

THE PARADOX OF EXPLOITATION:
A NEW SOLUTION.

Benjamin Ferguson

A thesis submitted to the Department of Philosophy, Logic and Scientific Method
of the London School of Economics and Political Science
for the degree of Doctor of Philosophy, April 2013.
Revised and final draft submitted September 2013.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work. The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 84,557 words.

A handwritten signature in black ink, appearing to read 'Benjamin Ferguson', written in a cursive style.

Benjamin Ferguson

Abstract

In this thesis I present a rights-based theory of exploitation. I argue that successful conceptions of exploitation should begin with the ordinary language claim that exploitation involves ‘taking unfair advantage’. Consequently, they must combine an account of what it means to take advantage of another with an account of when transactions are unfair. Existing conceptions of exploitation fail to provide adequate accounts of both aspects of exploitation.

Hillel Steiner and John Roemer provide convincing accounts of the unfairness involved in exploitation, but because they fail to provide restrictive ‘advantage taking’ conditions, their theories of exploitation include cases that we would not ordinarily describe as exploitations. Ruth Sample and Robert Goodin provide accounts that place a stronger emphasis on the attitudes involved in exploitation and the processes that bring it about. Unfortunately, these theories do not properly incorporate the unfairness aspect of exploitation. Consequently, they are either self-frustrating or incomplete.

I provide a conception of exploitation that combines both aspects. I claim that a bilateral, strictly Pareto improving, and voluntary transaction ϕ between A and B is exploitative just in case (1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that (2) reduces B’s bargaining power by altering her status quo point, and (3) the degree to which B’s preferences are satisfied in ϕ is less than it would be had B’s bargaining power not been reduced (by the alteration of her status quo point). Further, (4a) A’s egoistic preferences are better satisfied in ϕ because B’s bargaining power is limited, and (4b) if the previous conditions also hold for A, then the reduction in A’s bargaining power is less than the reduction in B’s bargaining power. Finally, (5) A truly believes the previous conditions hold.

Contents

	Page
List of Cases	7
List of Tables and Figures	8
Acknowledgements	9
I What Is Exploitation?	10
1 Introduction: The Paradox of Exploitation	11
1.1 Chapter 1 Introduction	11
1.2 The Sweatshop Case	13
1.2.1 The Ethics of Exploitation	14
1.2.2 Zwolinski's Argument	14
1.3 Chapter 1 Summary and Thesis Outlook	16
2 Rights-Based Exploitation	17
2.1 Chapter 2 Introduction	17
2.2 Steinerian Exploitation	18
2.2.1 Steiner's Conditions	18
2.2.2 Revised Conditions	21
2.2.3 Rights and Exploitation	23
2.2.4 The Scope of Rights	24
2.3 The Problem of Moral Responsibility	26
2.3.1 Unintentional Exploitation	27
2.3.2 The Problem	27
2.3.3 The Negligence Solution	32
2.3.4 The Badness Solution	34
2.3.5 Wrongfulness and Intentionality	36
2.4 Pervasiveness	37
2.4.1 Diachronic Pervasiveness	37
2.4.2 Synchronic Pervasiveness	38
2.4.3 The Intentional Condition	39
2.5 Chapter 2 Summary	41
3 Marxist Exploitation	43
3.1 Chapter 3 Introduction	43
3.2 Classical Marxist Exploitation	44

3.2.1	Surplus Labour and the Labour Theory of Value	45
3.2.2	Critiquing the Classical Account	48
3.3	Roemerian Exploitation	49
3.3.1	Unequal Exchange	49
3.3.2	Problems of Over-constraint	51
3.3.3	Property Relations	54
3.3.4	Contributions, Rights, and Transactions	57
3.4	Overextension	58
3.4.1	Dominance	59
3.4.2	Dependence and Interaction	60
3.4.3	The Process of Exploitation	63
3.5	Chapter 3 Summary	65
4	Duty-Based Exploitation	66
4.1	Chapter 4 Introduction	66
4.2	Goodin's Vulnerability-Based Theory	67
4.2.1	Exploiting Persons	67
4.2.2	Using the Vulnerable	68
4.2.3	Goodin's Normative Foundation	69
4.3	Sample's Degradation-Based Account	70
4.3.1	Critique of Goodin	70
4.3.2	Exploitation as Degradation	71
4.3.3	Objective Value and Non-commodification	74
4.4	A Dilemma for Duty-Based Accounts	75
4.4.1	Sample's Dilemma	76
4.4.2	Goodin's Dilemma	80
4.4.3	Diagnosis	82
4.4.4	Extension	82
4.4.5	Responses	83
4.5	Chapter 4 Summary	90
II	What Exploitation Is	91
5	The Two Faces of Exploitation	92
5.1	Chapter 5 Introduction	92
5.2	The Wrongfulness of Exploitation	94
5.2.1	Eliminating the Bottom-left and Upper-right Quadrants	95
5.2.2	Eliminating Wood's Non-moralised Exploitation	96
5.3	Unfair Transactions	100
5.3.1	Use, Contribution, and Rights	100
5.3.2	Fair Market Value	102
5.3.3	Coercion	104
5.3.4	Fraud	107
5.4	Taking Advantage	117
5.4.1	Advantage as (Strict) Pareto Improvement	118
5.4.2	Vulnerability-Based Advantage	119

5.4.3	Advantage as Awareness	120
5.5	Chapter 5 Summary	121
6	Refining the Working Conception	123
6.1	Chapter 6 Introduction	123
6.2	Assessing the Working Characterisation	123
6.2.1	Excluding Mutual Exploitation and Negligence	124
6.2.2	Solving Responsibility and Pervasiveness Problems	126
6.2.3	The Problem of Obligation	127
6.3	Unsuccessful Solutions to the Problem of Obligation	129
6.3.1	Piecemeal Fairness	129
6.3.2	A Change of Scope	130
6.4	Bargaining-Based Exploitation	132
6.4.1	The Bargaining-Based Solution	133
6.4.2	Characterising Bargaining Power	139
6.4.3	Revisiting Micro-Level Fairness	144
6.5	Chapter 6 Summary	147
7	Conclusion: Open Questions	149
7.1	Chapter 7 Introduction	149
7.2	Thesis Summary	149
7.3	Two Open Questions	151
7.3.1	Normative Transactions	152
7.3.2	Solving the Paradox	153
	Appendices	158
A	Kinds of Normative Transaction	158
A.1	Core Transactions	159
A.2	Other Normative Transactions	160
	Bibliography	161

List of Cases

	Page
2.1 Case 1: Steiner's Case	18
2.2 Case 2: The Trust Funder	20
2.3 Case 3: The Car Auction	22
2.4 Case 4: The Revised Car Auction	27
2.5 Case 5: Desert Island Rescue	33
3.1 Case 6: Desert Island Dependence	61
3.2 Case 7: Billions of Burgers	61
4.1 Case 8: The Lotteries (sequential)	77
4.2 Case 9: Rural Roads	80
4.3 Case 10: The Old Well	86
4.4 Case 11: The Lotteries (Simultaneous)	88
5.1 Case 12: The Drowning Woman	105
5.2 Case 13: The Lemon	108
5.3 Case 14: The Unemployed	119
6.1 Case 15: (Upsetting) The Apple Cart	128
6.2 Case 16: Stuck In The Mud	132
7.1 Case 17: The Revised Apple Cart	154

List of Tables and Figures

	Page
4.1 Variation in Expected Values	78
4.2 Outcome Matrix for Case 11	88
4.3 Outcome Matrix for Case 11b	89
5.1 Justice, Morality, and Exploitation	95
6.1 A Basic Bargaining Problem	140
6.2 Three Bargaining Solutions	141
6.3 Shifting the Status Quo	142
6.4 Adjusting for Property Rights Violations	144
6.5 Many Fair Transactions.	144
6.6 Fairness and Rationality.	146
7.1 The Wrongfulness of Exploitation	156
7.2 The Permissibility of Non-Transaction	156
A.1 Normative Transactions	158

Acknowledgements

I thank Alex Voorhoeve, Richard Bradley, and Mike Otsuka for their generous and supportive supervision. Mike carefully read many versions of chapter 4 and parts of chapter 5. He reined me in each time I veered towards over-abstraction and patiently kept me focused on the task at hand. Richard offered careful feedback on many chapters and encouraged me to press ahead with the bargaining-based solution to the problems that emerged late in the thesis. His reassuring calm in the face of what I always believed to be insurmountable problems helped preserve my sanity. Finally, Alex provided support throughout the process. His challenging comments and encouragement of a case-based approach proved to be incredibly fruitful. Both Alex and Richard were instrumental in my decision to pursue a PhD at the LSE, and I am indebted to them not only for their supervision over the past four years, but also for the help they provided during my MSc.

The faculty, staff, and students at the LSE provided an excellent academic atmosphere and were the source of many friendships. In particular, amongst the faculty I thank Jason McKenzie Alexander, Luc Bovens, Foad Dizadji-Bahmani, Christian List, Katie Steele, and Max Steuer. Amongst the research students, Simon Beard, Seamus ‘L^AT_EX’ Bradley, Susanne Burri, Chlump Chatkupt, Mareile Drechsler, Roberto Fumagalli, Johannes Himmelreich, Dean Peters, and Chris Thompson were great sources of inspiration and wonderful friends. Outside of the LSE, Paul Bou-Habib, Julien Fink, Serena Olsaretti, Ruth Sample, Hillel Steiner, and Zofia Stemplowska all provided helpful feedback on individual chapters.

Sebastian Köehler, Orri Stefánsson, and Roberto Veneziani provided detailed feedback on much of the thesis. Ross Masood first encouraged me to explore unjust transfers and helped me consider some of the basic economics behind the concept, and Shane Sandhoefner’s careful copyediting and helpful feedback saved me from an embarrassment of errors. Along with Ciaran Dunphy and James Spillett, their friendship and encouragement saw me through the difficult patches of research.

I received financial support from the Department of Philosophy, Logic and Scientific Method’s departmental scholarship, for which I am incredibly grateful.

I thank my friends and family for tolerating me while I wrote the thesis, especially my friend and partner Alice Obrecht, who has tirelessly supported me and put up with both my elation and misery over the past three and a half years. She pressed me on many issues throughout the thesis and without her advice, mentorship, and intellectual support, I would not have finished. Finally, I thank my parents, Richard and Teresa Ferguson, for providing me with the moral education that has guided my research, and especially my brother Ryan for helping me remember to have fun.

Part I

What Is Exploitation?

Chapter 1

Introduction: The Paradox of Exploitation

1.1 Chapter 1 Introduction

Many transactions are deeply unjust. Some, like sweatshop labour, seem to be unjust because they are unfair; others, like the renting of one's body, seem—if not unjust—at least morally amiss because they are demeaning or dehumanising. We often call these transactions exploitative and, at least in most cases, this label reflects our belief that there is something wrong with them.

Yet, attempts to identify what, exactly, is morally amiss in cases of exploitation—and indeed which transactions should be counted as exploitation—are often frustrated. At least part of the difficulty of determining the extension and properties of exploitation lies in the fact that its boundaries are vague and ill-defined. Exploitation is a 'family resemblance-like' concept¹ that we use in various ways, in various contexts. But concerns about scope are not the only complication for accounts of exploitation. The concept also brings many of our deeply held moral intuitions into conflict, embodying both process-based (non-consequentialist) and consequentialist moral concerns. These conflicts make it difficult to articulate just which moral principles are violated in cases of exploitation and what our *response* to exploitation ought to be. Exploitation presents us with a difficult moral paradox.

When someone steals my wallet, they take it without consent and make me worse off than I was. Conversely, when I buy an apple from the fruit stand on my street, the transaction is consensual and beneficial to both the seller and myself. Some accounts of exploitation allow its scope to include non-consensual and/or harmful transactions, while others limit the scope to consensual and "mutually advantageous exploitation."² At least regarding normative analysis, it matters little whether we extend the term beyond the consensual and mutually advantageous cases, for it is easier to say what has gone wrong when consent is lacking and harm has been done. Rather, the paradox of exploitation arises in its sharpest form in those cases in which we have strong intuitions that exploitation has occurred, yet consent has been given and benefits accrue to all parties. Here, defending the claim that something impermissible has occurred becomes more difficult. It seems, at least *prima facie*, that if the parties in the transaction are all better off and each freely consents to the transaction (and no one else has been harmed), then there can be no wrong done. The conflict is especially troubling for liberals, since their commitments to anti-paternalism and a subjective theory of value block some of the more obvious resolutions to the paradox. The paradox of exploitation, as we often confront it, concerns the claims that engaging in a consensual, Pareto improving

¹Wittgenstein (1953, §66)

²Wertheimer (1996, p.14)

transaction cannot be worse than not transacting, since everyone gains,³ and yet... some consensual, Pareto improving transactions (such as those involved in sweatshop labour) seem to be wrongful, or in some way morally amiss.

Although these claims *appear* contradictory, note that the impermissibility of sweatshop contracts (for example) does not conflict with the ‘non-worseness’ claim that transacting is better—not worse—than not transacting. Consider three possible acts χ , ϕ and ψ . We may say that both ϕ and ψ are impermissible (and χ is permissible), but still endorse the non-worseness premise. The non-worseness premise claims only that if ϕ is Pareto improving and consensual it is the better of the two impermissible options, not that ϕ must be permissible. Thus, as we ordinarily confront it, the ‘paradox of exploitation’ does not really involve conflicting claims. However, if the premises are recast in terms of permissibility, a contradiction can be generated:

The Paradox of Exploitation

1. In transaction, ϕ ,
 - (a) (Consent) to ϕ rather than to not ϕ is consensual and
 - (b) (Pareto) to ϕ rather than to not ϕ is a strict Pareto improvement.
2. (Impermissibility) ϕ -ing is impermissible.
3. (Non-worseness as permissibility) It cannot be impermissible to ϕ , provided:
 - (a) (Consent) to ϕ rather than to not ϕ is consensual and
 - (b) (Pareto) to ϕ rather than to not ϕ is a strict Pareto improvement.

In this form,⁴ the paradox concerns a tension between two claims. Premises 1 and 2, taken together claim that there exist Pareto improving and consensual transactions that are impermissible, while premise 3 claims that Pareto and consent are sufficient for the permissibility of a transaction. For exploitation to be possible, one of the three premises must be false. The paradox arises from a three-fold tension between the ‘factual’ claims about the nature of a particular transaction in premise one, the normative claim in premise two that some of these transactions are impermissible, and the inferential principle connecting the two in premise three, namely, that if the factual claims are true, the normative claims cannot be.

³This claim, often called the “non-worseness premise”, is outlined in Wertheimer (1996, pp.289–93.) Wertheimer also includes a prohibition against negative externalities in his formulation of the non-worseness condition. Although this clause is important in real-life cases, the presentation of the paradox is simplified by assuming such concerns are covered by an implicit *ceteris paribus* clause. Many resolutions of the paradox rely on more detailed specifications of this condition; it is, therefore, not to be read as complete, but rather as a *prima facie* intuition. Additionally, some interpretations of Pareto conditions read the condition as normative. In this presentation Pareto is purely descriptive, while the non-worseness condition contains the normative element. The transaction is a Pareto improvement in the sense that each parties subjective preferences are better satisfied after the transaction.

⁴Interpreting \diamond as ‘it is permissible that’, C as ‘consent’, and P as ‘pareto’:

1. C
2. P
3. $\neg\diamond\Phi$
4. $(P \wedge C) \rightarrow \diamond\Phi$
5. $\diamond\Phi$ (From 1, 2, 4)
6. $\diamond\Phi \wedge \neg\diamond\Phi$
7. Contradiction

Premise 1 is indisputable. There exist Pareto improving and consensual transactions. Premise 2, as we will see throughout the thesis is less straightforward than it might first appear. In some problematic transactions that *prima facie* seem to be impermissible, explaining just *why* transacting is impermissible becomes a complicated issue. Nevertheless, of the three premises, the third is the obvious target for revision. Indeed, the three existing approaches to exploitation that I outline in the following chapters each provide a different reason for rejecting premise 3.

1.2 The Sweatshop Case

In order to better understand the tensions in the paradox and the motivations for abandoning one of the premises, it is helpful to consider an archetypal case of exploitation: the sweatshop labour contract. Here, perhaps more than elsewhere, the term exploitation is used to condemn a practice that many argue is contemptible and morally amiss. Labourers work long hours for multinational corporations⁵ at near-subsistence wages in order to produce goods the workers cannot themselves afford to buy. Robert Ross's account of conditions at Chentex, a Nicaraguan garment firm subcontracted by the retail chain stores J. C. Penny, Kmart, Wal-Mart, and Khol's is, sadly, typical of this form of labour:

Though the smallest of the four retail chains with major orders at Chentex, Kohl's had a 1999 revenue stream of \$4.6 billion (\$6.1 billion in 2000)—more than double Nicaragua's 1999 and 2000 GDP. . .

Chentex workers earn less than 1 percent of the retail price of the jeans they stitch—between thirty and forty cents an hour. This compares to the 10 percent typical of the global north and the 5 percent ratio in U.S. sweatshops.

When we visit the workers' homes we can see the result. . . we meet a woman, Cristina, who was fired from Chentex, unjustly she thinks, for low production. Her sister, who lives across town, was fired for being a union member. Christina's home is a wooden frame, ten-feet square, hung with plastic sheeting for two of the walls and with cardboard boxes that once held shirts shipped from the freezone in Panama for the rest. Her shack has a dirt floor and holds one large bed and (barely) two chairs for herself, her husband, and their baby. Her toilet is a hole in the ground. . . Her husband works seven days a week at another of the free zone plants, but even with his overtime pay they can only afford this bare shelter.⁶

Christina's living and working conditions are appalling, her wages (when she was employed) would not have exceeded \$30 per week. Unionisation succeeded for Chentex workers, but when the union attempted wage negotiations and ordered a one hour work stoppage, union members were fired and union leaders faced criminal charges.⁷ Stories like the Chentex case are common, and are familiar, in fact, to many people in developed countries that purchase the clothes produced in the sweatshops.

Yet, workers in the garment industry, employed (indirectly or directly) by foreign firms, often receive pay that exceeds wages offered in other industries. Ross notes that per capita income in Nicaragua "is about \$470 annually," but garment workers can earn nearly twice

⁵In most cases, companies hire local production companies, which in turn, hire the labourers, thus granting the contracting companies a degree of plausible deniability about labour conditions and pay.

⁶Ross (2004, pp.116–18)

⁷Ross (2004, p.118)

as much.⁸ For this reason, “the paychecks of the free trade zone factories, now received by about thirty-five thousand workers, are highly valued—even when they are earned under sweatshop conditions.”⁹ Not only do these workers earn more than their compatriots, they also freely choose (indeed compete) to work in sweatshop conditions. Further, far from seeing sweatshops as moral failures, some argue that they are a mechanism for moving people in developing countries *out* of poverty by allowing developing economies to exercise their competitive advantage. Writing in the *New York Times*, Nicholas Kristof argues that “the central challenge in the poorest countries is not that sweatshops exploit too many people, but that they don’t exploit enough.”¹⁰

1.2.1 The Ethics of Exploitation

In evaluating the ethics of sweatshops and exploitation, one must be careful to distinguish between the two closely related issues conflated in the brief discussion above. One concerns the *general* role sweatshops play in the structure of global poverty. The second concerns the morality of the *specific* transfer of labour for wages. Sweatshops, if they are wrong, may be so in part because they are exploitative. But this does not mean that the subsequent evaluation of the exploitative element of sweatshop labour is itself to be evaluated by the relative merits or demerits of sweatshop labour in particular. The two considerations coincide only in act utilitarianism, for here the morality of the specific transaction is judged based on the consequences of that action for overall utility.¹¹ For others, however, there is an important distinction to be made between what Alan Wertheimer has called the moral weight and moral force of exploitation.¹² Questions about the moral weight of exploitation concern whether it is wrong, and if wrong, whether it is “*seriously* wrong.”¹³ Questions about the moral force of exploitation, Wertheimer claims, are legislative and concern whether we ought to prohibit exploitative transactions, enforce exploitative contracts, etc. But questions about the moral force of exploitation, I suggest, go beyond the important legal concerns Wertheimer highlights. For if exploitation is wrong, there are further questions about whether it ought to be *morally* permissible if it, in fact, mitigates other, greater, forms of injustice. Thus, applied to the particular case of sweatshop labour, we can see that whether a particular transaction is exploitative is, in principle, a separate question from whether it ought to be prohibited from either a moral or legal perspective.

1.2.2 Zwolinski’s Argument

Matt Zwolinski (2007) takes this distinction on board in his evaluation of sweatshop labour, considering the moral issues internal to the labour transaction separately from the moral force these conclusions have for the general debate about sweatshops. Since our discussion concerns the paradox of exploitation, it is Zwolinski’s internal analysis that is of most interest. Zwolinski notes that the wage agreements in sweatshops are “the product of workers’ consent”

⁸Ross (2004, p.114)

⁹Ross (2004, p.114)

¹⁰Kristof (2009, p.A35). See also, Maitland (1996)

¹¹That is, if sweatshop wages, although demeaning or unfair for the individual laborer, are the best means of development for an impoverished population, they cannot be impermissible, and provided it is the case that exploitation is necessarily impermissible, then provided sweatshop wages lead to the greatest utility, they cannot be exploitative. Of course, this is not to say that the act utilitarian is powerless to condemn sweatshop wages. She can quite plausibly claim that, assuming decreasing marginal utility, a distribution where labourers receive *more* than sweatshop wages maximises utility.

¹²Wertheimer (1996, p.278)

¹³Wertheimer (1996, p.279)

and further, that sweatshops provide their workers with some benefit.¹⁴ Thus, he grants that the first premise of the paradox is true of sweatshops.

Zwolinski considers and rejects one challenge to the non-worseness condition found in Ruth Sample's (2003) analysis of exploitation. For Sample, exploitation "involves interacting with another being for the sake of advantage in a way that degrades or fails to respect the inherent value in that being."¹⁵ Thus, Sample advocates adding an additional sub-clause to the non-worseness condition: it cannot be morally worse to transact than to not transact provided the transaction is consensual, Pareto-improving, *and does not degrade or fail to respect* the persons involved. But, Zwolinski counters, although the employers are "treating their employees in some respects as means" by benefiting from their labour, he argues that "such treatment is ubiquitous and generally untroubling."¹⁶ *True* failures of respect occur when people are used "without their free consent."¹⁷ He concludes that "further argument would be required to demonstrate that such agreements are degrading and disrespectful even in those cases where they are freely consented to."¹⁸ Thus, for Zwolinski, a necessary condition for disrespect or degradation is a lack of consent.¹⁹

Zwolinski grants that certain extreme cases of sweatshop labour, in which actual or threatened violence is present, are cases of exploitation. The difference in such cases, he claims, is that they involve a violation of rights. Thus, Zwolinski does reject the simplistic formulation of the non-worseness condition, implicitly arguing that we must add a rights violation condition. Nevertheless, he rejects the claim that the wage labour contract involved in sweatshops involves such a rights violation: "Since...providing *no* monetary benefit does not violate anyone's rights, and since a contract whereby sweatshops agree to provide *some* benefit does not in itself violate anyone's rights, it follows that such contracts are not exploitative."²⁰ The argument is condensed here, but the way in which 'it follows' is that since neither consent, nor Pareto, *nor rights* are violated, Zwolinski's modified version of non-worseness is satisfied. And, since modified non-worseness is satisfied, the transaction cannot be impermissible. But, since exploitation is impermissible by premise three of the paradox, Zwolinski concludes that sweatshops (as a particular example of transaction ϕ) cannot be exploitative.

This conclusion seems odd. Sweatshops are, as I noted at the outset, paradigmatic cases of exploitation—if anything is exploitative, it seems that sweatshops must be. In his work on moral paradoxes, Saul Smilansky (2007) identifies a form of paradox he terms an 'existential paradox'. He writes,

Some cases of moral paradox disclose... that a segment of moral reality... is absurd. In this sense the "existential paradox" is constructive: we do not need to back track desperately and examine how we got to the conclusion, in order to dispose of it, but on the contrary—the paradoxical result is a revelation of how things are.²¹

¹⁴Zwolinski (2007, p.710–11)

¹⁵Sample (2003, p.57)

¹⁶Zwolinski (2007, p.710)

¹⁷Zwolinski (2007, p.710)

¹⁸Zwolinski (2007, pp.710)

¹⁹Shane Sandhoefner asks whether we can really say that a Nicaraguan working in a sweatshop is 'freely consenting' to do so since their economic situation limits their freedom of choice. They consent to the job, but this seems an empty sort of consent, insufficient to avoid exploitation. In one sense, it seems clear that US companies are exploiting the fact that economies in certain poor countries are such that they can open factories which pay workers very little and have sub-standard conditions. Here I can offer only a promissory note in response: issues of consent and coercion are addressed in 5.3.3, while questions about the motivation and classification of firms' behaviour are addressed in section 7.3.1.

²⁰Zwolinski (2007, p.711)

²¹Smilansky (2007, pp.4–5)

In the end, it may turn out to be the case that we were mistaken to believe sweatshops, in particular, were exploitative. This result would, indeed, be odd, for it would show that we are wrong about one of our primary uses of ‘exploitation’. Yet, even if true, it does not show that the concept itself is incoherent, for there may be other cases in which it can be employed. And, of course, we needn’t draw Zwolinski’s conclusion from the paradox. We might, for example, wish to argue that, despite appearances, sweatshop contracts are not consensual. Or we may wish to further amend or reject the non-worseness condition. And finally—although perhaps more revisionary than Zwolinski’s conclusion—we may wish to reject the idea that exploitation is wrongful or impermissible.

1.3 Chapter 1 Summary and Thesis Outlook

My interest in Zwolinski’s argument lies not in his particular conclusion about sweatshops, but rather in the way his argument engages with the paradox of exploitation. He rejects the simple formulation of the non-worseness condition, arguing that exploitation is possible if there are rights violations. But, he argues, since there are no rights violations in most cases of sweatshop labour, these cannot be cases of exploitation. As I noted, there are other ways of dealing with the paradox. Various theories of exploitation, can be grouped—at least broadly—by the reasons they give for rejecting the third premise of the paradox. In the first part of the thesis, I present three main approaches to exploitation, framing each as a particular solution to the paradox. In chapter 2 I present Hillel Steiner’s (1984; 1987; 2010) classical liberal account of exploitation that is centred on the notion of property rights violations. Chapter 3 features Marxist accounts of exploitation, and in particular, John Roemer’s (1982a; 1982b; 1986) analytical Marxist theories of exploitation. In chapter 4 I consider two related approaches to exploitation that share the common theme of being duty- or respect-based: Ruth Sample’s (2003) theory that grounds exploitation in degradation and Robert Goodin’s (1987; 1986) vulnerability-based model.

In the second part of the thesis I outline my own account of exploitation, drawing on the lessons learned from the first four chapters. In chapter 5 I argue that in order to avoid problems encountered in Sample’s and Goodin’s non-judicial accounts, exploitation must, as Roemer and Steiner claim, be a form of injustice. Yet Steiner and Roemer’s characterisation of exploitation as merely a matter of maldistribution is insufficient to distinguish exploitation from unfair transactions. Thus, I also argue that exploitation must involve an element of ‘taking advantage’. In chapter 6 refine the account of exploitation developed in the previous chapter and consider many forms of unjust transaction. There I argue that only one particular form should be labelled exploitation, and consider the relationships that exist between exploitation and other forms of unjust transaction.

Chapter 2

Rights-Based Exploitation

2.1 Chapter 2 Introduction

In a series of articles, Hillel Steiner (1984; 1987; 2010) has presented and defended a liberal theory of exploitation, aiming to explain the “core characteristic of the concept of exploitation” within “classical liberalism’s psychologically and sociologically ‘thin’ framework of discourse.”¹ He interprets classical liberalism as a morally minimalistic framework with “one or a very few basic moral principles” that imply a set of rights, “to be construed as property rights.”² These rights imply a “‘historical entitlement’ conception of just titles.”³ Further, the framework eschews “any conception of objective value or human needs, [is] agnostic as between different tastes and preferences. . . [and] commits itself only to the primacy of personal rights and liberties and to individual choice.”⁴ What emerges from these quotes is an interpretation of classical liberalism that is synonymous with libertarianism, broadly construed. Of course, both the extension of property rights and principles of appropriation vary under different forms of libertarianism and, since Steiner’s account of exploitation is grounded in property right violations, the extension of his theory will depend on the specification of the “classical liberal” (hereafter libertarian) framework, in particular, on the scope of property rights within a given libertarian framework.

Steiner’s stated goal is to make sense of the core characteristic of exploitation, which, he claims, is a “mutually self-interested, consensual exchange in which what one party transfers is—but need not have been—of greater value than what is received in return.”⁵ The account is limited to mutually advantageous consensual cases of exploitation as a matter of convention. Thus, he assumes the conditions stipulated in the first premise of the paradox of exploitation obtain. Like Zwolinski, Steiner modifies the non-worseness condition to include a provision for property rights: a property rights violation is a necessary condition for Steinerian exploitation.

Thus, both Steiner and Zwolinski resolve the paradox of exploitation with the inclusion of a third conjunct in the non-worseness condition.⁶ For both, consensual, Pareto-improving transactions *can* be morally worse than a failure to transact when rights violations are involved. However, unlike Zwolinski, Steiner would defend the claim that—at least in many

¹Steiner (1987, p.132)

²Steiner (1987, p.132)

³Steiner (1987, p.133)

⁴Steiner (1984, p.225)

⁵Steiner (1987, p.132)

⁶Note that for the libertarian, rejecting the non-worseness condition with the addition of a rights disjunction is one of only a few options for resolving the paradox. The libertarian cannot, for example, appeal (as Ruth Sample and Marx do) to an objective value theory in order to reject non-worseness.

cases—sweatshop labour is exploitative.

In this chapter I consider Steiner’s solution to the paradox of exploitation. Section 2.2 outlines Steiner’s original account of exploitation and my modified version that addresses a problem with the counterfactual that underpins his account. Here I also show why Steiner would disagree with Zwolinski’s conclusion that sweatshops are not exploitative. Sections 2.3 and 2.4 outline two further problems for Steiner’s account. The first addresses a tension between unintentional exploitation and moral responsibility that arises in Steiner’s framework and the second highlights two ways in which the scope of Steinerian exploitation is wider than our ordinary use of the concept.

2.2 Steinerian Exploitation

Steiner’s theory can be stated as set of necessary and jointly sufficient conditions that characterise it within a libertarian framework where the only moral considerations are a respect for property rights.⁷

2.2.1 Steiner’s Conditions

In his (1987) paper, Steiner presents the following case:

C1: Steiner’s case.

Consider an auction. B wishes to sell her right in X , and A and C each wish to purchase it. C bids but A outbids him and, so, gets X . Two conditions are required for this transaction to be a stark-budget liberalism exploitation. One is that C would have outbid A. The other is that the reason why C did not outbid A is that someone’s rights have been previously violated. And the result is that B gets less than she would have received for X and is, to that extent, exploited by A.”⁸

When this case is combined with Steiner’s claim that the core characteristic of exploitation is a “mutually self-interested, consensual exchange in which what one party transfers is—but need not have been—of greater value than what is received in return”, we can isolate the following conditions for Steinerian exploitation.⁹ A exploits B in the bilateral and voluntary transaction¹⁰ ϕ just in case:

- (S1) There exists an unrectified property rights violation at some point in a history of exchanges leading to the transaction,
- (S2) there is a counterfactual transaction, ψ , between A and B in which the violation is not present,
- (S3) B values ψ over ϕ ,
- (S4) A values ϕ over ψ .

⁷As we saw, libertarianism admits “one or *very few* basic principles” based on property rights (Steiner, 1987, p.132, emphasis added). While various nonstandard forms of ‘libertarianism’ might plausibly include other moral principles, I will limit my analysis to an interpretation of the framework that includes only property rights for two reasons: first, although Steiner’s “very few” might be taken to imply the existence of other principles in the framework, his *theory* relies only on property rights; second, it is interesting to see the extension of a theory of exploitation based *only* on property rights.

⁹Steiner (1987, p.132)

¹⁰Explicitly stated here, these domain conditions are assumed throughout the remainder of the thesis, unless otherwise stated.

The antecedent property rights violation (S1) means that B receives less (S3) and A more (S4) than each would have in another transaction ψ that is identical to ϕ , save for the absence of the rights violation (S2).

Steiner arrives at this characterisation by considering the properties that distinguish exploitation from four other forms of transfer: exchange, theft, benefit, and donation. Unlike unilateral transfers such as theft or donation, exploitation is a bilateral transfer—it involves goods moving in both directions. And unlike theft, exploitation is a voluntary transfer.¹¹ Thus, exploitations are members of the set of voluntary, bilateral transfers.

It remains to distinguish exploitation from benefit and exchange. Steiner argues that while exchanges are ‘equal’ transfers, exploitation is a form of unequal transfer. The equalisandum here is not immediately clear, which is an obscurity that has created some confusion. David Miller has argued that if what is equalised is the “standard” or “objective” value of the items, then Steiner’s theory will be inconsistent with its underlying libertarian framework, which presupposes a subjective approach to value.¹² While Miller’s observation is true, it is false that the equalisandum for Steiner is some form of objective value. Rather, exploitation occurs when the subjective values involved in the transfer in the actual world are not equal to the subjective values attached to a transfer in the possible world where no rights violations occur. The distinction between exchange and exploitation introduces the core characteristic of Steiner’s theory: property rights violations. Since property rights form the basis for libertarianism’s moral principles, it is no great surprise that, on Steiner’s account, exploitative transfers require property rights violations. After all, an account of exploitation’s wrongness must follow from the normative considerations of the framework in which it is embedded. Thus, in addition to the voluntary and bilateral domain conditions, we can add two conditions to distinguish exploitation from exchange: (S1) there is an unrectified property rights violation in the history of exchanges leading to the transfer that causes one party of the transfer to receive less than they would have from a transfer where no rights violation occurs (S2).

Finally, we require a distinction between benefit and exploitation. Steiner argues that the distinction here lies in a “difference in their counterfactual presuppositions,” claiming that “a benefit is a bilateral transfer of unequally valued items such that the possessor of the higher-valued item would not voluntarily make the transfer if the items to be transferred were of equal [subjective] value. An exploitation. . . is a bilateral transfer of unequally valued items such that the possessors of both items would voluntarily make the transfer if the items to be transferred were of equal [subjective] value.”¹³ Thus, in the case of exploitation, in the possible world where there is no rights violation, the transfer still lies on the contract curve, while in the case of benefit, if the beneficiary’s contribution increases to match the benefactor’s subjective valuation of the token gift, the transfer will not occur. As an example, consider a benefit dinner where seats can be obtained for \$1000. The diners do not (presumably) value the meal at \$1000; rather, they place that value on the combination of the value of the meal *plus* the donation they make. If the quality of the meal were increased so that the diners valued the meal *alone* at \$1000, they would not make the transfer, provided their goal was to enjoy the dinner and make a donation. However, this way of distinguishing benefits from exploitations encounters problems. Some transfers may come about *only* because of a past rights violation. Consider the following case.

¹¹At least when we restrict our analysis to mutually advantageous exploitation.

¹²Miller (1987, p.152)

¹³Steiner (1984, p.227)

C2: The trust funder.

B owns an investment that pays her a dividend of \$100,000 per annum, an amount that allows her to realise a fairly comfortable standard of living. B values her leisure time, but is not opposed to work. Given the dividend, B is willing to work only if she can secure a salary greater than \$40,000. However, no such jobs are available, so B is unemployed. One day C steals B's investment, putting an end to her dividend. Having lost her income, B must now find work. The only job she can secure is to work for A. Although A usually pays employees in B's position \$30,000, knowing that B needs a job and that the next best job B could obtain pays less than \$20,000, A pays B only \$20,000.

Here is a case where B is worse off, and A better off than either would have been because of a rights violation. But, importantly, there is no person with whom B could have transacted in absence of that rights violation that would have paid her more than the \$20,000 that she receives from A. Now, we might want to claim that since no such counterfactual transaction exists the case cannot be an instance of exploitation. But it seems that the case does embody the elements of exploitation Steiner's theory is attempting to capture. We have here a voluntary, bilateral transaction in which what one party gains from the transaction is less than they ought to have and what the other party gains is greater because of an antecedent rights violation.¹⁴

If Steiner accepts that 'the trust funder' is a case of exploitation that his theory is trying to capture, then it cannot be the case that the distinction between exploitation and benefit lies in their counterfactual presuppositions in the way he claims. Stopping to reflect on this counterfactual failure shows it to be a problem for the distinction Steiner draws between exchanges and exploitation as well. Here we saw that exploitation can be distinguished from exchange because in the case of exploitation, a property rights violation leaves someone accepting a transaction in which they receive less than they would have from a transaction in absence of the property rights violation. Yet, here again there may be no such transaction; the rights violation may bring about the need to transact.

The problem arises because benefits are, in fact, complex transfers. Although Steiner notes that benefits are a "more complex type of transfer," he, nevertheless, seems to treat them as a primitive (non-decomposable) form of transfer, stating only that "there is one clear sense in which the type of transfer involved here is an exchange, and another in which it is a donation. It is not a typical donation because, so to speak, there are items traveling in both directions."¹⁵ But benefits are not "in one sense" like exchanges and in another like donations; they are quite literally complex transfers composed of a donation and an exchange. Consider again the person who pays \$1000 for a plate at a charity dinner. The transaction can be straightforwardly decomposed into (for example) an exchange of \$30 for a meal and a donation \$970 to the charity. That is, it seems clear that when asked how much he donated, the businessman would consider how much he thought the meal was worth and report a figure that deducted this amount from the \$1000. Taking the claim that benefits are complex transfers seriously means that we need only distinguish exploitation from donation

¹⁴One response to the trust funder case is that the *reason* it is exploitative is that A could have paid B more, but altered her price in light of her unjust disadvantage. I address such cases in section 5.2. Although I do not think A's change of salary is a necessary condition for exploitation, if it *was* a necessary condition Steiner's approach would still face a problem, for conditions S1–S4 make no reference to A's ability to pay more or his alteration of price upon learning of B's need. And, more importantly, even if a 'price changing' condition were added, the counterfactual would *still* need to be generalised in order to capture cases where the transaction between A and B would not have occurred but for a rights violation.

¹⁵Steiner (1984, p.226)

and exchange. If benefits are complex transactions comprising exchanges and donations, then distinguishing exploitation from both exchanges and donations suffices to distinguish it from benefits as well. If exploitation is not a member of the set of donation transactions, nor an element of the set of exchange transactions, then it cannot be in the intersection of these two sets either. The counterfactual problem concerns the distinction between exploitation and exchange. It arises in the benefit case only because benefits include an exchange transaction. So, although this discussion achieves our original goal of distinguishing exploitation from benefit, it has also caused a bit of backtracking for we must revisit the distinction between exploitation and exchange.

What we require is a currency for the counterfactual that is not affected by the removal of the rights violation. We cannot use another transaction as a baseline because, as we saw, some transactions occur only because of antecedent rights violations. However, we *can* use the preferences of the parties involved in the transaction. As Steiner notes, the subjective conception of value “takes each person’s actual preference ordering as given”. Therefore, a presupposition of Steiner’s framework is that preferences cannot be counterfactualised.¹⁶

2.2.2 Revised Conditions

Prior to the rights violation, B might have had the following preference order among the following alternatives: $\langle \$40,000 \text{ and dividend; leisure and dividend; } \$30,000 \text{ and dividend; } \$20,000 \text{ and dividend; } \$40,000; \$30,000; \$20,000; \text{leisure (unemployment)} \rangle$. The rights violation eliminates those options that include the payment of the dividend, thus it is only the rights violation that puts B in a position where working for A is preferable to leisure. After the rights violation, B’s most preferred feasible element is taking the job for \$20,000, but before the violation her most preferred feasible element was ‘leisure and dividend’. In both cases, the full preference order remains the same, but the rights violation makes some previously feasible elements unfeasible; it is this sense in which B is ‘worse off’ in the trust funder case. So, exploitation is an unequal transfer in the following sense: the degree to which the agents’ preferences are satisfied post-rights violation is not equal to the degree to which they would have been satisfied had the property rights violation not occurred. In other words, A exploits B in the bilateral and voluntary transaction ϕ just in case:

- (1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that
- (2) makes at least one element of B’s original (and finite) feasible set, γ_B , unfeasible, leaving her with a new feasible set, δ_B (where $\delta_B \subsetneq \gamma_B$),
- (3) B prefers the unfeasible option in γ to her most preferred option in δ ,
- (4) A’s egoistic preferences:
 - (a) are better satisfied in ϕ because γ is restricted to δ ,
 - (b) and if (1)–(3) also hold for A, then the positive change in the degree of satisfaction of A’s preferences, given (1)–(3) hold for B, is strictly greater than any positive change in the degree of satisfaction of B’s preferences, given (1)–(3) hold for A.

