such an uncertain position when a clear helping hand is available.

Commercial care, while it leaves the victim in the hands of professionals, may not necessarily be a more optimal solution. Commercial care may saddle the victim and their family with heavy costs that they will struggle to repay, placing them under financial stress. In a study by Maja Kuharic and others on the burdens that care recipients perceived themselves to have placed on their caregivers, financial burdens were a major concern, with participants including medical bills, medication costs, and transportation expenses as major financial harms alongside lost income from caregiving. The subjective appraisal of these objective financial burdens by the care recipients was found to lead to greater distress about their current and future financial situations.⁴⁸

Of course, the impact of these financial burdens should not be stretched too far. It is, of course, possible that victims may feel more comfortable in a commercial care facility where more of their needs are taken care of by professionals without having to saddle their relatives and friends with those obligations of care instead, especially if their costs can be covered by a claim for damages all the same. This appears to be supported by the fact that many dimensions of self-perceived patient burden stem from the physical, social, and emotional aspects of informal, voluntary caregiving.⁴⁹

Yet, even if the victim and their family are able to afford commercial care, it is not necessarily the case that the victim will prefer it to care voluntarily provided by a relative. Victims may prefer a familiar home environment to a clinical, unfamiliar, and potentially frightening hospital or nursing facility.⁵⁰ Although the results were mixed, a study has found that chronic illness patients displayed lower rates of anxiety and depression, higher satisfaction, better quality of life, and slightly better morbidity rates when placed in hospital-at-home programmes, compared to in-hospital stays.⁵¹ The findings from the hospice sector may also be instructive: Hospice UK found that a vast majority of hospice services were delivered at a patient's place of residence, as opposed to an inpatient unit, across the UK (although no reasons for opting for home care were provided).⁵²

(ii) Greater Damages for Insurers

The Law Commission further identified that pushing a victim to opt for commercial care, rather than voluntary care by the tortfeasor, would ironically also be disadvantageous to

⁴⁷ Law Commission (No 262) (n 35) para 3.71.

⁴⁸ Maja Kuharic and others, 'Care Recipient Self-Perceived Burden: Perspectives of Individuals with Chronic Health Conditions or Personal Experiences with Caregiving on Caregiver Burden in the US' (2024) 5 SSM - Qualitative Research in Health <<https://doi.org/10.1016/j.ssmqr.2024.100398>> accessed 12 December 2024.

⁴⁹ *ibid.*

⁵⁰ Carlos Echevarria and others, 'Home Treatment of COPD Exacerbation Selected by DECAF Score: A Non-Inferiority, Randomised Controlled Trial and Economic Evaluation' (2018) 73 Thorax 713, 717, 720.

⁵¹ Geneviève Arsenault-Lapierre and others, 'Hospital-at-Home Interventions vs In-Hospital Stay for Patients with Chronic Disease Who Present to the Emergency Department: A Systematic Review and Meta-Analysis' (2021) 4(6) JAMA Network Open <<https://doi.org/10.1001/jamanetworkopen.2021.11568>> accessed 19 March 2025.

⁵² See 'Statistics about Hospice Care in Each UK Nation' (*Hospice UK*, October 2024) <<https://www.hospiceuk.org/about-us/key-facts-about-hospice-care/UK-nations>> accessed 12 December 2024.

insurers, despite *Hunt* being ‘conventionally viewed as a victory for insurers’.⁵³ Rather than allowing insurers to avoid compensating victims for voluntary care rendered by a tortfeasor, as more patients chose to opt for commercial care, insurers would find themselves paying the higher commercial rates.⁵⁴

It is recalled that section 2(4) of the Law Reform (Personal Injuries) Act 1948 states that:

In an action for damages for personal injuries (including any such action arising out of a contract), there shall be disregarded, in determining the reasonableness of any expenses, the possibility of avoiding those expenses or part of them by taking advantage of facilities available under the National Health Service Act 2006 or the National Health Service (Wales) Act 2006 or the National Health Service (Scotland) Act 1978, or of any corresponding facilities in Northern Ireland.

Therefore, if the victim opts for private healthcare providers to meet their recovery needs, the court will be able to hold an insurer liable to pay for the costs of treatment from the private healthcare provider, notwithstanding that NHS care might have been available. Insurers will likely want to avoid this expense where possible.

(iii) Evidential Difficulties with Multiple Voluntary Caregivers

Reed identified that encouraging a wider range of family members and friends than the tortfeasor-caregiver to provide care would also lead to evidential difficulties.⁵⁵ If relatives and friends chip in to provide the victim with care of their own accord out of love and concern, in tandem with the tortfeasor, it may become hard to disentangle the tortfeasor from the rest of the victim’s sources of support.⁵⁶ There is thus potential for knotty evidential questions to arise regarding the responsibilities and contributions of individual caregivers in looking after the defendant.

C. THE PROBLEM OF CONTRIBUTION

An additional problem with the *Hunt* approach arises where the caregiving tortfeasor is not solely causative of the victim’s loss. David Kemp QC rightly noted that many road accidents involve passengers being injured by relatives or friends, with the driver being partly to blame.⁵⁷

One could picture a scenario where there are two defendants, who are equally to blame, and each holds 50 per cent responsibility: D1 and D2. Ordinarily, D1, if sued, would be entitled to recover 50 per cent of the total damages from D2 in contribution. However, D1 provides voluntary care to D2 and ceases to be liable to pay damages under *Hunt*. D1 can therefore not claim any contribution from D2 and must bear the entire cost of providing the

⁵³ Law Commission (CP No 144) (n 3) para 3.63.

⁵⁴ *ibid.*

⁵⁵ Reed (n 45).

⁵⁶ *ibid.*

⁵⁷ David Kemp, ‘Voluntary Services Provided by Tortfeasor to His Victim’ (1994) 110 LQR 524, 525–26.

care, whereas, if the victim had opted for professional care, D1 could recover 50 per cent of the care expenses from D2.⁵⁸ This result is unsatisfactory.

Likewise, if D2 is sued, D2 will not be able to recover any contribution from D1, having to bear the entire value of D1's voluntary care. This is the case even though D2 is only 50 per cent responsible—an excessively onerous consequence for D2. Either way, one of the defendants will suffer an unjust outcome.

Of course, if D2 is a far greater cause of the victim's injuries than D1 (for instance, if D2 is 90 per cent responsible), then the victim may simply choose to sue D2 directly and recover damages for the value of D1's voluntary care to be held on trust for D1. The victim is therefore not at risk of losing out on the value of D1's voluntary care. Meanwhile, the injustice to D2 is lessened, as the proportion of damages that cannot be attributed to their fault, but which they must still pay, is lower. There is therefore a smaller difference between the damages that D2 must pay and the amount that they are actually responsible for. Still, it should be questioned whether this is really satisfactory, if a lacuna in the attribution of compensation remains.

D. THE EXISTENCE OF INSURANCE

In personal injury claims, insurance frequently constitutes an 'invisible third' lurking behind the tortfeasor. An insurer is not party to the proceedings, which are between the plaintiff and the tortfeasor only. Yet, when they are involved, it is undeniable that the insurer finances the payment of damages by the tortfeasor. The English law of torts has traditionally taken an insurance-blind approach to determining a tortfeasor's liability. This was emphatically restated in *Hunt* by Lord Bridge: 'At common law the circumstance that a defendant is contractually indemnified by a third party against a particular legal liability can have no relevance whatever to the measure of that liability.'⁵⁹ The Law Commission echoed his view in their Consultation Paper:

Although the existence of insurance was a vital part of the factual picture, and indeed the litigation would scarcely have made sense if the defendant had not been insured against such liability, we do not consider that the existence of insurance should be permitted to affect liability. We believe that the question of the defendant's liability must necessarily precede the one of the defendant's insurer's liability.⁶⁰

The Law Commission were right to note that the litigation in *Hunt* would have been absurd had the tortfeasor not been insured. The damages would have proceeded from Severs's own finances to Hunt, and then back to Severs again—a clearly circular outcome. However, recognising the existence of insurance could have been a clear and realistic way of curing that absurdity and giving the otherwise circular direction of compensation an additional dimension.

Reed deemed it 'unrealistic' to dismiss the existence of insurance. He noted that, had the care been rendered by a third party or another relative, the insurer would simply have to

⁵⁸ Law Commission (CP No 144) (n 3) para 3.64; *ibid* 526.

⁵⁹ *Hunt* (n 1) 363.

⁶⁰ Law Commission (CP No 144) (n 3) para 3.65.

pay the costs of voluntary care as part of the overall damage award. On this basis, he argued that it was unjustifiable for the insurance company to reach an ‘undeserved benefit’ just because it was the tortfeasor who provided the care voluntarily, when that same tortfeasor had paid ‘handsomely’ through insurance premiums.⁶¹ Reed further argued that public policy militated against ‘curtail[ing] the ordinary willingness of one human being (whether a tortfeasor or not) to care for another’, with the ‘main beneficiary’ being ‘an unaware insurance company’.⁶²

IV. CONCEPTUAL ISSUES WITH THE TRUST MODEL AND A PROPOSED NEW WORKING MODEL

A. DIFFICULTIES WITH THE TRUST MODEL

Given the complexity of the trust-related criticisms of *Hunt*, a full section will be dedicated to examining these issues. This section will also propose a different working model targeted at resolving a number of the problems with the original *Hunt* rule.

One problem faced by the ‘damages on trust’ model that *Hunt* had supported was categorising the exact type of trust involved. This problem had been carried over from Lord Denning’s conceptualisation of that model in *Cunningham*. In that case, Lord Denning had not made fully clear exactly what sort of trust he was referring to. None of the Lords in *Hunt* said anything to clarify that position.

Paul Matthews and Mark Lunney concluded that the trust involved was a constructive trust, imposed by law independently of the intentions of the parties.⁶³ However, it was not clear who exactly the beneficiary of the trust would be. Given that the voluntary provider of services was the one destined to receive the damages held on trust, the beneficiary of the trust would seem to be either (i) the specific third-party caregiver designated by name or (ii) any person in the role of the voluntary caregiver.⁶⁴

Neither of these formulations, however, appears to be fully satisfactory. Both cases make the provider of services, rather than the recipient, the beneficiary of a trust—an awkward and unusual position.⁶⁵ Furthermore, for subsequent or alternative caregivers to be considered would raise complex evidential issues, but allowing only the specific caregiver’s voluntary expenditure to be recoverable would be unfair in cases where the initial caregiver is no longer able to provide services and another caregiver takes their place.⁶⁶

However, it is submitted that the conceptual issues with imposing a constructive trust on the parties need not be as impenetrable as Matthews and Lunney suggest. The law does allow constructive trusts to be imposed as a remedy where one of the parties has made some performance that justifies the imposition of a trust. The imposition of the trust runs into little conceptual difficulty where it is done to reflect the value of a party’s contributions, rather than said contributions being part of the services to be performed under the trust itself.⁶⁷ This is the case in land law, where the factors to be considered in imposing a common intention

⁶¹ Reed (n 45) 140.

⁶² *ibid.*

⁶³ See Paul Matthews and Mark Lunney, ‘A Tortfeasor’s Lot Is Not a Happy One?’ (1995) 58 MLR 395.

⁶⁴ *ibid.* 401.

⁶⁵ *ibid.*

⁶⁶ *ibid.*

⁶⁷ *Stack v Dowden* [2007] UKHL 17, [2007] 2 AC 432 [69].

constructive trust may include the parties' contributions to the purchase and upkeep of the property and other expenses, such as bills.⁶⁸

The contexts may be different: the common intention constructive trust involves inferring shared intentions that a constructive trust would be created in relation to the property,⁶⁹ while voluntary care may not necessarily be provided with the intention on the part of the caregiver to recoup those expenses through a claim for damages. Yet, with the plaintiff having made a claim for damages, and with said claim including the costs of voluntary care, it can be said to weigh on the plaintiff's conscience that the costs of voluntary care should be returned to its provider. This would provide the basis for the imposition of a trust in this context.

With that said, there is a further difficulty with the trust model: a lack of clarity on the recovery of future caregiving costs. These difficulties were not considered in *Hunt*, where the tortfeasor's liability was dismissed outright, or in *Cunningham*, where the damages recovered were only for past care. In Scotland, statute provides that the plaintiff may claim expenses for past care⁷⁰ as well as for anticipated future care,⁷¹ but the injured person is only under an obligation to account to the voluntary caretaker for damages recovered in the case of past care.⁷² The exact way in which the trust model applies to the recovery of future care expenses is therefore likely to be a pertinent question for future courts.

It has been suggested that, instead of a trust existing between the plaintiff and the caregiver, the caregiver may have a restitutionary claim against the plaintiff for services rendered, on the basis that the plaintiff had freely accepted those services.⁷³ However, the law as relates to such a claim existing is uncertain, with no cases on the point. Any such claim in restitution would have to depend on the plaintiff-victim knowing (or at least being in a position to know) that the services were not being provided gratuitously or, alternatively, on an intention on the part of the plaintiff-victim to repay the caregiver.⁷⁴ Establishing either the required intention, or the required knowledge (or position of expected knowledge), though possible, may not be easy for the tortfeasor,⁷⁵ given that it will generally be known to both the care recipient and the caregiver that the care recipient will not be in a position to provide stable financial support during the period when they are under their care.

B. A PROPOSED ALTERNATIVE APPROACH

Here the author proposes an alternative formulation of the principles of *Hunt* that makes Severs liable to pay damages, consistent with past authority, without any reference to insurance. Severs should still be seen as being factually responsible and legally liable for Hunt's injuries. Severs certainly caused the accident and any other tortfeasor in his position would be found liable. Severs's liability to pay damages should not be seen as 'erased' just because he was both the tortfeasor and the caretaker. Severs simply completed the loop between the payment of damages and the caregiver's need for compensation of his own accord, by stepping into the role of the caregiver at the usual endpoint of that compensation. Yet, that

⁶⁸ *ibid*; *Jones v Kernott* [2011] UKSC 53, [2012] 1 AC 776.

⁶⁹ *Stack* (n 67) [60].

⁷⁰ Administration of Justice Act 1982 (UK), s 8(1).

⁷¹ *ibid* s 8(3).

⁷² *ibid* s 8(2).

⁷³ *Matthews and Lunnay* (n 63) 400.

⁷⁴ Law Commission (CP No 144) (n 3) para 2.29.

⁷⁵ *ibid*.

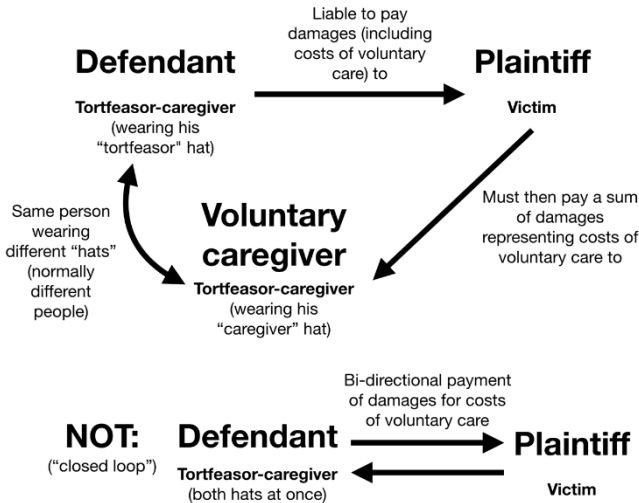
does not mean that the loop collapsed into a single point. Severs as tortfeasor, and Severs as caregiver, occupied two distinct roles in the traditional compensatory chain found in voluntary-carer cases.

The author’s working model therefore views a tortfeasor-caregiver, such as Severs, as wearing two ‘hats’: the hat of tortfeasor, and the hat of caregiver. This tortfeasor-caregiver puts on his ‘tortfeasor’ hat and steps into the tortfeasor’s position in the initial determination of liability to pay damages: he is liable to pay damages, just as any other tortfeasor would normally be. These damages include the costs of voluntary caregiving.

These damages go to the plaintiff, and the tortfeasor-caregiver then crosses over to the position of ‘caregiver’, like a batter moving between bases. He dons his ‘caregiver’ hat. He gets in position to catch the ‘ball’ thrown to him by the plaintiff-victim: the sum of damages awarded for the costs of his voluntary caregiving. The tortfeasor-caregiver is therefore able to receive the costs of his voluntary caregiving in the damages claimed, nominally, from himself (but, in practice, from his insurer).

There is still a chain of legal relations involved, and it is one in which compensation may flow, even if the ultimate destination of compensation is to return to its starting point: Severs himself. This chain of legal relations should not be closed into a ‘closed loop’ as a way of denying damages, when it is consistent with previous cases in which the defendant tortfeasor and the caregiver were different people. The important part of this reasoning is that the tortfeasor-caregiver does not wear both ‘hats’ at once, but rather one at a time, being consistent with older cases like *Cunningham* where the tortfeasor and caregiver were different people.

The author’s ‘compensation chain’ model is summarised by the diagram below:



Looked at in this way, it is hard to see how concerns about the existence of insurance affecting liability could have had any force in this case. Lord Bridge could have found room for Severs-as-tortfeasor to provide compensation down the chain to Severs-as-caregiver without insurance coming into the process of reasoning at all. Insurance coverage would then only become relevant after the case had been determined, when damages were being doled out, and, as

Reed properly noticed, the final result could have become much fairer: Severs could have relied upon his insurance to cover the costs of the voluntary caregiving he had rendered, through Hunt's claim.

How, then, does this model deal with the problems of the trust approach adopted in *Cunningham*? As explained above, imposing a constructive trust on the parties as a remedy seems to be a feasible option. This would establish a clear obligation for the plaintiff to repay the costs of voluntary care to the tortfeasor-caregiver. In the alternative, if the constructive trust formulation is not feasible, the loss claimed by the plaintiff can simply be considered to be in compensation for the plaintiff's needs as a whole, including all aid rendered by friends and relatives, as Teh J held in *Rajandran*.⁷⁶ This does not provide a legal obligation for the costs of voluntary care to be repaid to the tortfeasor-caregiver, but those costs can be repaid by the plaintiff on her own initiative.

As for the problems of 'future care', the proposed model could follow the above-mentioned Scottish statutory model in providing for damages for past and anticipated future care, but with the plaintiff only being under an obligation to account to the tortfeasor-caregiver for voluntary care that has already been provided. Any unfairness can be mitigated in practice by the fact that the plaintiff (or her litigation representative) can simply wait for her to return to a stable state of health before bringing a claim for damages.

It is accepted that the flow of compensation still appears to be circular and a little absurd if the proposed model of a 'compensation chain' is adopted. Certainly, it may not be the most intuitive metaphor at first sight. However, it is submitted that it is a logical approach to the problem of compensation, and a natural extension to *Cunningham*, which provided for the same chain of compensation (albeit without the caregiver and tortfeasor being one person wearing two 'hats').

V. RAJANDRAN: THE SINGAPOREAN COURT'S APPROACH

Similar facts to *Hunt* would emerge 30 years after that case in Singapore, in *Rajandran*'s case. This would provide the High Court of Singapore with the opportunity to reconsider the outcome and underlying reasoning behind *Hunt*.

Ms Rajina Sharma d/o Rajandran was a 32-year-old police officer. On 2 November 2016, she, like Hunt, had been riding pillion on a motorcycle with the defendant, Mr Mani, who was her husband. The second defendant had been riding in front of the couple on his own motorcycle when he skidded and fell from his motorcycle. Mr Mani, hoping to avoid a motorcycle, quickly applied his brakes, but still collided with the back of the second defendant's motorcycle, causing Ms Rajandran to be flung from Mr Mani's motorcycle. Ms Rajandran incurred a host of severe and debilitating injuries.

After Ms Rajandran was discharged from the hospital and returned home, Mr Mani took unpaid leave to care for her and resigned after four months of unpaid leave to care for her full-time. Ms Rajandran sought, by her litigation representative (who was also Mr Mani), to claim damages for Mr Mani's loss of earnings in ceasing work to care for her. Mr Mani and the second defendant were both covered by insurance.

In the High Court of Singapore, Teh J refused to apply *Hunt* to the case before her. Rejecting arguments of double compensation and circularity of payment, she awarded Ms

⁷⁶ *Rajandran* (n 3) [193].

Rajandran damages representing the cost of voluntary care by Mr Mani.⁷⁷ Teh J also did not encounter any conceptual difficulties in awarding Ms Rajandran damages for the costs of future care.⁷⁸

Notably, Teh J observed that Singaporean law had embraced the rationale in *Donnelly* in the case of *Lee Wei Kong v Ng Siok Tong*,⁷⁹ diverging from the House of Lords' preference for *Cunningham*. In that case, the Singapore Court of Appeal referred to *Donnelly* to hold that a plaintiff was entitled to claim for a caregiver's loss of income because it was the plaintiff's loss, 'being the reasonable cost of meeting the need created by the tort'.⁸⁰ The *Donnelly* rationale would subsequently find support in two Singapore High Court cases: *Christian Joachim Pollmann v Ye Xianrong*,⁸¹ and *AOD (a minor suing by his litigation representative) v AOE*.⁸²

With the *Donnelly* rationale in mind, Teh J was able to dispose of the concern of circularity of payment. She held that, since the award of damages was meant to be for the plaintiff's loss, it would return to the hands of Ms Rajandran to fund her needs and likely go towards aiding her family in caring for her, rather than directly back to Mr Mani's hands.⁸³ This was fair because an act of benevolent care like that of Mr Mani was done without any expectation of reimbursement.⁸⁴

Teh J likewise dismissed concerns of double recovery by holding that Mr Mani's gratuitous services could not be characterised as 'compensation or payment in the discharge of legal liability'.⁸⁵ She noted that Lord Bridge had recognised 'the fruits of the benevolence of third parties motivated by sympathy for the plaintiff's misfortune' as an exception to the rule against double recovery in *Hunt*,⁸⁶ and deemed Mr Mani's gratuitous services to be 'of the same nature'.⁸⁷ Like Lord Bridge, Teh J did not take the insurance position of either defendant into consideration in coming to her decision. However, she did acknowledge that 'the defendants' insurers will get a windfall by saving on a pay-out that they would otherwise be obligated to make', even if it was not a relevant factor to her determination of the payable damages.⁸⁸

VI. STRENGTHS AND WEAKNESSES OF THE SINGAPOREAN APPROACH

A. POTENTIAL WEAKNESSES

Certainly, the Singaporean approach is not perfect and it has its weaknesses. Although the *Donnelly* rationale has been well-accepted in Singapore, *Rajandran* does not eliminate its conceptual difficulties. It is true that the award of damages for the expenses of voluntary care

⁷⁷ *ibid* [205].

⁷⁸ *ibid* [162].

⁷⁹ [2012] SGCA 4.

⁸⁰ *ibid* [53] (Chao Hick Tin JA).

⁸¹ [2021] SGHC 77 [255].

⁸² [2015] SGHC 272 [89].

⁸³ *Rajandran* (n 3) [193]-[194].

⁸⁴ *ibid* [193].

⁸⁵ *ibid* [194].

⁸⁶ *Hunt* (n 1) 358.

⁸⁷ *Rajandran* (n 3) [194] (Teh J).

⁸⁸ *ibid* [205].

is meant to ‘fund the plaintiff’s needs’, as Teh J held,⁹⁰ but it is hard to see how it is not meant to address the expenses of the voluntary caregiver. This is the case even if the damages do ‘go towards aiding the family in their care for the plaintiff’, as she held, and not solely towards the voluntary caregiver.⁹⁰

For this reason, it is submitted that the author’s ‘closed loop’ model, in Section IV(B) of this article, potentially presents a better solution to the problem. Damages claimed for voluntary caregiving would flow from the tortfeasor-caregiver (or, in reality, his insurer) to the plaintiff and then back to the tortfeasor-caregiver. This could be expressed candidly and clearly without needing to hold that the amount of such damages was for care provided by the entire family and not just the voluntary caregiver. It is noted that Teh J’s explanation will be especially artificial where it is really just the voluntary caregiver who is providing care, and other family and friends are mostly absent.

On the other hand, this is purely a conceptual distinction, and not likely to affect the overall results of such a case, now that the Singaporean courts have accepted that a plaintiff may recover damages for the costs of voluntary care from a caregiver who is also the tortfeasor. Surely such a result will accord more with the natural sense of fairness of many plaintiffs and their caregivers. *Hunt v Severs*-type scenarios, after all, represent unique examples of cases where the plaintiff and tortfeasor are united in a common goal, and giving effect to the developments in *Rajandran* would allow that goal to be achieved.

One further weakness of the *Rajandran* approach falls to be considered here—its potential for harshness in cases where a tortfeasor is not covered by insurance and pays damages out of pocket. In such cases, the ‘invisible third’ that lies beyond the tortfeasor is absent; there is only a two-party relationship between tortfeasor and plaintiff. If the plaintiff proceeds against the tortfeasor for the costs of voluntary care in such a scenario, the tortfeasor will find himself paying out of pocket for the costs of his own voluntary care, construed as the plaintiff’s own loss. Because Teh J envisioned that ‘the damages will in all likelihood go towards aiding the family in their care for the plaintiff’ rather than ‘flowing back to the hands of the defendant’,⁹¹ there is a chance that the tortfeasor will not see the damages he has paid return to him.