¹⁶Steiner (1987, p.137). The liberal may want to distinguish between an agent’s ‘true’ preferences and her ‘adaptive’ or post-rights violation preferences, by for example taking ‘meta-preferences’—preferences over preference orderings—as given.

Although the nature of the conditions is similar to Steiner’s original conditions—(1) is a rights violation condition, (2) a counterfactual condition, (3) a loss¹⁷ condition and (4) a gain condition—the revised conditions alter the currency of the counterfactual, making it ‘degree of preference satisfaction’ rather than ‘transaction’.

Condition (1), which remains unchanged, makes exploitation an injustice within a libertarian framework since it involves the violation of one of the basic moral principles of the libertarian framework; (2) and (3) serve to distinguish exploitation from exchange and capture the intuition that exploitation makes people worse off than they might have been. Condition (4), which captures Steiner’s claim that what B transfers *to* A is “of greater value than what is received in return,” requires a little more clarification.¹⁸

Condition (4a) requires that A does better—his egoistic preferences better satisfied than they would have been—*because of* the rights violation and that this gain occurs ‘in the right way’. The move to ‘degree of preference satisfaction’ opens the door to gains from rights violations not typically associated with exploitation.¹⁹ For example, when the only benefit A receives from a past rights violation affecting B is the satisfaction of his preference that he transact with the victim of a rights violation in such a way that they lose (relative to a no rights-violation baseline), we would not ordinarily say that when A gains *sadistically* he *exploits* B. There is certainly something odd—perhaps even morally amiss—about his preferences, but he does not exploit. Condition 4a requires that A gain from the transaction and that his gain excludes all non-self-interested preferences. Condition (4b) rules out those cases in which both A and B are victims of a rights violation and each are equally disadvantaged by the violation.

The following example will help to clarify these conditions.²⁰

C3: The car auction.

Suppose B wants to sell her car at an auction that includes, among others, the potential buyers A and C. B’s (selling) reserve price is \$500, A’s (buying) reserve price is \$800, and C’s is \$1000. As it happens, A and C have the highest two reserve prices at the auction. The auction begins, with bids accepted at \$10 intervals. When C bids, he discovers his wallet has been stolen. Consequently, A outbids C and wins the car for \$600 (which is at least \$10 greater than the *third* highest reserve price at the auction). Had C been able to bid, the car would have sold for at least \$810, but the rights violation that left C without a wallet also meant that B received only \$600 for her car.²¹

In this example, B is exploited by A because of a violation of C’s property rights. This rights violation satisfies condition (1). Assuming that all participants prefer more money to less, the fact that the payment B receives from A is less than she would have received from C, had C not been robbed, satisfies conditions (2) and (3), and the fact that A gets the car at all (or pays less than he would have had to had he wanted to win the car when C was present) satisfies (4). Finally, the transfer is bilateral and nothing in the case indicates that it was anything other than voluntary, so the domain is that of a voluntary and bilateral transaction.

¹⁷Loss relative to a non-exploitative baseline.

¹⁸Steiner (1987, p.132). Alan Wertheimer also defends a gain condition, arguing “A does not actually exploit B unless A gains from the interaction, even if A *seeks* to gain from the transaction. We can distinguish between the claim that A *exploits* B and the claim that A *acts exploitatively* towards B. (Wertheimer, 1996, p.209).

¹⁹Thanks to Orri Stefánsson for pointing this out to me.

²⁰This case is adapted from Steiner (1987).

2.2.3 Rights and Exploitation

Above I claimed Steiner would disagree with Zwolinski's conclusion that sweatshops are not exploitative. According to both, rights violations can generate exploitation. The non-worseness condition tells us that it cannot be worse to transact than to not transact provided the transaction is consensual, Pareto-improving and—Zwolinski and Steiner add—the transaction does not violate anyone's rights. So, their disagreement over whether sweatshops are exploitative must be a disagreement about what constitutes a rights violation. An examination of this disagreement is instructive, for the scope of Steiner's theory of exploitation is dependent on what it is that we take a (property) rights violation to be. Conditions (1)–(3) provide the core content of the theory, grounding exploitation on a framework of property rights, specifically those property rights violations that restrict the feasible elements of one party's choice set. Thus, the extension of Steinerian exploitation depends on the scope of the underlying theory of rights. The greater the scope of rights, the greater the scope of rights violations and, thus, the greater the scope of exploitation.

Zwolinski's argument that there is no rights violation, and thus no exploitation, in the particular case of sweatshop labour is the following:

- Z1 If providing no benefit to the poor does not violate their rights, then providing some benefit to the poor does not violate their rights. (Premise)
- Z2 Providing no benefit to the poor does not violate their rights. (Premise)
- Z3 Contracts for sweatshop work provide some net benefit. (Premise)
- Z4 A necessary condition for exploitation is the existence of a rights violation. (Premise)
- Z5 Providing some benefit to the poor does not violate their rights (From Z1, Z2)
- Z6 Since contracts for sweatshop work provide a benefit, they do not violate workers' rights. (From Z3, Z5)
- C Sweatshop contracts are not exploitative. (From Z4, Z6)

The argument is valid, but I will argue that it is not sound. Some libertarians might want to challenge the second premise on grounds that (for example) the poverty in developing countries is a result of the over-appropriation of individuals in wealthier countries; however, here we can grant Zwolinski this premise, along with premises three and four. The deeper problem lies with premise one. Steiner's theory of exploitation can be seen as a response to those arguments against libertarian theories of exploitation that depend on premise one. On a historical entitlement theory of just titles, contracts where sweatshops provide beneficial wages *can* be seen as a form of rights violation when they fail to satisfy principles of rectification of injustice. If a worker is disadvantaged by past property rights violations, then contracting with workers for wages lower than they would have accepted without these violations *does* violate their rights and, therefore, *is* exploitation. Not only do workers have a right to the property that was lost in the earlier violation, they also have a right to the greater advantage this property would have afforded them. The rights-based approach to exploitation employed by both Steiner and Zwolinski embodies a historical principle of justice which "creates differential entitlements or differential deserts to things."²² In the case

²²Nozick (1974, p.155)

of sweatshops, employers are under a correlative obligation to honour the worker’s rights to more favourable contracts, provided these workers are in a worse position because of previous rights violations. Transactions that give the worker less than she would have received had the rights violation(s) not occurred (but more than under the current status quo) are exploitative. Note that this treatment does not mean that sweatshop contracts are *necessarily* exploitative. Labelling a sweatshop contract exploitative depends on the existence of antecedent rights violations. If workers’ weak bargaining positions are not due to past rights violations, then sweatshop contracts are not exploitative. One may argue that those who work in sweatshops have not been the victims of past rights violations; in this case, they are not exploited. However, the reason they are not exploited is not because providing a benefit cannot represent a rights violation; rather, they are not exploited because the absence of any past rights violation means they are not entitled to higher wages. Thus, in one sense Steiner and Zwolinski can agree: Pareto optimal and consensual transactions that still leave one party impoverished are not necessarily exploitative, but Steiner points out that this is not because small benefits cannot be rights violations.

Zwolinski’s premise appears appealing only because it fails to compare the transaction to the appropriate baseline transaction where property rights have not been violated. The rhetorical thrust of Zwolinski’s arguments for premise one invites us to imagine scenarios like the following: B’s house has been destroyed by lightning. If it is not wrong for A to ignore B’s plight, then surely it is not wrong for A to help B. We are much more willing to accept premise one when framed in this context than we are in the following: B would have only sold her family heirloom to A for \$10,000, but since her home has been destroyed by arson (a rights violation committed by someone other than A), she is willing (or forced) to sell it to A for \$100. Of course B still receives some benefit when she sells the heirloom to A in the second case—if it were, all things considered, *worse* for B to sell the heirloom, then B would not sell it. The second scenario is a case where A takes advantage of a rights violation to increase his gains from the transfer. Using the baseline of what B is entitled to, we see that A’s ‘benefit’ is no benefit at all. Even if B is desperate for the \$100, she has a right (but lacks the power) to claim much more.

So, while Steiner and Zwolinski agree that a necessary condition for exploitation is the violation of a (property) right, they disagree about when rights violations entitle persons to a claim on greater benefits in exchange. The above analysis shows that, *contra* Zwolinski, on an entitlement theory, persons do have a claim to greater benefits when they are the victims of rights violations because principles of justice in holdings require past rights violations to be rectified. However, although this analysis explains why persons can be said to be exploited when they are the victims of rights violations, it does not explain *which* actions can be classified as a property rights violations.

2.2.4 The Scope of Rights

Steiner’s theory is situated as a general libertarian theory. He does not fully specify the libertarian framework underlying his theory, but offers an analysis of exploitation within the historical entitlement framework that is amenable to many forms of libertarianism. All forms of libertarianism include a robust right to self-ownership²³ along with some form of ownership

²³Most libertarians endorse *full* rights to self-ownership. However Otsuka’s formulation is a notable exception since his form of self-ownership includes “robust and stringent rights of control” rather than full self-ownership (Otsuka, 2003, p.2).

rights in artefacts and natural resources.²⁴ Additionally, as we've seen, libertarian approaches include principles of just transfer and principles of rectification. Although the forms of these principles vary, it is the principle of acquisition that serves to most sharply differentiate the scope of property rights under each framework. Approaches to appropriation can be distinguished according to whether the appropriation requires consent or can be unilateral, if unilateral, whether there are limits on appropriation, and if the appropriation is limited, whether these limits include a tax on the benefits of appropriation.

Perhaps the best-known libertarian framework is Robert Nozick's (1974) form of right-libertarianism, but there are a number of left-libertarian frameworks, including both Michael Otsuka's (2003) approach and Steiner's (1994) own account. Each of these three frameworks employs a different principle of appropriation, resulting in variations in the scope of property rights and, consequently, variations in the scope of Steinerian exploitation. A consideration of the differences between Nozick's right-libertarianism and Steiner's left-libertarianism shows how the extension of Steinerian exploitation depends on the principle of appropriation and scope of property rights employed in the underlying libertarian framework.

Under Nozick's form of libertarianism, natural resources are assumed unowned and unilateral appropriation is permissible, but limited, with no taxation on the benefits of appropriation. Individuals may freely appropriate until the position of those "no longer at liberty to use the thing is thereby worsened"²⁵... [but this proviso] does not include the worsening due to more limited opportunities to appropriate... and it does not include how I 'worsen' a seller's position if I appropriate materials to make some of what he is selling and then enter into competition with him."²⁶ Additionally, an individual may appropriate beyond the point where others are worsened, provided "he compensates the others so that their situation is not thereby worsened."²⁷ Thus, Nozick's theory allows for the appropriation of natural resources, provided others are compensated for any direct negative impact the appropriation has on their situation, but does not require appropriators to neutralise the benefits they reap from appropriation.

Steiner's libertarianism, like Nozick's, allows for limited unilateral appropriation without taxation of benefit; however, he places greater limits on appropriation in the following four ways. First, he argues that appropriators must compensate those who have not appropriated by paying the value at which the goods would trade in a competitive market rather than an amount that merely brings others back to their original pre-appropriation state. Second, like all left-libertarians, Steiner argues that, prior to appropriation, natural resources are owned by all. Third, once appropriated, these resources remain appropriated only while the agent is alive, returning to egalitarian ownership upon the agent's death.²⁸ Finally, Steiner expands the set of appropriable resources to include 'germ line genetic material.' This expansion of the concept of property, coupled with egalitarian ownership and stronger limits on appropriation, increases the scope of Steinerian exploitation. A consequence of Steiner's approach is that those who inherit "golden genes" must contribute the market value of their talent differential to a social fund, to be distributed amongst the (relatively) genetically less well-endowed.²⁹

²⁴Following Peter Vallentyne, I take these to be "produced non-agent resources such as buildings" and "unproduced, non-agent resources such as unproduced land" respectively (Vallentyne and Steiner, 2000, p.5)

²⁵While what is meant by worsened depends on a number of details—the interpretation of the baseline, the approach to welfare, the scope of welfare, etc., these details need not detain us here.

²⁶Nozick (1974, p.178)

²⁷Nozick (1974, p.178)

²⁸Steiner argues that since the dead have no rights, rights to bequest are not morally valid (but may still be honored for practical reasons) (Steiner, 1994, ch7, §c).

²⁹Steiner (1999)

In general, since restrictions on appropriations are more stringent in Steiner’s framework, the set of actions that can qualify as property rights violations is larger than in Nozick’s framework. Since property rights violations are a necessary condition for Steinerian exploitation, the extension of Steinerian exploitation under Steiner’s left-libertarian framework is greater than the extension of Steinerian exploitation under Nozick’s framework. Not only can the actions of others make one worse off, but by including germ-line genetic information in the set of appropriable goods, Steiner’s framework cordons off a portion of natural events that can be treated in the same way as the actions of others. Practically speaking, under Steiner’s framework the extension of Steinerian exploitation is increased to cover cases involving individuals who are disadvantaged due to disabilities issuing from relatively poor genetic endowments.³⁰

Finally, note that if the rights underpinning Steiner’s model are not construed as property rights, but rather rights of a more general form, the account may be extended to a broader range of theories of justice (e.g. Rawls (1971)). However, it should be noted that Steiner’s commitment to *property* rights is based on more than his motivation to show that even classic liberals must be concerned with exploitation. In Steiner (1977) he shows that “any possible set of rights contains a set of titles to objects.”³¹ Sets of non-title-based rights—“action-based rights”—cannot always be mutually satisfied, they are not “compossible” because “there is no reason why two persons’ actions, each of which is informed by certain kinds of permissible purpose, cannot have at least one identical physical component in common. Thus it is logically possible for one person’s exercise of his rights to interfere with another’s exercise of his rights within the same such set.”³² However, sets of title-based (property) rights *are* always mutually satisfiable. The compossibility of property rights give us a reason to prefer theories of justice based on property rights rather than action-based rights.

I have characterised Steiner’s original account of exploitation, offered revised conditions that allow the original account to capture the ‘trust funder’ case, shown how, contra Zwolinski, small benefits may nevertheless be rights violations, and considered variations in the extension of Steinerian exploitation. I will turn, in the following two sections, to two problems for Steiner’s account.

2.3 The Problem of Moral Responsibility

According to (my generalised form of) Steiner’s account of exploitation, A exploits B iff: (1)–(4) are satisfied. And, as we’ve just seen, the scope of this account depends upon the scope of property rights: when the scope of property rights is wide, many actions will count as property rights *violation* and when the scope of those actions that can generate exploitation—property rights violations—is wide, the scope of exploitation is wide as well. But note that Steiner’s account says nothing about the potential exploiter’s epistemic status with respect to the rights violation; A may not know that B has suffered from a rights violation. The following section addresses a tension that arises between the claim that exploitation is a negatively moralised concept and the existence of unintentional exploitation.

³⁰To be disabled in this framework means that the market value of an individual’s skills is below the mean value for society. Thus, the concept is relative to the value society places on one’s skills. Further, it includes the *total* bundle of genetically endowed skills and talents: e.g., blind geniuses may not be disabled in the present sense if the positive market value of their mental abilities outweighs the negative value of blindness.

³¹Steiner (1977, p.770)

³²Steiner (1977, p.771)

2.3.1 Unintentional Exploitation

Consider the following variation on the original car auction case:

C4: The revised car auction.

As before, B is selling her car at a large auction and C would have purchased it for at least \$810. And because C's wallet is stolen, A purchases the car for \$600. Here, however, we can add the following details: C's wallet is pick-pocketed by a passer-by, D, while C is on his way into the auction. Further, D has no knowledge of the auction, or of individuals A and B, and only knows C as the victim of his pick-pocketing. C does not realise he has been pick-pocketed. When the auction begins, C and many others bid against A for the car, but while bidding, C checks his pocket and discovers his wallet is missing. He immediately leaves to contact the police. Meanwhile, unaware of C's departure, A continues bidding and wins the car.

In this revised case, four parties are involved: A, B, C, and D. On Steiner's theory of exploitation, the analysis of the case remains the same as in the original car auction case: A exploits B and C's rights are violated by D. However, in the revised case, it becomes clear that A *doesn't need to know* that C's rights were violated by D in order to be considered an exploiter. In fact, here A is non-culpably ignorant of C's past rights violation, yet A's purchase of the car still satisfies conditions (1)–(4). The possibility of cases of exploitation involving three or more individuals makes room for situations in which an individual may be a non-culpably ignorant exploiter, that is, the revised car auction shows that Steinerian exploitation allows for non-culpably ignorant exploitation.

And if ignorant exploitation is possible, so is unintentional exploitation. In the revised car auction, A does not know he exploits B. If A does not know he exploits B, then A cannot intend to exploit B by this transaction. Similarly, while we may assume that D knows that it is a violation of C's rights to steal his wallet, D does not know of either the auction or the existence of individuals A and B. This knowledge is necessary for D to form the belief that the characterising conditions (1)–(4) obtain. It follows that D cannot know that his theft of C's wallet will lead to A's exploiting B. Lacking this knowledge, D cannot intend that A exploits B. However, since the characterising conditions remain satisfied, it is still the case that A exploits B on Steiner's theory, even if the exploitation is intended by neither A or D. In short, since on any plausible approach to intentional action, knowledge of the action is necessary on the part of the agent, and in the revised case both A and D have no knowledge about the conditions surrounding the transaction, we can conclude that Steiner's theory also allows for unintentional exploitation.

2.3.2 The Problem

We might dismiss this feature of the theory as a mere oddity, if not for the fact that Steiner also purports to give a negatively *moralised* account of exploitation. There is an apparent tension between the existence of unintentional exploitation and the claim that exploitation is a negatively moralised concept. That unintentional exploitation may occur within Steiner's theory is not particularly surprising: the 'disconnect' between rights violator and exploiter ensures that cases (like the revised car auction) where the latter is unaware of a rights violation are fairly easy to generate. While the intuitive sense that there is a tension between these two features is fairly strong, the following presentation makes the inconsistency explicit.

P1 Steiner's theory permits non-culpably ($\neg Cu$) unintentional ($\neg In$) exploitation (Ex).

P2 On Steiner’s theory, exploitation is a negatively moralised concept, which means if exploitation occurs, then it is wrong (Wr).

P3 If an action is wrong, then there is at least one person who is morally responsible (Rs) for the action.

P4 If a person is morally responsible for an action they performed, then they acted either intentionally or from culpable ignorance.

But since the persons (A and D in the revised case) were non-culpably ignorant of the fact that exploitation occurred, they could not be described as intentional exploiters. If they acted neither intentionally nor in culpable ignorance, then they could not have been morally responsible for the exploitation. But if exploitation is a negatively moralised concept, someone must have been morally responsible.³³

Clearly, these four premises generate a contradiction. If we think that unintentional and ignorant exploitation seems counterintuitive, then this intuition may arise from the problem of ascribing moral responsibility presented in the four inconsistent premises. If we accept the arguments about moral responsibility in premises 2 through 4, then premise 1 must be wrong, that is, we would be committed to the claim that exploitation cannot be unintentional. Yet, as we’ve seen, Steiner’s theory of exploitation *does* admit such cases. Since premise 1 follows from Steiner’s theory, if we wish to preserve the theory, consistency requires us to alter or abandon one of the remaining premises 2 through 4. Let’s consider each, beginning with premise 3.

Premise 3 The third premise states that if an action is wrong, then there is at least one person who is morally responsible for the action ($Wr \rightarrow Rs$); to deny premise 3 is to hold that there are wrongful actions for which no person is responsible.

Premise 3 can be decomposed into the following three claims:

3a only agents perform actions,

3b only actions³⁴ are wrong,

3c if an action is a member of the set of wrong actions, it is a member of the set of actions for which an agent can be held morally responsible.

Taken together, 3a and 3b entail that the only source of wrongness is agents.³⁵ In other

³³The argument can be represented in premise form (using the parenthetical letters from above) as follows:

1. $Ex \wedge (\neg Cu \wedge \neg In)$ (P1)
2. $Ex \rightarrow Wr$ (P2)
3. $Wr \rightarrow Rs$ (P3)
4. $Rs \rightarrow (In \vee Cu)$ (P4)
5. $Ex \rightarrow (In \vee Cu)$ (From 2, 3, 4)
6. $In \vee Cu$ (from 1, 5)
7. In (From 1, 6)
8. $In \wedge \neg In$
- C. \perp (From 8)

³⁴Here I mean actions construed broadly and thus will gloss important problems pertaining to omissions and negative acts.

³⁵That is, where A is an action, I is an agent, P is performs, and W is a wrong, (3a) $\forall x[Ax \rightarrow \exists y(Iy \wedge yPx)]$ and (3b) $\forall x(Wx \rightarrow Ax)$ entail $\forall x[Wx \rightarrow \exists y(Iy \wedge yPx)]$. Support for such claims can be found in Scanlon’s (1998, p.272) account of moral criticism, which, he claims, “applies only to rational creatures, since only they are capable of the kind of reflective self-governance in question [in his account].”

words, a necessary condition for wrongness is the performance of an action by at least one agent. Thus, the first two claims restrict the domain of premise 3 to the actions of agents, but they do not supply the core content of premise 3. This is done by 3c, which claims that wrongness is a subset of responsibility. Together the three claims give us premise 3: since only agents generate wrongness, and wrongness is a subset of responsibility, then if an action is wrong, there is at least one agent who is responsible, that is, $Wr \rightarrow Rs$.

Are these three claims plausible? The first I take to be definitional, at least for the sake of this chapter. Actions are a sub-class of events that are performed by agents.³⁶

The second claim, too, is definitional. In order to understand the claim, we must distinguish wrongness from badness. Badness is a negative evaluation of a state of affairs brought about by an event; wrongness, on the other hand, implies that an act violates the standards of a given ethical theory.³⁷ Wrongness is a property adhering only to acts and, therefore, requires an agent. Badness does not require an agent, but it may also be applied to events that can be described as the actions of agents. Thus, while it may be bad that my home is destroyed by a fire started by lightning, it is not wrong that it is so destroyed. It is, however, both bad and wrong if it is destroyed through an act of arson. Those who are unhappy with claim 3b will also (likely) reject my distinction between badness and wrongness. While the definitions in 3a and 3b are not beyond debate, 3c is a more substantial claim, so it is here I focus my attention.

Claim 3c argues that wrongness is a subset of responsibility: all wrong actions are actions for which (accepting the truth of 3a and 3b) agents are responsible. A counterexample to claim 3c would be an action that is wrong, but for which no agent is responsible. In order to determine whether such a counterexample exists, we must be clear about what is meant by responsibility and wrongness. Wrongness is the easier of the two; for my purposes here we can take the account offered above, where wrong actions are those that, according to a given moral framework,³⁸ ought not to be performed. Responsibility, on the other hand, can be understood in a number of ways. One can be *causally* responsible for an action without being *morally* responsible: I knock over a cup while scratching my arm and set off a Rube Goldberg-like chain of events that brings down the building. While I am causally responsible for the collapse—since there exists a causal chain connecting the building’s collapse to my upsetting the mug—I am not morally responsible. Nevertheless, the two concepts are related: all events for which an agent is morally responsible are also events in which their actions play some causal role. Moral responsibility, then, is a subset of causal responsibility. So, which is the form of responsibility employed in claim 3c of premise 3? The answer is that both forms of responsibility apply to premise 3, but only moral responsibility is non-trivial. It is straightforwardly true that if an action is a member of the set of wrong actions it is something for which an agent bears causal responsibility; this is just part of what it means to be an action. But the substantive claim of 3c is that if an action is wrong, it is a member of the set of actions for which an agent is morally responsible. Thus, a counterexample to premise 3 must be a morally wrong action for which an agent (or group of agents) is not morally responsible.

Yet, as Scanlon (1998) points out, moral responsibility, too, can be understood in two

³⁶Lowe (2010) and Davidson (2001). The exact specification of the identity conditions of actions need not concern us here, for what follows will be consistent with a wide number of positions regarding the relations between events and agents.

³⁷Strictly speaking, if we believe there is one true ethical theory, then wrongness implies that an act violates the standards of the *true* ethical theory.

³⁸Again, this is wrongness in a framework relative sense (see above footnote).

ways: either as ‘substantive responsibility’ or ‘responsibility as attributability’. Claims of substantive responsibility “express substantive claims about what people are required...to do for each other”; an agent is attributively responsible when “for a given action...it is appropriate to take it as a basis of moral appraisal of that person.”³⁹ Such an appraisal is warranted when an action engages with an agent’s judgment-sensitive attitudes. Note that, for Scanlon, “having such an attitude involves not only being disposed to judge in certain ways, but also being disposed to various patterns of unreflective thought”⁴⁰, so holding someone attributively responsible does not require that they have *actually* rationally reflected on a particular action, but only that a particular action follows from habits or behaviours that are themselves the kinds of things for which agents can offer a reason for holding. Thus, judgment-sensitive attitudes cover more than intentional action, narrowly construed. They extend to, for example, the foreseeable side-effects of intentional action. Although it is important to determine whether an agent is substantively responsible for exploitation, this inquiry is secondary to the establishment of attributive responsibility.⁴¹ The problem of moral responsibility we are considering is a problem concerning responsibility as attributability. If exploitation is morally wrong, then there must be someone to whom moral responsibility for exploitation can be *attributed*. So, finally, a counterexample to premise 3 is a case in which an action is morally wrong, but there is no agent (or group of agents) who are attributively responsible for the action. With these three sub-premises in mind, it does not seem likely that a counterexample to premise 3 can be produced, but let’s consider two possible contenders.

Suppose John, an epileptic, has a seizure in a shop selling expensive chinaware, which causes extensive damage to the goods in the shop.⁴² Suppose we say that John cannot be attributed responsibility for this action. Of course John may be excused for the seizure itself; the seizure is an episode that happens to John, not an intentional action. But since judgment sensitive attitudes can be wider than narrow intentional actions, in order to maintain the supposition that John is not attributively responsible, it must also be the case that John had good reason to believe that the seizure would not occur. If John knew he was prone to destructive seizures, then he ought not to have entered the china shop. In order to excuse John from attributive responsibility we might suppose John did not know he was an epileptic; the seizure in the china shop was the first he experienced. In such a case it seems clear that John cannot be attributed responsibility for the damage done to the merchandise.⁴³ But once the case is described in a way that exonerates John from attributive responsibility, it is clear that it cannot be the case that the action is wrong. Under no description of the event—even widely construed so as to incorporate John’s decision to enter the shop—can the breaking of the china be described as intentional. Thus, since actions are intentional under some description, the event cannot be described as an action. But since only actions (broadly construed) are wrong, the breaking of the china must be described as bad, not wrong.

The November 17, 2009 issue of the *Guardian* newspaper reported an unusual murder trial where the *prosecution* had asked the jury for a not guilty verdict.⁴⁴ The defendant, Brian Thomas, suffered from a chronic sleep disorder which caused his mind to have “no control

³⁹Scanlon (1998, p.249)

⁴⁰Scanlon (1998, p.21)

⁴¹Whether an agent is substantively responsible for exploitation is, I take it, an analogue of Wertheimer’s concerns about the moral weight of exploitation. See Wertheimer (1996, Ch.9)

⁴²Thanks to Erik Curiel for this example.

⁴³Although, of course, from a legal perspective he may still be liable for the damage.

⁴⁴Thanks to Zophia Stemplowska for this example.

over what his body was doing” while he was asleep.⁴⁵ While dreaming he was attacking an intruder inside his camper van, Thomas inadvertently strangled his wife. Thomas accepted that he had *caused* the death of his wife, but he argued that he had in no way intended her death. Through a series of tests, sleep disorder specialists confirmed that Thomas’s behaviour “was consistent with the legal concept of automatism.”⁴⁶ The Guardian reported that “the prosecution [did] not seek a verdict of guilty to murder or manslaughter.”⁴⁷ While the news report makes clear that Thomas did not intend the death of his wife, one might argue that if he knew he suffered from such episodes, he ought to have taken care to avoid such a tragedy. Failure to do so might allow us to attribute responsibility to Thomas; legally he would, perhaps, face charges of manslaughter. But since we are interested in a case involving wrongness, but not attributive responsibility, we must suppose that Thomas *had* taken care to avoid such a situation. In such a case, the death would not engage with Thomas’s judgment-sensitive attitudes. But supposing that Thomas had restrained himself, yet escaped whilst asleep and inadvertently killed his wife it seems—for the same reason as above—that there is no wrong done. The situation is only a bad state of affairs. The difficulty of constructing a counterexample to premise 3 seems to shift the burden of proof to those who would argue that wrongness does not imply attributive responsibility.

Premise 4 Premise 4 claims that persons are attributively responsible when they act either intentionally, or are culpably ignorant ($Rs \rightarrow (In \vee Cu)$). Suppose this premise is false; then someone may be responsible without acting intentionally or in culpable ignorance. How, then, might they be attributed responsibility? In order to answer this question, we might first consider *who* could be attributed responsibility for exploitation in the revised car auction, and from there determine why such an ascription might be justified.

If we follow this strategy, there will be five logically possible ascriptions of responsibility for the exploitation of B in the revised case:

- A, the exploiter, is attributively responsible.
- B, the exploited, is attributively responsible.
- C, the person who’s rights are violated, is attributively responsible.
- D, the rights violator, is attributively responsible.
- E, some other party, is attributively responsible.

It is implausible to ascribe attributable responsibility in the auction to B or C. Both of these individuals have been *harmed*⁴⁸ in the situation. Clearly, they are innocent victims. Ascribing responsibility to a party not described in the example (party E) is also implausible. As the example is described, no other agent’s actions were relevant to the situation. Thus, if anyone is responsible, then either A or D is attributively responsible.

A cannot be attributively responsible for exploiting B. Consider an analogy to another form of transfer, theft. For Steiner, property rights violations generate exploitation. Every theft is a property rights violation as well. If we suppose that classical liberalism admits only moral principles based on property rights, then the moral wrongness of both theft and

⁴⁵Morris (2009)

⁴⁶Morris (2009)

⁴⁷Morris (2009)

⁴⁸Harmed at least from a certain perspective.

exploitation flows from the violation of this same moral principle. With this in mind, consider a situation where a person is ignorant of a past theft. Suppose they purchase stolen goods, but are non-culpably ignorant of the fact that they were stolen. Have they done anything wrong (in the libertarian framework)? Provided they have taken care to ensure the goods were legitimate, did not engage with disreputable retailers—in short, were non-culpably ignorant of the purloined nature of the goods—we can say they do not act wrongly in purchasing the goods, at least as ‘wrongness’ functions in Steiner’s moral framework.⁴⁹ Were they to learn that the goods *were* stolen, their obligations would change. Depending on the details of the situation, we might require the person to return the goods to the original owner, or turn them over to authorities. Some form of rectification of injustice in holdings is required, the details of which depend upon a particular framework’s principle of rectification. Prior to their learning the goods were stolen, *the thief* was attributively responsible for the theft and resultant ownership of the goods; but, when the new owner learns the goods are stolen, they acquire moral obligations. They become attributively responsible for any failures to respect principles of rectification—but, importantly, not for the original theft. Likely, such principles of rectification will imply some form of substantive responsibility for the holder of the goods as well. Applying this response to exploitation, we can say that just as A does not act wrongly when he (non-culpably) ignorantly purchases stolen goods, he does not act wrongly when he exploits B. However, if A learns about past rights violations that made B worse off, and so comes to learn that he has exploited B, then A inherits rectificatory moral obligations for which he is attributively and (likely) substantively responsible.⁵⁰ Thus, although A may act wrongly if he learns of the exploitation and does not respect the inherited rectificatory moral obligations, A is not attributively responsible for the exploitation of B in the revised car auction.

2.3.3 The Negligence Solution

Thus, if anyone is responsible, D must be responsible. How is it that a person can be both ignorant of the exploitation and not intend it, and yet be attributively responsible? Persons can be attributively (and substantively) responsible in such cases when they act negligently. If someone drives a car under the influence of alcohol, they act negligently.⁵¹ If in so doing, they hit a pedestrian, then they have done something wrong for which they can be held morally responsible. Their negligence makes them attributively responsible even if they did not know the pedestrian was present and did not intend to hit her. Thus, premise 4 must be amended to include negligence.⁵² A person can be attributed responsibility for a wrong action if the action can be said to be intentional, if the person was culpably ignorant, or *if the person acts negligently*. Therefore, D is attributively responsible for A’s exploiting B because, in stealing C’s wallet, D acts negligently. Thus, premise 4 can be amended to read

⁴⁹Of course, they also do not act wrongly according to many other moral frameworks.

⁵⁰We might, for example, say A should compensate B for the difference in their exchange that was brought about by the rights violation to C.

⁵¹Similarly, we might claim that if John, the epileptic, enters a china shop, or Brian, with his sleep disorder, fails to ensure he is restrained, they act negligently.

⁵²One point of confusion here concerns the relationship between negligence and rights violations. As we have seen, our minimal libertarian framework admits only property rights violations as the basic moral principle. Thus, for an act to be morally wrong, it most involve a property rights violation. Yet our (possible) solution to the problem of moral responsibility is that D is attributively responsible because D is negligent. It may seem that we are invoking an additional moral principle; however, this is not the case. We must distinguish between criteria for *attributing moral responsibility* and the *source of moral wrongs*. D acts negligently when he violates C’s rights. The source of the moral wrong is the rights violation, but the negligence of the rights violation means that consequences of the rights violation—such as A’s exploitation of B—can be attributed to D.

$Rs \rightarrow [In \vee (Cu \vee Ng)]$. It now appears we have a solution to the problem, for when we run the argument, we are not forced by the existence of non-culpability in premise 1 to choose intentionality from the disjunction in premise 4. We may choose negligence, that is, that D is morally responsible because he acts negligently. The moral responsibility argument no longer leads to a contradiction. More generally, provided the original rights violator acts in negligence, in culpable ignorance, or intentionally, they can be attributed responsibility for ensuing exploitations. In some cases, the exploiter, A, may know of the rights violation or may consort with the original rights violator to exploit B. In such cases A, too, can be held (at least) attributively responsible.

However, there is a problem with the negligence solution: just as we can construct cases of unintentional exploitation in Steiner's theory, we can also construct cases of non-negligent exploitation. So, although in some cases responsibility for exploitation can be attributed to a person because they act negligently—just as in some cases it can be attributed because they act intentionally or in culpable ignorance—there are cases in which none of the three conditions for the attribution of responsibility are satisfied. Consider the following example.

C5: Desert island rescue.

Suppose D and C are on a desert island with no other inhabitants. D has good reason to believe C is dying and that when C dies, C's possessions will be bequeathed to him. Nevertheless, while C is on his deathbed, D steals C's wallet. Soon after D steals the wallet, the pair are unexpectedly rescued by a passing ship. C goes on to make a full recovery, and, months after the rescue, attends a car auction where, sadly, he discovers his wallet has been stolen. Consequently, he is unable to purchase the car, but would have been the highest bidder, had his wallet not been stolen.

Here D has good reason to believe the consequences of the rights violation will not go beyond their island. Given the information he has, it does not seem he acts negligently in stealing C's wallet. In such a case, if we accept that D did not act negligently, then there would be no way to attribute responsibility for the exploitation to him.

Perhaps this particular case is not very plausible; it is certainly far-fetched. But the general point is that while negligence may solve the problem of moral responsibility in many practical cases, we should not consider it an acceptable *theoretical* solution. Negligence seems to be a 'broader' concept than culpable ignorance, but we must be careful not to construe negligence so broadly that it applies to all cases where negative consequences are possible. If the concept is satisfied in all cases, it can do no work for us. Thus, in cases such as the desert island rescue, the negligence solution will not solve the problem of moral responsibility. So, although in its first form, premise 4 was false, its repair does not offer a theoretical solution to the problem of moral responsibility. We must now consider the final premise, premise 2.

2.3.4 The Badness Solution

Premise 2 This premise asserts that exploitation is a negatively moralised concept. According to premise 2, that exploitation is negatively moralised means that to exploit is to do something wrong ($Ex \rightarrow Ws$). Wood notes that “if someone proposed a non-moralised account of exploitation, that would not necessarily preclude saying that all acts of exploitation, simply as such, are morally wrong. . . It would merely deny that the moral wrongness was built into the very meaning of the term ‘exploitation.’ ” Of course, as Wood’s quote implies, non-moralised accounts of exploitation are possible. In fact, Wood goes on to offer his own non-moralised account, arguing that “exploitation is not unjust by definition.”⁵³ Thus, for Wood, although A exploits B, it is nevertheless possible that A does not act wrongly in doing so. Wood’s approach breaks the chain from exploitation to moral responsibility at the first step. Yet, although a complete rejection of premise 2 is one solution to the problem, Wood’s non-moralised account is not very convincing, as we will see in chapter 5. However, one modification to premise 2 offers a promising alternative that is less drastic.

Wood’s own definition of (negatively) moralised concepts is that they have “wrongfulness or moral badness” built into their meaning.⁵⁴ Yet here it is apparent that the form of premise 2 must be altered to include badness ($Ex \rightarrow (Wr \vee Bd)$). As I argued in premise 3, badness is a negative evaluation of a state of affairs brought about by an event; wrongness, on the other hand, is a property of acts that implies that an act violates the standards of a given ethical theory. A state is *morally* bad—the term Wood employs—when it generates a *pro tanto* reason for action.⁵⁵ More precisely, we may say that a concept is negatively moralised iff the occurrence of an act or event within the scope of the concept implies wrongdoing or moral badness.⁵⁶ The disjunction in the consequent of the modified formulation of premise 2 seems to offer a way out of the paradox, while still allowing us to say that exploitation is a negatively moralised concept.

Does this move succeed? There are two questions here. The first is whether moral badness, like wrongness, implies moral (attributive) responsibility. If it does, then the antecedent of premise 3 would take the form of a disjunction ($(Wr \vee Bd) \rightarrow Rs$) and the inconsistency would be reinstated, leaving the paradox unresolved. However, if badness does not imply moral responsibility, there is still a second question, namely, whether exploitation can be described as morally bad. For although it may be the case that negatively moralised concepts *in general* include those that imply badness or wrongdoing, it does not follow that individual moral concepts are also both morally bad *and* morally wrong. It may turn out to be the case that the badness solution is inappropriate for the concept of exploitation.

In the analysis of premise 3 we saw that, according to Scanlon, attributive responsibility may be ascribed when an act engages with an agent’s judgment-sensitive attitudes. Badness is an evaluation of a state of affairs brought about by an event. An event may be called an action under some descriptions; yet, there are also bad states of affairs brought about by events that are not actions under any description. Events that bring about *morally* bad states of affairs and can be re-described as actions can (often) also be described as wrong.⁵⁷ Thus,

⁵³Wood (1995, p.2)

⁵⁴Wood (1995, pp.3–4, emphasis added). Wood’s ‘or’ here should be understood as an ‘inclusive’ or.

⁵⁵See, for example, Williams (1985, pp.140–41)

⁵⁶Negatively moralised concepts are special cases of moralised concepts. A concept may be moralised iff its characterising conditions contain any (positive or negative) evaluative terms.