However, we must be careful not to detach ourselves too far from practical reality. We must keep in mind, as the Law Commission noted, that *Hunt v Severs*-type litigation only makes sense at all if there is insurance behind the tortfeasor. Given that such litigation is also characterised by the tortfeasor and plaintiff having a common goal in the recovery of insurance monies, it hardly makes sense for the plaintiff to proceed against the tortfeasor if it makes them both worse off, while incurring inconvenience and legal costs along the way.

It must also be considered that, in the road-accident context from which both *Hunt* and *Rajandran* sprung, the nonexistence of insurance coverage is essentially a non-risk. Like in the UK, users of any motor vehicles in Singapore are required by law to be insured, under section 3 of the Motor Vehicles (Third-Party Risks and Compensation) Act 1960.⁹²

One could imagine cases where the tortfeasor initially voluntarily provides gratuitous care, but the tortfeasor and plaintiff’s relationship subsequently sours due to the pressures of caregiving, and the tortfeasor ceases to provide care. The plaintiff, with a little animosity, then brings proceedings against the tortfeasor, seeking to recover all the losses she has suffered, including past expenses of voluntary care and possibly the future expenses of being cared for

⁹⁰ *ibid* [193].

⁹⁰ *ibid*.

⁹¹ *ibid*.

⁹² Motor Vehicles (Third-Party Risks and Compensation) Act 1960 (Singapore), s 3.

by someone else. It could be hoped that caregivers in a *Hunt v Severs*-scenario are likely to commit to maintaining their relationship with the plaintiff, but this is not guaranteed. Studies appear to show that caregivers' motivations become less intrinsic and altruistic, and more egotistical and extrinsic, as they provide care over a longer period.⁸³ Indeed, some studies note that extrinsic motivations may lead to less positive outcomes than intrinsic motivations and that caregivers may find relief in relinquishing caregiving duties to third-party services.⁸⁴ The concern that this presents for *Hunt v Severs*-type cases appears to be mitigated primarily by the fact that, once again, it will be rare for the tortfeasor to be uninsured.

However, for the most part, proceedings in *Hunt v Severs*-type scenarios where the tortfeasor is not insured are likely to be extremely rare. Although the application of *Rajandran* to such scenarios has the potential to generate gross unfairness and give rise to greater fears of double compensation, it is submitted that such fears are hyperbolic and largely unfounded.

B. STRENGTHS OF RAJANDRAN'S APPROACH

Rajandran, on the other hand, has a number of notable strengths. By avoiding the 'damages held on trust' model of *Cunningham* entirely, it avoids invoking any conceptual difficulties with the trust approach. This includes the difficulty of recovery for future care. If we proceed from the starting point that recovery of the expenses of voluntary caregiving is designed to compensate the plaintiff's loss, then it is conceptually straightforward to proceed from that to recovery of expenses for the costs of future care. This is what Teh J was able to do, by evaluating the requested damages and refusing a discount for hired domestic care, and awarding the plaintiff SGD 400,848 for future caregiver expenses.⁸⁵

Similarly, problems with contribution simply do not arise at all on this approach, if the tortfeasor-caregiver can be made liable for the costs of their own voluntary care. Having been sued and found liable in the normal way, a tortfeasor-caregiver can then claim contribution from another tortfeasor, or the other tortfeasor can then claim contribution from the tortfeasor-caregiver. Allowing contribution to work as usual allows it to achieve its aim of distributing the losses caused by a tort fairly and proportionately, recovering from each wrongdoer their appropriate share. Notably, *Rajandran* was a two-defendant case of the exact type that Kemp had warned of—one where the tortfeasor-caregiver and another motorist were both responsible for the plaintiff's loss.

In addition, it is not ever suggested that any kind of legal relationship between the caregiver and the victim is required in order for damages to be claimed. Damages are claimable simply due to the existence of a need for care provision, which was then met by the voluntary caregiver satisfying the 'proper and reasonable cost' of that care.⁸⁶ Merely establishing that such a state of affairs exists at trial will be sufficient for damages to be claimable to remunerate the costs of that voluntary care. This is a far less intuitively objectionable approach than the requirement of a contract between tortfeasor-caregiver and victim which *Hunt* implicitly laid down.

Although insurance was not a major determinative factor in the outcome of *Rajandran*, its presence forms an implicit driving force behind Teh J's decision. Her willingness to

⁸³ Zarzycki and Morrison (n 42) 640.

⁸⁴ *ibid* 641, 651.

⁸⁵ *Rajandran* (n 3) [165].

⁸⁶ *Donnelly* (n 8) 462.

recognise a caregiver's expenses spent on gratuitous care as the fruit of the caregiver's benevolence, and not as compensable costs of caregiving, allowed her to evade the concerns of circularity of payment and double compensation.⁹⁷ However, these classifications do not make much sense in the absence of insurance.

Instead, Teh J's conception makes the most sense when read in the light of the existence of insurance behind Mr Mani. A finding of liability to pay damages would allow Mr Mani to discharge his legal liability, while supporting the fulfilment of his moral responsibility through his voluntarily care for her. Indeed, although she left that factual element aside, Teh J acknowledged once, and quite rightly, that the insurers would receive a windfall by 'saving on a pay-out that they would otherwise be obligated to make' if Mr Mani were not found liable.⁹⁸

It may be noted here that the law of tort does not typically consider factors like benevolence and the tortfeasor's willingness to atone, having a strictly compensatory function. Her statement that 'it will be wholly unfair for the plaintiff to be denied her claim only because Mr Mani has acted faithfully, and out of his love for his wife'⁹⁹ therefore marks an interesting and somewhat unorthodox step for the law of tort. It appears to be harsh and unbending at first glance, as well as inconsistent with her earlier statement that the claimed compensation 'should not be considered as flowing back to the hands of the defendant', given that he undertook the voluntary care without the expectation of payment or reimbursement.¹⁰⁰ However, by drawing a comparison with the existing exception to the rule against double-recovery of 'the benevolence of third parties motivated by sympathy for the plaintiff's misfortune',¹⁰¹ Teh J has reconciled *Hunt v Severs*-type cases with an existing exception recognised by the law of tort. Indeed, it makes little sense, once the 'circular compensation' problem is dispensed with, to draw a distinction between the benevolence of a third party and the benevolence of the tortfeasor in the law of damages.

It is therefore clear that *Rajandran* avoids all of the aforementioned problems with *Hunt*, while raising relatively inconsequential issues of its own. Such an outcome is clearly to be welcomed. It is, of course, not a doctrinally flawless solution, but surely those flaws can be accommodated by the Singaporean courts, in view of their longstanding support for *Donnelly* and the desirability of the ultimate outcome.

VII. CONCLUSION

Recovery of damages for the expenses of voluntary care incurred by a caregiver represent an anomaly in the law. Such recovery provides a rare example in tort law of damages recovered, not for loss suffered directly by the plaintiff, but by a third party. As the clash between *Cunningham* and *Donnelly*, and the resulting difficulties from the concept of the plaintiff holding damages on trust for the caregiver, have shown, the law has struggled to formulate an adequate rationale to permit plaintiffs to recover this head of damages.

The tension came to a head in *Hunt*, a case that turned upon the determination of which rationale was the more appropriate one in order to arrive at a principled solution where the tortfeasor was himself the voluntary caretaker. The House of Lords, having concluded

⁹⁷ *Rajandran* (n 3) [188], [193]-[194].

⁹⁸ *ibid* [205].

⁹⁹ *ibid*.

¹⁰⁰ *ibid* [193].

¹⁰¹ *ibid* [194], citing *Hunt* (n 1) 358 (Lord Bridge).

that it made more sense for the recovery of this head of damages to be based on compensating the voluntary carer, then foundered in the next step of their reasoning: the application of this principle to knock out the claim as an absurd one.

As analysed above, while the *Cunningham* rationale was coherent and sensible, its application in *Hunt* was erroneous and certainly not the only way that the Lords could have approached the problem. As proposed above, an alternative formulation that saw the circular chain of compensation in its entirety, and as not having collapsed into a ‘closed loop’, could have produced a fairer outcome. Instead, *Hunt* spawned further difficulties of its own, while trivialising, disregarding, and even disincentivising the benevolence behind the decision of a voluntary caregiver to provide gratuitous care to an accident victim suffering from debilitating injuries.

The Singaporean approach, adopting a different formulation from that proposed by the author, continued on its long march away from *Cunningham* along the path of *Donnelly*. Facing a similar fact pattern to *Hunt*, it saw recovery for expenses of voluntary care as compensating the victim’s own loss and, following *Donnelly*, allowed such damages to be claimed when the tortfeasor was himself the caregiver. Such an approach still presents some logical incoherence and has the potential to generate unfairness in a limited range of circumstances. However, these fears are far outweighed by the increased logical coherence and greater fairness of the outcome of the case in other respects, avoiding the same pitfalls that *Hunt* created.

It remains to be seen whether the law of England and Wales chooses to remedy these pitfalls. Legislation appears to be unlikely at the moment, and it is not certain when, or even if, a case similar to *Hunt* will once again arise for determination before the UK Supreme Court. Until an alternative model, such as that proposed by the author in this article, is adopted, the search for coherence and a more principled solution in this area of law will continue unabated, troubling future courts in more cases to come.

With Love and Affection: Rethinking the Fairness of Proximity of Relationship in Secondary Victim Claims

SANIYA MEHMOOD AKC*

ABSTRACT

Since the 1990s, English tort law has recognised the limited circumstances in which a claimant is owed a duty of care for psychiatric injury caused by an accident that they were not directly involved in. In particular, the eponymous *Alcock* control of ‘proximity of relationship’ mandates that a secondary victim must have a sufficiently close relationship with the participant in the accident (the primary victim). This article examines the fairness of proximity of relationship as a means of restricting claims in secondary victim cases, ultimately arguing that it requires reform. Section II provides a doctrinal analysis of *Alcock v Chief Constable of South Yorkshire Police* [1992] 1 AC 310, identifying three weaknesses within the judgment. Section III provides a quantitative analysis of the case law on secondary victims since *Alcock*, revealing how the courts have not always conformed to the requirements of proximity of relationship, as well as analysing the role of the ‘sudden shock’ requirement and gender stereotypes in exacerbating the unfairness of this control mechanism. Thereafter, Section IV explores and compares alternative avenues for reform of this control mechanism, including both conservative and radical changes.

Keywords: negligence, psychiatric injury, secondary victims, Alcock controls, proximity of relationship

I. INTRODUCTION

In the tort of negligence, for a defendant to be held liable for a wrong committed against a claimant, the claimant must prove that the defendant owed them a duty of care, that this duty was breached, and that the breach caused the harm suffered by the claimant (and that this damage was not too remote).¹ For a claim of pure psychiatric harm, the same must be proved, but claimants face additional hurdles as ‘secondary victims’. As opposed to ‘primary victims’, who are subject to less stringent controls on the basis that they were in the zone of danger and

* BA (King’s College London), MA Law (Conversion) (The University of Law). I am grateful to the reviewers for their thoughtful comments. Any remaining errors are my own.

¹ *Donoghue v Stevenson* [1932] AC 562 (HL); *Overseas Tankship (UK) Ltd v Morts Dock & Engineering Co Ltd (The Wagon Mound)* [1961] AC 388 (PC).

directly involved in the accident,² secondary victims are classified as such because they suffered psychiatric injury as a result of witnessing the injury, death, or imperilment of the primary victim.³ Thus, although secondary victims must make a separate claim, their victimhood is dependent on there being a primary victim, in that their (recognisable) psychiatric injury must be caused by witnessing the harm suffered by the primary victim or its aftermath.⁴

Whilst the recoverability of damages for *consequential* psychiatric harm—that is, psychiatric injury arising as a result of the tortfeasor’s breach of duty which caused physical injury to a claimant—is readily accepted by the courts, a standalone claim for psychiatric injury is more fraught.⁵ This is because of the imposition of policy-driven control mechanisms by the House of Lords in the leading case of *Alcock v Chief Constable of South Yorkshire Police*.⁶ It is important to note that there were at least 150 similar claims made at the time of the initial proceedings, and the potential for countless more as millions of people watched the Hillsborough tragedy unfold via televised broadcasts.⁷ Their Lordships were therefore concerned with strengthening the existing checks (i.e. reasonable foreseeability) on those bringing a claim as a secondary victim,⁸ most likely in an attempt to prevent the opening of the ‘floodgates’, as suggested by Lord Oliver’s fear of extending the law ‘in a direction... in which there is no logical stopping point’.⁹ Indeed, as the Supreme Court recently noted in *Paul v Royal Wolverhampton NHS Trust*,¹⁰ the control mechanisms were informed by ‘practical and policy considerations rather than purely analogical development of the law’.¹¹

Thus, in *Alcock*, 10 plaintiffs sought damages for psychiatric harm as a result of witnessing (either at the time or after) their loved ones perishing in the Hillsborough tragedy. In their judgment, their Lordships set down certain controls to limit the class of persons who could bring such claims. In doing so, they required that a claimant must enjoy a sufficiently proximate relationship to the victim, be proximate in time and space to the event or its immediate aftermath, and have suffered a sudden shock upon witnessing or hearing the event or its immediate aftermath with their unaided senses.¹² This was all in addition to the usual requirement of reasonable foreseeability, which is modified for the secondary victim so that it must be foreseeable that the injury would have been suffered by a person of ‘normal nervous strength’¹³ or ‘customary phlegm’.¹⁴ By contrast, a primary victim is not required to prove any of the above, merely needing to satisfy the court that the *risk* of injury was reasonably foreseeable.¹⁵

The *Alcock* controls have long been criticised for leaving the law on secondary victims in an unsatisfactory state.¹⁶ However, instead of presenting a holistic analysis of the controls, as many publications have done, this article is instead solely concerned with the fairness of

² *Page v Smith* [1996] AC 155 (HL) 184 (Lord Lloyd).