⁵⁷Sebastian Köhler points out to me that this is not always the case. It depends, in part, on whether the badness is ‘prima facie’ badness or ‘all things considered badness’, and, more importantly, whether we adopt a consequentialist or non-consequentialist framework. For example, a non-consequentialist can say in a Kantian vein that telling the Nazis the truth about the Jews hiding in the cellar is not wrong, even though

these events will still generate a paradoxical result. Yet these are not the events in which we are interested; we already know from premise 3 that wrongness implies responsibility. We are interested in whether events that cannot be re-described as actions also imply responsibility. Clearly, the answer is no, for if an event cannot be re-described as an action it cannot engage with an agent's judgment-sensitive attitudes. Thus, provided exploitation can be described as morally bad, the inclusion of the badness disjunct in the consequent of premise 3 would resolve the problem of moral responsibility.

This brings us to the second question: is exploitation morally bad, or morally wrong? My claim that wrongness pertains to actions and moral badness to events (when not re-describable as wrongness) means the second question can be restated, 'does the use of the word exploitation describe an event, or an action?' There are two separate answers to this question. The first is in terms of ordinary use, that is, when we ordinarily speak of exploitation are we speaking of an action or an event. The second concerns Steiner's account: is exploitation *for Steiner* an action or an event? If, on Steiner's account, exploitation is an event, then the problem of moral responsibility—a problem of internal consistency—is resolved.

In his analysis of agency, Davidson writes, "very often a sentence will record an episode in the life of the agent and leave us in the dark as to whether it was an action. Here are some examples: he blinked, rolled out of bed, turned on the light, coughed, squinted. . ."⁵⁸ Might we add 'he exploited' to this list? It seems we can. In the pre-theoretical, everyday sense, the sentence 'A exploited B' is very often taken to impute agency, but not as strongly as those verbs that, as Davidson describes them, "cannot be anything but intentional", such as "taking a square root".⁵⁹ For after all, on the classical Marxist account (outlined in the following chapter) a sentence such as 'the capitalist exploited the worker' does not necessarily imply intention. One may argue that exploitation here is being used in a different sense, but that is just the point: the verb exploited *can* be used in both senses, even if it is more natural to assume it implies agency.

Steiner's use, too, is ambiguous. In places he writes as if exploitation is an action, claiming that an individual "exploits"⁶⁰, is an "exploiter"⁶¹, and can be "exploited by"⁶² another and claims that there are three principle roles in an exploitation, the "exploited, exploiters, and sufferers of rights violations."⁶³ Yet, elsewhere he writes of "the occurrence of exploitative circumstances"⁶⁴, viewing exploitation as an unjust event that happens to the exploited. Although the first set of quotes is perhaps more naturally interpreted as implying intention, while the second seems to lack it, all of the above quotes allow for either interpretation.

Cases of unintentional (and non-negligent) exploitation are cases where the term 'exploitation' does not imply agency, since there is no true description under which the event can be said to be intentional. Thus, these cases can only be described as morally bad. As we saw, moral badness does not imply responsibility. Thus, no one can be held attributively responsible for the exploitation. Here the 'exploiter' is merely an innocent conduit for an unfortunate outcome. Of course, the moral badness does mean that there is a *pro tanto* reason for dealing with—removing, avoiding—the exploitation. Morally bad situations may

it has really bad moral consequences.

⁵⁸Davidson (2001, p.45)

⁵⁹Davidson (2001, p.45)

⁶⁰Steiner (2010)

⁶¹Steiner (1984, p.227,232,234)

⁶²Steiner (1984, p.237)

⁶³Steiner (1984, p.235)

⁶⁴Steiner (1984, p.228)

generate substantive responsibility; however, although interesting, an answer to the question of who bears the burden of this form of responsibility will have to wait until the part two.

The ambiguity in Steiner's use serves him well because it also allows for cases of intentional exploitation. When A knows about a past rights violation that makes some (more preferred) elements of B's choice set unfeasible and A uses this information to take advantage of B, A intentionally exploits B and, in so doing, acts wrongly. If the rights violator acts negligently, or violates rights in order to exploit, A and the rights violator share responsibility for the exploitation. When A is non-culpably ignorant, he may help to bring about a situation that is morally bad, but he does not act wrongly.

In the assessment of the first question we saw that badness does not imply attributive responsibility. In the answer to the second we saw that both ordinary use and Steiner's account are consistent with interpretations of exploitation as both a bad event and a wrong action. Thus, it seems the badness solution does offer an acceptable way out of the problem of moral responsibility.

2.3.5 Wrongfulness and Intentionality

Premise 1 Although the badness solution does offer an acceptable solution to the problem, there are those that may disagree with this approach, not because the solution itself fails, but rather because the problem of moral responsibility sets out in the wrong direction even before premise two. For some, the 'obvious' solution to the paradox will be to abandon the first premise. From this perspective, the unintentional exploitation that Steiner's theory allows represents a failure of sufficiency for his theory. The solution, then, is the addition of some form of intentionality sub-condition to the definition of exploitation in subsection 2.2.2 along the following lines:

- (5) A intentionally takes advantage of the disadvantageous effect(s) the rights violation has on B.

By explicitly ruling out cases of unintentional exploitation, the intentional condition dissolves the problem by fiat. This solution to the problem of moral responsibility has the added benefit of solving two other problems of pervasiveness for Steiner's account, discussed in the following section. Since the 'badness' solution has already dissolved the tension, we may, for the time-being, take it as a (relatively) unproblematic solution to the problem of moral responsibility, however various forms of the intentionality solution will reappear in the following section and chapter 5. As we will see, there may be other reasons to prefer this solution to the badness solution.

In this section I explored a tension between the existence of unintentional exploitation in Steiner's theory of exploitation and the claim that exploitation is a negatively moralised concept. While the two claims appear to be inconsistent, I argued that following Wood's claim that a negatively moralised concept "already has wrongfulness *or* moral badness built into its very meaning" dissolves the apparent tension.⁶⁵ Although wrongfulness requires a person to whom moral responsibility can be attributed, badness carries no such requirement. Along the way I outlined two additional solutions: the intentional condition solution and the negligence solution. Although the negligence solution allowed for the attribution of responsibility in some cases, the paradox could be reinstated for cases of non-negligent exploitation. The intentional condition solution offers a solution not only to the problem of moral responsibility,

⁶⁵Wood (1995, p.3, emphasis added)

but also (as we will see in the following section) to problems of pervasiveness for Steiner’s account. The advantages and disadvantages of this approach will be discussed in the following section.

Although Steiner’s theory of exploitation emerges relatively unscathed in my analysis, the discussion does have wider consequences for theories of exploitation in general. The ambiguity between exploitation as an action and exploitation as an event in Steiner’s theory allows his approach to capture the ambiguity of the concept in ordinary use; however, eliding the difference comes at a cost. Having identified a case of exploitation, we are unsure of whether the exploiter acts wrongly: person A is an exploiter regardless of his intention. The ambiguity is unproblematic in the case of B, the exploited, since regardless of whether A’s action was intentional, B is the victim of an injustice. But from the perspective of substantive responsibility it seems to matter a great deal whether A intentionally exploited B. A theory of exploitation that tells us someone has been exploited, but not whether this exploitation is the result of a *wrongful* action seems incomplete.

The general lesson of the analysis is that accounts of exploitation that allow for unintentional exploitation and narrowly and negatively moralised accounts that take exploitation to be necessarily wrong (and not simply morally bad) are incompatible. Of course, this is relatively unsurprising and follows from my claim that wrongness is a property of actions and not events that cannot be described as actions. Yet both a narrowly negatively moralised account and a systemic account are desirable. Narrowly negatively moralised accounts allow us to easily assign both attributive and substantive responsibility to the exploiter. Accounts allowing unintentional exploitation offer an explanation of what is morally amiss (morally wrong or morally bad) with certain transactions even when there is no person to whom moral responsibility can be attributed. While Steiner’s account offers a marriage of both accounts, there may be reason to think that our ordinary concept of exploitation encompasses (at least) two concepts, one captured by narrowly negatively moralised accounts and the other captured by systemic accounts. An examination of this idea appears in chapter 6.

2.4 Pervasiveness

The first two sections explored how exploitation can arise from unrectified past rights violations and failures to respect limits on appropriation. In the discussion of Zwolinski’s argument we saw that even when a transfer is beneficial and consensual, there are cases where the benefiter may violate rights by not benefiting enough *even when* not transacting at all is not a rights violation. The third section addressed a tension that arises from the existence of unintentional exploitation in Steiner’s account. In this fourth and final section on Steiner’s theory I address two problems of pervasiveness the account faces.

2.4.1 Diachronic Pervasiveness

Steven Walt (1984) points out that it is possible for many transactions to be ‘infected’ by exploitation from a single rights violation. Since Steiner’s theory is built on a historical entitlement conception of just titles, the theft of C’s wallet in the car auction examples can lead to more than just B’s exploitation. Commenting on Steiner’s (1984) paper, Walt writes:

[M]ost if not all titles have a far from “clean” causal ancestry... hence if a title to objects resulting from a rights violation is invalid and exploitation involves a rights violation, then it follows that most if not all unequal bilateral exchanges

are cases of exploitation. But this is implausible. The presence of exploitation is unquestionable. Its *omnipresence* is questionable.⁶⁶

Walt identifies a problem of overextension for Steiner's theory that stems from the historical entitlement theory of rights. Since this pervasiveness occurs over a series of transactions, we can call it 'diachronic pervasiveness'. Diachronically pervasive exploitation is certainly a problem if one is looking for a theory of exploitation that closely approximates the extension of ordinary use; it seems that, given the number of known past property rights violations, Steinerian exploitation *is* omnipresent. But while overlap with ordinary use is a reasonable desideratum for a theory of exploitation, it is not the only desiderata. One may, for example, wish to capture certain normative characteristics, even when this involves a departure from ordinary use.

Additionally, note that Walt's criticism of overextension is a problem not only for Steiner's theory of exploitation, but for historical entitlement accounts of justice in general. If just holdings of property depend upon persons being entitled to those holdings "by the principles of justice in acquisition and transfer, or by the principles of rectification of injustice" as Nozick claims, then it is likely that most, if not all, holdings are also unjust.⁶⁷ Yet, both the proponent of a libertarian theory of just distribution and a proponent of a libertarian theory of exploitation may respond that yes, exploitation (or injustice) is more widespread than ordinary language would indicate, but there is nothing in ordinary language that dictates our normative use of the concepts. If we accept the normative force of the principles of an entitlement theory, then despite what ordinary language might suggest, we do have a reason to think that exploitation is pervasive and that this pervasiveness is morally problematic. Of course, I do not wish to offer a full defence along these lines here, but only to suggest how such an argument might begin. Thus, there are two main responses one can give to the critique of diachronic pervasiveness: first, that it does not seem to be especially problematic for exploitation, but rather a general problem with the historical entitlement approach and second, without further argument there does not seem to be a reason that we should, upon previously accepting the normative force of the historical entitlement approach, abandon it because it does not conform to ordinary use. This is not to say that the diachronic critique cannot be further developed; however, there is another problem of pervasiveness which, in combination with the diachronic form, dissolves any practical difference between exploitations and ordinary transactions.

2.4.2 Synchronic Pervasiveness

In section 2.2.3 I claimed that conditions (2)–(4) capture the intuition that exploitation involves a case where one person gains at the expense of another, and also serve to distinguish exploitation from exchange. The problem of diachronic pervasiveness is that, given enough time, nearly all transactions will be 'downstream' from some property rights violation. The problem of *synchronic* pervasiveness is that the outcome of nearly all property rights violations satisfies (2)–(4). If C steals B's x , then the next time B engages in an ' x -related transaction' she is likely to be disadvantaged vis-à-vis A's position. B's loss of her x will always place her at a disadvantage. . . except when it does not. Her loss of x will *not* place her at a disadvantage only when not having x is, for some reason, advantageous in a transaction. But in this case, A, not B, will be at a disadvantage *and* his disadvantage will have been

⁶⁶Walt (1984, p.242, emphasis added)

⁶⁷Nozick (1974, p.153)

brought about by a past property rights violation as well. There are only two cases in which a rights violation will not lead to an exploitation: either when B never engages in an x -related transaction, or when a past property rights violation leaves neither B nor A with a relative advantage over the other (that is, when (4b) is not satisfied). Synchronic pervasiveness makes exploitation (nearly) as widespread as rights violations; the only remaining difference between the two (ignoring the above two rare cases) is that exploitation is the *consequence* of a rights violation. Its cause is the generalisation of the currency of the counterfactual from transactions to the degree of preference satisfaction.

There are a number of responses Steiner might give to this form of pervasiveness. He may attempt to refine (2)–(4) in order to make the conditions more restrictive. However, it does not seem that this approach is likely to be successful since the generalisation to the domain of preferences is required to capture the trust funder case. Steiner may of course deny that the trust funder case is a case his theory should capture, but this too seems unlikely since the case is one in which one party is made worse off by a rights violation in a way that is very similar to ‘ordinary’ cases of Steinerian exploitation where there is, in fact, a transaction in the absence of the rights violation. Finally, Steiner may adopt a line of defence similar to the approach outlined in the discussion of diachronic pervasiveness, arguing that as it happens, the concept of exploitation in ordinary language is too restrictive to capture those cases that we have reason to care about. However, the combination of both forms of pervasiveness makes this response less promising. For if almost every transaction includes an unrectified rights violation in its history and almost every rights violation makes someone worse off in the sense of (2)–(4), then it will be nearly impossible not to be an exploiter or an exploitee in any given transaction. But surely we want a theory of exploitation that tells us something important about our *paradigmatic* cases of exploitation. There seems to be a strong difference between every-day transactions (dog sitting for my neighbour in exchange for a bottle of wine) and paradigmatic cases of exploitation (sweatshop labour), but if both are cases of exploitation on Steiner’s theory then, in lieu of addressing the pervasiveness of his account, he owes us an account of the difference in these kinds of transactions. Ordinarily—and as Steiner himself attempts to show—we take exploitations to be a particular kind of wrongful (or bad) transaction that is distinct from ordinary exchanges, thefts, donations, and benefits. But if the above two pervasiveness objections are correct, then most transactions are not, in fact, exchanges of these varieties. Most exchanges, according to Steiner’s account, are exploitations.

2.4.3 The Intentional Condition

A solution to both synchronic and diachronic pervasiveness can be found in the intentional condition (5) outlined in section 2.3.5. As we’ve seen, in addition to the constraints of the domain (bilateral voluntary transactions) (1)–(4), limit the scope of exploitation to transactions in which a rights violation (1) limits asymmetrically (4b) and negatively (3) affects the degree to which one transactor’s preferences are satisfied (2) and this transactor’s loss is the other’s gain (4a). But the point of the synchronic pervasiveness critique is that almost every property rights violation alters the feasibility of elements of someone’s choice set. And the point of the diachronic pervasiveness critique is that these rights violations can infect otherwise innocent streams of transaction. These facts make the scope of exploitation too wide.

Above I argued that we may have reason to diverge from ordinary use when we wish

to capture certain normative characteristics of a concept. Normatively speaking, the same elements that make sweatshop labour exploitative may be present in a seemingly innocent every-day transaction, the difference being one of degree rather than kind. However, even if the normative point is granted, the descriptive difference remains unexplained: what is the difference between my dog sitting and Chentex's garment production that leads us to *call* the latter exploitation when the former is not? If the only difference is one of degree, why do we not have a greater number of vague cases of exploitation? And what is the unit of measurement of these degrees? And finally, what licences the use of the 'exploitation' label in these cases? Can't the normative status of the transaction be preserved without calling the transaction an exploitation? These questions are not meant as an argument against the claim that the difference in our ascription of exploitation is not one of degree. Rather, the point is that claiming that it is one of degree is not the end the matter; important questions remain.

One important distinction between certain transactions is, of course, the intention of the would-be exploiter. Even if the criteria for Steinerian exploitation are satisfied by both my dog sitting and Chentex's garment production, it seems that, unlike in my every-day transaction, Chentex *intends* to make use of the labourer's unjustly weakened bargaining power. That market rate for Nicaraguan labour is less than American labour is *why* Chentex is based in Nicaragua. The addition of the intentionality condition (5) to Steiner's list of conditions would allow his theory to capture this difference and thus limit its scope. Condition (5) provides a response to both forms of pervasiveness. It limits diachronic pervasiveness because one must intend to take advantage of historical rights violations in order to exploit, and it limits synchronic pervasiveness because although any rights violation that excludes more preferred elements of an individual's choice set may generate exploitation, fewer of these exploitation-generating violations will be converted into exploitations. Rather than placing a limit on the 'candidate' causes of exploitation, (5) places a restriction on how these causes may be converted into an exploitation. The previous set of conditions (1)–(4) can be seen as containing two main requirements, that there be a past rights violation and a future transaction, and two sub-requirements: that the past rights violation resulted in a restriction of some more preferred element of an individual's choice set, and that the other individual gained from this restriction. (5) adds a sub-requirement not to the past rights violation side of the main requirements, but rather to the transaction side: the future transaction must be one in which the party whose advantage is increased by the past rights violation (the exploiter) *intends* to make use of this strategic advantage for gain. Of course, in addition to solving the two pervasiveness issues for Steiner's account, the intentional condition solution (in subsection 2.3.5) also dissolves the problem of moral responsibility by explicitly ruling out the problematic cases of unintentional exploitation. So, the intentional condition solution offers three benefits. Unfortunately, there is also a disadvantage of adopting the condition.

Cases of exploitation seem to derive their *wrongness* from the fact that agents intended to exploit. If this is the case, (5) must be read normatively. But if the intentional condition is a normative condition, further work must be done to establish its relationship with the property rights principle of Steiner's libertarian framework. At best, the intentional condition might be derivable from the property rights principle, although here I am sceptical. It seems likely that if the intentional condition is to be adopted as a normative principle it will not only require a revision to Steiner's theory of exploitation, but also a revision to his underlying ethical framework. Thus, even though the intentional condition offers a solution to both

the pervasiveness issues and the problem of moral responsibility, it seems that its success is purchased at a strongly revisionary price. Although (5) warrants more discussion, further analysis must be saved for the second half of this thesis, where these issues will be revisited.

2.5 Chapter 2 Summary

In this chapter I have presented Steiner's solution to the paradox of exploitation, which involves a modification to the 'non-worseness' condition. Transactions can be impermissible, Steiner argues, if they violate the rights of one of the transacting parties. However, premise one of my reconstruction of Zwolinski's argument claims that if not transacting is not a violation of rights, then a beneficial transaction is not a violation of rights, a premise that frustrates attempts to label sweatshop labour a case of exploitation. Yet, an analysis of the entitlement theory of property rights that grounds the Steinerian approach shows that premise one is false. If one of the transacting parties is made worse off by a rights violation, then a transaction that offers her less benefit than she would have received in the non-rights violating case *does* violate her rights. Thus, Steiner's theory does capture paradigmatic cases of exploitation, such as sweatshop labour. We saw that the extension of Steiner's theory depends on the scope of property rights in the underlying libertarian framework. Right-libertarian accounts like Nozick's contain limits on appropriation that are weaker than left-libertarian accounts such as Steiner's. Thus, the set of violations of limits on appropriation will be smaller for right-libertarians and, consequently, the scope of Steinerian exploitation on such accounts is more limited than it is for left-libertarians.

The third section outlined a problem for Steiner created by the existence of unintentional exploitation within his theory. The claim that exploitation is a negatively moralised concept seems at odds with the existence of unintentional exploitation: if one cannot be responsible for exploiting without knowing one is exploiting, and wrongness requires responsibility, then how can unintentional exploitation be wrong? The solution I explored in this section involves the response that in cases of unintentional exploitation there is no wrongdoing; rather, the exploitation is morally bad. Negatively moralised accounts are not limited to the claim that the concept in question is wrong; rather, as Wood points out, they are those that have "wrongfulness *or* moral badness" built into their meaning.⁶⁸

Finally, in the fourth section, I argued that there are two problems of pervasive exploitation for Steiner's theory that generate a gap between the extension of the concept in ordinary use and the extension of Steinerian exploitation. While there are some responses that might be offered to the problem of diachronic pervasiveness, synchronic pervasiveness erodes the distinctive nature of Steinerian exploitation, leaving it—in most cases—a mere downstream consequence of rights violation. In my discussion of diachronic pervasiveness (and again in the discussion of the synchronic form) I briefly noted that we may want a theory of exploitation to capture certain normative characteristics of exploitation, even at the expense of coextension with ordinary use. Of course, these two desiderata must be balanced against one another—we don't want a theory of exploitation that picks out clearly unrelated actions. And the combination of synchronic and diachronic pervasiveness seems to do just this for Steiner's theory, undermining the relevance of the account. But we need not respond by retreating to a theory that seeks to maximise the fit with ordinary use. Other approaches may be more successful at capturing certain normative characteristics without overextending the theory.

⁶⁸Wood (1995, pp.3–4, emphasis added)

One response to the pervasiveness issue—and to the problem of moral responsibility—is the inclusion of condition (5) in Steiner’s characterising conditions. While this condition solves three problems for Steiner, it also restricts the scope of exploitation and it is not clear that it coheres with Steiner’s normative framework.

Chapter 3

Marxist Exploitation

3.1 Chapter 3 Introduction

Steiner and Zwolinski agree that the non-worseness condition, absent any mention of rights, is false. What these two *disagree* about is whether there are, in fact, rights violations involved in sweatshop labour. There are, however, other ways in which the non-worseness condition might be false. Historically, the most prevalent reason for rejecting the basic form of the non-worseness condition is that—to put it simply—transactions may be impermissible if something is sold for more than it is *worth*. For the classical liberal who accepts a subjective theory of value, such a condition is nonsensical; provided the transactors consensually transact, things are simply worth what people will pay for them. However, although the subjective approach to value is dominant in economics and prevalent in moral and political philosophy, alternative ‘objective’, or ‘just price’, theories of value are by no means scarce. Aquinas’ just price, Marx’s labour theory of value, and even recent approaches that base a fair price on the cost of production¹ all draw, at least in part, on an objective approach to value. Objective value theories have also been used to mount a different critique on the non-worseness condition that should not be confused with Marx’s labour theory of value. Ruth Sample argues that exploitation may be brought about, in part, by “commodifying, or treating as a fungible object of market exchange, an aspect of that person’s being that ought not be commodified.”² Sample’s claim is based on the argument, defended recently by Michael Sandel, that the “extension of markets and of market-oriented thinking to spheres of life once thought to lie beyond their reach . . . is by and large a bad thing, a development that should be resisted.”³

Although the classical Marxist account of exploitation relies on the labour theory of value, which is an objective theory, the ‘analytical Marxists’ such as Cohen (1979) and Roemer (1982a,b) have shown that many of the concerns central to the classical account can be captured by accounts of exploitation that do not depend upon the labour theory of value. Roemer’s property relations model of exploitation attempts to incorporate the normative elements of the classical Marxist approach while abandoning any notion of an objective approach to value. Much of the present chapter focuses on Roemer’s property relations account of exploitation. Before examining Roemer’s account, however, it is helpful to consider the classical account, not only because it is this account that informs Roemer’s approach, but also because some of the failures of the labour theory of value that underpins

¹Reiff (2013)

²Sample (2003, p.57)

³Sandel (1998, pp.93–94)

the classical account will be relevant to the plausibility of other objective value-based accounts of exploitation. An analysis of Sample’s objective value theory will have to wait until chapter 4, where I discuss duty-based accounts of exploitation.

The remainder of the chapter is organised as follows. In section 3.2 I present the classical Marxist account of exploitation and argue that the failure of Marx’s labour theory of value means we should reject the classical account. Section 3.3 presents Roemer’s unequal exchange and property relations accounts of exploitation. In the first part of the section I sketch the unequal exchange account and show why certain problems of ‘over-constraint’ motivate a move to his more general property relations account. The second part considers the property relations account and the relationship between this model of exploitation and Steiner’s account. In section 3.4 I show that without a properly specified dominance condition, Roemer’s account, like Steiner’s, faces an over-extension problem. Finally, section 3.5 concludes.

3.2 Classical Marxist Exploitation

The classical Marxist approach to exploitation differs significantly from modern liberal theories both extensionally and normatively. Extensionally, Marx was concerned with a narrower set of interactions—namely, the economic relationship between the capitalist and proletariat classes.⁴ Marxist exploitation is limited to interactions similar to those in the sweatshop case. It does not provide a general account of exploitation. The normative difference between the classical Marxist account and modern liberal accounts is not one of kind, but of degree. While many modern writers offer accounts not only of what exploitation is, but also why it is wrong,⁵ Marx *seems* to have conceived of exploitation as wrong,⁶ but his primary concern was to offer an explanation of the cause of exploitation.⁷

However, both lines of inquiry share a common motivation: the paradox of exploitation. While the source of expropriation in Feudalism was clear—the serf was coerced into providing *corvée* and *demesne* labour—explaining how non-coerced labourers would submit to expropriation was less so. As Roemer notes, “The riddle for Marx was: How can one explain the systematic expropriation of the surplus product... when the institution for labor exchange is not coercive?”⁸ Thus, while modern liberals are concerned with the normative question of how a Pareto-improving consensual exchange can be impermissible, Marx was concerned with why workers would consent to some transactions in the first place.

Marx offers his own outline of the paradox of exploitation in *Capital, Volume I*. He notes that, *prima facie*, it appears the conditions of Freedom, Equality, Property, and Bentham

⁴As Elster writes, “it is a special case of a more general notion of exploitation that also encompasses non-Marxist varieties.” (Elster, 1986, p.27)

⁵See, for example, Ruth Sample’s (2003) book, *Exploitation: What It Is and Why It’s Wrong*.

⁶This claim is controversial, many Marxists deny that the Marxist account of exploitation is moralised, arguing that the account is purely descriptive. See, for example, Wood (1981). As G.A. Cohen notes, these interpreters reject the claim that “One reason for overthrowing capitalism is that it is a regime of exploitation (and exploitation is unjust),” either because there is no *reason* to overthrow capitalism, or because injustice is not the reason to overthrow capitalism (Cohen, 1979, p.342). However, like Cohen, Wolff, (1999) and others, I believe Marx’s conception is moralised. See Geras (1985) for an overview of the debate.

⁷Both Elster and Cohen make note of these two dimensions of exploitation, with Elster writing that the underlying the concept of exploitation is the idea that “to exploit someone means to *take unfair advantage* of him, a notion that embodies a causal and a normative aspect” (Elster, 1982, p.364). Max Steuer has pressed me on my claim that Steiner, Sample, and even myself offer a *theory* of exploitation. Theories of exploitation, Steuer argues, should provide an account not only of what exploitation is, and why (if at all) it is impermissible, but also an explanation of its causes and persistence. Characterisations of exploitation we might simply call ‘accounts’. In Steuer’s sense, Marx, more than any other thinker considered in the thesis, offers a *theory* of exploitation.

⁸Roemer (1982a, p.6)

are satisfied in the labour market.

Freedom, because both buyer and seller of a commodity. . . contract as free agents, and the agreement they come to, is but the form in which they give legal expression to their common will. Equality, because each enters into relation with the other, as with a simple owner of commodities, and they exchange equivalent for equivalent. Property, because each disposes only of what is his own. And Bentham, because each looks only to himself. The only force that brings them together and puts them in relation with each other, is the selfishness, the gain and the private interests of each.⁹

He acknowledges that in exchanges between the labourer and capitalist the conditions of consent (freedom and equality) and Pareto improvement (Bentham) are satisfied. And yet, Marx claims, while these conditions are ‘satisfied’ in one sense, a closer look at the mechanisms at work will disclose something altogether different—the exchange is not, in one sense, equal. Elster writes, “The meaning of [Marxist] exploitation can be stated, somewhat simplified, as follows: workers are exploited if they work longer hours than the number of labour hours embodied in the goods they consume.”¹⁰ In order to understand this definition we must first understand Marx’s notions of surplus labour, and the labour theory of value.

3.2.1 Surplus Labour and the Labour Theory of Value

Since Marx employs a labour theory of value, the currency of exploitation is embodied labour hours. The labour theory of value states that the value of all commodities is determined by the amount of labour-time necessary for a commodity’s production and reproduction.¹¹ So, for the commodity of workers’ labour-power, value is determined by the amount of time needed for the labourer herself to be produced and reproduced, that is, to eat, find shelter, acquire skills, and procreate.¹² When capitalists employ their capital to produce a product, they must purchase raw materials and machinery, as well as labour, to form the materials into a marketable product; the price they pay for these commodities is also given in terms of the amount of labour time necessary to reproduce the commodity.

So, Marx tells us, suppose a capitalist who wants to make a profit producing yarn purchases 10 pounds of cotton for \$10 and machinery for \$2. If the amount of labour necessary to reproduce (i.e., sustain) labourers is $\frac{1}{2}$ a day’s labour, that is, \$3, then, the value of labour for a whole day is \$6. Thus, the \$12 spent on materials represents 2 days embodied labour. And finally, if spinning $1\frac{2}{3}$ lbs. of cotton into the same of yarn takes 1 hour, then “the [10 pounds of] cotton absorbs six hours’ labour” when it is transformed into yarn.¹³ At the end of the process we have 2 days labour from the raw materials and machinery and $\frac{1}{2}$ a day of labour from the labour. Since the value of labour-power is \$6 per day and $2\frac{1}{2}$ days are embodied in the yarn, then “the value of the product [\$15] is exactly equal to the value of the capital advanced” and no profit has been made for the capitalist.¹⁴ The capitalist can only make a profit by working one of his inputs harder. But, Marx claims, buying more raw cotton will simply result in more produced yarn, but no profit. The machinery too cannot provide him with profit. Working the machinery harder causes it to wear out faster, leaving

⁹Marx (1887, p.121)

¹⁰Elster (1986, p.121)

¹¹Marx (1887, p.118)

¹²Marx (1887, pp.118–20)

¹³Marx (1887, p.131)

¹⁴Marx (1887, p.131)

him in the same position as the yarn. And, while he may buy machinery that is technologically superior to the present machines, the development of the technology represents an increase in embodied labour and thus price, according to Marx’s theory of value, once again thwarting profit. The remaining possible source of profit is labour.

In the example, although labour is paid \$3, which is enough to subsist for 24 hours (and reproduce), they work only for 6 hours. They *could* however, work 12 hours for \$3, converting \$20 worth of cotton and \$4 of machine use into \$30 of yarn¹⁵ since the \$3 is sufficient to sustain and reproduce the worker for the next day. Here the capitalist realises a profit of \$3. This profit Marx calls ‘surplus value’ and the “degree of exploitation (D) of labour-power by capital” is given by this value divided by the time required to produce the worker (socially necessary labour time), i.e., in the present example:¹⁶

$$D = \frac{3}{3} = 1 \tag{3.1}$$

Thus, to reiterate Elster’s claim, for Marx exploitation occurs when workers work longer hours than the hours embodied in the goods they consume.

It is easy to see why, with this characterisation of exploitation, Marx was concerned with the cause of exploitation. Why would the worker consent to working a 12 hour day for the same price as a 6 hour day; why would he continue to work once he earned enough to sustain himself? The answer, Marx claims, lies in the private ownership of the means of production. The worker is characterised *as a worker* because he lacks capital: “In order that a man may be able to sell commodities other than labour-power, he must of course have the means of production, as raw material, implements, &c [*sic*].”¹⁷ Yet, in order to survive, the worker must eat; he “requires also the means of subsistence.”¹⁸ And in order to eat, he is forced to sell his labour to the capitalist. Thus, it is the inequality in the ownership of capital—the means of production—that forces the worker to sell his labour, and it is the acuteness of his poverty that forces him to sell it cheaply. Here the parallels with the sweatshop case are striking. The worker is made better off by working for the capitalist since by so doing he will not starve,¹⁹ and he freely chooses this alternative. Thus, here again, we see that the Pareto and consent conditions of premise one obtain.

Although Marx confronts a version of the paradox of exploitation, as I stated above, he is less concerned with the normative premise. Rather, the paradox for Marx, is as follows.

The Paradox of Capitalist Exploitation

- C1 Transaction P is a Pareto improvement over the status quo (Q).
- C2 However there is another transaction that is a *greater* improvement than transaction P. Call this transaction G.
- C3 Workers make choices that maximise their income.
- C4 Workers consensually chose transaction P over G or Q.

Premises one and two show that the preference order for a worker is $\langle G, P, Q \rangle$. Premise three says that workers chose that option from their choice set (G) that maximises their income, but this contradicts premise four, which claims workers consensually chose P.

¹⁵Since the yarn contains 20 cotton +4 machinery +6 labour, its value is \$30.

¹⁶Marx (1887, p.150)

¹⁷Marx (1887, p.117)

¹⁸Marx (1887, p.117)

¹⁹Indeed, the careful capitalist will ensure the worker is paid well enough to work effectively and to reproduce himself, just as he would care for his machinery, since a failure to do so may lead to the erosion of his capital stock.

Roemer writes, “Marx’s effort was to resolve this paradox by providing a theory of value which claimed that *despite* the non-coerciveness of the institution of labour exchange, that exchange was exploitative.”²⁰ In particular, an appeal to Marx’s labour theory of value shows that there is a possible distribution in which the worker works only long enough to reproduce himself, the minimal limit on the length of the working day.²¹ There is also a maximal limit to the working day, which depends both on the “physical bounds of labour-power” and on what Marx calls ‘moral’ (but what we might call ‘social’) limits, since “the labourer needs time for satisfying his intellectual and social wants”.²² It is this range then over which the worker and capitalist bargain. The measure of exploitation is the difference between the work hours (actually) set in transaction P, and the minimal possible hours, represented by the greater Pareto improvement afforded by (from the worker’s point of view) transaction G. If (by C3) workers make preference maximising choices, why don’t they choose the transaction G solution to the bargaining problem?

The answer, quite obvious by now, is that they lack the power to do so. The distribution represented by transaction G is not feasible (for the worker acting alone). It is this infeasibility that leads to exploitation. And, Marx claims, this infeasibility is caused by inequalities in the means of production.²³ Thus, Marx uses the labour theory of value to identify a particular baseline—transaction G—against which exploitation can be measured. He then shows that it is the inequality of ownership in the means of production that causes a shift from transaction G to P, since workers’ weak bargaining power makes G unfeasible. In terms of the paradox of capitalist exploitation, C4 is false. The worker does not choose P over options G or Q; G is not feasible. Rather the worker chooses P over Q.

We can see how Marx solves his own puzzle about exploitation, but how does his approach solve the paradox of exploitation outlined in chapter 1? Marx accepts the claims that the transaction is consensual and Pareto-improving. Although debatable, he may also be seen as accepting the claim that exploitative transactions are impermissible—at least insofar as exploitation provides us with a reason for condemning capitalism. In this case, the labour theory of value is read not only as a theory of exchange or economic value, but as a more general—and prescriptive—theory of value. On such a reading, Marx, like Steiner, can be seen to be rejecting the non-worseness condition. For Marx, the cause of exploitation is unequal ownership of the means of production. If workers had such ownership, they would not be exploited. Insofar as Marx can be read as claiming that exploitation is impermissible, it must so because this inequality is impermissible. The labour theory of value provides us with a non-exploitative baseline G. Transactions would be non-exploitative if they occurred at G and they would occur at G if there was equal ownership of the means of production. Thus, Marx can be read as tacitly rejecting the non-worseness claim of the paradox of exploitation and replacing it with G.²⁴

²⁰Roemer (1994, p.16)

²¹Marx argues that the working day tends towards this limit, but, given “the basis of capitalist production . . . the working day itself can never be reduced to this minimum.” (Marx, 1887, p.159)

²²Marx (1887, p.159)

²³Marx is not as concerned to explain the exact cause of this inequality, noting that “one thing, however, is clear—Nature does not produce on the one side owners of money or commodities, and on the other men possessing nothing but their own labour-power. This relation has no natural basis. . . [but] is clearly the result of a past historical development, the product of many economic revolutions.” (Marx, 1887, p.118)

²⁴Roberto Veneziani reminds me, however, that Marx is not primarily concerned with individual actions or transactions, but about the fundamental structure of society. It is possible to read Marx as claiming that individuals are not obliged to refrain from exploitative relations even if it would be better for everyone and morally appropriate to replace capitalism with a society in which workers can choose transaction G. Such obligations are not logically contradictory, Veneziani claims. I am not convinced of Veneziani’s claim that such obligations are logically consistent. But his general assertion—that applying Marx’s comments to the

3.2.2 Critiquing the Classical Account

There are two main types of critique that the classical Marxist account faces. The first and most common concerns the well-motivated claims that the labour theory of value is false. As Nozick writes, “with the crumbling of the labor theory of value, the underpinning of [the Marxist] theory of exploitation dissolves.”²⁵ The second form of critique, presented most succinctly by G.A. Cohen, is that “the relationship between the labor theory of value and the concept of exploitation is one of mutual irrelevance.”²⁶ This second critique is, of course, logically prior to the first. We can first ask whether Marx’s account of exploitation depends upon the truth of the labour theory of value, and if the answer to this question is ‘yes’, then we can ask whether the labour theory of value succeeds (or, more accurately, if and why it fails). While I agree with Cohen’s claim that the spirit of the Marxist account does not depend essentially on the labour theory of value, it is worth noting the reasons for rejecting the theory.

In his dismissal of the labour theory of value, Nozick provides us with a laundry list of problems and counterexamples: “found natural objects (valued above the labor necessary to get them); rare goods. . . that cannot be reproduced in unlimited quantities; differences in value between identical objects at different places; differences skilled labor makes; changes caused by fluctuation in supply and demand; aged objects whose producing requires much time to pass (old wines), and so on.”²⁷ While Nozick’s disdain for the theory may not be misplaced, some of his critiques are, uncharitably, levelled at what Cohen calls the “popular surrogate” of Marx’s account: the claim that “labour creates value.”²⁸ According to Marx, however, *socially necessary* labour time creates value.²⁹ Nevertheless, the lack of a convincing (non-circular) definition of socially necessary labour time, the theory’s inability to reliably explain the relationship between value and prices (known as the ‘transformation problem’), and the arbitrariness of labour as the value numéraire mean that the labour theory of value’s problems are not only numerous, but deep as well.³⁰

Like Nozick, Cohen argues that the labour theory of value does indeed fail to deliver. But, he claims, “what really animates Marxists, whatever they may claim. . . is not that the capitalist gets some of the value the worker produces [as the labour-theory implies], but that he gets some of the value *of what* the worker produces.”³¹ Workers, Cohen argues, “do not create *value*, but they create *what has value*.”³² While Nozick is correct to claim that the particular form of exploitation found in the classical account is formulated in terms of the labour theory of value, and that this theory of value is false, it is not the case that amended versions of the account must also fail. In terms of the paradox of exploitation, we may say that, since Marx’s formulation of the baseline G depended on an unacceptable approach to value, we must reject G. However, this does not mean that we must accept the non-worseness premise. We may identify a new baseline in the Marxist spirit. We may still ask how it is that capitalists get some of what the workers produce if it is indeed true that workers are the

paradox of exploitation represents an extension of Marx’s *systemic* focus—is certainly correct.

²⁵Nozick (1974, p.253)

²⁶Even though, Cohen hastens to add, Marx’s theory of value *was* false. (Cohen, 1979, p.338)

²⁷(Nozick, 1974, p.258)

²⁸Cohen (1979, pp.357–58)

²⁹Marx (1887) See Elster (1986) for a discussion.

³⁰For a discussion of the arbitrariness of labour as the value numéraire see Roemer (1982a). The transformation problem is discussed in Samuelson (1971) and an overview of the debate between Samuelson and Marxist economists is provided in Bronfenbrenner (1973).

³¹Cohen (1979, p.354, emphasis original)

³²Cohen (1979, p.354, emphasis original)

only ones who create what has value. Although the alternatives in the bargaining problem outlined for classical Marxist exploitation do depend on Marx's formulation of the labour theory of value, new (and similar) alternatives may be formulated that do not. We may also ask normative questions about whether such distributions are impermissible and what our response to them ought to be. And finally, we may also wish to confront cases that seem, in an ordinary sense, to be exploitative, but that are not covered by Marx's theory. The rehabilitation, normative justification, and expansion of Marx's approach are the tasks Roemer sets himself in his analysis of exploitation.

3.3 Roemerian Exploitation

Roemer offers two accounts of exploitation: the unequal exchange model and the property relations model. The former explores the relationship between labour and exploitation. Here Roemer abandons Marx's labour theory of value and shows that what is *really* necessary for exploitation is asset inequality. Unfortunately, Roemer's unequal exchange models are unrealistically over-constrained. Once generalised to dynamic environments with heterogeneous labour, the strength of these results becomes compromised. The complications brought by weakening some of the initial assumptions motivates a move to the second account of exploitation called the property relations model. Here Roemer abandons the labour theory of value altogether, focusing primarily on the normative properties of exploitation and offering a general theory of exploitation that covers a wider domain than either the unequal exchange model or the classical Marxist account.

3.3.1 Unequal Exchange

The first puzzle about the unequal exchange model is why we should consider it at all. If it is indeed true that the property relations model offers a more general account of exploitation that also does not suffer from the over-constraints of the unequal exchange account, why not simply focus on the property relations model? There are two reasons. First, the orthodox Marxist may be concerned that the rejection of the classical Marxist account of exploitation is the rejection of a straw man. Marx's original account, based on his labour theory of value, is not the most plausible unequal exchange version of exploitation. Roemer's account improves significantly on the classical Marxist account and represents a more plausible version of an approach to exploitation that is, nevertheless, similar to Marx's own account. Second, even if Roemer's unequal exchange account fails it is altogether clear—at least to Elster—that the different approach he takes in the property relations model is a step in the right direction. Elster argues, “[Roemer's] theory of exploitation as unequal exchange is strikingly original and illuminating, whereas the game theory account of exploitation in terms of hypothetical alternatives is seriously misleading.”³³ Even if the unequal exchange model Roemer presents ultimately fails, other approaches along the same lines might prove more successful than the property relations model.

Roemer notes that Marx was concerned with explaining how one class could expropriate surplus value from another “when the institution for labor exchange is not coercive”; Marx's answer, as we saw, was that private ownership of the means of production caused this exploitation.³⁴ However, Roemer notes, in some societies that have transitioned to socialism

³³Elster (1982, p.363)

³⁴Roemer (1982b, p.254)

“we observe certain systematic types of inequality and certain political behaviour that... we might wish to think of... as an indication of exploitation.”³⁵ This presents a problem similar to the one faced by Marx—“He required a theory of exploitation that was robust even without a coercive institution of labor exchange; we require a theory of exploitation that is robust even without private ownership of the means of production.”³⁶

Roemer begins his analysis by asking what is minimally necessary “for one producer to appropriate the labo[u]r of another” and, perhaps surprisingly, he finds that labour markets are not necessary.³⁷ He sets up a model of an economy with N producers who have varied endowments of produced goods and their own labour (which is homogeneous), which they use to produce goods to exchange in the market for goods that they can use to meet their subsistence needs. Each producer faces the same technology, which is assumed to be a Leontief input-output matrix, each has the same subsistence needs, and producers live for just one period (thus precluding the ability to save). Roemer shows that when the producers seek to minimise labour time—subject to the constraints that they meet their subsistence needs (which are homogeneous) and replace their endowments of produced goods³⁸—two solutions emerge: either the amount of time that each producer works is equal, or some producers work less and others more.

The inegalitarian solution occurs only when the initial endowments of produced goods are unequal. This is, in a sense, unsurprising: the only variation in the initial conditions is in the endowments of produced goods and those who are endowment-rich will have greater access to production than the endowment-poor; consequently, the endowment-poor will need to work longer than the endowment-rich in order to meet the same subsistence requirements. However, the importance of the simple model is its demonstration that expropriation of labour can occur “even in the absence of an institution for labo[u]r exchange.”³⁹ This is true in the following sense: suppose in a two person economy consisting of persons i and j , i worked more than the average labour time and j less. Then, “suppose j killed i and took i ’s endowment and then tried to reproduce herself in the economy where only she existed; she would have to work [the average labour time] to reproduce herself.”⁴⁰ Since j would have to work more without i being present, j is “somehow *expropriating* labour from i ”.⁴¹ And this form of expropriation—where one person was made to work more than the socially necessary labour time and one less (or not at all)—is central to the classical Marxist approach to exploitation. Roemer’s achievement is to show that this core component is possible in the absence of a labour market.

Roemer then shows that when labour markets *are* introduced, an interesting equilibrium emerges. In a second model economy that now includes labour markets, producers face three options for meeting their subsistence needs: they can produce their own goods, buy labour, sell labour, or optimise by combining one of the three pure options. Roemer proves that, at an equilibrium, the population can be partitioned into five groups, depending on how producers optimise: those who only buy labour, those who optimise through a combination of self-employment and buying labour, those who are strictly self-employed, those who are both self-employed and sell their labour, and finally those who only sell their labour. The

³⁵Roemer (1982b, p.254)

³⁶Roemer (1982b, p.254)

³⁷Roemer (1982b, p.256)

³⁸Along with the formal constraint that they work less than or equal to one period.

³⁹Roemer (1982b, p.259)

⁴⁰Roemer (1982b, p.259)

⁴¹Roemer (1982b, p.259, emphasis added)

initial endowments of each producer determine their optimal behaviour since, again, all other variables are fixed. Those with the greatest endowments optimise by only employing other labourers; those who have no endowments optimise by selling labour. These five partitions correspond to five classes: the pure capitalist, small capitalist, petit bourgeois, mixed proletarian, and proletarian. From these results Roemer derives the following theorem:

[The Class Exploitation Correspondence Principle (CECP)] states that *every agent who is in a labor-hiring class... is an exploiter, and every agent who is in a labor-selling class... is exploited*. The exploitation status of members of the petit bourgeoisie is ambiguous.⁴²

Roemer shows not only that the core of classical Marxist exploitation can exist without labour markets, but also that when such markets *are* introduced, both exploitation and the Marxist class structure can be derived (not merely postulated) from the models *without reference to surplus production*. Roemer continues to expand his models, showing that similar results can be obtained from the introduction of credit, rather than labour markets. Thus, Roemer demonstrates that “the entire constellation of Marxian ‘welfare’ concepts can be generated with no institution for the exchange of labor.”⁴³ In his final model Roemer considers the case of exploitation in an accumulation economy in which, rather than minimising their labour after reaching a subsistence level, producers seek to maximise their revenues. In this model, an agent is exploited if the revenues they receive from profits and wages are not sufficient to purchase *any* bundle of goods for which the embodied labour time is greater than the amount of time the agent works. Conversely, the agent is an exploiter “if the minimum of these embodied labo[u]r times is greater than the amount of time he worked.”⁴⁴ These models present a view of Marxist exploitation that is revisionary, yet (arguably) true to the spirit of the Marxist project and show that labour is not necessary for Marxist exploitation. What *is* necessary is asset inequality.

3.3.2 Problems of Over-constraint

Roemer’s results are obtained in idealised analogue economies that are simple representations of reality, an approach that worries some philosophers such as Nancy Cartwright (2009). Cartwright argues that—in many cases—the results of models such as Roemer’s “are over-constrained. They include all the constraints necessary to isolate the cause properly so that it is the only cause at work. But there are a good many more constraints involved beyond that.”⁴⁵ Cartwright’s concerns are about the general use of models in economics, about when we are justified in making an inductive inference from a model to the real world. However, we needn’t accept her claim that idealised analogue economies as a *type* are problematic in order to apply Cartwright’s concerns to Roemer’s particular models. Of course, economists, too, are aware of the fragility of their models—hence the plethora of statistical robustness tests. Roemer’s progression from a simple economy lacking labour markets to an accumulation economy is a progression towards more realistic model economies. Nevertheless, his models still employ simplifying assumptions: production technologies are limited to Leontief input-output matrices, labour is assumed to be homogeneous, producers live for one period, etc. As we will see, relaxing some of these assumptions alters the results obtained in Roemer’s

⁴²Roemer (1982b, p.262, emphasis original)

⁴³Roemer (1982b, p.266)

⁴⁴Roemer (1982b, p.269)

⁴⁵Cartwright (2009, p.49)

models. Roemer is not unaware of the fragility of his models and, in fact, he identifies two concerns of over-constraint himself—the restriction of production sets to Leontief matrices and the assumption of homogenous labour.

Leontief Production Sets. First, Roemer notes (echoing Cartwright), “the Leontief technology is terribly special, and a theory that is only true for that technology is a fragile theory.”⁴⁶ The generalisation of the CECP to a model employing standard neo-classical production functions with constant returns to scale “verifies that the correspondence is robust, and not an artefact of the special Leontief assumptions.”⁴⁷ However, in order to maintain the CECP in a generalised environment, we must assume embodied labour time is defined as the “minimum amount of direct labour used in producing [a bundle] c , minimised over the set of maximally profitable processes.”⁴⁸ While the technical details of how this definition maintains the CECP are unimportant here, what is significant is that *this* definition of labour value makes labour value dependent on price.⁴⁹ Had the CECP held only for Leontief production technologies, we would have had reason to suspect the result was ‘assumption generated’; that the results hold only when labour values are made dependent on prices is a concession (from the orthodox Marxist perspective), but not one we need find worrying. Given the rejection of the traditional labour theory of value in section 3.2.2, this concession is desirable. Thus generalising the production technology to models that assume constant returns to scale results in the ‘rehabilitation’ (from the neo-classical perspective) of the Marxist theory of exploitation, achieving one of the three tasks stated above. The first ‘problem’ of over-constraint turns out to be a problem for the orthodox Marxist only.

Homogenous Labour. The second problem occurs with the relaxation of the homogeneous labour assumption, allowing for producers with varying amounts of labour. In this case, although the CECP remains true, the appealing (from a Marxist perspective) correlation between wealth and exploitation no longer holds. As Roemer explains, “There may be very wealthy and skilled producers who sell labo[u]r power and are exploited, and there may be poor, unskilled producers who hire labo[u]r and are exploiters.”⁵⁰ Though Roemer offers a brief defence of the strict version of the homogeneous labour assumption where labour delivery is uniformly distributed, it is unconvincing. In the models comprising the unequal exchange account, a uniformly distributed commodity is required to produce a theory of exploitation that maintains the desired correlation between poverty and exploitation, and “according to the Marxian theory of capitalism, there is just one such commodity, and that is labour power.”⁵¹ The assumption that labour is so distributed is, Roemer assures us, appropriate “for the study of capitalism at one level of abstraction.”⁵² While an approach to exploitation that vindicates the Marxist interpretation of history as a clash between workers and capitalists is desirable for those already committed to Marxism, the response that Marxist theory assumes homogeneous labour is not appropriate from an external perspective. An appreciation of these concerns motivates Roemer’s move to the property relations model of

⁴⁶Roemer (1982b, p.271)

⁴⁷Roemer (1982a, p.147)

⁴⁸Roemer (1982b, p.272)

⁴⁹Labour value is defined with reference to those processes that are maximally profitable; determining maximal profit requires knowledge of equilibrium prices. Therefore labour value is dependent on equilibrium prices.

⁵⁰Roemer (1982b, p.274)

⁵¹Roemer (1982b, p.274)

⁵²Roemer (1982b, p.274)

exploitation. Although the results obtained in the unequal exchange models are not robust to the relaxation of these two assumptions, the takeaway lesson of the unequal exchange approach is not found in the way the approach fares in realistic models. Rather, the most important results are obtained in the sparsest of Roemer’s models, for here he shows—contra the classical theory—that “differential distribution of property and competitive markets are *sufficient* institutions to generate an exploitation phenomenon, under the simplest possible assumptions.”⁵³ The models’ simplicity allows them to act as “Galilean experiments” that “[isolate] the cause under study so that it operates ‘without impediment.’”⁵⁴

Static Models. However, as Roberto Veneziani (2007; 2013) shows, even the results from the simplest of Roemer’s models depend upon over-constraint. Roemer’s results are driven by the static nature of his models; when the models are extended to dynamic subsistence economies, exploitation as unequal exchange no longer persists (though both asset inequalities and classes do).⁵⁵ The inter-temporal nature of Veneziani’s models allows for two interpretations of exploitation: producers may be exploited—that is, work longer than socially necessary labour time—within a period (WP_t exploited), or over the course of their whole lives (WL exploited). Although Veneziani argues (correctly) that the former concept is of greater normative import, the choice of exploitation concepts can make a difference in determining, for example, class membership, since over the course of their lives individuals may belong to many classes. Veneziani’s introduction of dynamic models allows for the possibility of savings.

If agents save, then a “change in WP status does not necessarily reflect genuine inter-class mobility and may simply be the product of inter-temporal trade-offs with little normative content.”⁵⁶ Further, these shifts between classes complicate the interpretation of the CECP: “potentially large portions of society” can be classified as members of a class that has no clear Marxian interpretation and, as in Roemer’s relaxation of the homogeneous labour assumption, the correlation between wealth and exploitation breaks down.⁵⁷ On the other hand, if agents do not save, then the dynamic model can be seen as a generalisation of Roemer’s original theory; however, in such cases, in the long run

profits and WP ⁵⁸ exploitation decrease over time and disappear in the limit. The WP class structure tends to become, loosely speaking, more just, due to the tendential decrease in the dispersion of WP classes around the petty bourgeois. . . and to the absence of big capitalists.⁵⁹

Additionally, the CECP fails in those cases where agents do not save between periods. Roemer’s claim that the jointly sufficient institutions for the persistence of exploitation are differential distribution of property and competitive markets is shown to be false in a dynamic setting. Veneziani concludes that “asset inequalities are proved to be normatively secondary, though causally primary in explaining exploitation and the normative relevance of asset inequalities per se is put into question”, along with Roemer’s account of class.⁶⁰

⁵³Roemer (1982a, p.42, emphasis added)

⁵⁴Cartwright (2009, p.47)

⁵⁵Roemer’s theory can be seen as a special case of Veneziani’s dynamic model “under the *assumption* that $s_v^t = \mathbf{0}$, all v [that is, there are no net savings for any agent at period t].” (Veneziani, 2007, p.197)

⁵⁶Veneziani (2007, p.199)

⁵⁷Veneziani (2007, p.199)

⁵⁸Veneziani notes that WL exploitation does not disappear in these models.

⁵⁹Veneziani (2007, p.202)

⁶⁰Veneziani (2007, p.204)

Although Veneziani’s results are the most damaging for Roemer’s unequal exchange approach, the relaxation of the homogeneous labour assumption is a significant problem on its own. (The remaining issue—the dependence of labour value on price—is, as we saw, problematic only for orthodox Marxists.) Roemer’s own response to tensions in the unequal exchange approach led him to embed it within a more general account of exploitation, which he called the ‘property relations’ account. Although he took the unequal exchange account to be a specific case of capitalist exploitation within the scope of the general account, Veneziani’s results raise questions about the relationship between the two unequal exchange and property relations approaches, since unequal ownership is taken to be the basic moral concern of the property relations account as well. Despite these concerns, Roemer’s general theory is worth considering. Although the insufficiency of asset inequality may make the link between the two accounts less clear, both may also be seen as stipulative (conventionalist) accounts of exploitation. There are two responses we may take to Veneziani’s results. We may conclude, as Veneziani does, that the necessary and sufficient conditions for exploitation as unequal exchange are not as they seemed to be in Roemer’s models. Or, we may alter our definition of exploitation just enough to avoid the problems Veneziani identifies. For example, Veneziani and Yoshihara (2011) show that the necessary and sufficient conditions for the persistence of exploitation are asset inequalities, competitive markets, *and* that agents have a time preference factor that is less than 1, that is, that agents are not indifferent between receiving income in the present period and later periods. The unequal exchange approach can provide a ‘correct’ account of exploitation if the time preference condition is incorporated in the definition of exploitation. Yet this move lacks theoretical justification. The required notion of time preference—pure time preference—is not normatively significant.⁶¹ Thus, although a conventionalist definition may be altered to account for problematic results, we desire normatively justifiable alterations. Taking the latter approach is *ad hoc* and lacks theoretical justification. However, in the same vein as the latter approach, we may also produce a new theoretical account of exploitation that departs from the unequal exchange of labour values, but places an emphasis on asset inequality. Such is the position of the property relations account, which Roemer claims “makes clear what the ethical imperative of the [Marxian] theory is.”⁶²

3.3.3 Property Relations

Roemer’s property relations account begins with the proposal that “a group be conceived of as exploited if it has some *conditionally feasible alternative* under which its members would be better off.”⁶³ This proposal can be captured in the following two conditions. A coalition S in a larger society N is exploited if and only if:

- (PR1) There is an alternative, which we may conceive of as hypothetically feasible, in which S would be better off than in its present situation.
- (PR2) Under this alternative, the complement to S , the coalition $N - S = S'$, would be worse off than at present.⁶⁴

⁶¹Veneziani and Yoshihara write, “the moral relevance of time preference is disputable even in non-Marxian approaches, and a theory of persistent inequalities that crucially depends on time preference seems objectionable” (2011, p.21). See also Rawls (1971) and Ramsey (1928).

⁶²Roemer (1982b, p.275)

⁶³Roemer (1982b, p.276, emphasis original)

⁶⁴Roemer (1982a, p.195)

Roemer notes that certain “bizarre” counterexamples may be produced in the absence of a third ‘dominance’ condition (but his formal analysis makes use of only (PR1) and (PR2)), namely,

(PR3) S' is in a relationship of dominance to S .⁶⁵

However, in the present section I will follow Roemer, focusing primarily on (PR1) and (PR2). Various concepts of exploitation can be specified in the property relations model by altering the withdrawal conditions that specify the hypothetical situations. Roemer focuses on three specifications that characterise feudal, capitalist, and socialist exploitation. The conditions are formally modelled as a cooperative game played by the agents in N , which is defined by what “any coalition can achieve on its own, if it ‘withdraws’ from the economy.”⁶⁶ Under specific withdrawal conditions there is a “utility frontier available to any coalition” and coalitions are exploited if there are frontiers, achievable by withdrawing, that dominate the frontier available to the coalition when it remains in the economy (PR1) *and* the coalition’s complement fares worse if the coalition withdraws (PR2).⁶⁷ Roemer defines the following three withdrawal conditions:

Feudal exploitation The coalition can take with it its own endowments.

Capitalist exploitation The coalition can take with it its per capita share of society’s alienable, or transferable, nonhuman property.

Socialist exploitation The coalition can take with it its per capita share of all endowments, alienable and inalienable (i.e., skills and talents).⁶⁸

A coalition is, for example, capitalistically exploited when it would do better and its complement worse if it withdrew with its per capita share of (alienable) assets. The property relations approach clarifies the ethical imperative(s) behind exploitation claims. One may agree with the general structure Roemer defines in the property relations approach while disagreeing that the content—i.e. a particular withdrawal condition—is normatively acceptable. The paradox of exploitation can thus be seen as a disagreement about the ethical content of the property relations approach.

The neo-classical economist (canonically) accepts the claim that property rights must be respected. For her, the withdrawal conditions for capitalist (and socialist) exploitation are not acceptable. If a coalition’s condition was such that the *feudal* withdrawal conditions were satisfied, then she could endorse the claim that the coalition was exploited. Shaping the problem in this way, Roemer claims, explains why many liberals accept the claim that there is no exploitation in a private ownership market economy. They are correct to claim that there is no feudal exploitation under capitalism (at a competitive equilibrium). But,

⁶⁵One critique of Roemer, as we will see below, is that the number and oddity of the counterexamples that arise without a dominance condition are greater than Roemer admits. This issue is discussed at length in Veneziani (2013). Julius Sensat (1984) offers the following example. Consider a two person society comprising Able and Unable (who cannot work). If Able works 1 day to turn 1 unit of seed corn into another unit of seed plus 2 units of daily subsistence for himself and Unable, then we can say that Unable is exploiting Able since if Unable withdrew with his per capita share he would be worse off and Able better off. Thus, by (PR1) and (PR2) he would be an exploiter. But this fails to distinguish between relevant differences in social power. Unable is an exploiter regardless of whether Able is his slave, Able is his employee, or Able and Unable are friends and Able acts out of concern for Unable. The difference in these cases, the argument goes, can be captured by the dominance condition.

⁶⁶Roemer (1982a, p.195)

⁶⁷Roemer (1982a, p.196)

⁶⁸Roemer (1982a, chapter 7)

the Marxist claims there is capitalist exploitation under capitalism, because she accepts the capitalist withdrawal conditions.⁶⁹

The dependence of a particular form of the property relations approach on a theory of justice can be further emphasised by noting that, in addition to the three withdrawal conditions Roemer identifies, many others may be produced. Proponents of the libertarian approach outlined in the previous sections would accept withdrawal conditions that allowed agents to take their just holdings. Libertarian withdrawal conditions occupy a space between the feudal and capitalist conditions and would form part of a theory of just rectification. For radical right-libertarians who place no limits on appropriation, just distributions are determined only by principles of justice in transfer. The withdrawal conditions of feudal exploitation that demand current endowments be respected correspond to the withdrawal conditions of radical right-libertarianism (assuming that current endowments were justly acquired). Libertarians who do place a limit on appropriation would specify withdrawal conditions that allow the worse-off coalition S to withdraw with its current endowments, plus its share of those endowments that the complement, S' , had gained through over-appropriation. A brief consideration of the various libertarian accounts of justice shows that there are many withdrawal conditions possible between the feudal and capitalist conditions Roemer specifies. A similar consideration of various Marxist approaches would reveal other conditions lying between the capitalist and socialist conditions.

The solution that Roemer's property relations account offers for the paradox of exploitation depends on the specified withdrawal conditions. Under Feudalism, it is false that the transactions occurring between lord and serf were consensual for the serf; thus, the consent premise (and likely the Pareto premise) are false. Here there is no confusion about why the serfs fare poorly in the transaction; the only puzzle is why we might want to call such cases exploitation rather than outright theft. As we saw, if one accepts the feudal withdrawal conditions, but not the capitalist withdrawal conditions, then it is clear that there is no feudal exploitation under capitalism. From the capitalist perspective the paradox is dissolved: the mistake is in thinking that it is impermissible to engage in Pareto optimal and consensual exchanges at a competitive equilibrium. Thus, here the second premise is dropped. From the Socialist and Marxist perspectives, morally superior alternatives can be identified, namely, the equality of alienable assets or alienable assets, skills, and talents. Thus, under both of these perspectives, the third premise, non-worseness, is abandoned.

Roemer's property relations approach can be seen as a general framework for thinking about exploitation that separates the structure of an exploitative transaction from the ethical content that allows us to condemn the transaction. The normative conditions of exploitation are specified by approaches to distributive justice. Because of this separation, debate about the normative acceptability of the withdrawal conditions requires debate about issues that reach beyond considerations of exploitation (and the scope of this chapter). However, we may still examine the acceptability of the structural conditions Roemer provides. In section 3.4 I examine the claim that, as in Steiner's account, the formal conditions of Roemer's property relations approach are insufficient to capture our ordinary understanding of exploitation. However, before moving to this critique, I will first consider the relationship between Steinerian and Roemerian exploitation.

⁶⁹Roemer (1982a, pp.206–07)

3.3.4 Contributions, Rights, and Transactions

Above I noted that libertarian withdrawal conditions may be specified for Roemer's account. In such cases it is tempting to conclude that Roemerian and Steinerian accounts coincide, but, as I will show in this section, the principle of fair transaction underpinning Roemer's account—CONTRIBUTION—differs from that which underpins Steiner's—RIGHTS. However, the most *plausible* versions of each principle coincide. Still, even when both accounts' principles of fair transaction coincide, the two are not fully coextensive since the domain of Roemer's account is not limited to bilateral transactions.

In his discussion of 'fair transactions' Alan Wertheimer notes that the strongly intuitive idea that the distribution of the gains from a mutually beneficial endeavour should be proportional to each individual's contribution to the endeavour is "at the heart of the Marxian theory of exploitation."⁷⁰ This principle is not only central to the Marxian theory, but also to Pigou's (1920) and Robinson's (1933) neo-classical accounts, whereby the exploited are those who receive less than their marginal contribution to production. That the principle—which we may call CONTRIBUTION—can unite such disparate approaches should make us suspicious that it bears some similarity to the normative elements of Roemer's account, and indeed it seems that this is the ethical imperative that underpins Roemer's account.

When left unqualified, the basic form of CONTRIBUTION is implausible; plausible versions "[do] not solely refer to what a factor of production contributes to the end product, but to its *moral* claim to be rewarded in accordance with that contribution."⁷¹ For example, we may argue that individuals should not be rewarded for contribution that is due to undeserved or morally arbitrary talents and skills. In this case, those who make larger contributions that are the result of above average skills and talents do not have greater *moral* claims to reward. Thus, CONTRIBUTION claims that a transaction is fair iff the following condition is satisfied:

CONTRIBUTION: the distribution of gains from transacting is proportional to the morally relevant contribution of each individual to these gains.

Roemer's three withdrawal conditions provide differing accounts of which factors should be counted in determining claims to reward. Of course, on Roemer's account the failure of a fully specified CONTRIBUTION principle is not sufficient for exploitation. Exploitation, for Roemer, requires not only that B receives less than her proportional (per capita) share, but also that A receives more.⁷²

The principle of fair transaction underpinning Steiner's account is RIGHTS. In general, it claims a transaction is unfair when what B receives is less than what she has a right to receive. As is the case with the former principle, RIGHTS may take many forms, depending upon the conception of rights that is employed. For example, as we will see in chapter 4 Ruth Sample claims that those whose basic needs are unmet have a right that these needs are met. However, as with CONTRIBUTION, plausible versions of RIGHTS must be restricted to certain forms of rights. As we saw in chapter 2, Steiner (1977) has shown that in order for a set of rights to be compossible, and so to avoid problems of self-frustration, the set must consist only of property rights. With this restriction in place, we may say that a transaction is fair iff the following condition is satisfied:

⁷⁰Wertheimer (1996, p.227)

⁷¹Wertheimer (1996, p.228, emphasis added)

⁷²It is possible for B to receive less than her per capita share without A gaining if, for example, a portion of B's share is simply destroyed. Thus, we require also that A receives more.

RIGHTS: there does not exist an unrectified property rights violation in the history of exchanges leading to a transfer ϕ that makes one party worse off than they would be.

According to my revised version of Steiner's account, a party is 'worse off than they would be' because of a property rights violation when the violation means that ϕ satisfies their preferences to a lesser degree than they would have been satisfied absent the rights violation.

We have then four principles of fair transaction, two of which are the general rights- and contribution-based accounts, and two of which are the more specific RIGHTS and CONTRIBUTION principles. The most general of the four is the general rights principle which claims a transaction is unfair when individuals receive less than that to which they have a right. The general contribution principle is a subset of this principle, claiming that what individuals have a right to is their proportional contribution to the transaction. While it seems clear that the specific principles will be subsets of the general contribution principle, the relationship between the two individual principles themselves is less straightforward. And this is because each of the more specific principles may be further specified, depending on the scope of property rights and what is taken to be a 'morally relevant' contribution. As we saw above, these do coincide in cases where the withdrawal principle is specified by a libertarian principle. Thus, the most plausible and fully specified versions of RIGHTS and CONTRIBUTION can coincide in certain cases. Any apparent discrepancies between the principles that arise do so because of differences in the accounts of 'morally significant' or 'property rights' and not because of differences in the structure of the principles of fair transaction themselves.

However, agreement between Steiner and Roemer in what counts as a fair transaction does not entail that their accounts of *exploitation* are coextensive. The difference in the scope of the accounts arises because of a difference in the domains for the characterising conditions of each account. For Steiner, exploitation is one kind of bilateral, voluntary exchange. Because it is a form of exchange, exploitation is a relational concept. When Steinerian exploitation occurs, the participants A and B are in an asymmetric relation: A exploits B. Yet, as Elster points out, Roemer's account "does not enable us to define the relation of exploitation, only the predicates 'is exploited' and 'is an exploiter.'" ⁷³ Being exploited, for Roemer, is a condition of certain members of a society and it is not necessarily a kind of bilateral voluntary exchange. Since Roemer does not place the same domain restrictions on exploitation as Steiner does, the extension of the two accounts does not coincide even in cases where both accounts use the same notion of fair transaction. Roemer's omission of these domain conditions causes some notable problems for his account; in particular, like Steiner, his conception of exploitation outstrips ordinary use.

3.4 Overextension

Many disagreements about the adequacy of a theory of exploitation are normative in nature. Roemer's property relations approach avoids these debates, situating the normative issue within theories of distributive justice. Nevertheless, complications for Roemer arise at the structural level. In this section I will focus on Jon Elster's (1982) and Julius Sensat's challenges to Roemer's property relations account. In 3.4.1 and 3.4.2 I identify and evaluate three features of cases that might allow us to claim Roemer's account is insufficient to capture exploitation. I argue that although Roemer can meet two of these challenges, without an operationalisation of the dominance condition, his account is overextended.

⁷³Elster (1982, p.363) Jonathan Wolff also notes that, for Roemer, exploitation "is not a relation between individuals" (Wolff, 1999, p.199).

3.4.1 Dominance

Exploitation in the property relations approach is specified by a “doubly hypothetical” situation in which we assume both that a group withdraws from society and that it suffers no incentive problems.⁷⁴ This approach, Elster claims, “fails to capture the idea that exploitation occurs because of *interaction* between the exploiters and the exploited”, that is, the idea that the exploiters cause the exploitation of the exploited.

Elster begins his critique by presenting his own working definition of exploitation, where exploitation is a theoretical concept that (as in Roemer’s account) obtains much of its normative content from a broader theory of distributive justice. To exploit someone, Elster claims, means “to *take unfair advantage*⁷⁵ of him, a notion that embodies a causal and a normative aspect.”⁷⁶ The normative content—the unfairness—is derived from a theory of distributive justice. This supplies the normative content of (PR1) and (PR2). The causal element—the taking of advantage of—is provided in Roemer’s theory by the structure of conditions (PR1) and (PR2).

The problem with Roemer’s account, Elster claims, is that “causality essentially cannot be captured by counterfactual statements [since]... the truth of ‘if [e_1] had not been present, [e_2] would not have been present’ is neither necessary nor sufficient for the truth of ‘[e_1] caused [e_2].’”⁷⁷ The non-necessity of the counterfactual claim can be shown by the existence of pre-emptive causation, the non-sufficiency by what Elster calls cases of “epiphenomena”, that is, cases where e_1 and e_2 are both caused by a common cause, e_c . Elster then uses these general problems for counterfactual accounts of causation to produce counterexamples to Roemer’s theory where, although S and S' satisfy (PR1) and (PR2), it is nevertheless clear that S' does not cause the ‘exploitation’ of S . Since, according to Elster, an account of exploitation must include both the normative and causal components, in failing to deliver on the second, Roemer’s account fails.⁷⁸

To illustrate this, Elster presents what we might call the ‘non-dominance problem’. He asks us to consider two groups that have ideological interactions, but no economic interaction. One group, S , has (religious) beliefs about the other’s work habits that negatively influence its own productivity; the other, S' , has beliefs that positively affect its productivity. If S' did not exist, then S would be better off. If S did not exist, then S' would be worse off. Conditions (PR1) and (PR2) are satisfied. But it seems absurd to claim that S' exploits S precisely because S' does not appropriately cause S to be worse off.⁷⁹ By identifying the general causal problem that leads to cases like the above, Elster provides a recipe for creating any number of counterexamples to Roemer’s property relations account.⁸⁰ Elster notes that,

⁷⁴Elster (1982, pp.365–66)

⁷⁵Found in Wertheimer (1996), Sample (2003), Goodin (1987), and elsewhere, the phrase “take unfair advantage” has become the *de facto* ordinary language definition of exploitation.

⁷⁶Elster (1982, p.364, emphasis original)

⁷⁷Elster (1982, pp.366–67)

⁷⁸Elster’s claims about counterfactuals imply that he favours what Ned Hall calls the productive concept of causation, which we evoke “when we say of an event c that it helps to generate or bring about or produce another event e ” (Hall, 2004, p.225). Hall distinguishes this concept of causation from what he calls the notion of causation as “dependence”. Although not necessary for causation as production, the counterfactual ‘ c causes e just in case, had c not occurred, e would not have occurred’ is necessary for causation as dependence. The additional conditions for Roemer’s account that Elster’s critique prompts can be seen as conditions that attempt to capture (at least a portion of) a *productive* causal relationship between exploiter and exploitee. However, some of the concerns Elster seems to want to capture by appealing to a productive concept of exploitation can be, arguably, better captured by including conditions like my (A5) (see chapter 5) which explicitly outline the productive role of exploiter. With the addition of (A5), the (relatively simpler) dependence concept of causation can be employed.

⁷⁹Elster (1982, p.368)

⁸⁰For example, in addition to the example provided above, one might imagine a dual desert island case,

although restricting the scope of Roemer’s theory to ‘situations of arm’s-length economic transactions’, along with the inclusion of (PR3), may handle epiphenomena, both “appear *ad hoc* and are, in any case, much less precise than the other criteria.”⁸¹ He continues with the (unargued) claim that “any attempt to make them precise would come up against new counterexamples that formally satisfy them and yet would not count as exploitation.”⁸²

This last quote from Elster is somewhat puzzling, if his concern about Roemer’s theory is that it cannot assure the right kind of interaction between S and S' . For if Roemer’s conditions are restricted to arm’s-length economic transactions and (PR3) is true of the transaction between S' and S , then it seems possible to say that one group exploits the other. Of course, these modifications may still be *ad hoc* since, as Kymlicka notes, “they are disconnected from the ‘ethical imperative’ he identifies as the basis of exploitation theory.”⁸³ But while an *ad hoc* solution might be *unacceptable*, it is not necessarily inadequate. There are two other reasons that Elster may think that Roemer’s theory cannot succeed.

3.4.2 Dependence and Interaction

Putting aside the general weaknesses of an *ad hoc* solution, one reason for thinking that precisifying (PR3) cannot adequately handle Elster’s case is that it involves more than a mere failure of dominance. Roemer’s (PR1) and (PR2) conditions fail to restrict exploitation to cases where the following occur:

Individual Interaction: Particular members of S' and S ‘interact directly’.

Dependence: S' ‘depend on’ S for their advantage.

The clauses in quotes—‘interact directly’ and ‘depend on’,—are underspecified, but I take it that, as they stand, these descriptions of cases are intuitively clear. In Elster’s counterexample, none of these conditions seem to be satisfied (depending on how we understand ‘depend on’). The counterexamples produced when these three conditions are not met are similar to those that would be produced if Steiner’s theory did not contain condition (4).⁸⁴ What we want from (4) and from a properly specified (PR3), arm’s-length condition, etc., is a restriction of exploitation to cases where not only does A *gain* from the transaction, nor merely does A gain *from B* in the transaction, but A gains from B *in the right way*. In order to emphasise what is intuitively missing from Roemer’s account, consider another counterexample presented by Julius Sensat (1984), which highlights a case where the lack of dependence features prominently.⁸⁵

Our intuition that this result is wrong is driven by the fact that the wealthy islanders do not depend on the poor island for their wealth in any way. There may be something unjust about the distributions from an egalitarian perspective, but it is surely not the case

where S inhabits one island and S' another. A story can then be told about how the mere presence of one group positively or negatively affects the other.

⁸¹Elster (1982, pp.368–69)

⁸²Elster (1982, pp.368–69)

⁸³Kymlicka (1990, p.204, fn.13). Even Roemer himself notes that, since his goal was to produce a purely distributive notion of exploitation, the inclusion of an undefined notion of dominance is *ad hoc* (Roemer, 1982b, p.277, fn. 15). Thanks to Roberto Veneziani for bringing this point to my attention.

⁸⁴Recall that condition (4) states: ‘A’s egoistic preferences are better satisfied in ϕ because γ is restricted to δ , and if (1)–(3) also hold for A, then the positive change in the degree of satisfaction of A’s preferences, given (1)–(3) hold for B, is strictly greater than any positive change in the degree of satisfaction of B’s preferences, given (1)–(3) hold for A.’

⁸⁵The case also features a lack of dominance because non-dominance can be seen as a subset of non-dependence. If A and B do not interact with each other at all (non-dependence), then their interaction cannot be of the wrong kind for exploitation (non-dominance).

C6: Desert island dependence.

Suppose there are two fully independent, equally populous desert islands, one wealthy and one poor. The existence of each island is unknown to the other. We may assume resources are equally distributed *within* each island. If the total wealth of those on the rich island is \$15 million, and the wealth of those on the poor island is \$5 million, then the rich exploit the poor by (PR1) and (PR2). Taken together as N , the total wealth is \$20 million; each island could withdraw with \$10 million, which makes the rich island worse off and the poor island better off than before.

that the wealthy *exploit* the poor in the example. Exploitation, at least in ordinary use, is not merely an unjust distribution or an unfair transaction. It is a particular *kind* of unfair transaction. However, Roemer's (PR1) and (PR2) do not allow us to draw any distinction between exploitation and maldistribution.

Roemer offers the following condition to deal with a lack of dependence: "capitalist exploitation can be said to exist only if one coalition is gaining *at the expense of* another"⁸⁶, that is, "if the would-be exploited withdrew with *their own* endowments, the complementary coalition would be worse off."⁸⁷

Sensat argues that the dependence condition is problematic. Although required to assure the sufficiency of Roemer's account, the dependence condition is not necessary for exploitation. In order to motivate this claim he presents the following case.⁸⁸

C7: Billions of burgers.

Suppose I force you to prepare hamburgers for me every day without providing you with anything in return. If this is exploitation, then clearly it is exploitation even if I own a freezer filled with a lifetime supply of prepared hamburgers that I need only pop into my microwave oven to bring them up to the condition of your prepared ones (suppose this takes no more time, energy or trouble than forcing you to work for me does). I am not dependent on you in the sense of the requirement, and yet I am exploiting you. . . mere expansion of the resource base cannot turn an exploitative state into a non-exploitative one.

Sensat draws a distinction between what he calls 'Framework I' theories of exploitation, in which exploitation is seen as a property of a state of affairs arising from a maldistribution of resources and 'Framework II' theories, in which exploitation can be a property of a social system and can include social features such as power and dominance. He concludes that all 'Framework II' theories violate the condition that "the mere expansion of the resource base cannot turn an exploitative state into a non-exploitative one."⁸⁹ While I am sympathetic to his concern that the focus Roemer's and similar theories place on what Sensat calls the "material conditions" of exploitation fail to capture important social factors that also play a role in our willingness to label cases exploitative, his charge that the dependence condition generates a failure of necessity is mistaken.

We may grant Sensat the claim that forced hamburger preparation is exploitation⁹⁰, but his claim that "I am not dependent on you in the sense of the [dependence] requirement"

⁸⁶Roemer (1982c, p.300, emphasis added)

⁸⁷Sensat (1984, p.28, emphasis original)

⁸⁸I quote Sensat directly here from (Sensat, 1984, p.29), but have titled the example myself.