³ Law Commission, *Liability for Psychiatric Illness* (Law Com No 249, 1998), paras 2.12, 2.19.

⁴ *McLoughlin v O’Brian* [1983] 1 AC 410 (HL) 431 (Lord Bridge).

⁵ Law Commission (n 3) paras 2.13–2.14.

⁶ [1992] 1 AC 310 (HL).

⁷ *ibid* 316.

⁸ *ibid* 396 (Lord Keith).

⁹ *ibid* 416 (Lord Oliver).

¹⁰ [2024] UKSC 1, [2024] 2 WLR 417.

¹¹ *ibid* 434 (Lord Leggatt and Lady Rose).

¹² *Alcock* (n 6) 402–05 (Lord Ackner).

¹³ *Bourhill v Young* [1943] AC 92 (HL) 109 (Lord Wright).

¹⁴ *ibid* 117 (Lord Porter).

¹⁵ *Page* (n 2) 197 (Lord Lloyd).

¹⁶ See for example Law Commission (n 3) para 1.3.

proximity of relationship as a control mechanism which, it is argued, has been overlooked by scholars. Though Peter Handford argues that proximity of relationship ‘has assumed centre stage’ in the case law in recent years, relegating the roles of proximity of perception and time and space to an afterthought, this article contends that it has *not* received the scholarly attention befitting its role as a control mechanism.¹⁷ For example, one of the most scathing critiques of proximity of relationship remains one of the earliest,¹⁸ yet the fact that it forms a mere paragraph within an article on the landscape of tort law indicates how this control mechanism deserves more attention. Other works have sought to analyse proximity of relationship within the wider context of the *Alcock* controls or have devoted more attention to the other proximities.¹⁹ The exception is, of course, Handford’s analysis of proximity of relationship within a book dedicated to pure psychiatric harm.²⁰ However, a fresh perspective is needed on the matter given that the majority of secondary victim claims have been heard in the near-20 years since the publication of Handford’s book.

This article therefore seeks to gauge the extent to which proximity of relationship can be said to be a fair control mechanism through both doctrinal and quantitative analyses of the case law. This article submits that fairness should be measured in two ways: first, by its depth in the case law (that is, the extent to which it is actively considered by the courts); and secondly, by whether it has proved effective in restricting the class of claimants in line with the rationale of *Alcock*. For example, if the courts have spent little time devoted to this control mechanism, and it has not prevented more remote relations from bringing forth claims, then it is submitted that the continued existence of proximity of relationship as a control mechanism is unfair and needs reform.

This article is therefore structured in the following way. Section II analyses *Alcock* and highlights three concerns with their Lordships’ approach of imposing proximity of relationship as a control mechanism. These are the following: the existence of what we call ‘prima facie presumptions’ (that is, categories of relationship that are rebuttably presumed to have ‘close ties of love and affection’²¹); the requirement for claimants to provide evidence of their love and affection for other categories of relation; and their Lordships’ contradictory positions on the recoverability of damages by bystanders. Thereafter, Section III analyses the case law using a sample of 30 cases heard between 1991 and 2024. In doing so, this article highlights how proximity of relationship has played a limited role, in that reference to it has been omitted from most judgments and it has not prevented certain relations other than parents and spouses from being deemed sufficiently proximate. Finally, Section IV sets out the arguments for, and evaluates the strengths of, different types of reform, considering, in particular, the impact of the Supreme Court’s judgment in *Paul* on the role of the control mechanisms in secondary victim cases.

¹⁷ Peter Handford, *Mullany & Handford’s Tort Liability for Psychiatric Damage* (2nd edn, Lawbook Company 2006) 188.

¹⁸ In the form of Jane Stapleton, ‘In Restraint of Tort’ in Peter Birks (ed), *The Frontiers of Liability*, vol 2 (OUP 1994) 95.

¹⁹ See for example Bela Bonita Chatterjee, ‘Rethinking *Alcock* in the New Media Age’ (2016) 7 *Journal of European Tort Law* 272; Lee Eng Beng, ‘Vindication of the Three Proximities: *Alcock & Ors v Chief Constable of the South Yorkshire Police*’ (1992) *Singapore Journal of Legal Studies* 528; David Oughton and John Lowry, ‘Liability to Bystanders for Negligently Inflicted Psychiatric Harm’ (1995) 46 *Northern Ireland Legal Quarterly* 18.

²⁰ Handford (n 17) ch 7.

²¹ *Alcock* (n 6) 397 (Lord Keith).

II. SETTING THE PRECEDENT: *ALCOCK* AND ITS INHERENT FLAWS

This section focuses on the principles governing proximity of relationship that arose from *Alcock*, exploring the origins of this control mechanism. In particular, it critiques three aspects of this control mechanism: the existence of prima facie presumptions; the evidential burden; and the treatment of bystanders.

However, before commencing this critique, it is necessary to discuss the case of *McLoughlin v O'Brian*,²² as this was the leading authority on secondary victim claims before *Alcock*. In this case, which concerned the recoverability of damages by a woman who suffered psychiatric harm from hearing about the imperilment of her family and later witnessing their injuries, the House of Lords unanimously found in favour of the plaintiff, but for varying reasons. On the one hand was Lord Wilberforce's test—an early variation of the control mechanisms in *Alcock*—which represented the majority view. In his judgment, the recoverability of damages for 'shock' would be limited by reference to 'the class of persons whose claims should be recognised; the proximity of such persons to the accident; and the means by which the shock is caused'.²³ On the other hand was Lord Bridge's simple test of reasonable foreseeability: rejecting the 'arbitrary'²⁴ factors put forward by Lord Wilberforce, Lord Bridge argued (in the minority) that whether a duty of care was owed to a victim of psychiatric harm rested upon whether the harm was reasonably foreseeable.²⁵

Indeed, subsequent judgments appeared to favour Lord Bridge's approach over that of Lord Wilberforce.²⁶ For example, in *Attia v British Gas plc*, the Court of Appeal relied on the reasonable foreseeability test to find in favour of the plaintiff, who sought damages for psychiatric harm that stemmed from witnessing her house burn down.²⁷ In particular, Bingham LJ viewed Lord Wilberforce's test as being conditioned on satisfying certain 'grounds of policy', but did not view this 'suggested boundary line' as being 'fair or convenient'.²⁸ Thus, discussion of proximity of relationship, for example, was circumented in favour of disparaging the floodgates argument (perhaps because, if this control *was* considered, it would have been impossible for the plaintiff to succeed, as one cannot have a relationship with one's home).²⁹ In this regard, we can see how, at the time of *Alcock*, the control mechanisms were not concrete, which suggests that the House of Lords could have adopted a more flexible—and fairer—approach than the one ultimately taken had it not been for their concern with 'grounds of policy', such as limiting the number of potential claims.

²² *McLoughlin* (n 4).

²³ *ibid* 422 (Lord Wilberforce).

²⁴ *ibid* 442 (Lord Bridge).

²⁵ *ibid* 441–42.

²⁶ *Beng* (n 19) 530.

²⁷ [1988] QB 304 (CA) 312 (Dillon LJ).

²⁸ *ibid* 319–20 (Bingham LJ).

²⁹ *ibid* 320 (Bingham LJ).

A. PRIMA FACIE PRESUMPTIONS AND THE EVIDENTIAL BURDEN

The case law before *Alcock* was skewed towards recognising only two types of relationship: marital and parental.³⁹ For example, in *Hambrook v Stokes Brothers*,³¹ the plaintiff brought a claim under the Fatal Accidents Act 1846 following the death of his wife, which was allegedly caused by the nervous shock she suffered from witnessing the defendant, who was driving a lorry, ‘coming rapidly round the bend in her direction’, and from seeing her daughter, who *had* been hit by the lorry, in hospital.³² Likewise, in *Bourhill v Young*,³³ the appellant was an eight-months-pregnant woman who suffered nervous shock as a result of witnessing a collision between a motorcyclist and a car and later suffered a stillbirth. In *King v Phillips*,³⁴ a mother had suffered nervous shock from witnessing, through her window, a taxi reverse into her son;³⁵ similarly, in *Boardman v Sanderson*,³⁶ a father had suffered nervous shock after seeing his son’s foot trapped under the wheel of car, which had reversed into him.³⁷ The relevant relationship in *Hambrook*, *Bourhill*, and *King* was that between a mother and child, whilst in *Boardman*, it was that between a father and child. Meanwhile, in *Hinz v Berry*,³⁸ the plaintiff sought damages for nervous shock arising from witnessing the death of her husband and injury to her children, with Lord Denning MR remarking that such damages were available ‘to a close relative’.³⁹ Finally, in *McLoughlin*, Lord Wilberforce commented that, while the law recognised claims of ‘the closest of family ties’, it denied those of ordinary bystanders.⁴⁰

In this context, it is clear why the House of Lords sought to impose the presumption that only parental and marital (including betrothal) relationships were sufficiently proximate. As Donal Nolan argues, the ‘retreat from *Anns*’ and policy concerns about opening the floodgates dictated that their Lordships codify the preceding chain of authorities in a fairly conservative judgment.⁴¹ Thus, in *Alcock*, Alexandra Penk had a close tie of love and affection with her fiancé, as did Agnes and Harold Copoc with their son, while Brian Harrison did not have such a tie with his two brothers, nor did Robert Alcock with his brother-in-law.⁴² In defence of *Alcock*, then, its treatment of proximity of relationship was merely the consequence of attempting to clarify an uncertain area of law in the context of a tragedy viewed by millions and the potential for countless claims.

However, although their Lordships were at pains to emphasise that more remote relationships should be examined on a case-by-case basis,⁴³ it is argued that the imposition of the prima facie presumptions (and, by extension, the evidential burden) was unnecessarily restrictive. By way of example, one wonders why love and affection are more inherent in a

³⁹ *Alcock* (n 6) 411 (Lord Oliver).

³¹ [1925] 1 KB 141 (CA).

³² *ibid* 142.

³³ *Bourhill* (n 13).

³⁴ [1953] 1 QB 429 (CA).

³⁵ *ibid* 437.

³⁶ [1964] 1 WLR 1317 (CA).

³⁷ *ibid* 1319.

³⁸ [1970] 2 QB 40 (CA).

³⁹ *ibid* 42.

⁴⁰ *McLoughlin* (n 4) 422.

⁴¹ Donal Nolan, ‘Alcock v Chief Constable of South Yorkshire Police (1991)’ in Charles Mitchell and Paul Mitchell (eds), *Landmark Cases in the Law of Tort* (Hart Publishing 2010) 304–05.

⁴² *Alcock* (n 6) 398 (Lord Keith).

⁴³ See for example *ibid* 404 (Lord Ackner), 415 (Lord Oliver).

relationship of betrothal than in a sibling bond. At the very least, it suggests that the lifespan of a relationship is not relevant: for example, Ms Penk knew her fiancé for four years, compared to a childhood (presumably) spent by Mr Harrison with his brothers. In the author's view, the length of time one spends in a relationship with the primary victim should have been considered relevant by their Lordships when they imposed the prima facie presumptions. Indeed, this was considered by Lord Oliver in obiter remarks referencing the hypothetical plaintiff who lived with the primary victim in an invalid marriage for 40 years. In his view, '[t]he source of the shock and distress in all these cases is the affectionate relationship which existed between the plaintiff and the victim'.⁴⁴ Further, marital and parental bonds can be just as susceptible to estrangement as the sibling bond, so it is unfortunate that the latter was excluded from their Lordships' assessment of prima facie close ties in an attempt to keep the floodgates closed to an influx of *potential* claims.

This is further emphasised by the fact that, in the decision at first instance, Hidden J treated the sibling relationship as being as proximate as marital and parental relations.⁴⁵ In his generous, yet paradoxically conservative judgment (as he closed the doors of recoverability on other types of relationship), Hidden J extended the realm of foreseeability to siblings on the basis that they are 'the human beings [plaintiffs] know best apart from their parents' due to the nature of an 'ordinary' family unit.⁴⁶ Lee Eng Beng agrees with this analysis, arguing that, being part of the nuclear family, the sibling relationship should not have been differentiated from the husband-wife and parent-child bonds.⁴⁷ Aside from the antiquated concept of the nuclear family, this approach was not so revolutionary that it necessitated overturning. Although Handford argues that not every sibling bond can be as deep as that between lovers, and warns against making 'class presumptions' of this kind,⁴⁸ it was illogical for the Court of Appeal to overturn the presumption of the sibling relationship whilst preserving that of the marital and parental relationships (albeit now making it rebuttable).⁴⁹ If the presumption was to be rebuttable anyway, this article submits that the Court of Appeal (and, later, the House of Lords) should have extended it to all types of relationship.

Despite this, a claimant being a sibling to the primary victim, or having any other type of relationship except a parental or marital one, does not automatically fail this control mechanism; instead, they must prove that their relationship was sufficiently proximate. Exactly how they can do this is unclear, as the House of Lords failed to outline the type of evidence that was required to prove a close tie of love and affection.⁵⁰ It is therefore both a potentially difficult and unnecessarily callous task for the claimant to measure their bond with the primary victim, and this article submits that their Lordships should have provided greater clarity on this matter to aid claimants.

Indeed, Jonathan Morgan argues that this evidential requirement is 'distasteful', and it is easy to see why.⁵¹ For one, it trivialises claimants' grief: a victim does not exist in isolation; they are a person, with relationships and bonds. It is common sense to foresee that there exists *someone* who would suffer devastation at the death or injury of a relation, as

⁴⁴ *Alcock* (n 6) 416 (Lord Oliver).

⁴⁵ Handford (n 17) 196.

⁴⁶ *Alcock* (n 6) 338.

⁴⁷ Beng (n 19) 534.

⁴⁸ Handford (n 17) 197–98.

⁴⁹ *Alcock* (n 6) 385 (Nolan LJ).

⁵⁰ Handford (n 17) 202.

⁵¹ Jonathan Morgan, *Great Debates in Tort Law* (Bloomsbury Publishing 2022) 234.

acknowledged by Lord Oliver.⁵² It is not extreme to suggest that a defendant should foresee that this devastation may convert into psychiatric injury—indeed, this was readily accepted by Lord Keith.⁵³ It is unclear, then, why, in the same breath, he sought to impose the requirement that such ties must be proved with evidence. If it is indeed ‘common knowledge’ that close bonds can exist between lovers, family members, and friends, and that the real risk of psychiatric harm is reasonably foreseeable, then this should be enough for the courts.⁵⁴ There is little to gain from requiring claimants to prove their bond, except merely adding yet another qualification to the policy box-ticking exercise. The other control mechanisms provide ample opportunity for the courts to restrict the ability of claimants to recover damages, as reaffirmed by the Supreme Court in *Paul*.

In these conjoined cases, the separate claimants were each close relatives of the primary victims. More specifically, in *Paul*, the claimants were the children of the deceased; in *Polnmeier v Royal Cornwall Hospitals NHS Trust*, they were the parents of the primary victim; and in *Purchase v Ahmed*, the claimant was the mother of the deceased.⁵⁵ However, this seminal Supreme Court decision was not concerned with whether proximity of relationship was satisfied; rather, the Court clarified the recoverability of damages by secondary victims in clinical negligence cases.⁵⁶ In doing so, Lord Leggatt and Lady Rose dismantled the requirement that the claimant suffer a sudden shock upon witnessing a horrifying event, arguing that *Alcock* had been misinterpreted.⁵⁷ As such, proximity of relationship, proximity in time and space, and proximity of perception were reaffirmed as the main control mechanisms. Nevertheless, the necessity of physical and temporal proximity cannot be understated, with the majority upholding the Court of Appeal’s judgment in *Taylor v A Novo (UK) Ltd*,⁵⁸ confirming that a claimant must be ‘present at the scene of the accident or its immediate aftermath’.⁵⁹ As a result, the ability to recover damages as a secondary victim is restricted only to circumstances where there has been an accident (defined as ‘an external event’ which causes or can cause injury⁶⁰). This restricts the recovery of damages by claimants who, for example, suffered psychiatric harm upon witnessing the deteriorating condition of the primary victim in a clinical setting, but were not privy to the event that caused the injury, as was such in cases like *Taylor v Somerset Health Authority*⁶¹ and *Liverpool Women’s Hospital NHS Foundation Trust v Ronayne*.⁶² Thus, the need to witness the accident itself is now the defining factor of secondary victim claims, which greatly limits the context in which one can claim damages.

The unfairness of proximity of relationship, as established in *Alcock*, is further highlighted by the fact that the subsequent case of *McCarthy v Chief Constable of South Yorkshire Police*,⁶³ arising from the same Hillsborough tragedy, recognised the relationship between the plaintiff and his two half-brothers as being sufficiently proximate. Although Handford argues that the relationship between Mr McCarthy and his half-brothers, on the one hand, and Mr

⁵² *Alcock* (n 6) 408–09.

⁵³ *ibid* 397.

⁵⁴ *ibid*.

⁵⁵ See *Paul* (n 10).

⁵⁶ *ibid* 456 (Lord Leggatt and Lady Rose).

⁵⁷ *ibid* 438–39 (Lord Leggatt and Lady Rose).

⁵⁸ [2013] EWCA Civ 194, [2014] QB 150.

⁵⁹ *Paul* (n 10) 446 (Lord Leggatt and Lady Rose).

⁶⁰ *ibid*.

⁶¹ [1993] PIQR P262.

⁶² [2015] EWCA Civ 588, [2015] PIQR P20.

⁶³ *McCarthy v Chief Constable of South Yorkshire Police* (QBD, 17 December 1996).

Harrison and his brothers and Mr Alcock and his brother-in-law, on the other, were ‘clearly distinguishable’ as the latter relationships were ‘much less close’,⁶⁴ it is difficult to accept this view. It is only because Mr Harrison and Mr Alcock failed to plead in the manner that the judgment required (as it was impossible for them to know that they were required to present evidence) that their relationships to the primary victims were not recognised.⁶⁵ As the House of Lords made its judgment based on the absence of evidence, it is impossible to adduce that Mr Harrison and Mr Alcock were less close with their brotherly relations than Mr McCarthy was with his. The fact that Mr Harrison attempted to find his brothers in the aftermath of the disaster and ‘stopped up all night waiting for news’ should have been evidence enough for the House of Lords.⁶⁶ As Jane Stapleton eloquently lamented: ‘Is it not a disreputable sight to see brothers of Hillsborough victims turned away because they had *no more* than brotherly love towards the victim?’⁶⁷

Therefore, in accepting that parental and marital relationships can be rebuttably presumed to involve close ties of love and affection, the House of Lords created a problematic hierarchy of relationships. By dividing claimants into those who did not need to produce evidence of their love and affection and those who did, their Lordships arguably imposed an arbitrary barrier upon claimants. The evidential requirement only burdens both claimants seeking to quantify their relationship to the primary victim and defendants seeking to rebut it. If, indeed, the ‘source of the shock and distress in all these cases is the affectionate relationship which existed between the plaintiff and the victim’,⁶⁸ then it is meaningless to enforce such presumptions and evidential requirements in an effort to keep the floodgates closed. Such requirements need reform.

B. BYSTANDERS

The logic of imposing proximity of relationship was that it could be ‘readily foreseen’ that a claimant would be more likely to suffer ‘nervous shock’ where they were particularly close and affectionate with the primary victim.⁶⁹ At the same time, their Lordships, in obiter remarks, did not find it necessary to exclude a ‘bystander unconnected with the victims’ from being ‘within the range of reasonable foreseeability’ where the accident was ‘particularly horrific’.⁷⁰ Although this was an improvement upon the previous case law, which merely contemplated the possibility of bystander recovery,⁷¹ it was contrary to the rationale of imposing proximity of relationship to limit the class of claimants. By acknowledging the recoverability of damages by bystanders in limited circumstances, *Alcock* suggested that, far from being a strict control mechanism, proximity of relationship was merely a factor that did not need to be applied in certain cases. For example, the evidential burden was made irrelevant to a bystander, as they were defined by their non-relation to the victim. Thus, if such a requirement could be circumvented by someone unrelated to the victim, then, in the author’s view, it is

⁶⁴ Handford (n 17) 204.

⁶⁵ Donal Nolan and Ken Oliphant, *Lumley & Oliphant’s Tort Law: Text and Materials* (7th edn, OUP 2023) 370–71.

⁶⁶ *Alcock* (n 6) 352 (Parker LJ).

⁶⁷ Stapleton (n 18) (emphasis in original).

⁶⁸ *Alcock* (n 6) 416 (Lord Oliver).

⁶⁹ *ibid* 422 (Lord Jauncey).

⁷⁰ *ibid* 397 (Lord Keith).

⁷¹ Oughton and Lowry (n 19) 21.

difficult to justify compliance with this control mechanism by family members, further emphasising its unfairness.

This apparent error was soon remedied by the Court of Appeal in *McFarlane v EE Caledonia Ltd.*⁷² In this case, the plaintiff, who had witnessed a series of explosions on the Piper Alpha oil rig, was held to be a mere bystander. Rejecting submissions that relied on the obiter remarks in *Alcock*, Stuart-Smith LJ refused to extend a duty of care to the plaintiff on the basis that this would ‘base the test purely on foreseeability’ and instead emphasised the necessity of proximity of relationship.⁷³ Thus, in mandating that a plaintiff must have close ties of love and affection with the primary victim, the Court of Appeal reinforced the evidential burden upon family members by excluding the recoverability of damages by unrelated bystanders (and, perhaps, their Lordships agreed with this view, as leave to appeal was refused). That this was recently reinforced by the Supreme Court in *Paul*, which clarified that recovery is limited to those ‘who were present at the scene, witnessed the accident and have a close tie of love and affection with the primary victim’,⁷⁴ indicates how proximity of relationship is firmly entrenched in the law.

Nevertheless, another problem with this stance on bystanders is that, akin to the presumptions discussed above, the obiter comments created a hierarchy of horror, in the sense that it would force the courts to determine whether an event was sufficiently horrific to justify the foreseeability of a claimant’s psychiatric injury. Lord Ackner envisaged that the duty could extend to a person witnessing a ‘petrol tanker careering out of control into a school in session and bursting into flames’.⁷⁵ Yet, when a similar catastrophe occurred on the Piper Alpha oil rig in *McFarlane*, the Court of Appeal did not find the events that the plaintiff had witnessed to be horrifying enough, namely the multiple explosions setting the oil rig ablaze, men jumping overboard, and the capsizing of a lifeboat carrying survivors. Instead, Stuart-Smith LJ shied away from the responsibility by remarking that horror is subjective and elicits varied responses from people.⁷⁶

Therefore, in embracing a more flexible approach to bystander claimants, the House of Lords in *Alcock* (unintentionally) undermined their own approach to proximity of relationship. Given the aforementioned critiques of the evidential burden and categorisation of relationships, this flexible stance on bystanders is not unwelcome; however, the fact that the obiter comments were subsequently ignored by the courts, then, demonstrates how a close tie of love and affection is deemed a legal necessity for secondary victims to recover damages. Bystanders, therefore, have limited means of recoverability. This article contends that, for those who have suffered psychiatric harm and who comply with all other requirements, this is both an unsatisfactory and unfair state of affairs and is in need of reform.

III. FOLLOWING THE PRECEDENT? THE COURTS AND PROXIMITY OF RELATIONSHIP

Despite the preceding criticism of proximity of relationship, it is clear that this control mechanism is entrenched within the law governing secondary victims’ claims for negligently caused

⁷² [1994] 2 All ER 1 (CA).

⁷³ *ibid* 14 (Stuart-Smith LJ).

⁷⁴ *Paul* (n 10) 456 (Lord Leggatt and Lady Rose).

⁷⁵ *Alcock* (n 6) 403.