⁸⁹Sensat (1984, p.29)

⁹⁰Although, this is at odds with the paradox of exploitation in which exploitation is a consensual Pareto improvement that, nonetheless, is unjust. Sensat's example might be amended to allow for a very small payment to the hamburger flipper.

is false.⁹¹ If the hamburger flipper is providing some burgers to the exploiter, then the exploiter would be worse off without these burgers. Sensat's charge of non-necessity hinges on an equivocation in what is meant by 'worse off'. One may be worse off in a strict sense or in a practical sense. Strictly speaking, if the exploiter, A, has billions of burgers in the freezer and receives five burgers from the would-be exploited, B, then if B withdraws and A no longer receives burgers, we can call A worse off. His burger endowment decreased from billions plus five to mere billions. Of course, in a practical sense, this will make little difference to A. But clearly this is not the sense in which Roemer intends the condition. Sensat makes another attempt to explain the non-necessity of the condition:

Consider a state x in which according to the theory some individual is exploiting the rest of society, i.e., he would be worse off and his complement better off under coalition withdrawal with per capita shares of means of production. We may also suppose that the individual dominates his complement and that he satisfies the dependency condition. Now expand the resource base by providing him with skills which would enable him to maintain under the withdrawal circumstances the benefit level he enjoys in x . Relative to the expansion, he is no longer an exploiter in x , according to the theory."⁹²

It is unclear what this example is supposed to show. Suppose that the baseline case is one in which A's holdings are \$15 and B's are \$5. If each withdraws with their per capita holdings and dominance and dependence are satisfied, then A exploits B. Now, Sensat asks us to imagine a case where A's skills are such that he can maintain the same level of benefits he enjoyed pre-withdrawal, i.e., he can produce for himself \$15 worth of endowments. It is unclear why we are allowing skills to vary over the counterfactual situation. Surely A exploits B when A depends on B for his unjust holdings, and the situation is not rendered non-exploitative if A has the means for producing more resources to replace those he loses from B's withdrawal. Additionally, even if Sensat's counterfactual can be shown to be well-motivated, the move from endowments to skills confuses capitalist exploitation with socialist exploitation. Suppose we grant that Sensat's modification of A's skill set makes sense. Although it may be the case that A does not *capitalistically* exploit B, A can still be said to exploit B when *socialist* withdrawal conditions—which allow for the equalisation of skills and talents—are employed. Thus, even granting the odd counterfactualisation of skills, Sensat's example seems to rest on a slide between capitalist and socialist exploitation.

Although the desert island dependence case shows the need for a dependence condition, the billions of burgers example does not establish a failure of necessity for Roemer's model when the condition is added. The dependence condition successfully excludes non-dependence cases. When combined with the arm's-length economic transaction restriction, it seems that these restrictions can also rule out non-interaction cases. 'Being exploiters' on Roemer's account is a property of a coalition, not *necessarily* an individual. Coalitions may contain only one individual, but they may also contain many members. Thus, that the exploitation relation holds between the two *coalitions* S and S' does not assure us that it holds between particular individuals in each coalition.

This problem might be solved by simply restricting the language of (PR1) and (PR2) to individuals rather than groups so that, for example, A exploits B iff (in an arm's-length economic transaction) A would be better off (PR1) in the transaction and B worse off (PR2)

⁹¹Sensat (1984, p.29)

⁹²Sensat (1984, p.30)

if B had X and A's doing better depends on B not having X. But by broadening his scope to coalitions, Roemer is able to capture cases where both individuals and groups are worse off not because of the actions of any particular individual, but rather because of many actions of a group. Of course, since each member of the exploiting group may be unaware of the aggregate effect of the multiple actions on the exploited party, they will be unintentional exploiters. Thus the systemic scope of Roemer's theory is related to its inclusion of non-intentional cases. This feature bears a striking resemblance to the tension between unintentional exploitation and moral responsibility encountered in the previous chapter.

Despite Sensat's concerns, Roemer's dependence condition and restriction to arm's-length economic transactions seems to ensure that—if we restrict our attention now to individuals—not only does A gain, he gains *from B*, when he exploits. This leaves us with the original problem of a lack of dominance. We must ensure that A gains from B *in the right way*.

Consider the impact of the restrictions we have considered on a series of versions of the desert island dependence case. Recall that in the original case a wealthy and poor island are completely independent from each other, one is rich; the other poor. If we take the total population of both islands as society N , we can identify the poor island as exploited, according to Roemer's (PR1) and (PR2), but this is clearly incorrect. Roemer responds that exploitation exists only if coalition A's gain comes 'at the expense of' coalition B. This solves the problem with the first version of the case. Restricting the scope of the theory to arm's-length economic transactions and restricting the scope to individuals rather than coalitions avoids problems stemming from a lack of interaction between A and B. However, we may still encounter a problem like the case Elster outlines, in which, although A's gain comes at the expense of B, A does not extract his gain 'in the right way'. A's gain is merely the effect of his odd psychological beliefs about B. The answer to this problem is the addition of a dominance condition. We require that, in order for A to *exploit* B, he dominates her.

Above I mentioned Kymlicka's claim that this approach is *ad hoc* for Roemer, given his stated goal of providing a *purely distributive* account of exploitation. While it may be *ad hoc* for Roemer, the failure of purely distributive conditions to provide an adequate account of exploitation does not lie with the conditions, but the goal. Exploitation does not merely involve an unfair transaction, it involves an unfair transaction that comes about in a particular way. Roemer may have failed to meet his own goals, but the cases in which his account fails show that something more is required. Unfortunately, as stated, (PR3) does not provide an adequate solution to Roemer's problem, since it is a trivial solution. In order to offer a full account of exploitation, Roemer must offer an account of dominance. Only with such an account in hand can we determine, for example, whether cases of unintentional exploitation are possible on Roemer's account.

While Roemer uses the term 'dominance', what problematic cases like Elster's really seem to indicate is that Roemer's account only captures the unfairness of exploitation, not the 'advantage taking'. In chapter 5 I outline an account of 'advantage taking' that may also serve as an operationalisation of the dominance condition (PR3).

3.4.3 The Process of Exploitation

Both Steiner and Roemer offer systemic accounts of exploitation, that is, accounts that allow for cases of exploitation that arise in economic and social systems. The benefit of such accounts is that they allow us to say that, for example, the labourer is exploited not necessarily by the capitalist, but by the system of *capitalism*. This seems correct because in

many cases it is not clear that there is one person or one group who is directly responsible for exploitation. Rather it seems to be brought about by a series of morally bad events and morally wrong actions.

Unfortunately the breadth of scope allowed by systemic accounts encounters two problems. First, although it allows us to capture what we ordinarily call exploitation in some cases, the more permissive conditions that are required to capture these cases also seem to let in cases that we would not call exploitation, for example, diachronic and synchronic pervasiveness, desert island cases, etc. Nevertheless, that a particular account renders exploitation more pervasive than we thought is not in itself a damning charge against that account. If the theory identifies a phenomenon that, on reflection, we find normatively undesirable and it provides a causal account that allows us to identify what brings the phenomenon about and what would lead to its cessation, then that theory provides *pro tanto* reasons for action. It may turn out that there is more exploitation in the world than we once thought. However, as I noted in the discussion of pervasiveness in Steiner's account, the desideratum of co-extension with ordinary use must be balanced with the ability of a theory to capture important normative and causal characteristics of exploitation. The extension of a *viciously pervasive* theory may fully separate from our ordinary use of the term, or it may not act as a guide to action. And in both Steinerian and Roemerian exploitation, providing conditions that capture *only* those systemic cases we would ordinarily call exploitation without overextending the concept seems unlikely. At the very least, it seems that *purely distributive* conditions will not be adequate. An account of exploitation that does not include (much) more than what we ordinarily think of as exploitation must involve conditions that go beyond specifying when the outcome of a transaction is a maldistribution.

Roemerian exploitation, when morally amiss, is so because of unjust inequalities. Steinerian exploitation, when morally amiss, is so because of previous rights violations. These two normative cases amount to the same thing in certain situations. For both accounts, the conditions that lead to exploitation's impermissibility are prior to the systematic or institutional elements that bring it about. The 'normative aspect' of exploitation—a rights violation or inequality—occurs first. The 'causal aspect' of exploitation occurs when A's use (intentional or otherwise) of the vulnerabilities created by the rights violation or inequality allow A to extract some gain from B. The problems of overextension for Roemer and Steiner concern when it is appropriate to say that this second, causal aspect is satisfied. As I noted above, the accounts are trying to pin down when or how 'A gains from B in the right way' for exploitation to occur. The tension between systemic and more personalised accounts of exploitation concerns two different criteria for what it means for A to gain 'in the right way'.

Rather than 'taking on' concerns about the process by which exploitation occurs or an account of when a transaction is unfair (maldistributive), some accounts of exploitation abandon the approach taken by Roemer and Steiner altogether. Instead they place the *process* of exploitation in the forefront. According to these approaches, the attitudes of A towards B—the *way* the two interact—are more important than the individuals' pre-transaction states. In these accounts, the priority of causal and normative aspects is inverted. The systematic or institutional elements that allow for exploitation occur first; what makes a particular act exploitative is the individual's use of the vulnerabilities created by these elements to take advantage of another. Although the background conditions leading to a case of exploitation *may* involve injustice or wrongdoing, they are not necessarily themselves morally tainted. Here, the causal comes before the normative aspect of exploitation.

3.5 Chapter 3 Summary

Marxist exploitation is a broad church, comprising not only the classical Marxist account based on Marx's labour theory of value, but also Roemer's unequal exchange and property relations models of exploitation. In 3.2 I outlined the classical Marxist account and explained why problems with Marx's labour theory of value undermine its success. Section 3.3.3 presented two alternative accounts developed by Roemer: the unequal exchange and property relations account. Here we saw that all that is really necessary for the core components of Marxist exploitation is asset inequality. While important, Roemer's results from the unequal exchange models are limited, applying only to static models with homogenous labour. His property relations model offers a more general account of exploitation that is based on the CONTRIBUTION principle of fair transaction, which coincides with Steiner's RIGHTS principle, at least in special cases. In section 3.4 we saw that the basic version of property relations, characterised by (PR1) and (PR2), allows for exploitation in cases where A does not gain, does not gain from B, and does not gain from B in the right way. We saw that a dependence condition and a restriction of his account to arm's-length economic transactions solve two of these problems, ensuring that A exploits B only when he gains unfairly from her. While the addition of a fully specified (PR3) could, in principle, ensure that A exploits B only when he gains from her *in the right way*, Roemer's version of (PR3) is *ad hoc* and, more importantly, does not define dominance. Both Steiner and Roemer provide accounts that do not restrict exploitation to a particular kind of interaction between A and B, thus allowing for systemic exploitation. Thus, both permit cases of 'exploitation' where individuals are unaware they exploit (and are unaware they are exploited). Although stretching their conceptions to cover such cases allows Steiner and Roemer to identify as exploitation cases that other theories cannot, this wider scope also brings counterintuitive results. In my discussion of these cases in both the present and previous chapter I have remained agnostic about whether accounts *should* try to capture systemic exploitation or whether a narrower scope is more appropriate. The following chapter addresses a family of accounts that places a stronger emphasis on the attitudes involved in exploitation and the processes that bring it about. After exploring these theories we will be in a better position to adjudicate the trade-offs between individualised and systemic accounts of exploitation.

Chapter 4

Duty-Based Exploitation

4.1 Chapter 4 Introduction

In many cases it seems natural to say that A's exploiting B requires that he directly cause a maldistribution. Roemer's and Steiner's accounts, as we have seen, do not always assure us that A's actions are connected to the maldistribution in such a way that we can naturally call the case an exploitation. Sections 3.4 and 2.4 in the two previous chapters have addressed various ways in which the scope of these two accounts might be limited to cases in which A's actions and the maldistribution are appropriately connected.

But many accounts of exploitation find the distributive approaches of Steiner and Roemer misguided. According to accounts of exploitation offered by Ruth Sample and Robert Goodin, exploitation does not necessarily involve maldistribution—at least not in the same way as it does for Steiner and Roemer. Rather, according to their accounts it is possible (but not always necessary) to characterise exploitation by referring to the attitudes and/or duties involved in A's interaction with B. Goodin claims that exploitation involves A's failure to protect the vulnerable¹; for Sample, A exploits B by “interacting with [her] for the sake of advantage in a way that degrades or fails to respect [her] inherent value.”²

In the sections that follow, I outline both Goodin's and Sample's accounts. I argue that although they are correct to claim that A's attitudes are in some way connected to—or necessary for—his exploitation of B, their rejection of Steiner and Roemer-style maldistribution as a necessary component of exploitation creates insurmountable difficulties for their approaches. Section 4.2 considers Goodin's vulnerability-based account and his distinction between general exploitation and the exploitation of a person. In 4.3 I outline and endorse Sample's critique of Goodin's account and consider her own degradation-based approach. Section 4.4 presents a problem for both accounts. Here I show that both Sample and Goodin produce accounts that imply the existence of non-judicial exploitation; yet, purportedly non-judicial exploitation creates a dilemma for the accounts: either they are incomplete or self-frustrating. In order to expunge non-judicial exploitation from the scope of their theories, Sample and Goodin must prohibit prudential risks, but since such a prohibition is implausible, I argue we must instead abandon the ‘come-what-may duties’ on which their accounts are based. Finally, Section 4.5 concludes both the chapter and the first part of the thesis.

¹Goodin (1987)

²Sample (2003, p.57). A similar position is advanced by Allen Wood (1995, p.15), who claims that exploitation is the violation of “proper respect for others [by treating] their vulnerabilities as opportunities to advance our own interests.”

4.2 Goodin's Vulnerability-Based Theory

Goodin (1987) develops a theory of exploitation grounded in the duty to respect the vulnerable.³ The analysis of exploitation, Goodin argues, is “parasitic upon an analysis of [the] duty to protect the vulnerable.”⁴ Goodin begins by distinguishing what he calls “advantage₁” and “advantage₂”, which correspond to bargaining power and the benefits received from a cooperative venture respectively. He notes that exploitation does not merely involve strategically positioning oneself to maximise bargaining power, but only occurs when superior bargaining power is actually exercised to extract benefits. Nor, he argues, is every benefit gained through the use of superior bargaining power the result of exploitation; rather, he suggests that “exploitation refers to the taking advantage of an *unusual* situation”, either when the situation itself is non-standard, or when we make non-standard or non-paradigmatic use of ordinary situations.⁵ Finally, he notes that only agents can exploit. Thus, generally speaking, agents exploit when they use superior bargaining power to secure benefits in ways or situations that are non-paradigmatic. So, as an example, we exploit a concert when we use it to distribute advertising fliers at the gate. Concert goers attend to listen to music, not to collect fliers; thus our distribution of fliers is ‘non-standard’. However, Goodin claims that unlike the exploitation of *persons*, this form of exploitation is not always wrong.⁶

4.2.1 Exploiting Persons

Although exploiting a person is always, *pro tanto*, wrong⁷, Goodin argues that we must also distinguish between whole-person exploitation and exploiting a person's attributes: “We might exploit their strong backs or weak minds... their fears, ignorance, superstitions... love or trust... [but] we infer whole-person exploitation from the fact that certain of the person's attributes are being used if they are being used *unfairly*.”⁸ But, this notion of unfairness must be explicated. We need an account of what it means to use a person unfairly. Goodin considers four possible interpretations of the unfair use of persons: coercion, non-reciprocity, manipulation, and (mere) use, but ultimately rejects each.

The unfairness in exploitation is not to be understood as coercion, since as Steiner also mentions, this would make exploitation indistinguishable from theft.⁹ Nor is it to be understood as a lack of reciprocity, since although some cases where reciprocity is lacking may be exploitative, not all cases are (for example, gift-giving). Therefore, non-reciprocity is not a sufficient condition for exploitation. Nor, Goodin argues, is it a necessary condition, since “we can have such a thing as ‘reciprocal exploitation’.”¹⁰ According to Goodin, the third option, manipulation, has “three salient features”: it is active, i.e., intentional, it is deceptive or hidden, and it is “contrary to the putative will of the persons being manipulated.”¹¹ But although exploitation may satisfy the third feature, Goodin claims that there exist cases of exploitation that are unintentional and “out in the open”, and that exploitation cannot be

³Goodin establishes the foundational argument for this duty in his book *Protecting the Vulnerable* (1987).

⁴Goodin (1987, p.167)

⁵Goodin (1987, pp.169–70) Goodin emphasises that his use of ‘usual’ does not connote ‘frequent’ but rather refers to the socially accepted intended use of things and situations.

⁶Though it may sometimes be wrong. Goodin develops an account of the wrongness of exploiting mere things at the end of his analysis.

⁷Goodin (1987, p.173)

⁸Goodin (1987, p.171)

⁹Steiner (1984, p.227), Goodin (1987, pp.174–75)

¹⁰Goodin (1987, p.176). It should be noted that the existence of reciprocal exploitation—especially in Goodin's consequentialist framework—is dubious.

¹¹Goodin (1987, pp.177–78)

unfair because it is manipulative.¹² The remaining possibility is that exploitation is unfair because it involves the use of others to attain our own ends.

Now, we use people far more often than we exploit them, so the notion of ‘use’ that underpins the unfairness of exploitation must itself be restricted. Goodin considers three restrictions: mere instrumental use, non-consensual use, and harmful use. There exist cases of the first, such as a butcher and baker’s mutual (mere instrumental) use of each other, that are not cases of exploitation; mere instrumental use is thus not sufficient for exploitation. Nor does non-consensual use imply a case of exploitation: we do not, Goodin argues, exploit a larger spectator at an outdoor event when using their body’s shadow to shield our eyes from the sun. Although we are non-consensually using an attribute of the spectator, we are not engaging in whole-person exploitation.¹³ So, non-consensual use is not sufficient for exploitation. Nor is non-consensual use necessary, for we may exploit someone, claims Goodin, by renting their womb, even when they consent.¹⁴ Finally, we cannot explicate the notion of unfair use through an appeal to harmful use because some of the most troubling cases of exploitation are just those where the exploited do in fact benefit, at least when compared to a no-transaction baseline.

Goodin concludes that in the forms considered, neither the three elaborations on unfairness (coercion, non-reciprocity, and manipulation) nor the three senses of use (mere instrumental, non-consensual, and harmful) succeed as an account of the wrongness of whole-person exploitation. However, he remarks that use *simpliciter* is a necessary condition: “what more is required in order to make it sufficient is . . . some specification of the *ways* in which people are being used.”¹⁵

4.2.2 Using the Vulnerable

Here Goodin abandons the language of ‘use’, opting instead for the term ‘play for advantage’. He argues that the unfairness associated with exploitation lies “in *playing for advantage* in situations where it is *inappropriate* to do so” and “exploitation thus consists in playing games of strategy in circumstances which render them somehow inappropriate.”¹⁶ Goodin identifies four kinds of situation where playing for advantage is inappropriate: when playing against those who have “renounced playing for advantage themselves”¹⁷, “who are no match for you in games of advantage”, “who are unfit or otherwise unable to play in games of advantage at all”, and “when your relative advantage derives from others’ grave misfortunes.”¹⁸ But each of these situations are “manifestations of one particular kind of wrong,” namely, a failure to honour the moral norm of “protecting the vulnerable. . . *regardless of the particular source of their vulnerability.*”¹⁹

We can fail to protect the vulnerable in one sense by failing to aid them, but we can also fail them in a more grievous sense by using their vulnerabilities for our own personal

¹²Goodin (1987, pp.177–78)

¹³Goodin (1987, p.179)

¹⁴Goodin (1987, p.179). It should be noted that although Goodin makes this claim he does not show it can be incorporated by his theory of exploitation. Thus, despite this quote, it seems that the extension of Goodin’s account does not include those transactions that might fall under the remit of an anti-commodification argument. For further discussion see §4.3.2.

¹⁵Goodin (1987, p.180)

¹⁶Goodin (1987, p.184, emphasis original)

¹⁷Michael Otsuka has pointed out to me that it is not clear that playing for advantage in this situation is inappropriate. Is it inappropriate to compete with the wealthy robber-baron who has now renounced his cut-throat tactics (but nevertheless keeps his wealth)? This particular concern is related to the problems I will outline in section 4.4.

¹⁸Goodin (1987, pp.184–87)

¹⁹Goodin (1987, p.187, emphasis added)

gain. It is this flagrant violation of the norm²⁰ that Goodin claims “wholly accounts for the wrongness of interpersonal exploitation.”²¹ Goodin concludes his analysis of exploitation by arguing that the duty to protect the vulnerable implies not only a negative duty to refrain from exploiting them, but also a positive duty to “do whatever we can to prevent *anyone* (ourselves or others) from taking unfair advantage of people’s vulnerabilities.”²²

In short, Goodin argues that ‘exploitation in general’ involves taking advantage of an unusual situation. But the negatively moralised form of exploitation that we are concerned with involves the exploitation of *a person*. The wrongful use of others’ attributes is exploitation when we use their vulnerabilities, no matter the source, to secure for ourselves some form of benefit in games of strategy. We have a duty, Goodin claims, not only to refrain from exploiting people in this way, but also to do whatever we can to ensure that these vulnerable are not exploited by others.

4.2.3 Goodin’s Normative Foundation

The normative foundation for Goodin’s account depends on the duty to protect the vulnerable.²³ In his (1986) paper Goodin presents what he calls the ‘first principle of individual responsibility’ as follows:

If A’s interests are vulnerable to B’s actions and choices, B has a special responsibility to protect A’s interests; the strength of this responsibility depends strictly upon the degree to which B can affect A’s interests.²⁴

Goodin argues that responsibility for protecting the vulnerable means a responsibility to protect others from harm, regardless of whether the source of the harm is in human action or natural events. Further, Goodin argues that this responsibility falls upon “whoever has the ‘last clear chance’ to prevent harm. . . regardless of the causal history up to that point.”²⁵ For Goodin there is no important moral difference between acts and omissions.²⁶ Further, we should not conceive of the harm that might befall persons as merely material harm; rather, a person is harmed when their general welfare is compromised. Yet, Goodin also notes that material interests should not be altogether excluded: “there are certain sorts of things that the principle would require us to do for others no matter what particular standard of interests we adopt. Specifically we would be required to protect people’s ‘needs’ or ‘vital interests’: these are ‘primary goods’.”²⁷ This focus on the outcomes or consequences of vulnerability makes his account “fundamentally consequentialistic in form”²⁸ since “it is always responsibility for the consequences of one’s actions and choices that is at issue.”²⁹ Goodin equates vulnerability with harm and harm with a loss of welfare. Thus, the duty to protect the vulnerable can be expressed as a duty to minimise welfare loss, or conversely, to maximise welfare.

²⁰Goodin claims that exploitation is a “flagrant violation of a duty” (p.167) and that our obligations originate in a “norm” (pp.188–89). At the end of his paper he returns to the use of ‘duty’ (p.189). These shifts give us some insight into the sense in which Goodin is using ‘duty’: for Goodin the duty to protect the vulnerable depends on the existence of a commonly agreed upon norm.

²¹Goodin (1987, p.188)

²²Goodin (1987, p.189)

²³Goodin (1986)

²⁴Goodin (1986, p.118). Goodin also develops a “Principle of Group Responsibility” and a “Second Principle of Individual Responsibility”, both of which deal with special cases of distributed responsibility.

²⁵Goodin (1986, p.129)

²⁶Goodin (1986, p.110)

²⁷Goodin (1986, p.111)

²⁸Goodin (1986, p.116)

²⁹Goodin (1986, p.113)

4.3 Sample's Degradation-Based Account

Ruth Sample argues that Goodin's consequentialist normative foundation creates three problems for his account of exploitation. First, at least *prima facie*, there seems to be a tension in the idea that we have special obligations to intimates and the underlying consequentialism Goodin espouses; second, it is unclear how Goodin's account can solve the paradox of exploitation since (strictly) Pareto-improving transactions will entail welfare improvements. Finally, the account seems subject to the traditional demandingness objections levelled against consequentialist accounts.³⁰

Nevertheless, Sample acknowledges the importance of vulnerability, but seeks to limit the scope of our obligations to the vulnerable. Thus, in this sense, her account can be seen as a refinement of Goodin's. The normative foundations of Sample's account are non-consequentialist. She argues that exploitation involves interacting with another being for the sake of advantage in a way that degrades or fails to respect the inherent value in that being.

4.3.1 Critique of Goodin

Sample sees Goodin's consequentialist foundation as fundamentally mistaken, arguing that it "has all of the traditional drawbacks and counterintuitive results of consequentialism" and is ultimately unable to account for our intuition that we owe more to intimates than strangers.³¹ Although Sample critiques the tension between Goodin's consequentialism and his desire to incorporate special obligations into a vulnerability-based moral principle, she does not argue that the tension *cannot* be resolved, but only states that "it is unclear that the notion of obligation is consistent with Goodin's framework."³² While the existence of such a tension is clear, Goodin is not unaware of this tension; indeed, he spends a great deal of time in his book *Protecting the Vulnerable* (1986) addressing it. Nevertheless, an evaluation of Goodin's arguments here would take us far afield. Since Sample claims that Goodin's account suffers from two remaining problems. In the interest of space, I will turn to these concerns.

Sample worries that Goodin's consequentialism leaves him unable to provide a solution to the paradox of exploitation. Both Sample and Goodin agree that there are some consensual, Pareto-improving transactions that are, nevertheless, impermissible. They must, therefore, both abandon the non-worseness premise. Yet as Sample states, "if consequences are ultimately what matter, Goodin needs an argument to show that exploitation in general fails to bring about good consequences. . . [but] it is hard to see how mutually beneficial exploitation is a case of failing to bring about good consequences."³³ In short, Goodin's consequentialism seems to commit him to the *acceptance* of the non-worseness condition.

Sample notes that Goodin can claim we are not merely obliged to bring about good consequences to the vulnerable, but we ought to "optimally enhance" their welfare.³⁴ But in this case, she argues, Goodin's account is too morally demanding. Surely exploitation does not involve the failure to *optimally enhance* the welfare of those with whom we transact. But, Sample claims, without a successful argument against the non-worseness premise, Goodin is

³⁰Sample (2003, p.27)

³¹Sample (2003, p.47)

³²Sample (2003, p.46)

³³Sample (2003, p.51) Sample's argument here is weak. Sebastian Köehler reminds me that a consequentialist can be a value pluralist. For example, Goodin can attach value to welfare and also to how that welfare is created (e.g. whether it is created by harming the vulnerable). This might give a high amount of welfare a lower moral value than a lesser amount of welfare if there former was created by harming the vulnerable.

³⁴Sample (2003, p.53)

unable to provide a solution to the paradox of exploitation. Goodin seeks to limit the scope of our obligations by claiming that “on balance, persons relatively near to us in space and in time probably *will* be rather more vulnerable to us. . . . This fact saves my argument from the traditional *reductio* of requiring that we give everything we have to starving Asians. . . .”³⁵ Sample disagrees:

The *reductio* works. . . global communication has made it possible to collect, store, and transfer data about others quickly and accurately. Not only can we discern who is suffering from starvation, dehydration, and disease. . . but it is also possible to match organ donors to recipients with just the same technology. The vulnerability model implies enormous moral debt to countless unknown others.³⁶

My own view is that Sample and Goodin’s debate about whether the scope of Goodin’s theory extends to ‘unknown others’ is tangential to a much deeper problem. I believe that the *primary* demandingness objection that Goodin’s account faces is not that it might oblige us to persons half a world away, but that it might too easily oblige us to persons *wherever* they live. Goodin’s insistence that we have not only a negative duty to refrain from exploitation but a positive duty to do “whatever we can to prevent *anyone*” from taking advantage of vulnerabilities *regardless of the source* is extremely morally demanding and leaves scant room for personal responsibility, an issue I address in section 4.4.³⁷ For now though, it is clear that Goodin must provide additional argument, both in order to account for both forms of demandingness and to show how his account of exploitation can condemn Pareto-improving consensual transactions. It seems that an account of exploitation that incorporates Goodin’s concerns about the misuse of vulnerable others in a non-consequentialist framework could allow us to characterise the concept by considering the attitudes and duties involve in A’s interaction with B without the problems of consequentialism. Sample’s goal is to provide such an account.

4.3.2 Exploitation as Degradation

Sample recognises the need for a theory that sits “between the extreme level of obligation entailed by the views of Peter Singer and Robert Goodin, on the one hand, and the extremely deflationary account of obligation delivered by voluntarism. . . and the other views that emphasise negative rights, on the other.”³⁸ For Sample, it is neither the case that we are responsible only to those whose vulnerability is the result of our voluntary actions, nor is it the case that we are committed to doing whatever we can to prevent *anyone* from being exploited. While Goodin’s focus on vulnerability is well-motivated, Sample argues this vulnerability is not itself the source of our obligation to others. Rather, people may be vulnerable because their basic needs are not met, or because they are the victims of past injustice, but in either case, taking advantage of such vulnerability means we degrade or fail to appropriately respect the person. Vulnerability, Sample argues, “is thus dependent upon an account of what humans require for a good life.”³⁹ And exploitation stems from our failure to respect others by ensuring that their needs are met when we do choose to interact with them. According to Sample’s account, exploitation involves

³⁵Goodin (1986, p.121, emphasis original)

³⁶Sample (2003, p.48)

³⁷Goodin (1987, p.187, emphasis added)

³⁸Sample (2003, p.75)

³⁹Sample (2003, p.74)

... interacting with another being for the sake of advantage in a way that degrades or *fails to respect* the inherent value in that being. It is this lack of respect that explains the badness of exploitation... This accommodates the fact that moral agents can fail to demonstrate respect for persons not simply in the course of harming them, but also when improving their situation.⁴⁰

These failures of respect fall into three general domains. We can fail to show proper respect for, or degrade the value of other persons by:

- (D1) Taking advantage of an injustice done to him.
- (D2) Neglecting what is necessary for that person's well-being or flourishing.
- (D3) Treating as a fungible object of market exchange, an aspect of that person's being that ought not to be commodified.⁴¹

The first form of disrespect echoes the domain of exploitation found in Steiner's and Roemer's theories and, as we might expect, Sample's elaboration of (D1) bears a strong resemblance to Steiner's account. She claims, "If a person is in a weaker bargaining position because of a past injustice, we stand to gain disproportionately in virtue of that injustice."⁴² But Sample extends the domain of exploitation to cases covered by (D2) and (D3).

Sample argues that when we interact with others in a way that benefits them "but nonetheless would ignore [their] basic needs, then [we] must ask whether the maxim of [our] action would pass the test of the categorical imperative. If my action cannot pass this test, then it is prohibited."⁴³ When we transact with others whose basic needs are not met we fail to respect their value as persons and, thereby, exploit them.⁴⁴ Sample notes that there are many accounts of those goods that constitute basic needs or human flourishing, but she argues that the concept is best elaborated by Amartya Sen's capabilities approach,⁴⁵ and in particular, Martha Nussbaum's interpretation of basic capabilities.⁴⁶ Thus, Sample argues that we should interpret basic needs as capabilities and exploitative interactions as "those in which the capabilities of our interactors are ignored⁴⁷ in the pursuit of our own advantage."⁴⁸

Finally, we may also fail to respect others by commodifying an aspect of their person that ought not to be commodified. There are some transactions in which the "nature of the transaction itself" is degrading.⁴⁹ When we engage in such transactions for the sake of advantage we engage in exploitation. Sample notes that "just what sorts of things ought

⁴⁰Sample (2003, p.57, emphasis added)

⁴¹Sample (2003, p.57)

⁴²Sample (2003, p.82). One significant difference between Sample's account and Steiner's is that Sample argues that if a person's bargaining power is not "seriously compromised" (Sample, 2003, p.83) by an injustice, then they are not exploited; for Steiner, all unjust disadvantages can lead to exploitation.

⁴³Sample (2003, p.71)

⁴⁴It should be noted that this test of the categorical imperative, in its rudimentary form, is problematic. Also problematic is Sample's application of the test to exploitation. It doesn't seem that all actions falling within Sample's example fail such a test: when I buy a handmade wallet from a man at a craft fair but give no regard for whether his basic needs are met do I *really* fail to respect his value as a person? It does not seem that Sample means that exploitation is caused by mere neglect of what is necessary for well-being, since one might neglect the well-being of the well-off. Rather, more charitably she seems to mean something like the following: exploitation can be caused by the failure to constrain one's transaction with persons who are blatantly impoverished.

⁴⁵Sen (1999)

⁴⁶Nussbaum (2003)

⁴⁷Note that Sample does not mean 'ignored' in the literal sense, but rather something like 'not given proper weight' or 'not considered from a moral point of view'. In many cases, in order to exploit others we must in fact pay attention to their capabilities, if only to use them for our own gain.

⁴⁸Sample (2003, p.81)

⁴⁹Sample (2003, p.81)

not to be subject to market exchange is of course a controversy. Hence the exploitation claims made by appealing to such an anticommodification argument are also controversial.”⁵⁰ Sample refrains from endorsing a particular account of those transactions that fall within the domain of (D3), but, in places, she suggests that the sale of sex is a characteristic example.⁵¹

For Sample, most cases of exploitation involve a failure to meet the basic needs of the vulnerable party; many others occur when we take advantage of past injustice. She implies that relatively few cases will involve the anticommodification domain (D3). We can imagine her account as a nested account: all transactions⁵² in which the benefits of the transaction fall below a certain basic needs level for one person are exploitative. Some transactions above this level are also exploitative, namely, those involving historical injustice. Finally, some transactions that are neither below a basic needs level nor affected by past injustice are also exploitative because they are inherently degrading.

Note that, like Steiner, Sample considers the general domain of exploitation to be that of bilateral and voluntary transactions. Her account is clearly an individualised account that does not allow for systemic exploitation. Interestingly she also argues that unintentional exploitation is possible when exploiters “believe that their interactors are not deserving of respect. . . [when they are] mistaken in their accounts of what respect requires. . . , [and when they are] mistaken in their judgement of what it takes to fulfil the requirement of respect.”⁵³ Sample also claims that exploitation is a negatively moralised concept, but that it may at times be morally *bad* (as opposed to morally wrong). Although she does not explicitly address the tension between unintentional exploitation and the claim that it is a negatively moralised concept, Sample’s remarks about the moral badness of exploitation indicate that she would prefer this solution to the problem.

Sample’s move to a nonconsequentialist framework allows her to avoid the problem plaguing Goodin’s theory: nothing in her ethical framework prevents her from claiming that Pareto-improving consensual transactions *are* impermissible when they degrade one of the parties. Thus, she offers a consistent solution to the paradox of exploitation. An appeal to disrespect (degradation) allows her to reject the non-worseness premise. Further, her shift from Goodin’s claim that we are obliged to the vulnerable regardless of the source of their vulnerability to the claim that we are obliged to those who are vulnerable in (D1) and (D2) limits the demandingness of her theory. Whether Sample’s inclusion of the anticommodification domain (D3) is an improvement upon Goodin’s theory or a weakness is debatable.⁵⁴ Sample offers little in the way of a normative defence of (D3) and does not endorse “any particular conception of which features of persons are beyond exchange.”⁵⁵ Noting that it is “the most controversial” of the domains, she offers what seems to be a defence of (D3) on descriptive grounds: “whether one finds these anticommodification claims plausible, it seems clear that they constitute an important class of putative judgments of exploitation claims.”⁵⁶ We can accept the claim that people do speak about exploitation in this way, without agreeing these claims have *normative* force.

⁵⁰Sample (2003, p.83)

⁵¹Sample (2003, p.95)

⁵²Sample does offer a caveat here: if the only possible mutually beneficial transaction is one in which the basic needs of one person cannot be met then the transaction is not exploitative (Sample, 2003, p.75). The concerns that motivate this caveat will reemerge in chapter 6.

⁵³Sample (2003, pp.58–59)

⁵⁴Although Goodin agrees with Sample that “there are some uses to which people should not be put” and that it would be “exploitative to use a person in these ways”, it is not clear (nor does Goodin tell us) how such uses can be wrong based only on an appeal to vulnerability (Goodin, 1987, p.181).

⁵⁵Sample (2003, p.75)

⁵⁶Sample (2003, p.58)

4.3.3 Objective Value and Non-commodification

It is time to make good on chapter 3's promise to discuss objective value-based approaches to exploitation in more detail. In that chapter I noted that objective value theories may be used to support accounts of exploitation in two ways, either by identifying a moralised baseline that grounds the claim that a particular transaction does not give 'enough' to B, as Marx's labour theory of value does, or by claiming that a particular *kind* of transaction, such as the commodification of sex, is impermissible.

In chapter 3 I also provided a number of reasons for rejecting Marx's particular objective theory of value. While many other labour-based and cost-of-production-based objective value theories exist, there are a number of reasons to curtail discussion of this approach to exploitation. The first and most practical is that space and time do not permit an excursion into territory tangential to the current project. The second reason is that at least some of the problems that Marx's theory faces are not limited to his account, but rather are problems that exist for a whole family of objective value theories. These common problems give us some reason to see whether subjective value theories can deliver an adequate account of exploitation. The third and final reason is that, problems related to establishing anything like wide consensus about the plausibility of a particular objective value theory aside, objective approaches make the solution of the paradox of exploitation quite straightforward. Pareto-improving, consensual transactions are impermissible when the goods exchanged do not trade at their fair price. There is simply nothing morally perplexing about exploitation for the objective value theorist (once they develop their value theory). The 'hard' problem of exploitation arises only for the *subjective* value theorist. If goods are simply worth what people will pay for them, then it cannot be the case that a particular transaction is an exploitation because the goods don't exchange for what they are worth. This feature of subjective value theories has led many to claim that "liberalism is conceptually incapable of construing any consensual exchange as exploitative"⁵⁷ and thus to simply reject the concept of exploitation out of hand.⁵⁸ Thus, one reason for focusing on subjective theories of value is to see whether such dismissals of exploitation are justified.

However, these three reasons are not (at least directly) reasons for abandoning Sample's claim that certain transactions are exploitative because they involve non-commodifiable goods. One may claim, as Sample appears to, that most of the goods that we exchange are 'normal' goods that are not subject to the anticommodification argument. The value of these 'normal' goods is determined simply by individuals' subjective preferences. Those 'special' goods that form the basis of the anticommodification argument, on the other hand, have a value that is not determined by subjective preferences, and it is this special objective value that prohibits exchanges of these goods. This kind of pluralism does not escape the 'hard' problem of exploitation, for even if *some* exploitations involve trade in non-commodifiable goods, Sample must still explain—without appealing to objective values—what is wrong with those cases that involve Pareto-improving consensual transactions of 'normal' goods. From the pluralist perspective then, anticommodification exploitation is an extension of ordinary exploitation.

Sample's comments in the quotes from above indicate that she is aware that lists of non-commodifiable goods are "controversial" and indeed the difficulty of defending such a list

⁵⁷Steiner (1987, p.132)

⁵⁸See, for example, Nozick's claim that "Marxian exploitation is the exploitation of people's lack of understanding of economics." (1974, p.264)

seems to be behind her reluctance to endorse “any particular conception of which features of persons are beyond exchange.”⁵⁹ However, it is possible to make the weaker claim that certain aspects of life ought not to be commodified *without* also holding the view that there is anything *intrinsically* wrong with doing so. Rather, our worry about the commodification of, for example, child labour could be that it exists only because of past injustices and inequalities. “The problem with child labor, on this view,” writes Debra Satz, “is the whip of poverty and hunger that compels parents to put their children to work, not the market in child labor itself.”⁶⁰

This approach seems preferable for Sample since it would allow her to claim that (for example) *nearly all* exchanges of sex for money are exploitative without the need to defend the claim that such exchanges are intrinsically morally amiss. Such transactions would be exploitative when they violate either of the two remaining domains: past injustice and basic needs. Further, if nearly all of a particular kind of transaction involve violations of these two domains, we have good practical reasons to enact legal prohibitions or restrictions on these transactions.

Thus, while there are two ways in which we might base an account of exploitation on an appeal to objective values, neither seems well-motivated. First, subjective theories of economic value, such as Marx’s labour theory of value, face many challenges and explaining why even those who endorse a subjective value must worry about exploitation is a valuable project. Second, claims that some forms of exploitation involve trade in goods that we ought not commodify may be made without the need to appeal to the claim that these forms of transaction are intrinsically morally amiss. This approach reduces Sample’s domains of exploitation from three to two. In the following section I present a dilemma for both Sample’s and Goodin’s accounts that reduces Sample’s account of exploitation to the first domain of exploitation, leaving her with an approach equivalent to Steiner’s.

4.4 A Dilemma for Duty-Based Accounts

It has been commonly acknowledged that exploitation is—or involves—taking unfair advantage.⁶¹ Of course, as Richard Arneson points out, *useful* definitions must go further. If exploitation is simply taking unfair advantage, then there will “be as many competing conceptions of exploitation as theories of what persons owe to each other by way of fair treatment.”⁶² Steiner and Roemer go further by arguing that exploitation falls within the domain of injustice. Sample and Goodin argue that although exploitation can often be brought about by injustice, it may also involve a failure to be constrained by others’ disadvantage.