⁷⁶ *McFarlane* (n 72) 5-6, 14.

psychiatric harm. Nevertheless, to measure its usefulness, this section seeks to quantify proximity of relationship by reference to the case law since *Alcock*. In particular, it seeks to test Donal Nolan and Ken Oliphant's claim that proximity of relationship has been applied more liberally than was first suggested in *Alcock*. They claim that the decisions arising from *McCarthy* (discussed in Section II), *Shorter v Surrey and Sussex Healthcare NHS Trust*,⁷⁷ and *RE (a child by her mother and litigation friend LE) v Calderdale and Huddersfield NHS Foundation Trust*⁷⁸ suggest the generous application of proximity of relationship, contrary to what was envisioned in *Alcock*, as the courts found that there was proximity of relationship in relation to a half-brother, sister, and grandmother, respectively.⁷⁹ Although the particulars of *Shorter* and *RE (a child by her mother and litigation friend LE)* will be discussed in more detail below, it is worth noting that, if this claim is true, it would point towards the unfairness of the control mechanism. This is because, if the courts are willing to accept more remote relationships, then it would be clear that parents and spouses are not the only relationships that are readily accepted as having close ties of love and affection, which calls into question the continued need for a test that imposes significant hurdles on claimants.

The most effective way of quantifying proximity of relationship is by collecting data on the types of claimants in secondary victim cases. To this end, this article has taken a sample of 30 cases, starting from *Alcock* and ending with *Paul* (see Appendix A for the full table of cases). The aim of this data is to measure whether the relationship between the claimant and victim in these cases was regarded as sufficiently proximate in order to gauge the fairness of this control mechanism, with reference to its depth in the case law and efficacy in restricting the class of claimants. The data sample is entirely random, involving a mix of both personal injury and clinical negligence cases. Many of these cases are taken from the LexisPSL 'Psychiatric Injury—Secondary Victims—Case Tracker',⁸⁰ which helpfully summarises both successful and unsuccessful secondary victim claims. Others have been found on Westlaw UK by using the search term 'secondary victims'.

What follows is a visualisation of the results and a discussion of the patterns gleaned from the case law. In addition to offering some general observations, the rest of this section will be dedicated to analysing the prevalence of the 'sudden shock' control mechanism and commenting on the role of gender stereotypes in influencing some of the judgments to emphasise the unfairness of proximity of relationship. As a summary, Figure 1 details the type of relationship between the claimant and the primary victim and their frequency in the case law. The x-axis represents the relationship, while the y-axis represents its frequency:

⁷⁷ [2015] EWHC 614 (QB), (2015) 144 BMLR 136.

⁷⁸ [2017] EWHC 824 (QB), (2017) 156 BMLR 204.

⁷⁹ Nolan and Oliphant (n 65) 371.

⁸⁰ LexisPSL, 'Psychiatric Injury—Secondary Victims—Case Tracker' (*Lexis+ UK*) <<https://plus.lexis.com/api/permalink/cbc93188-a24d-49a5-9804-d61deb1e2f90/?context=1001073>> accessed 26 June 2024.

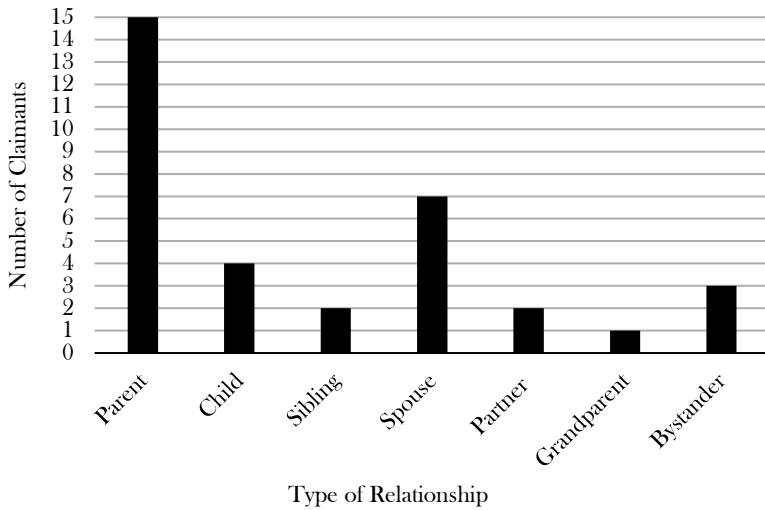


Figure 1: The claimant’s relationship with the primary victim.⁸¹

A. GENERAL OBSERVATIONS

It is clear that the long-recognised relationships have dominated the case law, with 15 claimants being the parent(s) of the primary victim, followed by seven spouses.⁸² This suggests that, far from the flexibility claimed by Nolan and Oliphant, the case law has stuck rigidly to the principles in *Alcock*. However, that most claims have been brought by parents and spouses begs the question of whether they have done so because—aside from the psychiatric harm that they have suffered—the existence of the prima facie presumptions makes it easier for them to pass this control mechanism. If so, it would serve to highlight the chasm between parents and spouses and more remote relations; though, of course, no firm conclusions can be drawn from this hypothetical.

However, although parental and marital relationships have been most frequent, this does not per se allow us firmly to conclude that proximity of relationship does not require reform. This is because, in the majority of cases, proximity of relationship was unproblematic for the claimant, in that their claim as a secondary victim was not negatively impacted by this

⁸¹ Where there are multiple claimants in the same action, with the same relationship to the primary victim (i.e. ‘parent’), they have been counted as one. Conversely, where there are multiple claimants but different relationships, or separate claims consolidated into one case, each relationship has been counted individually.

⁸² It should be noted that the case law is clear that, in the event of a stillbirth, the mother is the primary victim, as the child does not have a separate legal personality: see *Wild v Southend University Hospital NHS Foundation Trust* [2014] EWHC 4053 (QB), [2016] PIQR P3 [22]. As such, the claimants in such cases have been regarded as having a marital or partner relationship.

control mechanism. This can be explained by the fact that, at least where the prima facie presumptions were concerned, the defendant chose not to rebut them. However, even in cases where the relationship was more remote, proximity of relationship was largely accepted without question, being either not discussed at all or given a brief discussion. Instead, what has been more contentious in secondary victim claims is whether the claimant suffered a sudden shock that induced the psychiatric illness.

Before analysing this in more detail, however, it is worth discussing one case where proximity of relationship proved particularly problematic. In *Monk v PC Harrington Ltd*,⁸³ the claimant, a self-employed foreman, attempted to rescue fellow workmen from a construction accident. His claim as a rescuer (and therefore as a primary victim) was rejected. However, although the judge found him to be a 'secondary victim', it was determined that he was not entitled to recover damages. This is because, despite satisfying 'two of the three conditions', the claimant did not have a close tie of love and affection with the primary victims.⁸⁴ Likewise, it is noteworthy that the Court of Appeal in *McFarlane* failed to recognise proximity of relationship in the plaintiff's claim as a secondary victim, though the plaintiff's claim also failed on the ordinary fortitude test,⁸⁵ whereas, in *Monk*, it is clear that the claimant 'was a "secondary victim"' but for proximity of relationship.⁸⁶ Thus, the fact that *Monk* was the only case in the entire sample to fail solely because of proximity of relationship suggests the limited depth of this control mechanism as compared to the others, or, at least, that it has been considered a noncontentious issue. If so, then it is argued that to continue to recognise proximity of relationship as a valid control mechanism would be unfair due to the disproportionate role it plays in the case law.

Another case that elucidates the limitations of proximity of relationship is *RE (a child by her mother and litigation friend LE)*. There were multiple claimants in this case: RE; LE (the mother of RE); and DE (the mother of LE, and thus RE's grandmother). Both LE and DE sought damages for personal injury caused by nervous shock and, whilst LE was accepted as a primary victim, DE was limited to recovering damages as a secondary victim.⁸⁷ However, because LE was the primary victim, it is unclear whether DE's claim (which was successful) was defined by her relationship to her daughter or her granddaughter. Evidence that points to the former includes Goss J's statement that DE had 'a very close relationship with [LE]'; yet, he also points out that the defendant did not argue that 'she was not sufficiently close in terms of relationship to RE and to the event to be capable of being a secondary victim'.⁸⁸ Thus, this article has categorised this claimant as having both a 'parent' and 'grandparent' relationship in Figure 1 above.

More importantly, if DE's secondary victim claim was indeed characterised by her proximate relationship to her granddaughter, whom she knew for 15 minutes, then this elucidates the illogicality of the original reasoning in *Alcock* of including only parental and marital relationships in the prima facie presumption. The idea that a grandparent can have a close tie of love and affection for her grandchild, in a manner perhaps similar to that between a parent and child, indicates the ease with which proximity of relationship can be bypassed, especially as the defendant did not object to this. Although this is premised upon the condition that the

⁸³ [2008] EWHC 1879 (QB), [2009] PIQR P3.

⁸⁴ *ibid* [50] (G Leggatt QC).

⁸⁵ *McFarlane* (n 72) 14.

⁸⁶ *Monk* (n 83) [50] (G Leggatt QC).

⁸⁷ *RE (a child by her mother and litigation friend LE)* (n 78) 219.

⁸⁸ *ibid* 220.

relationship was grandparent-grandchild, rather than parent-child, it nonetheless points towards the possibility of the liberal application of proximity of relationship that was envisaged by Nolan and Oliphant. This also suggests the unfairness of maintaining a control mechanism that does not accurately reflect the reality of the case law, which, as is clear from the data, does not wholly object to recognising the proximity of remote relationships.

B. SUDDEN SHOCKS AND HORRIFYING EVENTS

As noted above, the crux of most of the sample was whether the claimant suffered a sudden shock during or in the immediate aftermath of a sufficiently horrifying event. In such cases, proximity of relationship was either not mentioned at all or was given the briefest of acknowledgements. For example, in *Brock v Northampton General Hospital NHS Trust*,⁸⁹ the claimants were the parents of the primary victim, who died of a brain haemorrhage while under the care of Queen Elizabeth Hospital (though the claimants sought damages from Northampton General Hospital NHS Trust, alleging negligence in its previous care of their daughter). They also sought damages for psychiatric harm; however, no explicit reference to proximity of relationship was made, despite the fact that observations were made on proximity of perception in time and space.⁹⁰ Instead, HHJ Yelton dismissed their claim on the basis that they suffered ‘mere’ grief instead of a sudden shock.⁹¹ While this could be attributed to the fact that the close ties of love and affection between parent and child were so obvious that they did not necessitate discussion, this article submits that even a brief acknowledgement of this control mechanism could have been made for clarity; the fact that there was no such acknowledgement therefore points to its limited relevance. Likewise, in *Tanner v Sarkar*,⁹² the claimant sought damages for psychiatric injuries after witnessing the death of her younger brother. The judgment made no reference to whether this relationship was sufficiently proximate, as the claim failed to establish that the claimant suffered a recognisable psychiatric illness. Nonetheless, the claim would have failed on the basis that there was no ‘shocking and horrifying event’, nor was there proximity in time and space.⁹³ However, the fact that the judgment made specific reference to proximity in time and space but not proximity of relationship, failing to provide clarity on this aspect, further emphasises this article’s claim that proximity of relationship has limited depth in the case law.⁹⁴

Such cases thus demonstrate the limited depth that proximity of relationship plays where the courts are concerned with establishing other control mechanisms. Though the requirement that a claimant’s illness must have been induced by a sudden shocking event has been criticised for being detached from ‘clinical reality’,⁹⁵ the fact that much of the courts’ time has been devoted to this requirement in secondary victim cases indicates its significance as a control mechanism—far more than proximity of relationship—due to its subjective nature.

As a contentious point of law, it is logical that the courts would devote greater time to the sudden shock requirement, which is emphasised by the high threshold that the courts have set themselves—nothing short of an ‘assault upon the senses’ which is ‘exceptional in

⁸⁹ [2014] EWHC 4244 (QB).

⁹⁰ *ibid* [79]–[87].

⁹¹ *ibid* [92].

⁹² *Tanner v Sarkar* (CC, 12 December 2016).

⁹³ *ibid* [80]–[81].

⁹⁴ *ibid*.

⁹⁵ Leonie Stüssi, ‘An Appeal for Abolishing the Sudden Shock Requirement’ (2022) 47 *Exeter Law Review* 44, 53.

nature' will suffice.⁹⁶ However, this article argues that proximity of relationship should have been regarded as being equally as contentious in those cases where the claimant did not have a marital or parental relationship to the primary victim, and therefore clearly did not meet the test set out in *Alcock*. Instead, the courts have sometimes omitted reference to the control mechanism, such as in *Tanner* when presented with a sibling relationship. Courts have also accepted proximity without question, such as in *Berisha v Stone Superstore Ltd*,⁹⁷ where there was no disagreement over the fact that the claimant 'was in a close relationship of love and affection' with the primary victim, despite their status as unmarried partners (where they were not even engaged so as to fall within the relationships accepted in *Alcock*).⁹⁸ This therefore points to the unfairness of proximity of relationship as a control factor because, despite the limited role that it plays in the case law as compared to other control mechanisms, it is still recognised as a valid check on claimants. Of course, now that the requirement that the claimant has suffered a sudden shock upon witnessing a horrifying event has been removed by the Supreme Court's judgment in *Paul*, it remains to be seen whether proximity of relationship will take a more prominent role in claims.

C. GENDER STEREOTYPES

Finally, it is important to recognise the role that gender stereotypes play in the case law, due to the historically detrimental—and thus unfair—impact that such stereotypes have had on women claimants. Although most cases in the sample made no reference to gender, there were two cases that raised an interesting point about gender stereotypes. For example, in *Shorter*, the claimant and primary victim were sisters. The evidence indicated that '[the] two sisters were particularly close to each other', and Swift J accepted that their relationship was sufficiently proximate.⁹⁹ Interestingly, he did so on the basis that their relationship was akin to being 'almost like mother and daughter', with the claimant playing the role of the mother.¹⁰⁰ This remark first suggests that only the most exceptionally close sibling bonds will be accepted as sufficiently proximate, in line with the reasoning in *Alcock*. However, it also suggests that the perceived power of maternalism simplifies the process of proving proximity of relationship for a woman.

To illustrate this further, take the case of *Berisha*, as outlined above. Though the claimant was the partner of the primary victim, particular reference was made to the fact that she was 'the mother of the Deceased's child'.¹⁰¹ That the learned judge felt it necessary to remark on this supports the view that perceptions of maternalism played a role in finding proximity, almost as if the claimant and primary victim's 'non-traditional' relationship was negated by the fact that she bore his child.

Of course, it would be remiss to draw broad conclusions from just two cases out of a sample of 30. However, one wonders whether the same decisions on proximity of relationship would have been reached had the claimants been men instead of women, and whether a paternal role would have been as readily accepted. The courts' historical views on gender in secondary victim cases have been analysed by Martha Chamallas and Linda K Kerber, in their

⁹⁶ *Ronayne* (n 62) P354–56.

⁹⁷ *Berisha v Stone Superstore Ltd* (CC, 2 December 2014).

⁹⁸ *ibid* [33].

⁹⁹ *Shorter* (n 77) 141 (Swift J).

¹⁰⁰ *ibid* 142.

¹⁰¹ *Berisha* (n 97) [33].

article on the ‘law of fright’.¹⁰² They note how the early case law on psychiatric harm was characterised by women claimants—mainly mothers—and, as such, the courts viewed their injuries through a perceived ‘biological’ lens, failing to attribute their injuries to anything other than motherly hysteria,¹⁰³ which meant that they failed the ‘customary phlegm’ test. Of course, societal views on gender have progressed in a way that means that few judges would think it appropriate to refer to a woman claimant as a ‘fishwife’¹⁰⁴ or to describe her as ‘timid and lacking in the motherly instinct’¹⁰⁵ for being fearful for her own—rather than her child’s—safety in the twenty-first century. Yet, both *Shorter* and *Berisha* suggest that traditional views on motherhood still hold some (perhaps subconscious) influence in the law governing the recovery of damages by secondary victims for negligently inflicted psychiatric injuries, which further highlights the unfairness of proximity of relationship, especially in relation to when claimants need to prove that their injury was foreseeable in a person of normal fortitude.

D. SUMMARY

As the data exemplifies, most claimants had marital or parental relationships with the primary victim, in line with the original vision in *Alcock*. On its own, this would suggest that proximity of relationship is working as intended and does not require reform. In the majority of cases, proximity of relationship has been unproblematic, with courts focusing on the other control mechanisms. However, it is also apparent that more remote relations, such as unmarried partners, siblings, and even grandparents, have been recognised as being sufficiently proximate. This is in addition to the limited conceptual role that proximity of relationship has played in the case law—only one case out of 30 failed solely because of proximity of relationship, with most judgments omitting reference to it and therefore accepting the close ties of love and affection without question—as the courts have preferred to devote much of their time to analysing the sudden shock requirement. Therefore, it cannot be said that there is depth to proximity of relationship with regard to its role in the case law, nor has it proven wholly effective in restricting the class of claimants in line with the reasoning in *Alcock*.

As such, this article submits that proximity of relationship is an unfair control mechanism that is in need of reform. The fact that the ratio of *Alcock* with regard to proximity of relationship has not been strictly adhered to by the courts substantiates this need for reform in order to better reflect this limited role and make the process of claiming as a secondary victim fairer. Furthermore, as the discussion on gender stereotypes suggests, the current existence of the prima facie relationships perhaps encourages the courts to take a traditional—and inherently unfair—view on gender roles that would otherwise not exist if claimants were free from the burden of proving their devotion to the primary victim. Therefore, it is argued that the aforementioned quantitative analysis proves the unfairness of this control mechanism and that it would be more satisfactory if claimants were free from its requirements through reform.

¹⁰² Martha Chamallas and Linda K Kerber, ‘Women, Mothers, and the Law of Fright: A History’ (1990) 88 Michigan Law Review 814.

¹⁰³ *ibid* 832, 837.

¹⁰⁴ *Bourhill* (n 13) 93.

¹⁰⁵ *Hambrook* (n 31) 151 (Bankes LJ).

IV. RESHAPING THE PRECEDENT: THE AVENUES FOR REFORM

The preceding sections sought to demonstrate—through both doctrinal and quantitative analysis—why proximity of relationship is a flawed and unfair control mechanism. Aside from the emotional callousness that this control mechanism presents, the data demonstrates the limited role that it has played in secondary victim cases. In this vein, it is argued that this control mechanism is in need of reform. This section will therefore outline and evaluate possible options of reform, including both conservative and radical changes, ultimately arguing that proximity of relationship should be wholly abolished to make the process of claiming as a secondary victim fairer.

A. ABOLITION

The abolition of proximity of relationship is a radical reform, especially in the wake of *Paul*, which has reaffirmed the role of this control mechanism. Nevertheless, it is worth considering this type of reform in order to better gauge the impact on claimants seeking to prove their secondary victimhood. This article therefore submits that the argument that the law would be better served if proximity of relationship were entirely abolished is premised upon two conditions: first, that tort law is based upon the principle of corrective justice; and secondly, that removing proximity of relationship would not per se make it easier for a claimant to recover damages for psychiatric harm.

First, it is broadly accepted that there are two types of justice in tort law: corrective justice and distributive justice. While both are restitutionary in nature, the former corrects the wrong done by the tortfeasor by holding them wholly and directly responsible for the victim's loss, whereas the latter is concerned with dividing the loss and compensation in a proportionate manner.¹⁰⁶ Looking at the case law, it may be argued that the law on secondary victims is based on distributive justice. For example, reflecting on the decision in *White v Chief Constable of South Yorkshire Police*,¹⁰⁷ Lord Steyn notes that both his and Lord Hoffmann's opinions 'invoked notions of distributive justice'¹⁰⁸ when they determined that the police officers were unable to recover damages based on the fact that, in *Alcock*, the 'bereaved relatives [were] sent away with nothing'.¹⁰⁹ In his opinion, Lord Hoffmann was at pains to distinguish the idealism of corrective justice, which would see the abolition of all the control mechanisms, and the pragmatism of distributive justice which, in his view, better reflected the realities of the tort of negligence, which sees many wrongs unremedied.¹¹⁰ This reluctance to compensate the few when so many have been denied compensation may help to explain the low success rate of secondary victim claims. However, such remarks only support the main point of this article: that there would have been no need to deny compensation to the claimants in *White* if the House of Lords had not imposed such arbitrary rules in *Alcock*. At the very least, drawing distinctions and denying liability between different classes of claimants offends the

¹⁰⁶ Ernest J Weinrib, 'Corrective Justice in a Nutshell' (2002) 52 *University of Toronto Law Journal* 349, 349–50.

¹⁰⁷ [1999] 2 AC 455 (HL).

¹⁰⁸ Lord Steyn, 'Perspectives of Corrective and Distributive Justice in Tort Law' (2002) 37 *Irish Jurist* 1, 5.

¹⁰⁹ *White* (n 107) 510 (Lord Hoffmann).

¹¹⁰ *ibid* 503–04.

interpersonal nature¹¹¹ of corrective justice, as it legitimises the denial of wrongdoing based on the identity of the victim.

Although this debate usually occurs within the context of causation (i.e. that a claimant can only pursue a claim against and receive recompense from the person who *caused* their injury),¹¹² there is no reason why it cannot be repeated here, within the realm of the tortious duty of care. If the courts, when applying the *Alcock* controls, bear in mind the obvious (yet overlooked) premise that ‘liability rectifies the injustice inflicted by one person on another’,¹¹³ then the law governing the recoverability of damages by secondary victims would no longer be characterised by arbitrary lines drawn between those who can and cannot claim. Abolishing proximity of relationship means that claimants would be treated as potential victims in their own right, instead of having their victimhood depend on their relationship to the person considered by the courts to be the ‘true’ (primary) victim. What this article means by this is that, though there would still need to be an original accident and fear of the death, injury, or imperilment of another person, a claimant would not need to prove the legitimacy of their case based upon whether they knew the primary victim in an intimate manner, as the sole focus would be on the remaining control mechanisms.

In one sense, this may not change the landscape of claims made by blood relations, friends, or lovers—it is because the claimant cares for the victim that they have suffered psychiatric harm and believe that it is worth pursuing a claim against their wrongdoer. The only difference is that they would no longer need to satisfy the courts as to the legal extent of this care and love, as the process of making a claim would be made fairer. This would drastically favour unrelated bystanders, as the test would then be whether it is reasonably foreseeable that *anyone* who witnessed the death, injury, or imperilment of the primary victim would suffer psychiatric harm, taking account of the surviving proximities of perception and time and space, and the normal fortitude requirement. Although they would still need to prove that the tortfeasor owed them a duty of care, a claimant could do so on the premise that they are an independent victim who has suffered harm and who deserves to be compensated for their loss, rather than on the premise that they suffered harmed *only because* they were proximate to the primary victim. Likewise, the notion that a defendant could do no harm to the claimant because they failed to prove their love and affection would become a fiction of the tort of negligence.

Of course, this begs the question of whether the pursuit of corrective justice would open the floodgates to litigation. As previously noted, the fear of floodgates motivated the House of Lords to impose the *Alcock* controls, with even Lord Oliver, who was particularly sensitive to the arbitrariness of proximity of relationship, remaining sceptical towards ‘pragmatic extensions’ of this control mechanism.¹¹⁴ Nevertheless, the floodgates argument has been repeatedly dismissed as being exaggerated, not only due to an absence of empirical evidence,¹¹⁵ but also because it is believed that ‘the courts are well capable of controlling any such flood of claims’.¹¹⁶ Thus, despite the Law Commission’s implication that proximity of relationship prevents the opening of the floodgates,¹¹⁷ it is argued that removing this control

¹¹¹ Allan Beever, ‘Corrective Justice and Personal Responsibility in Tort Law’ (2008) 28 OJLS 475, 476–79.

¹¹² Morgan (n 51) 53.

¹¹³ Weinrib (n 106) 349.

¹¹⁴ *Alcock* (n 6) 417 (Lord Oliver).

¹¹⁵ Chatterjee (n 19) 291–92.

¹¹⁶ *White* (n 107) 464 (Lord Griffiths).

¹¹⁷ Law Commission (n 3) para 6.10.

mechanism would not cause an avalanche of claims because claimants would still face significant hurdles before being able to recover damages. Even considering the Supreme Court's abolition of the sudden shock control in *Paul*, claimants would still need to satisfy proximity of perception in time and space, as well as to prove that the primary victim's death, injury, or imperilment was caused by an accident. In the author's view, this is sufficient to prevent an influx of secondary victim claims because, as discussed in Section II, the Supreme Court's judgment in *Paul* makes clear the necessity of the claimant witnessing the accident, which greatly limits—if not altogether bars—the recoverability of damages in clinical settings.