In particular, according to Sample’s and Goodin’s approaches, exploitation may be caused by taking advantage of *mere* poverty, vulnerability, or disability, that is by taking advantage of mere disadvantage. Goodin claims disadvantage can generate exploitation “regardless of the particular source”⁶³ of the disadvantage, and Sample claims that exploitation may occur when we neglect “what is necessary for [a] person’s well-being or flourishing.”⁶⁴ In

⁵⁹Sample (2003, p.75, 53)

⁶⁰Satz (2010, p.5) Satz herself disagrees, arguing that “we have reasons to block certain markets, to limit the domain of things that money can buy, even when such limits cannot be justified by considerations of economic desperation or by a prior unjust distribution of income and wealth.” (Satz, 2010, p.5)

⁶¹See, for example, Goodin (1987); Sample (2003); Steiner (1984, 1987); Tormey (1976); Wertheimer (1996); Wolff (1999); Wood (1995). Wood is an exception: on his account although taking unfair advantage is sufficient for exploitation, it is not necessary.

⁶²Arneson (1992, p.350)

⁶³Goodin (1987, p.187)

⁶⁴Sample (2003, p.57)

contrast, while the disadvantaged—when *unjustly* disadvantaged—may be called exploited under theories such as Steiner’s, which admit only injustice as a cause of exploitation, those who are not unjustly disadvantaged cannot be exploited.⁶⁵ In other words, for the present debate we may suppose the sources of exploitation fall within two domains: the domain of injustice (D1) and the domain of disadvantage (D2). Sample and Goodin claim that exploitation may be brought about by causes from either domain, while Steiner and Roemer deny that the subset of (D2) that does not intersect with (D1) can be a source of exploitation. The latter admit only that subset of (D2) that intersects with (D1). Although I will not consider his account in this section, Allen Wood summarises the juridical versus non-juridical debate nicely when he writes,

It is often thought that exploitation is bad because the exploiter takes *unfair* advantage of the exploitee. But [for Sample and Goodin] we have reason to consider exploitation bad even when it involves no unfairness or injustice. For if the badness of exploitation lies in the fact that it is base and dishonourable of the exploiter, and insulting and degrading to the exploited, for the former to make use of the latter’s vulnerability, then this badness seems to be present even where the exploitation involves no unfairness, injustice, or violation of rights.⁶⁶

Thus, while both accounts acknowledge that injustice can generate exploitation, they also argue—implicitly or explicitly—that various forms of disadvantage (poverty, vulnerability) may be independent sources of exploitation.

4.4.1 Sample’s Dilemma

Although the existence of non-juridical exploitation brought about by mere disadvantage seems plausible, Sample and Goodin face a common dilemma: all cases of mere disadvantage-generated non-juridical exploitation are either self-frustrating or incomplete. Since the accounts differ slightly, I present a case that highlights the dilemma for Sample’s account first before considering a case that presents a problem for Goodin’s account.

Consider the following sequential game, given in both strategic and narrative form. In the strategic form tree, ‘A’ and ‘B’ denote the players, ‘O’ a chance node, and ‘C’, ‘W’, and ‘E’ represent the options ‘cooperate with’, ‘work for’, and ‘employ’, respectively.

⁶⁵That is, they cannot be exploited in virtue of this particular disadvantage. They may at some time, past or future, be victims of injustice and thus face the possibility of exploitation.

⁶⁶Wood (1995, p.17)

C8: The lotteries (sequential).

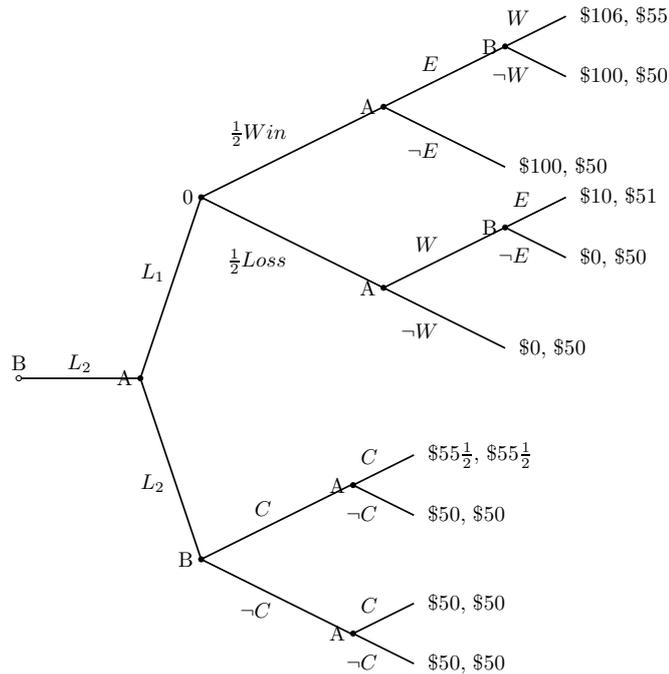
Stage 1: Consider the following two lotteries, L_1 and L_2 . Individual A may choose either, but B is (for whatever reason) committed to L_2 .

$$L_1 \ (p = \frac{1}{2}, \$100; 1 - p, \$0)$$

$$L_2 \ (p = 1, \$50)$$

Stage 2: A and B now face each other in the labour market. The wealthier of the two is the employer and pays wages out of the revenue he earns to the less wealthy, who is the employee. If they are equally wealthy they may work cooperatively, paying no wages and splitting revenue equally. Suppose the values for revenue, wages, and profit are as follow: $R = \$11$, $\omega = \$5$, $\pi = 6$. The revenue from a cooperative venture is thus $\$5.50$. Finally, suppose the following constraint, embodying Sample's particular approach to (D2), holds:

BN If a player chooses to transact with another, then all players' outcomes must be $\geq \$10$.



C8 in strategic form.

When B is committed to L_2 , which lottery⁶⁷ should A (rationally) choose, assuming players have full knowledge of the game structure? We can answer this question using backward induction. Let us first consider the bottom branch, supposing A chooses L_2 . In this case, the players will be equally wealthy, each holding \$50. They may choose to form a cooperative venture or to not cooperate, but clearly it is in the interest of each to cooperate, so the outcome of A choosing L_2 is (\$55.50, \$55.50).⁶⁸

Suppose A chooses L_1 . Then, there are two equiprobable possibilities: either A will win the lottery, or A will lose the lottery. Suppose A loses, then after the lottery A will be the poorest. He can choose to work for B or to remain unemployed, and B can choose whether to employ A, but both will choose employment. Since post-lottery A has \$0, BN holds. B must pay A \$10 in wages, but she keeps \$1 for herself. The outcome of a lottery loss is thus (\$10, \$51). What if A wins the lottery? Then A will be the wealthiest player. And again, it is rational for both players to choose employment. This time A is the employer and, since both players' holdings are greater than \$10, BN doesn't apply. A will pay B \$5, keeping \$6 for himself in profits. Thus, the outcome of a lottery win is (\$106, \$55). Now, a lottery win and a lottery loss are both equally likely, so the *expected value* of choosing L_1 is $\frac{1}{2}$ the outcome of a win and $\frac{1}{2}$ the outcome of a loss, or (\$58, \$53). Since the expected value of L_1 is greater for A than the expected value of L_2 , A will rationally choose L_1 .

In order to determine the effect of (D2) on the outcome of the case⁶⁹ we must consider two variants of the case⁷⁰: one with, and one without the (D2) constraints (BN) in place. When we consider the variant where BN does not hold, it is clear that a difference in the outcomes of A's choice is generated by BN. If BN does not hold, the outcome of A choosing L_1 , losing the lottery, choosing work, and B choosing employ (i.e., the fourth terminal node from the top) changes to (\$5, \$56), making the expected value of L_1 (\$55.50, \$55.50). Thus, if BN doesn't hold, the expected value of A's choice is equivalent for both players and for both lotteries. Table 4.1 lists the expected value of each lottery for A and B, with, and without BN.

		B			B
		L_2			L_2
A	L_1	\$58.00, \$53.00	A	L_1	\$55.50, \$55.50
	L_2	\$55.50, \$55.50		L_2	\$55.50, \$55.50

Table 4.1: Expected values with (left) and without (right) BN.

When BN is in place, if A chooses L_1 , A does better and B worse. Since this difference is brought about by A's actions, A gains \$2.50 of (expected) value at B's expense. With BN in place, players have an incentive to choose L_1 and to avoid L_2 . In other words, if given a meta-choice between being in the position of A or B, players would choose to be A because A's choice of L_1 affects the expected value of L_2 for B.⁷¹

Intuitively, and from an ordinary language point of view, when BN holds and A chooses

⁶⁷We may consider these lotteries as something like choices between alternative social positions.

⁶⁸I follow the convention of representing outcomes as: (A, B) and strategy profiles as $\langle A, B \rangle$.

⁶⁹Since the goal is to determine the effect of being bound by BN on the outcomes, we must consider variations in the presence of BN since "event 'e' depends causally on [possible cause] 'c' if and only if, if c had not been, e had never existed" (Lewis, 1973, p.563).

⁷⁰Strictly speaking, these are two different games, but here I refer to them as variations of one case.

⁷¹The inability of B to choose her own lottery is a topic I will return to below.

L_1 , A takes unfair advantage of—exploits—B. The presence of BN means that B subsidises the loss A experiences. If A loses, B must ensure that if she transacts with A, A's basic needs are met. Yet, A pays nothing for the chance of subsidy: if A wins the lottery, he does not share the gains with B. Since BN requires part of A's costs to be covered by B, but does not require that A share his gains with B, BN increases the expected value of choosing L_1 for A. Clearly, A takes advantage of B's being bound by BN.

Is A's advantage *unfair*? It seems it is. A's use of BN to secure greater expected value for himself is at odds with the spirit, if not the letter, of BN. Presumably the principle is meant to help those who are truly in need of aid—the *real* poor—and not those who, like A, began in a situation that is better than the situation of B and end up worse off through their own choice. Before A's choice, B could expect⁷² at least \$50, and A's choice set contained all options in B's choice set (L_2 , plus L_1). Thus, pre-choice, A is doing at least as well as B. Since A could have chosen L_2 , it is unfair for him to use BN to subsidise his losses if he chooses L_1 and loses. Thus, A not only takes advantage of B, but his doing so is unfair. If exploitation involves taking unfair advantage, then since A takes unfair advantage of B in the lotteries case, A exploits B in this situation.

Sample may respond that this means the choice of L_1 is impermissible. Since choosing L_1 leads to A's exploiting B, A may not choose the risky option. If this choice is impermissible, then in a deontically perfect version of the lotteries case, A cannot become disadvantaged and therefore, cannot be in a position to receive the BN subsidy. Therefore, he cannot exploit B. However, if A is not in a position to receive the subsidy, B is not obliged to abide by BN and there is no possibility of disadvantage-based non-judicial exploitation. This form of exploitation can only occur when B disrespects A by neglecting what is necessary for A's well-being or flourishing, but if A cannot lose the lottery, his basic needs will not go unmet. Thus, while we encounter no problems when prohibiting A's choice of L_1 , there is also no non-judicial exploitation in this case—at least when A abides by the obligation to refrain from choosing L_1 .

But what if either A does not abide by his obligation to refrain from exploiting by choosing L_1 , or there is no such obligation? In these cases, if A chooses L_1 and wins, it does not make sense to claim that he has exploited B because no exchange has occurred between A and B (and again here there will be no non-judicial exploitation). But if A chooses L_1 and *loses*, then when A receives BN from B, A exploits B because, as I argued above, in so doing A takes unfair advantage of B. Since exploitation is impermissible, A is obliged to refrain from taking the BN subsidy. But Sample's remarks about (D2)—cases involving a failure to meet basic needs—indicate that if B fails to meet A's basic needs, by failing to give him the subsidy when she employs him *B exploits A*. If A exploits B when he receives the subsidy and B exploits A when she fails to give it, we face a problem.

In order for each to avoid exploiting the other, they must allow themselves to be exploited. And conversely, in order to avoid exploitation, they each must allow the other to exploit. In a discussion of a related case, Hillel Steiner writes, "they are each obligated to employ all available measures necessary to remove the obstacle constituted by the other person's [not exploiting]. And this obligation persists until the point where either (i) one of them has actually prevailed... and the other is incapable of further 'attempting', or (ii) both are [so] incapable..."⁷³ Sample's account provides no way of determining who should concede. Short of an appeal to justice that would undermine the non-judicial nature of the case,

⁷²\$50 from L_2 plus whatever she gained from the second round.

⁷³Steiner (1994, pp.197–98)

there is no way to adjudicate between A's and B's attempts to not exploit. Sample's account of exploitation implies a self-frustrating moral code, "the rules of which at once make [non-exploitative] performances obligatory and, at the same time, are such as to require (in certain circumstances) the deliberate frustration of any [non-exploitative] performance *for the sake of no other principle or value.*"⁷⁴

Sample might claim that A's actions cannot be described as exploitation according to *her* theory because these actions do not satisfy (D1), (D2), or (D3). If A's actions are not exploitation, then the theory is not self-frustrating. Rather, it gives a unified guide to action. B is obliged to provide A with the BN subsidy. While this approach would avoid the self-frustrating objection, it leaves Sample's theory unable to capture a paradigmatic case of exploitation. A and B are in nearly identical situations—they are only *nearly* identical because A's ability to choose between the two lotteries means his opportunities are *better* than those of B. Yet through his own choice, A takes a risky option that may result in his becoming impoverished. Not only does A free ride, A free rides on B's willingness to abide by a moral principle. Surely we desire an account that is capable of identifying cases of free riding on moral principles as cases of exploitation.

The problem can be presented as a dilemma. When disadvantage-based non-judicial exploitation occurs, either:

Horn 1: The account is incomplete, or

Horn 2: The account is self-frustrating.⁷⁵

In order to avoid the horns, Sample must abandon the implicit claim in (D2) that "neglecting what is necessary for that person's well-being or flourishing" may serve as a separate source of exploitation above and beyond the source of injustice.⁷⁶ When combined with my argument against (D3) in 4.3.3, we see that Sample's account is reduced to (D1).

4.4.2 Goodin's Dilemma

A similar problem exists for Goodin's account of exploitation. Recall that Goodin argues exploitation is unfair because it involves the use of others to attain our own ends. In particular, exploitation is the objectionable use of others who are vulnerable: a failure to respect the duty to protect the vulnerable "regardless of the particular source of their vulnerability."⁷⁷ We exploit others when we use their vulnerabilities to secure for ourselves some form of benefit in a game of strategy. With this in mind, consider the following case:

C9: Rural roads.

Suppose A is returning home during a rainstorm. He faces a choice between two routes: rural route, and urban route. The first is much quicker, but there is a 0.50 probability he will become stranded (vulnerable) in the rugged and muddy terrain. The second route is longer, but certain to be passable with his car. However, A knows B's farm lies along the rural route and, since Goodin's vulnerability principle holds, he knows B is required to help him if he is stranded. Which road should A take?

⁷⁴Steiner (1994, p.199, emphasis original). Note that the self-frustrating objection does not depend on the denial of mutual exploitation, that is, it cannot be avoided by claiming A and B exploit each other. A's exploiting B and B's exploiting A in the lottery case are mutually exclusive actions. The performance of one action precludes the other.

⁷⁵Thanks to Christian List for pointing out this structure of the problem.

⁷⁶Sample (2003, p.57)

⁷⁷Goodin (1987, p.187)

Although less formal, the general structure of this case is similar to the lotteries case: A is faced with a choice between a risky and a certain option. Like Sample's claim that we should ensure basic needs are met, Goodin's duty to protect the vulnerable requires B to subsidise A when A loses a freely chosen gamble. However, unlike the previous example, here A is obliged to aid B when B is vulnerable to A.⁷⁸ Nevertheless as with Sample, Goodin faces a dilemma.

"The vulnerable", Goodin tells us, are "those whose interests are strongly affected by our actions and choices."⁷⁹ It is a "relational notion: a full specification will tell us who is vulnerable to whom with respect to what. Saying that A is particularly vulnerable to B with respect to X clearly fingers B as the agent who should be particularly responsible for seeing to it that A's interests in X are protected."⁸⁰ Since B is the only person able to help A in the rural roads case, A's interests are strongly affected by B's actions and choices. A is vulnerable to B with respect to being rescued because he is stranded close to B's home. Therefore, should B choose to aid A, B must not take advantage of A's vulnerability by charging him for rescue.⁸¹ However, by the same criteria, B is also vulnerable to A. B is vulnerable to A with respect to Goodin's duty to protect the vulnerable. In the rural roads case, if A chooses the risky option and becomes stranded, B must rescue A without profit, just as B must subsidise A in the lotteries case. When A takes the rural road and forces B to subsidise him B is vulnerable to A. If A is stranded and B profits from A's rescue, B exploits A, but if B *doesn't* profit, then A unfairly uses Goodin's duty to protect the vulnerable to extract a benefit at B's expense, namely, cheap rescue. If both A and B fail to protect the vulnerable, either A must allow B to profit for rescue, thereby making B an exploiter, or B must rescue without profit, thereby making A an exploiter. Goodin, too, faces a problem of self-frustration.

But perhaps, how the vulnerabilities have arisen is relevant. Perhaps A is not *really* vulnerable since A freely chose to take the rural route knowing B would be obliged to help him. Goodin disagrees: "how the vulnerabilities have arisen is not relevant: all vulnerabilities give rise to the same sorts of responsibilities."⁸² Thus, on Goodin's account of vulnerability we must conclude that A is indeed vulnerable to B when his car becomes stuck. Perhaps we must conclude then that B is not really vulnerable to A because vulnerability is such that only one of the two may be 'all-things-considered' vulnerable to the other. While this claim about vulnerability seems reasonable, it runs contrary to Goodin's arguments.

Goodin may, of course, alter his account of vulnerability so that we cannot claim B is vulnerable to A, but as with Sample's theory, this move only damages the ability of Goodin's theory to incorporate a paradigmatic case of exploitation. Surely when A takes advantage of B's vulnerability to Goodin's moral principle, A exploits B. Thus, the incompleteness objection that confronts Samples theory arises again for Goodin. Any theory that cannot capture this case of exploitation is incomplete. Goodin faces the same dilemma as Sample: when disadvantage-based non-judicial exploitation occurs, either the account is incomplete, or the account is self-frustrating.

⁷⁸Note that what is important here is that A is obliged to transact with B according to *Goodin's account of obligations*. Whether there is such a general obligation in these non-life-threatening cases is open to debate.

⁷⁹Goodin (1987, p.191)

⁸⁰Goodin (1987, p.118)

⁸¹Perhaps B is entitled to receive the cost of rescue from A, but he is presumably not entitled to *profit* from A's vulnerability. Unlike in the lotteries case, the rural roads case does not operationalise Goodin's vulnerability obligation in a constraint like BN.

⁸²Goodin (1986, p.142)

4.4.3 Diagnosis

In both theories we find the claim that mere disadvantage may generate wrongful non-judicial exploitation. It seems true that those who engage in (D1) exploitation, that is, those who take advantage of *unjust* disadvantage, are guilty of degrading behaviour. But who are those who would fall *only* in (D2)? They are those who are disadvantaged, but not unjustly disadvantaged.⁸³ It is possible to construct cases that generate a dilemma for both theories. Specifically, when individuals know that potential losses from risky options must be subsidised by others, they have an incentive to engage in behaviour that lowers the expected outcome of those who do not take risks. On Sample's account the independence of the duty to ensure basic needs are met from the cause of the poverty generates the dilemma; for Wood and Goodin the source of the problem is the duty to protect the vulnerable "regardless of the particular source"⁸⁴ or "even where the exploitation involves no unfairness, injustice, or violation of rights."⁸⁵

The common cause of the dilemma is an insensitivity to personal responsibility. Consider the original lotteries case. Note that the dilemma here would not exist if A's choice of L_1 were not exploitative. I argued above that it is indeed exploitative, but what is it that leads to this conclusion? It is the intuition that in choosing L_1 and accepting the subsidy, A takes unfair advantage. Mere advantage taking is permissible, so it must be the *unfairness* of A's choice that creates the problem. And, as I argued above, A's actions seem unfair because A and B were in analogous situations,⁸⁶ but A, through his own choice, took the risky option. Requiring B to subsidise A in such cases is unfair for two reasons. First, there is a failure of reciprocity: B shares in A's losses, but not his gains. Second, it violates B's autonomy, making B subject to the bad consequences of A's freely chosen actions. Note that this same unfairness, generated by the insensitivity to personal responsibility, occurs regardless of whether the requirement to aid A is based on an appeal to basic needs or vulnerability. Although the scope and normative foundation of exploitation differs slightly under each theory, the form of the problem remains the same. The unfairness is caused by the presence of what we might call a 'come-what-may' duty that makes the theory of exploitation insensitive to actions for which agents should be held responsible. This insensitivity, in turn, generates a dilemma for any theory of exploitation that is based on come-what-may duties.

4.4.4 Extension

As I noted in the discussion of the lotteries case, if Sample is forced to abandon the claim that exploitation can occur in (D2) *only*, and we abandon the claim that exploitation can occur in (D3), then she is left with an account that is quite similar to Steiner's: exploitation is generated by taking advantage of past injustice. Thus, it is clear why the existence of non-judicial exploitation is important for the uniqueness of Sample's account. Because non-judicial exploitation is important for Sample, avoiding the dilemma by prohibiting A from choosing the risky lottery does not seem an attractive option for her, for this precludes the possibility of non-judicial exploitation. However, it is not clear that non-judicial exploitation is important for Goodin's account. Goodin might claim that the existence of non-judicial exploitation implied by his account is simply a loophole, and one that should

⁸³Note that these arguments are compatible with any responsibility-sensitive account of justice. Which actions qualify as not unjust depends upon the account of justice employed.

⁸⁴Goodin (1987, p.187)

⁸⁵Wood (1995, p.17)

⁸⁶Indeed, since A has a choice of lotteries where B does not, A is arguably *better off*.

be closed. Thus, Goodin might quite happily concede that A should be prohibited from choosing the risky option.

Unfortunately, it does not matter much whether Sample and Goodin find prohibiting A's risk taking an *attractive* solution because it is an *unacceptable* solution. Avoiding the dilemma by prohibiting A's risk taking rather than through the abandonment of come-what-may duties amounts to a prohibition on what may often be prudent risks. Although this claim holds for both the lotteries and the rural road cases, I will illustrate it with reference to the lotteries case.

In the lotteries case, if L_1 were replaced with the more favourable lottery L'_1 : ($p = 0.99$, \$500; $1 - p$, \$0), A would still be required to forgo the risky option. In the case, A's choice of L_1 was deemed impermissible because it involved taking unfair advantage. As already discussed, A takes advantage because A gains at B's expense. A's advantage is unfair because A's opportunities were at least as good as B's and yet B must subsidise A in the event of a loss that is the result of A's own choice.

Yet, if it is unfair to subsidise A when A's options are better off than those of B, this unfairness must *increase* as the value of A's options increases. Thus, the justification for the prohibition against choosing L_1 also justifies prohibiting more prudential risks. Some reasonable risks may involve known, but unlikely, outcomes where we become impoverished. Prohibiting all such risks may seem particularly demanding, and taking many of these risks seems morally acceptable. More importantly, this prohibition exists only because the duties requiring B to subsidise or rescue A are not responsibility-sensitive. The fact that B may be required to unfairly aid A means that A must not put himself in a position to be unfairly aided. Were the duties responsibility-sensitive, B would not be required to aid A if A became disadvantaged through his own choice and, therefore, choosing the risky option would not be prohibited. Thus, even if Goodin and Sample abandon the claim that exploitation may be non-juridical, the come-what-may duties that underpin their accounts require A to forgo prudential risk, an outcome that seems less desirable than modifying the duties to incorporate responsibility.

The impermissibility of prudential risks highlights the fact that the fundamental problem with the accounts is their insensitivity to personal responsibility and not the claim that there may be cases of disadvantage-generated non-juridical exploitation. It is the responsibility=insensitive come-what-may duties that give rise to both problems. Thus, an attempt to avoid the dilemma by prohibiting A from choosing the risky option is merely a stopgap solution that does not address the root problem plaguing Sample's and Goodin's accounts.

4.4.5 Responses

There are a number of responses Sample and Goodin might give in defence of their positions. In what follows I present three. The first claims that the dilemma loses its force once we distinguish between conceptions of exploitation and moral codes prohibiting exploitation. The second grants that the dilemma is problematic in the lotteries and rural roads cases, but argues it does not apply in all possible cases or to all accounts employing 'come-what-may' duties. According to this response, modified cases and theories can avoid the dilemma. The third response argues that the dilemma's force is derived from the peculiarities of the cases, in particular from the asymmetry of strategic advantage between A and B.

Conceptions and Codes

The first response runs as follows:

The construction of a moral code and conceptual analysis are two separate endeavours, with different desiderata. That a particular moral code is self-frustrating may be reason for rejecting the code, but these are not reasons for rejecting a particular *conception* of exploitation. Conceptions must be judged by how well they handle our intuitions regarding particular cases. Adequate analyses provide necessary and sufficient conditions that capture all and only instances of a concept, measured by our intuitive grasp of the concept. Conceptions may be deemed inadequate in light of cases that show they fail to meet these demands.

The dilemma conflates the desiderata for adequate moral codes with the desiderata for adequate conceptions. The first horn attempts to provide a counterexample to two conceptions of exploitation while the second provides a problem for a moral code designed to eliminate exploitation. By conflating critiques aimed at separate endeavours, the dilemma fails to offer a unified argument against the *conceptual* analyses of exploitation presented by Sample and Goodin. For example, it does not speak against an analysis of the concept of murder that, one way or another, the analysis encourages more murder, but this result *would* speak against a moral code designed to prevent murder.

The apparent force of this objection is illusory because the purported distinction between conception and code cannot be maintained in the case of exploitation. If exploitation is a negatively moralised *concept* (as I argue in section 5.2), then exploitation entails wrongdoing or badness. Any analysis of a moralised concept must therefore include a normative element, that is, the analysis must contain elements that have action-guiding force. Thus, a conceptual analysis of exploitation implies a moral code. Problems for the implied moral code are also problems for the conception. It is true that, for presentational ease, the dilemma employs both the language of conceptual analysis and that of normative ethics, and this combination may prove misleading, but the dilemma may be presented in language appropriate to conceptual analysis.

The first horn is a claim at the conceptual level: the cases present possible counterexamples that the theories must incorporate. A failure to label A's choice-subsidy action 'exploitation' would leave the theories incapable of incorporating a paradigmatic case. But the second horn can also be made at the conceptual level. If exploitation is a negatively moralised concept, then conceptions must contain the normative claim that exploitation is impermissible. Yet, a conception that includes cases that are self-frustrating cannot consistently claim exploitation is impermissible, for in self-frustrating cases the elimination of one instance of exploitation brings about another. Such a normative account can provide no guide to action. It offers no principle by which we should judge which of A or B should back down. Surely this is a mark against the analysis of a moralised concept.

Consider the response's appeal to murder. While it may be correct that the fact that a particular conception encourages more murder does not count against the *conception*, this example mischaracterises the self-frustrating objection. The self-frustrating objection to MD theories is not that they provide incentives to exploit, but rather that they claim two logically exhaustive actions are both impermissible. The self-frustrating objection applied to murder would involve a case where A's failure to murder B results in A's being murdered by B.

Suppose A is at the bottom of a well and B is falling towards him. If A does not use his ray gun to disintegrate B, A will be killed.⁸⁷ Suppose A's use of the ray gun is permissible as an act of self defence; had he not used it, B would have murdered him. But in this case we would not say that A *murdered* B, for murder implies wrongdoing and the use of the gun was permissible. Then suppose A's use of the ray gun is *impermissible*: If A uses the gun, A murders B. But we would not then say that if A does not use the gun, B *murders* A. Of course, if A does not use the gun, B will *kill* A, but saying that B murdered A would mean that the conception of murder employed was self-frustrating. It would require that neither A nor B could avoid wrongdoing without being wronged. Thus, the self-frustration objection applies to conceptions of murder as well as conceptions of exploitation.

Repair and Restatement

The second form of response involves limiting the damage of the dilemma. Sample and Goodin may acknowledge that the dilemma can be generated by their accounts, but they might argue it can be avoided in other cases or that minor modifications to their accounts may dissolve the dilemma. In what follows I consider three such modifications, the intentionality approach, the demandingness challenge, and an attempt to weaken Sample's BN constraint.

The Intentionality Approach. Part of the initial intuition that A exploits B when he chooses L_1 and the rural route may stem from the fact that A wouldn't have made that choice if B were not present. In the lotteries case, it is B's subsidy that makes the choice of L_1 rational; in the rural roads case, the fact that B lives along the route gives A incentive to take the risky path. Because A knows B is there to help, A takes advantage of B. This approach makes A's intention criterial for exploitation. If A takes risks because of B's subsidy, A intends to take unfair advantage of B; if A would take these risks irrespective of B's presence, it does not seem that A intentionally takes advantage of B.

If instead of the original cases we consider cases where A would have taken the risky option *regardless* of whether B was present, we may no longer conclude that A exploits B when choosing the risky option since A does not intend to gain from B. If A does not exploit B when choosing the risky option, the dilemma can be avoided. As they are originally stated, neither Sample's nor Goodin's account makes an agent's intention criterial for exploitation. Thus this approach involves modifying the accounts to include an intentionality condition.

The strategy of the intentionality condition is to claim that exploitation requires intentionality, but that A lacks the relevant intention in some versions of the lotteries and rural roads cases and thus cannot be said to exploit B. If A does not exploit B, the self-frustrating objection can be avoided, permitting dilemma-free cases of non-juridical exploitation. The problem with this approach is that it seems difficult to maintain the claim that A lacks the relevant intention.

Whether A exploits B is not determined at the point of A's choosing the risky option. Suppose A chooses the risky option because B will subsidise him—that is, he intends to take unfair advantage of her—but A either gets lucky or upon losing chooses to forgo the subsidy. In either case A does not exploit B. Thus, A's intention *at the initial point of choice* between the risky and certain options is not *sufficient* for A's exploiting B. Now suppose A does not know that B will subsidise him, but chooses the risky option anyway. However, upon losing A discovers that B is there and that she will subsidise him. A accepts the subsidy, despite

⁸⁷See Nozick (1974, p.34)

knowing that he is responsible for his own loss. Here since A was better off than B, but lost through his own choice and still accepts the subsidy, A takes unfair advantage of B. A's intending to exploit B *at the initial point of choice* is not *necessary* for A to exploit B.

If A's intending to exploit B at the initial point of choice is neither necessary nor sufficient for A's actually exploiting B, why is it that part of our intuition that A exploits B seems to stem from the fact that A chooses *because* he knows B will subsidise him? There seem to be two reasons. First, in the case where A's choice depends on B's subsidy, we can say that A *intends* to exploit B. He intentionally puts himself in a position that may lead to his taking unfair advantage of B. While A only actually exploits B by accepting the subsidy, whether he finds himself in this position by design or accident may make A's action more grievous. Second, we may believe that A's intending an action makes that action more likely. A is more likely to actually exploit B if A intends to exploit B.⁸⁸

Since A's actually exploiting B occurs at the point of subsidy, then at the point of choice, A can only intend to exploit B. If intentionality is criterial for exploitation, and A exploits at the point of subsidy, then we must examine whether A's intentions at the point of subsidy are intentional. And in the two cases I have presented, it seems they always are. A cannot request cheap rescue or the BN subsidy from B without intending to take advantage of B; in other words, A cannot request subsidy from B without intending to receive the subsidy from B. The request for a better wage or a cheap rescue is an intentional action on A's part and, therefore, A never lacks the relevant intention. Thus, even if intentionality is criterial for exploitation and A would have taken the risky option without knowing of B's presence, we can still say that A exploits B when A loses a known risk and intentionally requests a subsidy or rescue. If A exploits B in such cases, then the self-frustrating objection remains.⁸⁹

The Demandingness Challenge. Consider the following case of what seems to be non-judicial exploitation:

C10: The old well.

Suppose that there is an old well on A's farm. A takes measures to close off the well, but he knows he should have it properly sealed and capped. One evening while out for a walk A's flashlight dies and, stumbling home, he falls into the well. The next day B discovers A in the well, offering to save A only if A pays \$10,000.

Surely, the old well presents a case where B exploits A. And since there seems to be nothing *unjust* about A's predicament, it must be a case of non-judicial exploitation. While B might be entitled to charge A the cost of rescue, B's \$10,000 request is simply too *demanding*. Although it is likely true that \$10,000 is too much, the conclusion that the old well presents a legitimate case of non-judicial exploitation moves too quickly. The parameters of the case are less well defined than those in the original dilemma generating cases.

First, as a preliminary, note that there is nothing *intrinsically* amiss with the \$10,000 figure itself. Were A and B billionaires, our intuition that \$10,000 was exploitative might

⁸⁸Note that when A chooses L_1 because he knows B must subsidise him, the incentive for this choice is provided by BN. We know this because when BN is not present, both L_1 and L_2 are rationalisable (rationally permissible). The response in 4.4.5 challenged whether we must abandon conceptions that encourage the kind of conduct they are designed to proscribe. Although encouraging self-undermining behaviour may not be sufficient for abandoning an account of exploitation, it must, all else equal, count against a conception that it is self-undermining.

⁸⁹Paul Bou-Habib has pointed out an alternative approach to me. We may grant the claim that intentionality is necessary for exploitation, but if A is vulnerable without intending to exploit B, then we may also say that it is unjust to fail to subsidise him.

be less strong. The figure represents, *inter alia*, the degree of sacrifice B asks of A; whether \$10,000 represents an impossible, large, small, or merely incidental sacrifice depends on the context and endowments of the transactors, an obvious fact that is, nevertheless, often obscured. Let us suppose then that the \$10,000 means the same thing to A and B that it does to most of us: it represents a significant sacrifice to A and a significant boon to B. In this case, the original conclusion that \$10,000 is too much retains its force.

Second, it is not clear that A is fully responsible for his having ended up in the well. We may presume he had a vague idea that there was a small chance of a mishap involving the well and may have known that his flashlight batteries were running low, but it is not the case that he chose between two options with well defined outcomes and probabilities. Had A clearly understood that he was taking a gamble involving, say, probability of $p = 0.10$ of finding himself in the well with the consequence that he would then have to pay B \$10,000 for rescue and a probability of $1 - p$ that he would enjoy his walk, the cost of catastrophe discounted by its probability (\$1000 in this case) might have led A to choose differently. If A would have chosen differently with full information, then it does not seem that he can be held *fully* responsible for his choice. Much of our intuition that B exploits A, I believe, arises from our belief that in the old well case A is not, in fact, fully responsible for his choice. If A would not have altered his choice with full information, then our intuition that B exploits A is diminished. In fact, if knowing the risks, A takes a chance that leads to a loss and, having lost, demands a better deal from B, it begins to look like A is the exploiter.

Third, there is an important question regarding the stakes of A's choice. As Serena Olsaretti points out, even if A is fully responsible for his choice, this does not mean that 'anything goes' with respect to its consequences. While the grounds of responsibility concern "the features of people that we can... hold them responsible for", there is a further question of stakes, which asks, "Just what costs should attach to whatever features constitute the justifiable grounds of responsibility?"⁹⁰ Olsaretti argues that in many cases the consequences of choices for which people are responsible are indeterminate; there is no clearly defined set of consequences that are the direct result of the responsible choice. However, while it is tempting to use Olsaretti's distinction to underpin the claim that B's \$10,000 claim is exploitation, it is not the solution the supporter of non-juridical exploitation is seeking. Although non-*laissez-faire* principles of stakes are possible, these principles form part of an account of justice.⁹¹ Thus an appeal to a principle of stakes is an appeal to justice, undermining the existence of non-juridical exploitation.

B's request of \$10,000 for rescue is too demanding if the old well case is taken at its (underspecified) face value. But the demandingness objection loses force with the requirement that A's choice is an informed and responsible choice and without an appeal to justice-based principles of stakes.

Simultaneity and Symmetry

The final response is that the dilemma's force is a product of the asymmetry of advantage in the lottery case. Perhaps the dilemma exists only because B is bound to L_2 in the first case, or her home's location in the second. This response is, in fact, correct. The dilemma *does* depend on asymmetry. But this response is not well-motivated, for if exploitation is taking unfair advantage, then one transactor must receive more than they should from the

⁹⁰Olsaretti (2009, p.170)

⁹¹Olsaretti (2009, p.165)

other. If there are only two transactors, then it cannot be the case that they *both* receive more than they should because (in the context of a subjective value theory in which goods are commensurable) gains and losses of advantage in two-person transactions are a zero-sum game. Thus, if A and B are in strategically equivalent situations, neither can take unfair advantage of the other. Some form of unequal advantage, in one respect or another, is necessary for the existence of exploitation. This fact alone is sufficient to answer the objection. However, it is worth considering a version of the lotteries case where A and B are in symmetric situations because the case presents a new problem for come-what-may duties, above and beyond the original dilemma. Consider the following simultaneous version of the lotteries case.

C11: The lotteries (simultaneous).

Suppose in the first stage B is no longer committed to L_2 and A and B face a *simultaneous* choice between L_1 and L_2 . Neither player knows the other's lottery choice. As before, the second stage remains sequential—the wealthy player may decide to employ the poor and if players are equally wealthy they may work cooperatively—and BN holds.

In order to determine the rational choice of lottery for each player, we must first determine the outcomes of the four possible strategy combinations: $\langle L_1, L_1 \rangle$, $\langle L_1, L_2 \rangle$, $\langle L_2, L_1 \rangle$, and $\langle L_2, L_2 \rangle$. We know from the discussion of the original case that if the player's strategies were $\langle L_2, L_2 \rangle$, the outcomes would be (\$55.50, \$55.50) and that the outcome of $\langle L_1, L_2 \rangle$ would be (\$58, \$53). Since neither player's situation differs from the other in the simultaneous lotteries case, the outcome of the reciprocal situation, $\langle L_2, L_1 \rangle$, is just the reverse: (\$53, \$58). All that remains is to determine the outcome of both players choosing L_1 , that is, $\langle L_1, L_1 \rangle$.

If both players choose the risky lottery we know there are four equiprobable possibilities: 1) A wins and B wins, 2) A wins and B loses, 3) A loses and B wins, and 4) both lose. From the original case we know that it is rational for the players to work, employ, and cooperate. In the first possibility both players win. Thus, the outcome for each is \$100 from the lottery, plus \$5.50 from the cooperative venture, i.e., \$105.50. In the second, A wins and B loses. The outcome for A is \$100 from the lottery and \$1 from employing B, i.e., \$101 and the outcome for B is \$10 in wages. The third possibility is the reciprocal of the second: A's outcome is \$10 and B's is \$101. Finally, in the fourth, both lose the lottery. Since neither has any capital, neither can work for the other, and a cooperative venture is not possible. The outcome is \$0 for each player. Since each of these four scenarios is equiprobable, the outcomes must be weighted by 0.25. Thus, the expected value of choosing L_1 for each player is \$54.125. We can now complete the payoff table for the simultaneous game shown in table 4.2 below.

		B	
		L_1	L_2
A	L_1	\$54.12, \$54.12	\$58, \$53
	L_2	\$53, \$58	\$55.50, \$55.50

Table 4.2: Outcome matrix for case 11.

In this game, the Nash equilibrium (in blue) is $\langle L_1, L_1 \rangle$. Thus, when BN is present, if both players were to choose simultaneously without knowing the other's choice, each would rationally choose L_1 . Structurally, the game shown in table 4.2 is a Prisoner's dilemma.

If we consider a similar version of the game, where BN does *not* hold (C11b), the payouts are as shown in table 4.3. Here, there are 3 Nash equilibria. Although both lotteries are ‘rationalisable’ in the sense that Nash equilibria exist under a choice of either lottery, $\langle L_2, L_2 \rangle$ weakly dominates all other strategies.

		B	
		L_1	L_2
A	L_1	\$54.12, \$54.12	\$55.50, \$55.50
	L_2	\$55.50, \$55.50	\$55.50, \$55.50

Table 4.3: Outcome matrix for case 11b.

When A and B are in equivalent situations, that is, when the game they play is a simultaneous game (in stage one), the presence of BN leads to a game with a Prisoner’s dilemma structure. When BN is not present (as in table 4.3), all Nash equilibria are strict Pareto improvements over the Nash equilibrium that occurs when BN is present. The presence of BN, on the whole, leads to outcomes that are worse for both A and B. So, although neither player exploits the other in the simultaneous lottery case because neither is at a relative disadvantage, a comparison of versions of the case in which BN is, and is not, present shows that the presence of BN leads to a worse outcome for everyone. In this sense, BN is non-universalisable—neither player can rationally will that BN apply.⁹² Therefore, although asymmetry in bargaining power is necessary to generate the dilemma cases, it is also necessary for any kind of exploitation. Further, the analysis of symmetric cases shows that the dilemma is not the only problem accounts based on come-what-may duties face. When the asymmetry is removed, the presence of BN causes both players to do worse than they would if they were not bound by BN. Here BN cannot be rationally willed.

Sample’s and Goodin’s accounts imply wrongful exploitation may be generated by sources independent from injustice, in particular that exploitation may be brought about by taking advantage of *mere* vulnerability, poverty, or disability. However, as I have argued, their accounts face a dilemma. In all purported cases of disadvantage-based non-judicial exploitation, the accounts are either self-frustrating or, in order to avoid the self-frustrating objection, they fail to capture a paradigmatic case of exploitation. Attempts to avoid the dilemma by prohibiting actions that cause it both preclude the existence of non-judicial exploitation and prohibit prudential risks. This strategy fails to recognise that the dilemma arises because the accounts are not responsibility-sensitive. This diagnosis suggests that the dilemma shows that exploitation must not only be located within the domain of injustice, but also that only responsibility-sensitive accounts of justice can adequately ground theories of exploitation.

⁹²Strictly speaking, neither player can rationally will that BN apply for *all* players. But a player might rationally will that BN apply for some players. A would presumably prefer that BN held when *he* lost the lottery, but not when B lost. However, generally we think of basic moral principles as universally applicable. Showing that all players would rationally prefer moral codes that *universally* exclude BN to those where BN applied universally is sufficient to say that BN is not universalisable. Nevertheless, it is an interesting question whether BN fails to be rationally will-able in this weaker (non-universal) sense. Although not included here, it can be shown that in a meta-game where A and B must (simultaneously) decide whether to bind themselves to BN and then play two kinds of sub-games like the simultaneous lottery case—one where BN holds and one where it does not—in almost all cases A and B would not choose to bind themselves.

4.5 Chapter 4 Summary

In this chapter I outlined two duty-based accounts of exploitation. Section 4.2 presented Goodin's account. Goodin (1986; 1987) claims that we have a duty to protect the vulnerable when transacting with them; failure to do so amounts to exploitation. In section 4.3 I considered Sample's critique of Goodin and her own account. Sample argues that Goodin's consequentialist foundations are incompatible with his theory of exploitation. This inconsistency leaves Goodin ultimately unable to provide a coherent solution to the paradox of exploitation. In place of Goodin's theory, Sample (2003) argues for an account of exploitation as degradation. According to Sample, we can degrade others when we transact with them by taking advantage of previous injustices, failing to ensure their basic needs are satisfied, and engaging in an exchange of a non-commodifiable good. In this section I also argued against Sample's claim that exploitation can be generated by trade in non-commodifiable goods.

The remainder of the chapter—section 4.4—focused on a dilemma for Goodin's and Sample's accounts. Here I presented two cases that create dilemmas for Sample's and Goodin's accounts (subsections 4.4.1 and 4.4.2), showing that the accounts are either self-frustrating or incomplete. In subsection 4.4.3 I argued that the cause of both dilemmas was the presence of come-what-may duties in each account that are insensitive to personal responsibility. In 4.4.4 I showed that Sample and Goodin can avoid the dilemma by prohibiting A from choosing risky options (at the expense of abandoning the implicit claim that non-judicial exploitation exists); however, this solution is unacceptable because its prohibition on prudent risk-taking is implausible. Abandoning the come-what-may duties that provide the foundations for their respective accounts is a far more plausible resolution. In the final subsection (4.4.5) I considered three forms of reply to the dilemma, none of which succeed. The first alleges that the dilemma depends on a conflation of conceptions with moral codes. However, because the concept of exploitation is negatively moralised, a proper analysis must *contain* a moral code. Modifications to the accounts that incorporate intentionality conditions, principle of stakes, or that weaken the come-what-may duties, also do not succeed in avoiding the dilemma. Finally, I showed that, although the dilemma depends on asymmetric advantage, the existence of exploitation also depends on asymmetric advantage. Removing the asymmetries in the dilemma generating cases dissolves the problem by fiat. Here I also showed another reason for rejecting the come-what-may duties: in a simultaneous version of the original lotteries case we can see that the come-what-may duties are non-universalisable.

The theories of exploitation advanced by Sample and Goodin are well-motivated. They are based on a desire to protect the disadvantaged in our societies from the harm of exploitation. However, in claiming that exploitation may occur apart from injustice, the duties that provide a foundation for their accounts reach too far. In failing to hold people responsible for their freely chosen actions, these 'come-what-may' duties lead to situations where the *truly* disadvantaged are exploited by the justly disadvantaged.

Part II

What Exploitation Is

Chapter 5

The Two Faces of Exploitation

5.1 Chapter 5 Introduction

Exploitation, it is claimed, involves taking unfair advantage.¹ Within this general definition are two separate overarching intuitions intended to account for our claims that an action is exploitative. According to the first intuition, which we may call the distributive intuition, exploitation is a matter of unfair distribution. It arises when the outcome of a particular transaction is the result of an unjust pre-transfer distribution. Both Steiner and Roemer provide conceptions of exploitation that prioritise the distributive intuition. Steiner claims that a transaction is exploitative when (*inter alia*) unjust property rights violations cause the outcome of the transaction to diverge from one in which the property rights violations did not occur. Similarly, Roemer claims that a coalition is exploited when (*inter alia*) there is a counterfactual situation in which the coalition would be better off, and its complement worse off, if the coalition withdrew with some externally specified set of resources. On both accounts, exploitation is a systemic phenomenon that results from a prior set of unjust conditions that generate unfair distributions.

According to the second intuition, which we may call the advantage intuition, exploitation involves the misuse of advantage. It arises when one transactor uses the disadvantage of another for their own gain. Goodin and Sample² each provide conceptions of exploitation that prioritise the advantage intuition. For these theorists, exploitation involves using another's disadvantage or vulnerability for personal enrichment in situations where it is inappropriate, degrading, or disrespectful to do so. The advantage intuition captures the claim that exploitation is not, or not only, determined by the distributive consequences of a transaction, but depends upon the attitudes, motivations, and intentions of the agents. Accounts of exploitation based on the advantage intuition take exploitation to be “not a matter of how things *end up*... [but] “a matter of how you got there.”³

Steiner and Roemer privilege the first intuition, claiming unjust pre-transfer distributions are necessary and sufficient for exploitation. But as we've seen, these conceptions extend beyond our ordinary use of 'exploitation'. For Roemer, both conditions of his property relations account may be satisfied in cases that are, intuitively, non-exploitative, as Elster's 'religious belief' case shows. Even with the addition of limiting conditions, Roemer's account may

¹This ordinary language definition enjoys explicit support in Elster (1982); Goodin (1987); Steiner (1984, 1987, 2010); Sample (2003); Schweickart (1997); Tormey (1976); Walt (1984); Wertheimer (1996); implicit or partial endorsement can be found in Miller (1987); Wolff (1999). Wood (1995) represents the most explicit rejection of the 'taking unfair advantage' approach. I consider his position in the following section.

²Similar accounts, which I do not assess in detail, are provided by Allen Wood (1995) and Nicholas Vrousalis (2013).

³Goodin (1987, p.181)

still overextend, without a properly specified dominance condition. For Steiner, A may be non-culpably ignorant of his role as an exploiter, a role he plays only because of someone else's previous wrongdoing.⁴ Both of these overextensions can be seen as a failure to fully incorporate the advantage taking intuition. According to both accounts, everyday transactions of ordinary goods between friends with the best of intentions may be exploitations. Again, while we may have reason to say that some of these 'innocent' transactions are in fact morally problematic, they are not natural cases of exploitation in ordinary use.

The lesson from the overextension of Steiner's and Roemer's conceptions is that our willingness to apply the label 'exploitation' to a transaction depends on something more than maldistribution brought about by A's interaction with B. Exploitation requires that A *take advantage* of B. Elster claims that we take advantage of another "by adding an option to his opportunity set that is better both for him and for [us] than the option preferred by him prior to the expansion."⁵ When 'taking advantage' is defined as a strict Pareto improvement, A can be said to take advantage of B (and vice versa) when the two merely transact. But, used in this sense, A and B 'take advantage' of each other in Steiner's exploitative transactions. Thus, Elster's sense of 'taking advantage' cannot be the 'something more' that is required to limit the extension of the distributive accounts.⁶ Rather, the required (stronger) sense seems to be that involved in the conceptions outlined by Goodin, Sample, and Wood: "Inappropriately seizing advantages"⁷, failing to constrain our behaviour towards others,⁸, and "playing on [or] manipulating"⁹ the vulnerabilities of others are all examples of this stronger kind of taking advantage. It is tempting—at least for Goodin and Sample—to think that, properly specified, the taking advantage condition *alone* can characterise at least some cases of exploitation. However, as I argued in chapter 4, conceptions that permit non-judicial exploitation are either self-frustrating or incomplete. While "playing for advantage in situations in which it is inappropriate to do so",¹⁰ but where no injustice is involved, may be morally amiss for other reasons, advantage taking *without* injustice is not equivalent to the 'taking unfair advantage' definition of exploitation.

An adequate conception of exploitation must take both intuitions contained in the ordinary language account seriously; both unfairness and advantage taking are necessary conditions for transactions to be exploitative. An analysis of exploitation that incorporates the advantage intuition and the distributive intuition can capture counterexamples that plague accounts based on only one of the two intuitions. Of course, we should recall Arneson's reminder in chapter 4 that without explication there will "be as many competing conceptions of exploitation as theories of what persons owe to each other by way of fair treatment."¹¹ The same can be said for 'taking advantage'. Thus, a complete characterisation of exploitation must provide some additional specification of these two conditions.

⁴Here I cite the problem of moral responsibility that Steiner's account faces, although in my discussion of this problem in chapter 2 I note that it may be solved by drawing a distinction between 'morally bad' transactions and 'morally wrong' transactions. However, unlike the addition of an intentionality or knowledge condition, the distinction-based solution to the problem of moral responsibility does not *also* solve the problem of synchronic pervasiveness. This problem, it seems, requires a reconsideration of the role of advantage taking.

⁵Elster (1982, p.364)

⁶Taken out of context, Elster's sense of 'taking advantage' seems trivial. However, his quote is taken from a passage in which he discusses problems of dependence and dominance for Roemer's property relations account. When 'taking advantage' in Elster's sense is added to (PR1) and (PR2) it does successfully limit the scope of Roemer's theory. However, as I argued in chapter 3, this limitation is not sufficient.

⁷Goodin (1987, p.184)

⁸Sample (2003, p.62)

⁹Feinberg (1988, p.179)

¹⁰Goodin (1987, p.184)

¹¹Arneson (1992, p.350)

In addition, this approach to exploitation takes the ordinary language definition of “taking unfair advantage” for granted, an assumption which is not altogether theoretically neutral, since the inclusion of the normative element of ‘unfairness’ commits us to a negatively moralised account of exploitation. While taking the ordinary language definition as a starting point enjoys a near consensus among existing accounts, Allen Wood dissents from this majority view, arguing that exploitation is *not* a negatively moralised concept. Wood claims that while exploitation is “*nearly* always a bad thing. . . its badness does not always consist in its being unjust.”¹² Although Wood is correct that the word ‘exploitation’ does admit to a non-moralised sense, the moralised and use-based senses can be sharply distinguished. Nevertheless, it is important to consider the reasons for rejecting Wood’s dissenting view before moving on to a characterisation of the two necessary conditions.

Thus, the remainder of the chapter will proceed as follows. In the next section, 5.2, I consider the relationship between injustice, exploitation, and moral permissibility. Here I outline and reject Wood’s argument against the claim that exploitation is a (negatively) moralised concept. In section 5.3 I consider the normative aspect of exploitation: unfairness. I outline the various principles of fair transaction that appeared in the first four chapters, introduce Alan Wertheimer’s FMV principle, and finally propose and defend a specification of the unfairness condition in terms of a modified version of Steiner’s rights condition. Section 5.4 provides a similar treatment of the advantage taking condition. In that section, I argue that the sense of ‘advantage taking’ relevant to cases of exploitation can be given by considering the epistemic state of the would-be exploiter. Finally, section 5.5 concludes with my first full characterisation of exploitation and a summary of the chapter.

5.2 The Wrongfulness of Exploitation

What is the relationship among wrongful action, injustice, and exploitation? Is exploitation necessarily morally amiss? Always unjust? The theories of exploitation examined thus far have provided three different positions.

Allen Wood argues that exploitation does not have “wrongfulness or moral badness built into its very meaning.” After all, he reasons, “there clearly do seem to be cases of exploitation which we regard as innocent or even admirable. Nobody thinks it is wrong for a chess player to exploit her opponent’s inattention in order to win the game.”¹³ Wood also argues that, even in those cases where exploitation *is* morally wrong or bad, it can be so “even when it involves no unfairness or injustice.”¹⁴

Ruth Sample disagrees with Wood’s first claim, arguing that exploitation is “an essentially moralised or thick concept,”¹⁵ but, as we have seen, she accepts his second claim that exploitation can be morally amiss even when it is not unjust. She argues that “neglecting what is necessary for [a] person’s well-being or flourishing” is a source of exploitation-generating disrespect that is distinct from the disrespect involved in taking advantage of past injustice.¹⁶

Finally, Steiner disagrees with both of Wood’s claims. He writes both of the “*injustice*

¹²Wood (1995, p.2, emphasis added). Wood’s claim is that, properly understood, the ordinary language definition of ‘exploitation’ is non-moralised. He is not claiming that our ordinary language definition is negatively moralised and that it is therefore mistaken. Thus, his disagreement is with philosopher’s *interpretation* of ordinary use.

¹³Wood (1995, pp.3–4)

¹⁴Wood (1995, p.17)

¹⁵Sample (2003, p.5)

¹⁶Sample (2003, p.57)

	Moralised	Not Moralised
Always Unjust	Steiner	—
Not Always Unjust	Sample	Wood

Table 5.1: Justice, morality, and exploitation.

of exploitation”¹⁷ and the “*wrongful* gains and losses of an exploitation.”¹⁸ And, of course, as we saw in the discussion of his characterising conditions, Steiner’s inclusion of the moralised ‘rights condition’ (S1) ensures that the truth conditions of his account depend upon a normative criterion.

The possible relationships between exploitation, injustice, and wrongful action are shown in table 5.1. Since Steiner claims that exploitation is negatively moralised and unjust, he falls in the top left corner. Sample also claims exploitation is negatively moralised, but because she believes it is not always unjust, she is placed in the lower left corner. Finally, Wood claims exploitation may, at times, be neither negatively moralised nor unjust. He belongs in the lower right. Although my goal in what follows is (primarily) to evaluate the viability of Wood’s heterodox position (bottom right), it is worth remarking briefly on the emptiness of the top right quadrant.

5.2.1 Eliminating the Bottom-left and Upper-right Quadrants

It is tempting to explain this emptiness with the claim that the position is a conceptual impossibility: it seems that unjust actions must be a subset of morally impermissible actions. Yet, as Steiner points out,¹⁹ this is not the case in a Kantian framework. For Kant, moral permissibility depends upon the first formulation of the categorical imperative—which is intensional—and justice depends upon the universal principle of justice—which is extensional.²⁰ “[I]t’s simply untrue that an action’s being a rights violation is a sufficient condition of its moral impermissibility. Its injustice is such a condition only if justice is one of our moral rules. But it needn’t be.”²¹ Thus, at least in a Kantian framework, the concepts of justice and morality do not rule out actions in any of the four quadrants of table 5.1. Of course, the fact that justice and morality do not, *a priori*, limit the possibilities for actions does not mean there are no other considerations that may be brought to bear.

In particular, although justice needn’t be one of our moral rules, moral codes that consist solely of action prescriptions, although not contradictory, can lead to situations where these prescriptions make “obligatory the performance of two actions that [are] mutually frustrat-

¹⁷Steiner (1987, p.132, emphasis added)

¹⁸Steiner (2010, p.27, emphasis added). Although we know from the problem of moral responsibility (outlined in section 2.3) that since Steiner’s account allows cases of unintentional exploitation, his account must claim that in such cases exploitation is not morally wrong, but morally *bad*.

¹⁹Steiner writes, “If I force you to work for a deserving charity or steal your money on its behalf or, indeed, force you to do something that’s in your own interest, I may be perfectly happy to universalize my benevolent intention in doing so. I may include benevolence among my moral rules. [Thus my action is morally permissible.] But that willingness in no way implies that my action is not a reduction of your equal freedom and a subordination of your purposes to mine. [Thus my action is an unjust violation of your rights.]” (Steiner, 1994, p.222)

²⁰In particular, Steiner notes that “whether an act is just doesn’t depend on the intention with which it’s done. It doesn’t depend on its having a certain intensional description. It depends, rather, on its having a certain extensional description: that is, on whether it’s compatible with a certain distribution of freedom, with a certain set of rights.” (Steiner, 1994, p.212). But the principle underlying the *permissibility* of action “depends on that *intensionally-described* action’s relation to whatever moral rules we accept...” (Steiner, 1994, p.222). Thus, since the criteria underlying the moral permissibility of actions is an intensional criteria and the criteria underlying the justice of action is extensional, the two can come apart.

²¹Steiner (1994, p.223)

ing.”²² When a moral code:

1. Consists solely of action prescriptions, i.e., contains no rules for rights and
 2. Prescribes the performance of X by two individuals, but
 3. The performance of X by each is obstructed by the other, then
- C. Their duties to perform X entail that “they are each obligated to employ all available measures necessary to remove the obstacle constituted by the other person’s acting. And this obligation persists until the point where either (i) one of them has actually prevailed... and the other is incapable of further ‘attempting’, or (ii) both are [so] incapable...”²³

The inclusion of justice as one of our moral rules provides “an arbitration rule that supplies a neutral, impartial and principled resolution of any deadlock [between other moral rules.]”²⁴ Thus, if we want to avoid conflicts of mutual duties to perform morally prescribed actions—surely a desiderata for good moral codes—our moral code must include justice as one of its rules. And, if the moral code includes justice as one of its rules, then an action’s being a rights violation *is* a sufficient condition of its moral impermissibility. Thus, if exploitation involves a rights violation (and we want moral codes with provisions for dealing with deadlock), then exploitation must be morally impermissible. Assuming we want moral codes with provisions for dealing with deadlock, we have eliminated the top right quadrant, since under such codes injustice implies wrongdoing.

However, this result is not particularly helpful since no one seriously defends an account of exploitation that makes it unjust but not morally amiss. Yet, when combined with chapter 4’s conclusion that accounts of purportedly non-juridical exploitation are either self-frustrating or incomplete, we obtain a stronger result. If exploitations are always unjust and injustice *implies* moral impermissibility, then exploitation must be impermissible. Neither the upper-right nor the bottom-left quadrants are live possibilities.

With this result in place, it might seem that there is no need to consider Wood’s claim that exploitation may be neither morally impermissible nor unjust. However, the result depends on a crucial assumption that Wood rejects. Chapter 4’s conclusion that exploitation is always unjust relied upon intuitions generated by the ordinary language definition that exploitation is ‘taking unfair advantage’. But Wood rejects the claim that exploitation involves unfairness, and so, he may reject both chapter 4’s conclusion and the derivative claim that exploitation is also morally impermissible.

5.2.2 Eliminating Wood’s Non-moralised Exploitation

Of course, when it comes to the *word* ‘exploitation’ Wood’s position is correct. It is obvious that uses of ‘exploitation’ do not always imply moral wrongdoing or injustice. ‘Exploiting a mine’, ‘exploiting a chess move’, and ‘exploiting machinery’ are all expressions that imply *use*, but not wrongful or unfair use. However, ‘exploitation’ used in this sense seems to denote a different concept from that referenced by uses like ‘Sweatshops are exploitative’ or ‘Price gouging during disasters is exploitative.’ *Prima facie*, at least, exploitation in the ‘use’ sense (call this S_u) and exploitation in the morally pejorative sense (call this S_m) seem

²²Steiner (1974, p.199)

²³Steiner (1974, pp.197–98)

²⁴Steiner (1994, p.207)

to be homonyms. While we may grant that Wood is correct that the word may be used in either sense, the concept we are interested in is the concept referenced in S_m and there does not seem to be any problem in drawing a distinction between the two.

The Nietzsche Argument

Wood disagrees, arguing that this distinction cannot be maintained. He asks, “why should we suppose that ‘exploitation’ has a special *meaning* when applied to cases of injustice or wrongdoing?”²⁵ After all, he reasons, other terms too—Wood lists “‘appropriation’, ‘transaction’, ‘seizure’”—apply to acts that are sometimes, but not always, impermissible.²⁶ Wood argues against the claim that there exists a sense of exploitation, S_m , that is negatively moralised by citing an example of Nietzsche’s use of the word. Nietzsche claims modern society’s economic life “represents a maximum in the exploitation (*Ausbeutung*) of humanity”²⁷, and, contrary to our own moral intuitions, he also affirms, this state as admirable. The difference in Nietzsche’s use and our own, Wood claims, cannot be semantic. Nietzsche is “using the word in the same sense as we do when we more conventionally hold that it is bad or wrong.”²⁸ Wood argues that both we and Nietzsche refer to the same concept, but Nietzsche’s values differ from our own when he affirms exploitation; our disagreement with the moral stance that Nietzsche expresses is only possible if he is “using the word in the same sense as we do when we more conventionally hold that it is bad or wrong.”²⁹

In order to evaluate Wood’s argument, let us first consider a similar case: suppose I say of an object ‘it is red’, but you disagree with me. In normal cases of disagreement, our disagreement about a certain property of the object requires that we have the same referent in mind. So, in a case where the object is a ball, and I say ‘the ball is red’, but you disagree, in order to be a disagreement about the ball’s colour, we must be speaking about the same thing: the ball. Our disagreement is about the properties of the same referent. On the other hand, in some cases, disagreement can come about because someone’s use is mistaken or nonstandard. So, if I say ‘light with a wavelength between 620nm and 740nm is red’, but you disagree, our disagreement comes about because you use red in some nonstandard way or because you do not know this fact, for on my use, having a wavelength between 620nm and 740nm is a necessary property of being red.³⁰ Here we disagree about the *constitutive* properties of the referent. So, the argument may be reconstructed as follows:

N1 If our disagreement with Nietzsche is of the first sort (disagreement about whether a particular case actually has a property), then we must have the same concept in mind.

N2 And, Wood implies, since our disagreement with Nietzsche *is* of the first sort, not the second (disagreement about the constitutive properties of the concept), we do have the same concept— S_m —in mind.

C But since Nietzsche claims that S_m is morally admirable, S_m is not necessarily morally amiss.

²⁵Wood (1995, p.4)

²⁶Wood (1995, p.4)

²⁷Wood (1995, p.5)

²⁸Wood (1995, p.5)

²⁹Wood (1995, p.5)

³⁰Of course, in some cases, for example, when the properties involved are vague (as in ‘X is tall’), the disagreement may also be because we use one word to stand for a range of properties in various contexts.

Wood is correct that in order to disagree about the properties of a case we must be speaking about the same referent.³¹ And it must also be the case that if a concept can be used in a positive or morally neutral way, it is not negatively moralised. So the first and third lines of the reconstructed argument are correct. However, Wood provides no argument for the second premise and there is no reason to believe it is true.

Consider a related case. Suppose that in his lyrics the singer of a death metal band claims that ‘murder is good’. This claim is, in a strict sense, contradictory, since ‘murder’ is, by definition, bad. But the singer’s lyrics are rhetorical. He uses the word ‘murder’, which ordinarily contains both descriptive and evaluative elements, in a merely descriptive—‘inverted comma’—sense.³² The intent of his lyrics is to affirm those *acts* characteristically associated with the ordinary use of murder, e.g., the brutal killing of innocents. We could restore meaning to the literal reading of the claim by rephrasing it as ‘what the folk call ‘murder’ is actually good.’ His lyrics are shocking because our evaluation of these events is strongly negative; we do not accept his claim that such things are good.

Similarly, when Nietzsche claims that the ‘exploitation’ of humanity by the *Übermensch* is good, he makes a rhetorical claim. His use of exploitation is in the inverted comma sense, picking out those particular cases characteristic of the negatively moralised sense in ordinary use. Our disagreement and disgust with Nietzsche is over his affirmation of the acts we associate with the exploitation of humanity (in the negatively moralised sense), e.g. sweatshops, price gouging, and the like. And again, we may restore meaning to the literal reading by rephrasing his claim as ‘what the masses call ‘exploitation’ (in the sense of S_m) is good.’ In providing a competing hypothesis of the linguistic data, I have shown that Wood’s case for his view is inconclusive at best. Wood’s Nietzsche example does not provide a decisive argument against the orthodox position that exploitation admits to two senses, S_m and S_u .

Wrongful Because Exploitative

The Nietzsche example is not Wood’s only argument against the claim that the concept denoted by exploitation in its S_m use is negatively moralised. He also claims that negatively moralised accounts make it hard to understand arguments that an act is wrongful *because* it is exploitative. He notes, “if I say that a practice is [S_m] exploitative... then my statement already means that the practice is wrongful or bad, and thus that before using the term (in that sense) I should already have shown that [it] is wrongful or bad.” The consequence of this implication, Wood claims, is that “such a statement could be understood as categori[s]ing the wrongfulness of the act... or as rhetorically emphasi[s]ing this wrongfulness to someone who already believes the act is wrong but is suspected of giving this belief insufficient deliberative weight.”³³ He concludes that “the whole argument form which says that a practice is wrongful because it is exploitative looks merely rhetorical, if not downright confused.”³⁴

Wood’s first claim is largely true. Of course, we may not always be *certain* an act is wrongful when we claim it is exploitative if we are making a *prima facie* ascription of exploitation, but it is true that when we claim an act *is* an all-things-considered exploitation, we are committed to the claim that it is morally amiss (that is, S_m exploitation \rightarrow morally amiss). The consequences of this entailment are that ‘wrong because exploitative’ claims

³¹Leaving complications of vague predicates aside.

³²Hare (1952)

³³Wood (1995, p.6)

³⁴Wood (1995, p.6)

can only be rhetorical or categorical, but, Wood claims, that an act is exploitative provides us with more than rhetoric or a way of categorising acts. It gives us a “substantive argument for considering the act wrongful.” Therefore, there is no negatively moralised sense of exploitation.

One person’s modus ponens is another’s modus tollens: Wood (using modus tollens) claims that if exploitation is negatively moralised then ‘wrong because exploitative’ claims cannot be substantive. But since these claims *are* substantive, exploitation cannot be negatively moralised. I disagree. I claim (using modus ponens) that because there is a negatively moralised sense of exploitation and this implies that ‘wrong because exploitative’ claims cannot be substantive, then these claims are *not* substantive. Wood implies both that if ‘wrong because exploitative’ claims are “at most” categorical or rhetorical, they are “confused” or purposeless and that the ability to make substantive arguments for considering an act wrongful because exploitative is important.³⁵ But these implications are simply untrue.

First, ‘wrong because exploitative’ claims are *not* “downright confused”.³⁶ Much of our reasoning about particular cases involves categorisation. We may reason as follows: act x is similar to act y , which we know to be a case of exploitation. Exploitation is morally amiss; therefore if those properties that x shares with y include the properties that make y exploitation, then x is morally amiss. Knowing that y is a case of exploitation tells us which properties to look for in x in order to determine whether it too is exploitation and, therefore, impermissible. Additionally, categorising the wrongdoing involved in an act is helpful in determining how to apportion responsibility, make retributive transfers, or in weighing *pro tanto* wrongs. Finally, many exploitation claims do, as Wood says, rhetorically emphasise previously underweighted wrongs involved in a particular act. Exploitation claims used in this way are not unimportant or confused.

Second, it is unclear why we should agree with Wood’s implication that it is desirable to make substantive claims about an act’s wrongness by saying it is an exploitation. If exploitation is negatively moralised, its wrongness is derived from one of its characterising properties; if exploitation is not negatively moralised, then when it is wrong, it is so because of the presence of a non-necessary property alongside the characterising properties. Suppose the property in question is unfairness. If we believe that ‘exploitation’ admits to the two senses S_u and S_m , then we might plausibly claim that ‘exploitation’ as S_u is taking advantage, while exploitation as S_m is taking *unfair* advantage. The presence of unfairness in the latter sense of ‘exploitation’ makes this sense of the word negatively moralised. Now consider Wood’s position: exploitation admits to only one sense, and is therefore not necessarily wrong, but when it is wrong, it is so because of unfairness. It is unclear why Wood’s approach is in any way more desirable than the negatively moralised approach. The claim that the action is wrongful still depends upon the presence of substantive moral considerations. The only difference is whether these are implied in the claim that the action is exploitation or are made alongside such a claim. If Wood’s approach is not essentially different from the negatively moralised account, and the categorising and rhetorical roles of exploitation claims are as significant as I argue, then my *modus ponens* should be preferred to his *modus tollens*.

Neither of Wood’s arguments against the orthodox claim that exploitation admits to two senses, S_u and S_m , is successful. Wood’s claim that interpreting Nietzsche’s use requires us to abandon the claim that there are two senses of exploitation is inconclusive and his claim that the negatively moralised sense makes ‘wrong because exploitative’ claims unintelligible

³⁵Wood (1995, p.6)

³⁶Wood (1995, p.6)

is mistaken. Perhaps the strongest indictment against the non-moralised position is the overwhelming consensus, based on common sense intuition, that a clear distinction *can* be drawn between S_u exploitation as ‘taking advantage of’ and S_m exploitation as ‘taking *unfair* advantage of’. Thus, we can eliminate Wood’s position in the lower-right quadrant. With only the upper-left quadrant of table 5.1 remaining, we can conclude that any acceptable theory of exploitation must offer an analysis in which the concept is both negatively moralised and a form of injustice. In the remainder of this chapter I will take the negatively moralised ordinary language definition of exploitation as ‘taking unfair advantage’ as my starting point.

5.3 Unfair Transactions

So far, we have established that ‘taking unfair advantage’ is an appropriate starting point for a conception of exploitation. Which is to say that we have not established very much, for there are many accounts of both unfairness and advantage taking. Indeed, the concept of fairness is likely thornier than the concept of exploitation. A full analysis of fairness is impractical and would take us far afield from exploitation, but assessing those accounts of fair transaction that have been used in past analyses of exploitation gives us a few starting points. In 5.3.1 I review the principles of fair transaction that we have already seen before introducing Alan Wertheimer’s ‘fair market value’ (FMV) approach in section 5.3.2.³⁷ Here I argue that not only does Wertheimer’s principle inherit the general complications of two-part, or ‘dualist’, principles of justice, it also lacks justification. Ultimately, FMV provides neither necessary nor sufficient conditions for fair transaction and must be rejected. In 5.3.3 and 5.3.4 I defend RIGHTS against two possible counterexamples: cases of coercion and cases of fraud. I conclude that RIGHTS provides the best principle of fair transaction.

5.3.1 Use, Contribution, and Rights

So far we have considered *three* accounts of fair transaction: RIGHTS (Sample and Steiner), CONTRIBUTION (Roemer and Marx), and Goodin’s claim that exploitative transactions are unfair because they involve *playing for advantage* in situations where it is *inappropriate* to do so.³⁸ Such situations, Goodin claims, are any in which the other is “vulnerable”. As we have already seen in chapter 4, Goodin’s unqualified use of vulnerable is implausible. His general claim that the unfairness in exploitation involves inappropriate playing for advantage is true, but trivial. Depending upon how we understand ‘inappropriate’, the claim is compatible with both RIGHTS and CONTRIBUTION. And indeed, the analysis in chapter 4 should assure us that the only non-self-frustrating interpretation of ‘inappropriate’ is one that makes Goodin’s account compatible with the remaining two principles. I have also argued that CONTRIBUTION is equivalent to RIGHTS when Roemer’s withdrawal principle is specified by a libertarian principle, that is, when a state in which there are no rights violations present is used as the egalitarian baseline for the CONTRIBUTION principle. At this point we might ask which we should prefer, CONTRIBUTION or RIGHTS, or indeed, whether it makes any difference. If the best versions of each are equivalent, then choosing one over the other makes no *normative* difference. However, there are some *practical* reasons for preferring RIGHTS.

In his assessment of CONTRIBUTION, Alan Wertheimer notes that we can ask two questions about the principle, first, “Is it morally sound?”, and second, “[i]f so, how is it to be

³⁷Wertheimer (1996)

³⁸Goodin (1987, p.184, emphasis original)

measured?”³⁹ Of course, whether CONTRIBUTION is morally sound depends upon how it is specified. Certain versions are more plausible than others, but Wertheimer believes that, properly specified, the principle “may still carry moral weight.”⁴⁰ Nevertheless, he believes problems arise for the principle when we turn to the second question.

Wertheimer argues that “difficulties in the way of measuring an individual’s contribution... suggest that the principle of [CONTRIBUTION] is unlikely to provide a stable basis for fair transactions upon which parties could reasonably concur” even if these measurement problems “do not show that the concept of an individual’s contribution is incoherent.”⁴¹

The problems to which Wertheimer refers occur in (but are not limited to) neo-classical accounts. A. C. Pigou argued that workers are exploited when “they are paid less than the value which their marginal net product has to the firms employing them.”⁴² Pigou’s detractors pointed out that production factors other than labour may also be exploited in the sense he describes. Nevertheless, the exploitation of labour may, as Cohen (1988) argues, be of primary *moral* concern. The central idea in the neo-classical definition is that workers should be paid according to what they add to the firm’s output. But in many cases setting the bar for exploitation at those rates falling below the marginal product of labour may seem inappropriate. Since the marginal contribution of labour decreases as more workers are added to the workforce, if B joins ‘late in the game’, her marginal contribution to output may be minimal, even if she and A work an equal number of hours. As Sen argues, “To read in that *counterfactual* marginal story what *would happen* if one more unit *were* applied, given everything else, an identification of who has ‘in fact’ produced what in the total output is to take the marginal calculus entirely beyond its purpose and depth.”⁴³ Substituting the *average* product of labour for the marginal product avoids the problem of decreasing returns, but does not solve the problem of assigning wages based on *individual* effort.⁴⁴

Of course, CONTRIBUTION also features prominently in Roemer’s account, and here too measuring individual contribution presents problems. In economies with increasing returns to scale, there may be no one individual that satisfies Roemer’s PR conditions. The output produced when A and B combine each of their $\frac{1}{n}$ shares is greater than twice the output of each producing on their own. In order to deal with this problem, Roemer introduces the idea of vulnerability and minimally exploited coalitions in place of exploitation *tout court*. Minimally exploited coalitions are groups that satisfy the PR conditions that contain no redundant members. In his assessment of Roemer, Miller writes, “This generalisation of the exploitation criterion appears likely to work successfully for large economies... on the other hand it may be more difficult to apply Roemer’s criterion to individual exchanges... How would we construct the relevant counterfactual here?”⁴⁵ The problems that Wertheimer and Miller outline for the Roemerian and Pigou-Robinson versions of CONTRIBUTION are twofold. In order to provide a theoretically complete account, those defending CONTRIBUTION must *define* what counts as a morally relevant contribution and, if CONTRIBUTION is to be of any practical use, they must also show how this morally relevant contribution can be *measured*.

Citing these measurement concerns, Wertheimer abandons CONTRIBUTION in favour of his own principle of fair transaction, ‘fair market value’ (FMV), a position I consider in

³⁹Wertheimer (1996, p.227)

⁴⁰Wertheimer (1996, p.229)

⁴¹Wertheimer (1996, p.229)

⁴²Pigou (1920, III.XIV.3)

⁴³Sen (1992, p.119, emphasis original)

⁴⁴When A works 5 hours and B 10 the average product is 7.5, more than A worked and less than B worked.

⁴⁵Miller (1987, p.154)

the following subsection. But the measurement concerns for CONTRIBUTION also give us a practical reason for pursuing RIGHTS. In one sense, property rights are easy to measure. Our modern legal apparatus is largely devoted to the definition and protection of alienable assets. Extended definitions of property that include inalienable assets, like talents and skills, present a greater challenge, but as Steiner argues, these skills and talents “are bought, sold, and leased with considerable frequency. Accordingly, they clearly have an assessable value.”⁴⁶ Thus, inalienable assets can be folded into frameworks designed to handle more conventional assets.

The most complicated measurement problem for RIGHTS involves the operationalisation of the counterfactual that underpins the principle. Recall that RIGHTS claims that a transaction is fair when there does not exist an unrectified property rights violation in the history of exchanges leading to a transfer that *makes one party worse off than they would be*. If RIGHTS is to be of practical use, then we require an account of what it means for ‘one party to be worse off than they would be’. One such account is provided by Steiner’s conditions (2) and (3). According to these conditions, B is worse off than she would be because of a past rights violation when the rights violation makes at least one element of B’s original choice set, γ , unfeasible, leaving her with a new choice set, δ (where $\delta \subsetneq \gamma$), and B prefers the unfeasible option in γ to her most preferred option in δ .⁴⁷

When we consider RIGHTS alongside Goodin’s use-based account of fair transaction and CONTRIBUTION, RIGHTS emerges as the most plausible of the three, both normatively and practically. However, it faces a challenge from Wertheimer’s fair market value approach, considered in the following subsection.

5.3.2 Fair Market Value

Wertheimer claims that, “for a certain range of cases”, fair transactions are those that take place at the “fair market value... the price that an informed and unpressured seller would receive from an informed and unpressured buyer if the [good] were sold on the market.”⁴⁸ The ‘range of cases’ Wertheimer has in mind is non-obligatory, strictly Pareto-improving transactions in which there are relatively few agents rather than perfectly competitive markets, since in the former (but not the latter) “there may be a large bargaining range.”⁴⁹ In perfectly competitive markets, all agents are price takers; there “A does not have the space to transact with B at a different price, even if A would prefer to do so.”⁵⁰ However, most actual markets provide “at least some space in which agents can choose among a set of prices.”⁵¹ According to Wertheimer, when agents find themselves in markets that are *not* perfectly competitive, they should act *as if* they were by transacting at the fair market value.

Wertheimer claims FMV is “a price at which neither party takes *special* unfair advantage of particular defects in the other party’s decision-making capacity or special vulnerabilities in the other party’s situation. It is a price at which the specific parties to this particular transaction do not receive greater value than they would receive if they did not encounter each other.”⁵² Wertheimer concludes that such an equilibrium price “may nor may not be

⁴⁶Steiner (1999, p.144)

⁴⁷In chapter 6, I consider a characterisation of exploitation based on this operationalisation of RIGHTS, along with an alternative version of these principles.

⁴⁸Wertheimer (1996, p.230)

⁴⁹Wertheimer (1996, pp.216–17)

⁵⁰Wertheimer (1996, p.233)

⁵¹Wertheimer (1996, p.233)

⁵²Wertheimer (1996, p.232)

a ‘just price’, all things considered, but it may well be a non-exploitative price, for neither party takes unfair advantage of the other party.”⁵³

FMV: A transaction is fair *inter alia* when it takes place at the “fair market value, . . . the price that an informed and unpressured seller would receive from an informed and unpressured buyer” if the good were sold in a competitive market with many buyers and sellers.⁵⁴

Without the ‘*inter alia*’ restriction, FMV is implausible. It is neither sufficient to assure us that a transaction is fair nor is it not necessary for a transaction to be fair.