Further, this article observes that there already exists another check on the recoverability of damages by secondary victim: an informal, economic control mechanism that, though unfair to those with limited financial means, would also help to mitigate the feared opening of the floodgates. Given the scarcity of successful secondary victim claims, it is difficult to calculate an average of damages. Nonetheless, from the few reported cases we do have, it is suggested that the damages awarded for psychiatric injury may present another hurdle for claimants seeking to pursue a secondary victim claim. For example, the Quantum Report for *C (In Her Own Right and on Behalf of the Estate of X, Deceased) v An NHS Trust*¹¹⁸ reveals that the claimant, who brought a secondary victim claim after witnessing the deterioration and death of her son, was awarded £16,000 for her injury in an out-of-court settlement. Similarly, the claimant in *Walters v North Glamorgan NHS Trust*¹¹⁹ was awarded £16,000 (£34,659.03 RPI) by the court for her pain, suffering, and loss of amenity ('PSLA') caused by the death of her son.¹²⁰ Of course, the quantum of damages for psychiatric harm does not depend on whether the claimant is a primary or secondary victim, but on the severity of the injury, with the maximum award of damages for PTSD, for example, being £122,850, while the minimum is currently £4,820.¹²¹ Moreover, although a higher award of £110,000 in PSLA damages (£145,551.76 RPI) has been recorded,¹²² the aforementioned figures suggest that a claimant with 'moderate' injuries may not expect to be awarded a vast PSLA sum, especially accounting for legal costs and the arduous litigation process. Therefore, even if proximity of relationship is removed, it is unlikely that the floodgates would be opened, not least because claimants would still be inhibited by the financial risks associated with pursuing a claim. The unfairness of such a reality is important to note, but beyond the scope of this article.

Thus, although removing proximity of relationship would not necessarily alleviate the burdens placed on claimants due to the continued existence of other controls, it *would* remove an unfair barrier to the process of claiming as a secondary victim, as there would be no distinction between claimants in the sense that it would not matter whether one is, say, the mother or cousin of the primary victim. Following this, the law on the recoverability of damages by secondary victims would no longer be wholly based on arbitrary rules born from anxieties about opening the floodgates, but would instead align with the principle of corrective justice. To argue in favour of abolition, then, is to recognise that it is fairer to remove a

¹¹⁸ *C (In Her Own Right and on Behalf of the Estate of X, Deceased) v An NHS Trust* (Out of Court Settlement, 6 May 2021).

¹¹⁹ [2002] EWHC 321, [2003] PIQR P2.

¹²⁰ William Norris and others (eds), *Kemp and Kemp: The Quantum of Damages*, vol 4 (release 171, Sweet & Maxwell 2024) para 5-002.

¹²¹ Judicial College, *Guidelines for the Assessment of Damages in Personal Injury Cases* (17th edn, OUP 2024) para 7.2.

¹²² (1) *A (In His Own Right and as Administrator of the Estate of the Deceased, X)*, (2) *B*, (3) *C*, (4) *D v HT* (Out of Court Settlement, 5 June 2020).

redundant control mechanism to widen the pool of potential claimants and to recognise their claim to justice than to keep it in place as a restrictive and unnecessary check.

B. ALTERNATIVE PROPOSALS FOR REFORM

Although Lord Steyn famously remarked that ‘the law on the recovery of compensation for pure psychiatric harm is a patchwork quilt of distinctions which are difficult to justify’, he also rejected dismantling these controls on the basis that it would be contrary to precedent, which the courts are bound by.¹²³ Thus, despite the preceding analysis indicating that doing so would not make it easier for claimants to access compensation, it is highly unlikely that proximity of relationship would be abolished. For one, except for Parliament, only the Supreme Court has the legitimacy to abolish this control mechanism, which would require an appropriate case to make its way through the courts—an extremely time-consuming and potentially distressing process for claimants. Even then, it is not guaranteed that the court would find it desirable to abolish this control mechanism, especially as it was given the opportunity to clarify proximity of relationship in *Paul* and omitted to do so. Therefore, the abolition of proximity of relationship would depend on the goodwill of Parliament, which has thus far expressed little interest in legislating on behalf of secondary victims.

If outright abolition of proximity of relationship is such an affront to precedent that it would never occur, then it is wise to analyse alternative, less radical approaches to reform. However, it is argued that such approaches must address two key problems that were highlighted in Section II—namely, the evidential burden and the exclusion of bystanders—which may prove problematic.

First, this article has emphasised that the evidential burden places an unnecessary emotional onus on claimants and may have also dissuaded others from pursuing a claim, though client privilege prevents one from knowing the exact number of abandoned cases. Thus, any reform short of abolition must dismantle this requirement, which also means reforming the prima facie presumptions. This can be achieved in two ways. First, the prima facie presumptions could be expanded to encompass every class of familial relationship so that, no matter whether a claimant is the sibling, cousin, or grandchild of the primary victim, a close tie of love and affection will be rebuttably presumed. On the one hand, this reform would be liberal, in the sense that any relation through blood or marriage would be recognised, from stepchildren to second cousins. However, to balance such liberalism, the defendant would still be allowed to rebut this presumption should they believe a claim to be too egregious (though, as the passivity of the defence in *RE (a child by her mother and litigation friend LE)* suggests, the task of disproving a claimant’s love and affection is unappealing). In this way, proximity of relationship would be preserved largely in its current form through the rebuttable presumption, but it would still address the emotional insensitivity of the evidential burden in a palatable way.

Alternatively, the evidential burden could be removed through the introduction of a closed list of relationships, akin to the Law Commission’s recommendation of introducing a fixed list of spouses, parents, children, siblings, and cohabitants.¹²⁴ On the one hand, this would remove the need for claimants to evidence their love and affection because they would automatically (and irrefutably) satisfy proximity of relationship by virtue of being a relation

¹²³ *White* (n 107) 500 (Lord Steyn).

¹²⁴ Law Commission (n 3) paras 6.24–6.35.

included on the list. It would therefore introduce another layer of certainty to the law. On the other hand, this reform would perhaps do more harm than good as it would bar certain classes of claimant from recovering—even if they satisfy all the other controls—which was something that Lord Oliver warned against in *Alcock*.¹²⁵ Moreover, this would only further embed the exclusion of bystanders, and so it can be seen how prioritising the removal of the evidential burden would be to the detriment of bystanders.

However, a reform that allows bystander recovery would not necessarily remove the evidential burden. For example, although outright abolition would best serve the interests of bystanders, a more conservative means of doing so would be to analyse the merit of each claim on a case-by-case basis. This would remove the prima facie presumptions entirely, so that no type of relationship is afforded special treatment. At the same time, bystanders would not automatically be excluded, as it would be up to the courts to decide whether the claimant's relationship to the primary victim is sufficiently close.

The problem with this approach is that the evidential burden would still exist, albeit in a slightly altered form. The courts would require some type of evidence from the claimant as to their relationship with the primary victim, which would be as emotionally insensitive as the current approach—if not more, as parents and spouses would now also be required to prove their love and affection. Therefore, it is not clear how any reform short of abolition could satisfactorily balance the two (competing) interests of removing the evidential burden and allowing bystander recovery. It is likely that one must be sacrificed, which would still leave proximity of relationship in an unsatisfactory state. Ultimately, then, the inadequacy of less radical reform only confirms this article's conclusion that proximity of relationship should no longer be recognised as a legitimate control mechanism and supports the argument in favour of abolition.

V. CONCLUSION

This article has demonstrated that the *Alcock* control of proximity of relationship is no longer fit for purpose. Through both qualitative and quantitative analyses, this article has argued that proximity of relationship is an unfair control mechanism and that it is therefore in need of reform. In doing so, this article has advocated for the abolition of this control.

Section II analysed the House of Lords' judgment in *Alcock*, highlighting three flaws inherent to proximity of relationship. First, the prima facie presumption that marital and parental relationships are ones of love and affection is unnecessarily restrictive. Secondly, the existence of the evidential burden is distasteful as it trivialises claimants' grief and presents a further barrier to their ability to recover damages. Thirdly, their Lordships' position on bystander recovery was contradictory, in that it suggested that proximity of relationship need not always be required where there was an extreme catastrophe. The fact that later courts did not follow this dictum, instead excluding bystander recovery, only highlights the imprecision of this control.

With these criticisms in mind, Section III evaluated the usefulness of proximity of relationship by reference to the case law. By taking a sample of 30 cases, this article demonstrated that proximity of relationship has limited depth in the case law, being outright ignored in some judgments, and that it has not restricted the class of claimants in the manner envisioned in *Alcock*. As such, Section IV argued that, given its limited role in the case law,

¹²⁵ *Alcock* (n 6) 415.

proximity of relationship requires reform, most ideally through abolition. Removing this control would unburden claimants who are not the parents and spouses of the primary victim, thereby helping to make the process of making a claim as a secondary victim fairer than it currently stands. However, even though this is unlikely to open the floodgates due to the preservation of the other controls, it is unrealistic to expect this to come to fruition. Yet, by evaluating the efficacy of less radical reforms, it was clear that only abolition can adequately redress the shortcomings of proximity of relationship in a holistic (and therefore fair) manner.

Establishing a duty of care in the tort of negligence is a burdensome process for secondary victims. Especially considering the Supreme Court's reluctance to depart from settled law in *Paul*, it is even harder to reform troublesome hallmarks of the law governing the recoverability of damages by secondary victims like the *Alcock* controls.¹²⁶ Nonetheless, in highlighting the futility of this control, it is clear that proximity of relationship is in need of reform—if not outright abolition—in order to make the law on secondary victim claims more certain, clear, and fair.

¹²⁶ *Paul* (n 10) 449.

APPENDIX A

Case	Relationship to the Primary Victim
<i>Berisha v Stone Superstore Ltd</i> (CC, 2 December 2014).	Partner
<i>Brock v Northampton General Hospital NHS Trust</i> [2014] EWHC 4244 (QB)	Parent
<i>Calascione v Dixon</i> (1993) 19 BMLR 97	Parent
<i>Doosey v Leeds Teaching Hospital NHS Trust</i> (CC, 15 August 2017)	Spouse
<i>Fagan v Goodman</i> (QB, 30 November 2001)	Bystander
<i>Galli-Atkinson v Seghal</i> [2003] EWCA Civ 697, (2003) 78 BMLR 22	Parent
<i>King v Royal United Hospitals Bath NHS Foundation Trust</i> [2021] EWHC 1576 (QB), [2021] PIQR P20.	Parent
<i>Kitchen v Hillingdon Hospitals NHS Foundation Trust</i> (CC, 4 May 2022)	Spouse
<i>Less v Hussain</i> [2012] EWHC 3513 (QB), (2013) 130 BMLR 51	Partner
<i>Liverpool Women’s Hospital NHS Foundation Trust v Ronayne</i> [2015] EWCA Civ 588, [2015] PIQR P20	Spouse
<i>McFarlane v EE Caledonia Ltd</i> [1994] 2 All ER 1	Bystander
<i>Monk v PC Harrington Ltd</i> [2008] EWHC 1879 (QB), [2009] PIQR P3	Bystander
<i>O’Connor v Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust</i> (CC, 13 July 2018)	Child
<i>Owers v Medway NHS Foundation Trust</i> [2015] EWHC 2363 (QB), [2015] Med LR 561	Spouse
<i>Paul v Royal Wolverhampton NHS Trust</i> [2024] UKSC 1, [2024] 2 WLR 417	Child Parent

	Parent
<i>RE (a child by her mother and litigation friend LE) v Calderdale and Huddersfield NHS Foundation Trust</i> [2017] EWHC 824 (QB), (2017) 156 BMLR 204	Parent/Grandparent Parent
<i>Shorter v Surrey and Sussex Healthcare NHS Trust</i> [2015] EWHC 614 (QB), (2015) 144 BMLR 136	Sibling
<i>Sion v Hampstead Health Authority</i> (CA, 27 May 1994)	Parent
<i>Tanner v Sarkar</i> (CC, 12 December 2016)	Sibling
<i>Taylor v A Novo (UK) Ltd</i> [2013] EWCA Civ 194, [2014] QB 150	Child
<i>Taylor v Somerset Health Authority</i> [1993] PIQR P262	Spouse
<i>Taylorson v Shieldness Produce Ltd</i> [1994] PIQR P329	Parent
<i>Tranmore v T E Scudder Ltd</i> (CA, 28 April 1998)	Parent
<i>Walters v North Glamorgan NHS Trust</i> [2002] EWHC 321 (QB), [2003] PIQR P2	Parent
<i>Ward v Leeds Teaching Hospitals NHS Trust</i> [2004] EWHC 2106 (QB), [2004] Lloyd's Rep Med 530	Parent
<i>Wells v University Hospital Southampton NHS Foundation Trust</i> [2015] EWHC 2376 (QB), [2015] Med LR 477	Parent
<i>Werb v Solent NHS Trust</i> (QB, 15 March 2017)	Parent
<i>White v Lidl UK GmbH</i> [2005] EWHC 871 (QB)	Spouse
<i>Wild v Southend University Hospital NHS Foundation Trust</i> [2014] EWHC 4053 (QB), [2016] PIQR P3	Spouse
<i>Young v Downey</i> [2020] EWHC 3457 (QB)	Child