Suppose A steals B’s car and then offers to sell it back to her at the equilibrium price of \$10,000. Here FMV is satisfied, yet the transaction is certainly not fair; a *fair* price for the car is \$0, or even some negative amount to account for B’s loss of use. In this case, a ‘simple’ FMV is not sufficient for the fairness of a transaction. In order to avoid the insufficiency, FMV must operate against a background of macro-level justice. There may be cases in which A owes B a price that deviates from the market price because of past injustice. If FMV is to be plausible, it must assume that there are no special reasons A and B owe each other non-equilibrium prices.

Suppose in a village of ten people, five are butchers and five are candlestick makers. One day, after months of research, an industrious butcher reveals that he has learned to bake bread and begins selling rolls in his shop. Because he is the only butcher-turned-baker in the village, his rolls sell for a higher price than they would if all of the other butchers also sold rolls. Here a price that deviates from that which would exist in a perfectly competitive market does not seem unfair. Exchange at FMV is not necessary for a transaction to be fair. But Wertheimer recognises this. He writes, “given that ski areas build lift capacity to accommodate *peak* (marginal) demand rather than *average* demand, it is not unreasonable that weekend skiers pay a higher fee. . . Given that taxi drivers often spend a long time waiting for a fare at the airport, a practice from which the passenger benefits by not having to wait for a cab, one must include that waiting time in any calculation of a fair fare.”⁵⁵ The quote indicates Wertheimer understands that deviations from FMV may be fair when an individual is entitled to the fruits of their labour. Of course, in order to determine which cases qualify for this exemption, we must again refer to a broader macro-level principle of justice.

The market equilibrium price provides us with a criterion of fairness that operates, as Wertheimer claims, in “a certain range of cases”, namely those in which other, broader, principles of justice are already satisfied. This implies that macro-level principles of justice are themselves insufficient to ensure the fairness of at least some transactions. Wertheimer’s FMV assumes a ‘dualist’ approach to justice.

According to the dualist view, personal principles of fair transaction are separate from more general principles of institutional justice and fairness. Representing this view, Rawls writes, “The principles of justice for institutions must not be confused with the principles which apply to individuals and their actions in particular circumstances. These two kinds of principles apply to different subjects and must be discussed separately.”⁵⁶ Thus, for dualists like Rawls and Wertheimer, the satisfaction of macro-level principles of justice is not sufficient to ensure the transaction is all things considered fair; it may fail to satisfy principles like FMV. The opposing view, monism, “reject[s] . . . that there could be a plausible fundamental

⁵³Wertheimer (1996, p.232)

⁵⁴Wertheimer (1996, p.230)

⁵⁵Wertheimer (1996, pp.226–227)

⁵⁶Rawls (1971, pp.54–55)

normative principle for the evaluation of legal and other institutions that does not apply in the realm of personal conduct.”⁵⁷ For monists, “the personal is political”, that is, principles “about the just distribution of benefits and burdens in society, apply, wherever else they do, to people’s legally unconstrained choices.”⁵⁸

If dualism is false, then FMV fails as a principle of fair transaction. Both Liam Murphy and G.A. Cohen argue fairly convincingly against the plausibility of dualist principles, but their arguments are not decisive. Nor would it be appropriate to wade into such a broad debate in the process of assessing FMV. An alternative approach to assessing FMV is to focus on the sufficiency of RIGHTS, which applies both at the macro- and micro-levels. If RIGHTS is sufficient to assure us that a transaction is fair, then there is no need to appeal to additional principles like FMV. If RIGHTS is *not* sufficient to assure us that a transaction is fair, then we may return to FMV in order to see whether it can bridge the sufficiency gap. In the following two subsections I consider two problematic cases in which, at least *prima facie*, it seems that RIGHTS cannot provide us with a plausible principle of fair transaction.

Before considering these cases, it is worth noting that as it stands FMV is in need of further justification. While we may, as a convention, accept equilibrium prices as fair prices, Sample claims this “is not the same as showing that they genuinely *are* fair. Wertheimer does not explain why the market price should be regarded as the fair price and why non-market prices are exploitative.”⁵⁹ There is no reason to believe that market prices are themselves *intrinsically* fair. Rather, it seems that if they *are* fair they are so because they embody other properties: reciprocity, equality, anonymity, and the like.⁶⁰ Thus, hypothetical markets appear to only embody other principles that operate on the micro-level. If the failure of RIGHTS warrants a return to FMV the latter must be justified in order to be convincing as a principle of fair transaction. For now though, I will turn to the first problem case for RIGHTS: coercive transactions.

5.3.3 Coercion

Challenges to RIGHTS concern primarily its sufficiency as a principle for fair transaction. In particular, cases involving coercion and fraud present problems for the principle. But before examining these potential problem cases, recall a case the principle handles easily: the car auction case that appeared in chapter 2. There B wants to sell her car at an auction that includes buyers A and C. Although C had the highest reservation price and *would have* paid up to \$1000 for the car, he discovers his wallet has been stolen and A wins the car for \$600 instead. B now receives \$600 for her car, rather than \$810 (that is, one bid-interval more than the second highest reserve price). Intuitively, this result is unfair to B.⁶¹ This intuition is easily captured by RIGHTS, which claims a transaction is fair iff it is not the case that an unrectified property rights violation in the history of exchanges leading to the transfer makes at least one (more preferred) element of one party’s choice set unfeasible. Because an unrectified property rights violation *does* make a more preferred element of B’s choice set

⁵⁷Murphy (1998, p.254)

⁵⁸Cohen (1997, p.3)

⁵⁹Sample (2003, pp.23–24)

⁶⁰David Miller, who endorses his own version of FMV, also fails to provide justification for the principle. Like Wertheimer, Miller argues that the benchmark for a fair transaction “is one which occurs at an equilibrium price.” (Miller, 1987, p.156). Miller claims this is determined both by “natural facts about the world such as the amount and type of labour needed to extract raw materials[.] . . . the tastes of the population, . . . and the availability of skills and talents” and by “the basic entitlements of individuals participating in the market, including entitlements to natural resources and personal assets.” (Miller, 1987, p.156)

⁶¹Of course, as we’ve seen, answering the question of who is *responsible* for this unfairness is a more complicated affair.

unfeasible, the transaction in which A receives B's car for \$600 is unfair. Now let's turn to the first of the three difficult cases for RIGHTS: coercion.

Consider the following intuitively unfair transaction.

C12: The drowning woman.

Suppose B is about to drown, but A sees her and promises to rescue her only if she first agrees to pay him \$10,000.⁶²

While the case has many features in common with the car auction, it also contains an extra feature: not only does A get more than he ought, but in the drowning woman case—unlike in the car auction—A is also *obliged*⁶³ to transact with B. There are two mistaken applications of RIGHTS to this case that must be resisted.

First, we may be tempted to claim that if A is obliged to save B, then B has a correlative *right* to be saved. The violation of the right to be saved is the way in which the *principle* RIGHTS is violated in the drowning woman case. This is too quick. Although it may be true that B has a right to be saved, B's right to be saved has *not* been violated when A exploits her in the case. Since *unfair* transactions are *transactions*, anytime A transacts unfairly with B, A has already discharged an obligation to save B. In other words, A cannot both fail to save B and charge her too much for saving her.

Second, we may be tempted to claim that in the drowning woman case B has a right to be saved at a *lower price*—perhaps splitting the cost of rescue with A. When A saves B for more than this price, the transaction is unfair. This is also incorrect. Although it may be true that B has a right to a better price, this is simply another way of claiming that B has a right to a fair transaction. If we have a right to a fair transaction, then all unfair transactions will trivially *involve* rights violations because they are *themselves* rights violations. But Steiner's rights condition is not a vacuous condition; it explains why a transaction is exploitative: the outcome is unfair because of some *other* (and previous) rights violation. Neither a right to be saved nor a right to fair transaction shows that RIGHTS can capture the drowning woman case. In order to show that the principle can explain the unfairness of the drowning case, we must identify some *other* rights violation.

Note that this is also true of the car auction case. There, the RIGHTS principle is violated, not by the violation of a right to a lower price, nor by the violation of a right to transact, but by the violation of C's (property) right to his wallet. Because the drowning case provides no indication of whether B's dire situation was brought about by a past property rights violation, as it stands, we cannot determine whether the transaction is unfair by reference to RIGHTS.

Assessing the Drowning Woman Case

Because this indeterminacy is not present in the intuitive claim that the transaction is unfair, it may appear that RIGHTS has failed to adequately account for the first problem case. This conclusion would be premature. Although the drowning woman case is intuitively unfair, this intuition implicitly *assumes*—so I claim—the presence of a past rights violation. But B may, or may not, be in the water because of a past right violation. Her dire straits may have come about in one of three ways: B may be in the water through her own fault, through someone else's fault (either A himself or a third party, C), or through no one's (nature's) fault, that is, her situation is a case of bad brute luck.

⁶³At least within certain bounds—A is not obliged to save B if the cost to A is extremely high, for example.

First, when B is drowning through the fault of another individual—either A or C—identifying the *rights violation* is straightforward. If A (or C) pushed her in the water, then A’s charging \$10,000 for rescue is both intuitively unfair and unfair by RIGHTS. In cases where a third party is the rights violator, determining which actions are *permissible* for A becomes more complicated, as we saw in chapter 2’s ‘problem of moral responsibility’, but RIGHTS still shows the transaction to be unfair.

Second, as we’ve seen in chapter 4, if B is fully responsible for her predicament and aware of the stakes involved, then it cannot be the case that the transaction was unfair. For example, if B has some reason for venturing into the water, but knows both that the cost of rescue is \$10,000 and that the odds of needing rescue are, say, $p = 0.5$, then, contrary to our initial intuitive assessment, it does not seem at all unfair for A to charge her \$10,000 for rescue.

Finally, if B’s predicament is the result of bad brute luck, incorporating the drowning case into a liberal theory of exploitation is more difficult. Nevertheless, it is clear that with an expansive account of property rights, RIGHTS gives results in the drowning woman case that are in line with our ordinary intuitions.

While this analysis shows how RIGHTS may account for the *unfairness* of the drowning woman case when it is unfair, it does not explain why the case seems to differ morally from the car auction case. In order to explain the moral difference we must distinguish two questions about the cases: ‘*Why* are the cases unfair?’ and ‘*How* is the unfairness used in each case?’

The reason (why) we may say cases are unfair is the same for the car auction and the drowning woman case. B is entitled to do as well as she would have done had her rights not been violated. It is the answer to the question of *how* A uses B’s unjustly disadvantaged situation that generates the difference between the two cases. In the car auction, A uses B’s unjust disadvantage to make an unfair offer—an offer lower than B would have accepted, absent a past rights violation. In the drowning woman case, A uses B’s unjust disadvantage to make an unfair *threat*. Unlike the car auction, this is a case of *coercive* exploitation.

Coercive Exploitation

In the most general sense, coercive proposals are those in which A attempts to bring about a state of affairs where B does β by saying that if B β ’s, A will α , but if B does not β , A will ω . Of course, as many writers have argued, this is not sufficient for coercion, since this structure also holds for ordinary exchanges. Wertheimer notes that coercion is often taken to require that A’s proposal is a threat, not an offer: “the dominant philosophical view about coercion is... that threats coerce whereas offers do not [and]... the crux of the distinction between threats and offers is [that]... A *threatens* B by proposing to make B *worse* off relative to some baseline; A makes an *offer* to B by proposing to make B *better* off relative to some baseline.”⁶⁴ Of course, according to this distinction, whether a particular proposal is a threat or an offer will depend upon the baseline. From one perspective, we might say that A makes an offer to make B better off by saving her, since B’s situation will, of course, be markedly improved after A rescues her. On the other hand, if A is *obliged* to rescue B, then compared to a baseline that includes A’s rescue, he threatens to make B worse off. Wertheimer, following Nozick (1969), notes that these two perspectives mark a distinction between moral and statistical baselines:

⁶⁴Wertheimer (1987, p.202)

Under the statistical test, whether A is making an offer or a threat will depend on what is ‘normal’ in their society. . . In some societies it may be likely that A will rescue B [in the drowning woman case]; in other societies, not. Under the moral test, whether A is making an offer or a threat will depend on whether A is morally required to rescue B. If he is, then B’s baseline *includes* A’s beneficial intervention, and A’s proposal is a threat. If he is not, then A’s proposal is an offer.⁶⁵

In order to clarify the baseline that is being employed, we may call those coercion claims that employ moral baselines ‘wrongful coercions’. If A is, by hypothesis, *obliged* to rescue B, then when A proposes to leave B in the water unless she pays him \$10,000 he coerces her into an unfair transaction. It is the coercive nature of A’s proposal in the drowning case that generates the moral difference between that case and the car auction. The drowning woman case involves an unfair transaction that is also wrongfully coercive, while the unfair transaction in the car auction is of the ‘ordinary’ variety.⁶⁶

The primary lesson of the analysis is that whether a transaction is obligatory does not determine whether it is unfair. Rather, if a transaction is obligatory and the terms are unfair, then it is very likely that the advantaged party also used coercion as a means of securing these unfair terms. Wrongfully coercive unfair transactions involve two forms of wrongdoing—unfairness and coercion—both of which depend upon underlying violations or abuses of rights, but neither of which imply the other. Thus, a failure to discharge an obligation and coercion both may be *used* in unfair transactions, but, the final test of whether the transaction is unfair is specified by the RIGHTS principle.

5.3.4 Fraud

The second potential problem case for RIGHTS is fraud. Fraud may be seen as a form of imperfect consent, so, before discussing fraud itself, it is worth considering various forms of imperfect consent. At the limit, imperfect consent is *non-consent*. Unilateral transactions that are non-consensual are generally referred to as thefts. Of course, that a transaction is unilateral and non-consensual is not sufficient for it to be a theft: thefts also require that the dispossessed had a *right* to the transferred goods. The recovery of stolen property is not theft. But showing that RIGHTS conforms to our intuitions in non-consensual transactions is not difficult; more difficult cases of imperfect consent involve transactions where consent is *partial*. Partially consensual transaction may come about in an number of ways.

Partial Consent

One approach to partial consent ties consent to the properties of the transaction. We may say that although a transactor consents to the transaction, there is another, closely related transaction, the terms of which she prefers to those of the actual transaction. Her consent to the related transaction is greater than her consent to the actual transaction. But if this is all that is meant by partial consent, most transactions are partially consensual, since we can almost always imagine another situation with better terms.

⁶⁵Wertheimer (1987, p.207)

⁶⁶We might be tempted to think that although threats do not generate the unfairness of exploitative transactions, they are at least indicative of exploitation. But this is incorrect. A may threaten B without actually exploiting her because B may not respond to the threat. If A allows B to drown, he has done something very wrong, but he has not exploited her, he has only *attempted* to exploit her.

Another approach ties consent to a transactor's information set. We may say that although a transactor consents to the transaction, if she had known some piece of information (*i*), she would not have consented to the transaction. Some, but not all, cases of imperfect information involve *asymmetric* information where one person knows more than the other. When cases involve asymmetric information, we often call the transaction unfair. For example, suppose A, who is risk averse, trades a lottery ticket with odds $p = \frac{1}{2}$, \$100; $1 - p$, \$0 to B for \$45. While B consents to the transaction when she doesn't know the outcome of the lottery, she would not consent to the transaction if she were to learn that the ticket was not a winning ticket. There is nothing unfair about the transaction when neither A nor B know the outcome of the lottery. However, we would say that the transaction was unfair if A knew the outcome of the lottery was a loss.

But not all transactions involving asymmetric information are unfair. Suppose A, who has invested time and effort developing a skill at recognising valuable baseball cards discovers a card he knows collectors would pay \$20 for while he browsing at his local flea market. The seller, B, who does not know of the card's potential value to collectors, has priced it at \$10. While B would not sell the card to A for \$10 if she knew what he knew, it does not seem unfair for A to use his extra information to reap a benefit from the effort he invested in developing his skill. So, some transactions may be partially consensual in the sense that the transactor's giving consent depends on their not knowing *i*. Were they to learn *i*, they would withdraw their consent. So, some transactions involve imperfect information, some cases of imperfect information involve asymmetric information, and some cases of asymmetric information are unfair. When are these cases unfair?

According to RIGHTS, asymmetries of information are unfair iff B doesn't know *i* because of a past property rights violation. So, if B would have priced the baseball card higher but didn't because A stole her price guide, the transaction is unfair. But if B simply doesn't care very much about baseball cards and it is her own preferences that have made it the case that she does not know about the value to a collector, the transaction with A is not unfair.⁶⁷ So far, so good: RIGHTS delivers intuitively acceptable results for many cases of partial consent in cases involving asymmetric information. However, cases of asymmetric information involving fraud present a problem for the RIGHTS principle.

The Problem of Fraud

Consider the following case:

C13: The lemon.

A is selling his car, which he knows is a lemon. B offers A \$1,000 for the car, on the condition that it is in good working order. A lies and assures B that it is. B purchases the car and it breaks down on the drive home.

As with disadvantage, a lack of information may be brought about by three sources: ourselves, others, or the world. In the baseball card case, this lack of information is B's fault. She simply doesn't care about baseball cards. Here it is not unfair that she does not receive the gains of A's skill. On the other hand, had A stolen her price guide, B's lack of information would have been unfair and the unfairness would have been A's fault. The situation presented in the lemon case is slightly different.⁶⁸ A does not *prevent* B from obtaining information

⁶⁷Steiner (1987)

⁶⁸A rather interesting real-life 'lemon case' actually involves baseball cards. The "Gretzky Honus Wagner

that would alter her consent. Rather, he *misinforms* her; he gives her bad information. The lemon case presents a problem for RIGHTS because, while A's misinformation places B at a disadvantage, it is not straightforwardly clear that A violates B's property rights when he misinforms her. Thus, *prima facie*, RIGHTS cannot condemn fraud.

This is the conclusion reached by James Child (1994), who argues that "the basic moral principles of libertarianism do not support a prohibition of fraud. . . Not only does a no-fraud requirement not follow from libertarian first principles, it is inconsistent with them."⁶⁹ Child concludes that "the only doctrine compatible with a libertarian theory of transactions is the most 'hard boiled' version of caveat emptor."⁷⁰ Child argues libertarians are committed only to two basic moral axioms: self-ownership (A_1) and ownership of private property (A_2). These two fundamental rights⁷¹ imply two second-order rights of defence: first, the right to defend oneself, and second, the right to defend one's property against force, threats of force, coercion, and seizure. Finally, the domain to which these rights and their correlative obligations apply is the domain of moral agents who possess what Child calls 'general competence', including the capacities:

MA1 "to acquire, understand and appraise information, which includes considering its probability of truth or falsity and its relevance;

MA2 to entertain a stable set of preferences by which choice among various options with various payoffs can be made; and

MA3 in light of this information and these values, to consider choices and weigh possible risks, costs, and benefits of these choices. This includes the risk and cost of acting on false information."⁷²

For moral agents who fulfil (1)–(3), illegitimate transfers include transfers using force, transfers with threats of force, and theft by stealth⁷³, but according to Child, fraudulent transfers, which do not fall into one of these three categories, are not prohibited by the fundamental libertarian rights: "in [the lemon case, B's] right to self-ownership, management, or control was not compromised. [S]he was not forced to yield the [car], . . . was not threatened by force or violence, . . . was aware a transaction was proposed, . . . had an opportunity to engage [her] market competence fully, . . . had every opportunity to be critical, even s[c]eptical of the story[.]. . . [but she] was not."⁷⁴ Thus, Child concludes, "libertarians have no moral machinery for deriving the . . . fraud standard."⁷⁵

T-206" baseball card holds the distinction of bringing the highest price at auction for any baseball card. It last sold September 6, 2007 for \$2.8 million. However, controversy surrounds the card since—it is alleged—the card was altered by memorabilia trader Bill Mastro. Mastro allegedly cut down from its original size in order to disguise flaws along the card's border. Mastro was indicted by the FBI for defrauding bidders July 25, 2012. According to the FBI press release, the indictment claims that "Mastro allegedly failed to disclose that he had altered the Wagner T-206 card by cutting the sides in a manner that, if disclosed, would have significantly reduced the value of the card" (U.S. Attorney's Office, 2012).

⁶⁹Child (1994, p.722)

⁷⁰Child (1994, p.723)

⁷¹Or we may say 'one fundamental right' if the right to the ownership of private property is derived from self-ownership, but nothing in the following arguments depends on how we count these 'fundamental principles'.

⁷²Child (1994, p.730)

⁷³Child (1994, pp.731–33)

⁷⁴Child (1994, pp.733–34)

⁷⁵Child (1994, p.734)

Three Flawed Responses to the Lemon Case

There are two simple responses the libertarian might give to the lemon case. The first response to Child's argument is to claim that the lemon case involves the destruction of property, in violation of A_2 . If information is a form of private property, then A's destruction of B's justly held information is a violation of A_2 . Provided the status of information as private property does not depend on the medium in which it is stored, then this applies not only to information B may have stored externally (on a piece of paper, for example), but also information stored in her mind. Thus, when A destroys information in B's mind, he violates A_2 .

Unfortunately, it is not clear⁷⁶ that fraud involves the destruction of information. When A gives B *bad* information he does not destroy B's information by force, but rather provides B with new information that she must assess. Whether B decides A's information is credible and should replace or supplement her previous information is B's own choice. Neither A_1 nor A_2 are sufficient to condemn the transaction that takes place in the lemon case. Thus, these basic libertarian moral axioms must be augmented by a 'no-fraud' clause. As Child remarks, "every libertarian I know of introduces a fraud standard, or what... amounts to the same thing, a requirement for fully voluntary exchange."⁷⁷

The second response involves the addition of a 'no-fraud' clause to the fundamental libertarian principles. For example, Nozick argues that the state is justified in using its power to protect citizens "against force, theft, *fraud*, enforcement of contracts, and so on..."⁷⁸ For Nozick, fraud is a form of injustice in transfer. Child argues this approach is not open to the libertarian. Not only can libertarians "not derive a fraud standard from their moral first principles, but also... it is *inconsistent* with those principles."⁷⁹ Child does not argue that a no-fraud clause is *directly* inconsistent, but rather that it is inconsistent when combined with other considerations. His argument contains two steps. First, he argues that because the no-fraud clause is *ad hoc* and the libertarian's commitment to self-ownership acts as a 'presumptive trump' on axioms that would limit it, the burden of proof is shifted to the libertarian. In the second step, Child argues that if the admission of A_3 is *ad hoc*, then "we have no non-arbitrary criteria for keeping out A_4, A_5, \dots, A_n ."⁸⁰ And, while one additional axiom is not inconsistent with A_1 and A_2 , "the conjunction of A_3 through A_n *is*, that is, you cannot chip away at A_1 and A_2 too much... without destroying A_1 and A_2 , and with them goes the whole game."⁸¹ Not only are A_1 nor A_2 insufficient to condemn the transaction that takes place in the lemon case, but the addition of a no-fraud clause seems to open the door to a slippery slope of unacceptable limitations on libertarianism's fundamental commitments. Neither simple response appears successful.

A third, more nuanced response is that if fraud becomes widespread, markets for true information would appear, with an equilibrium between truthful information and fraudulent information emerging. Suppose that this does in fact occur. Initially both fraudulent and

⁷⁶In certain cases we may say that when A provides B with new information he destroys her old information. Suppose B is trying to remember a phone number for a short period of time by repeating it under her breath until she writes it down. A mischievously forces B forget the number by loudly repeating a sequence of random numbers. In this case, A destroys B's information by overwriting it by force. Here we may claim that A violates A_2 ; however, this case differs significantly from the lemon case. Indeed, the phone number case is not one we would typically call a case of fraud.

⁷⁷Child (1994, p.738)

⁷⁸Nozick (1974, p.5, emphasis added)

⁷⁹Child (1994, p.735, emphasis added)

⁸⁰Child (1994, p.736)

⁸¹Child (1994, p.736, emphasis added)

truthful markets exist. The uncertainty involved in fraud would drive prices down in fraudulent markets. Buyers could decide whether they wanted to pay more for a guarantee of quality, or less and risk possible fraud. Of course, the truthful markets would require an enforcement mechanism, guaranteeing the quality of their goods. Absent an enforcement mechanism, buyers in the truthful markets would have no recourse for reversing purchases they made from invading fraudulent sellers. Fraudulent sellers would thus have an incentive to invade the truthful markets and uncertainty about quality introduced by the invading fraudulent sellers would drive prices down in the (formerly) truthful markets. Without an enforcement mechanism, such markets cannot be maintained. One possible enforcement mechanism is a contract. Child writes,

... [suppose] private parties will prevent fraud by contract. For certifying the truth of information given upon sale would be a kind of warranty, and warranties are contracts. Thus, a libertarian state can enforce warranted truth in pre-transaction communication. Given this, soon no one would deal with those who refuse to so warrant their representations. This is a promising move for the libertarian.⁸²

This is promising perhaps, but ultimately the appeal to contracts is unsuccessful because it is question-begging. Child himself acknowledges this failure, asking, “Whence comes the authority to enforce *contracts* if a breach of contract involves no violation of the two libertarian rights?”⁸³ Suppose A and B sign a no-fraud contract. A has an incentive to defraud B by failing to honour the contract, as the game theoretic analysis shows. Suppose he does so. This breach of contract is a form of fraud. When B accepts a contract from A, she engages in an exchange with him—B promises to provide x and A promises to provide y; however, in breaching the contract A does not provide y. B would not have accepted the contract had she known A would not honour it and her belief that A would honour the contract is based on A’s lie that he would do so. Yet, as we have already seen, lying is not prohibited by either the right to self-ownership or the right to property-ownership.

At most, we may claim that A has done something *wrongful*. That is, he has violated a principle of general morality, but they cannot claim that A has violated libertarian political morality. So while, it might be possible to say that B herself is permitted to try to enforce the contract with A, libertarians cannot also endorse the claim that the state is permitted to help B reclaim her property. Nor can they allow that B might threaten to violate A’s rights to self- or resource-ownership in order to reclaim her property. At most she might threaten to damage A’s reputation as a transactor by informing others of his behaviour. This may somewhat reduce the occurrence of fraud, but without the ability to justify state enforcement of contracts, these threats lose much of their force.

Without a prohibition on fraud, contracts certifying the truth of information in a sale cannot themselves be enforced. And without such contracts, independent markets for honest and fraudulent dealings cannot be sustained. These markets require an independent enforcement mechanism that allows those who are defrauded to reverse their fraudulent transaction. Such independent enforcement mechanisms require collective state action. Thus, this response also fails.

⁸²Child (1994, p.734)

⁸³Child (1994, p.735)

A Successful Response

Child claims libertarians (1) do not and (2) cannot prohibit fraudulent transactions. While the first claim is correct, the second is an open question.

Claim (2), Child's inconsistency argument, actually contains two problems and one implicit assumption. The first problem—the basic inconsistency thesis—is also the most explicit. Child implies that fraud cannot be meaningfully *morally* distinguished from a large number other situations in which an individual's responsibility for the outcome of the transaction may be excused. Therefore, any acceptable justification for the addition of a no-fraud clause also justifies prohibitions on an intolerably large number of other self-ownership-limiting principles. However, the plausibility of this argument depends on an implicit assumption that the justification for the inclusion of a no-fraud clause must take the form of a *single* distinguishing property.

This calls for some elaboration. It is likely that cases of fraud can be morally distinguished from, for example, non-fraud lying in the marketplace by appealing to some property of fraud that is not present in non-fraud lying. It is also likely that cases of fraud can be morally distinguished from non-market lies by, again, appealing to a property of fraud not present in non-market cases. However, it is much less likely that the *same* property will allow libertarians to distinguish fraud from not only these two cases, but all other closely related cases (A_n). If the only acceptable justification for the no-fraud clause must avoid the inconsistency thesis by appealing to *one* property that distinguishes fraud from related cases, then it seems unlikely that libertarians will be able to meet the challenge presented by the inconsistency thesis. However, if libertarians can provide a justification for the inclusion of a no-fraud clause by piecing together a conjunctive list of distinguishing properties that apply to all possible A_n , then it is much more likely that they will be able to avoid the problems of the inconsistency thesis.

In other words, Child's implicit assumption is that what I will call a 'conjunctive justification' would not be an acceptable justification for a libertarian prohibition on fraud. If a justification contains n conjuncts because different properties are necessary to distinguish each additional A_n , then the assumption that a conjunctive justification is unacceptable is less controversial.⁸⁴ However, between the 'unified' case in which one property distinguishes fraud from all other related cases and the fully 'disaggregated' case in which n properties distinguish fraud from n cases, lies another possibility. It may be that an appeal to *relatively few* properties may distinguish a no-fraud clause from *most* other, but not all, self-ownership-limiting principles. In this case, libertarians could prohibit fraud and a few closely related acts without intolerably limiting self-ownership and property rights. Of course, it remains for libertarians to provide an acceptable conjunctive justification.

Yet, even if libertarians do so, they are not in the clear. Lurking within his remarks about the inconsistency argument is a second problem that Child does not emphasise. Not all added principles are equally demanding—some are more demanding than others. As Child correctly notes, "a general duty of veracity enforceable by the state . . . is surely a libertarian nightmare."⁸⁵ This 'nightmare' arises not from the addition of many minor self-ownership-limiting prohibitions, but from one very demanding prohibition. This we may call the 'demandingness challenge': If the justification for a prohibition on fraud also entails

⁸⁴Although, moral particularists would presumably reject the assumption since, according to many forms of particularism, any justification of a principle that forbids fraud in all cases, would always require an infinite conjunction because, for particularists there can be no one principle that forbids fraud in all cases.

⁸⁵Child (1994, p.737)

the addition of another limiting principle, then if this principle is overly demanding for a theory of political morality, the libertarian cannot justify prohibiting fraud without being committed to an implausibly demanding account of political morality.⁸⁶ Thus, it is not enough for libertarians to show they can provide a justification for a prohibition on fraud that avoids the inconsistency thesis. If the justification for prohibiting fraud does not apply narrowly to fraud, then libertarians must also ensure that any other prohibitions entailed by the justification are not overly demanding.

The Distinctiveness of Fraud In defence of the inconsistency thesis, Child argues that fraud is difficult to distinguish “from non-market truth telling”, “non-fraud lying” and “other moral problems.”⁸⁷ Yet, as I argue below, distinguishing fraud from these related cases is rather straightforward.

The first of Child’s three categories—non-market truth telling—is the most problematic. Although a general state-enforced duty to veracity would prohibit fraud, it would also fail the demandingness challenge. Child is correct that there is, *prima facie* no reason to limit the prohibition of misinformation to market contexts. Injustice resulting from fraud may occur, not only in market-based economic transactions, but also in political and personal domains. But the move from non-market truth-telling to worries about a state-enforced duty to truthfulness is too quick. A successful no-fraud clause would, *inter alia*, distinguish those cases of lying—inside and outside of markets—that would justify state intervention from those that would not, thus limiting state intervention to an acceptable number of cases. Child shows the libertarian must provide a justification that avoids the demandingness challenge, but he does not provide a reason to believe that the libertarian cannot do so. Worries about non-market truth telling simply provide an *example* of a related prohibition that a successful justification of a no-fraud clause must exclude.

Second, consider non-fraud lying. Child claims that “not all lying in the sale of a product is fraud. . . in negotiations a query ‘what is your lowest price?’ will almost always be met with a lie, at least early on,” but, he wonders, “how can fraud be differentiated from these sorts of market misrepresentations?”⁸⁸ He argues that “no libertarian would say that the state is justly empowered to see that you tell the truth in this case.”⁸⁹ Here Child elides an important distinction between the relative bargaining positions of the transactors and the properties of the goods being exchanged. Unlike the latter, lies about the former do not have the negative practical consequence of threatening market stability. Indeed, such behaviour is common in most markets. Practically speaking, the libertarian has no need to prohibit A from lying about his bargaining position. More importantly though, even if the libertarian grants that there are no *moral* grounds for prohibiting fraud without also prohibiting lies about bargaining positions, it does not follow that the state is thereby “justly empowered to see that [A] tell[s] the *truth*” about his bargaining position.⁹⁰ A may be entitled to withhold strategic bargaining information from B even if he is not permitted to *misinform* her. Child fails to distinguish between a right to lie, a right to conceal information, and a duty to tell

⁸⁶Much depends on what counts as implausibly demanding. What is over-demanding for the libertarian might be tolerable for those who do not subscribe to a libertarian theory of political morality. Answering the demandingness challenge is a more acute problem for the libertarian than it is for other accounts of political morality. However, regardless of the demandingness standard, certain duties—such as a general duty to veracity—will be overdemanding for any account.

⁸⁷Child (1994, p.737)

⁸⁸Child (1994, p.737)

⁸⁹Child (1994, p.737)

⁹⁰Child (1994, p.737, emphasis added)

the truth. The libertarian may prohibit lies without granting the state powers to enforce truth-telling.

Finally, Child lists a number of additional cases he claims a ban on fraud would prohibit:

- a. “What about innocent mistakes by either party resulting in a ‘bad bargain’ for one?”
- b. What about transactions which transpire at exorbitant and unfair prices?
- c. What about transactions stemming from unequal bargaining power, including those entered under economic duress?
- d. Why not rescind transactions the consequences of which do not work out as planned by one party? How can you ban fraud without banning these transactions?”⁹¹

To each of these cases the libertarian can provide a response. In the case of (a), we must ask *why* the transactor made an innocent mistake. If it was because she was worse off through some previous property rights violation, then the transaction is unjust; if the mistake her own fault, then she is responsible for the ensuing ‘bad bargain’. For (b), we must ask whether the worse-off transactor has a right to a better price. If not, there are no libertarian grounds on which such transactions can be prohibited. If a price is unfair in libertarian terms, then it must be the result of previous (property) rights violations. Most importantly, however, transactions falling in (b) do not involve misinformation and may be distinguished from fraud in this way. A similar response may be given to (c). We must ask whether the worse-off had a right to different bargaining power or a right to not be under duress. If they did, then the transactions that follow are not morally binding; if they did not, then there are no libertarian grounds for prohibiting these transactions. Finally, fraud differs from type (d) transactions because, unlike in the purchase of most lottery tickets for example, the uncertainty involved in fraud is asymmetric and introduced by the defrauding party through deceit.⁹² Prohibitions on fraud do not extend to all transactions that are unattractive to one party *ex post*. Fraud is closely related to transaction types (a) through (d), but can be easily distinguished from each, primarily because none of these cases involve deceit. The slippery slope is slippery only if we are blind to important differences between the cases.

Fraud can be distinguished from each of these related cases by appealing to the deceit involved in fraud. Thus, a prohibition on fraud does not invite a large number of self- or property-ownership limiting principles. The inconsistency thesis is not the central problem a prohibition on fraud presents for the libertarian. Rather, the hard problem for the libertarian is the demandingness challenge. Libertarians must find a way to prohibit fraud without prohibiting deceit *tout court* because a general prohibition on deceit—even if it does not imply a state enforceable duty to veracity—is overly demanding for a theory of political

⁹¹Child (1994, pp.736–37, lettering added)

⁹²Of course, the sale of lottery tickets can involve asymmetric information introduced by the seller, along with a form of imperfect consent (since, if buyers knew their ticket was a losing one they would not agree to purchase it). This may occur when A manufactures scratch cards with serial numbers and knows which card is the winning card. While such lotteries can be distinguished from fraud because they do not involve deceit, it is worth pointing out that they also involve a different form of consent than do cases of fraud.

In lotteries, ticket buyers know that there may be such asymmetries and yet they consent to the purchase. Although they do not know whether the ticket they buy is a winner or loser, they know they lack this information. And, importantly, those who buy lottery tickets agree to this situation—it is constitutive of lotteries. Lottery buyers know they don’t know and consent to this state. However, when A lies to B about the quality of the car he sells her, he intends to put B in a situation where she does not know that she lacks full information about the car and this is the informational state to which B would not consensually agree to be placed in. Unlike lottery ticket buyers, the defrauded don’t know that they don’t know relevant information and (wouldn’t consent to not knowing).

morality. Yet, the fact that deceit is not only necessary for fraud, but also is what seems to make fraud *unjust* complicates the task of finding a justification for a prohibition on fraud that does not also apply to deceit.

Further, it does not follow from the *moral distinctiveness* of particular cases that those properties that *justify* a state imposed prohibition of one case will not entail a prohibition of related cases. As the demandingness challenge makes clear, the *number* of related cases is relatively unimportant; what matters is the collective demandingness of the related cases, be they one or many. I have shown that none of the above examples definitively establish either part (inconsistency or demandingness) of Child's second claim. But whether libertarians can justify a prohibition on fraud without intolerably undermining their basic commitments remains an open question, to which I will now turn.

Justifying a Prohibition on Fraud In what follows I argue that by appealing to both the wrongfulness of fraud and its negative social effects libertarians can provide a justification for a prohibition on fraud that applies narrowly to fraudulent transactions. A prohibition on fraud need not entail any further restrictions on the rights to self- or property-ownership. Since libertarians have already (incorrectly) assumed that their basic principles entail a prohibition on fraud, the exclusion of the ability to defraud others from the scope of permissible actions does not represent a restriction on individuals rights to self- or property ownership that is greater than restrictions already in place in existing libertarian accounts.

The demandingness challenge is the most difficult problem for the justification of a prohibition on fraud. It seems impossible to justify a prohibition on fraud without appealing to the wrongfulness or unfairness of deceit; yet a state imposed prohibition on deceit is overly demanding for libertarians. However, it may be possible to prohibit fraud without *prohibiting* deceit *tout court*. Libertarians may justify the inclusion of a prohibition on fraud by claiming that fraud is an act that combines wrongfulness in what I have called 'general morality' with consequences that undermine social stability. This is a 'conjunctive justification' of the sort that I claim Child implicitly assumes is unacceptable; yet, as I will show, it manages to avoid both a new form of the inconsistency thesis and the demandingness challenge.

According to many general moral theories, lying is, *pro tanto*, wrongful. Most libertarians (and their opponents) will concede this claim. However, the mere fact that an act is prohibited by general moral principles does not provide sufficient reason to prohibit the act in a theory of political morality. Ordinarily, state enforced prohibitions on acts like adultery, betrayal, and deceit seem inappropriate and heavy handed. While we may grant that such acts are indeed wrongful, we are nevertheless alarmed by reports of states that have imposed sanctions on such acts. Of course, some acts prohibited by theories of general morality also merit state enforced prohibitions. Murder, rape, and theft are not only wrongful, they are unjust, and prohibited by theories of political morality.

In order for an act that is prohibited by general morality to be prohibited by political morality, it must satisfy some additional criteria. Although libertarians can appeal to the violation of self-ownership in order to justify prohibitions of rape and murder, the fundamental question is why the proper domain for *self-ownership* is political morality. A rough answer to the question can be found in the contractarian tradition: a respect for rights to self-ownership and property ownership allows us to avoid a 'state of nature'. Only when we can own and improve land and resources without constant threat of theft and violence is

a stable society possible. Although both historical and contemporary contractarians⁹³ have outlined similar justifications for principles of political morality, perhaps none has been more influential on modern libertarianism than John Locke, who argues that “government, being for the preservation of every man’s right and property, by preserving him from the violence or injury of others, is for the good of the governed: for the magistrate’s sword being for a terror to evil doers, and by that terror to [e]nforce men to observe the positive laws of the society, made conformable to the laws of nature, for the public good”.⁹⁴ The important social effects of these rights (and their enforcement) justifies their addition to the domain of political morality. I am not claiming that rights to self-ownership and property ownership are necessary (or sufficient) for the emergence of a stable society, after all, contractarianism takes many forms. Rather, the argument is that an appeal to the important effects of the basic libertarian principles provides the libertarian with a justification for including these principles in the domain of political morality and this justification may, in part, also justify a prohibition on fraud.

Appeals to stability and social import allow us to differentiate sub-classes of violations of general morality. So, although a general prohibition on betrayal would be over-demanding for a theory of political morality, a prohibition on *treason*—the betrayal of the state—would not. Although the kind of wrongdoing involved in both treason and more pedestrian forms of betrayal is the same, we can single out treason for prohibition at the political level by noting that it threatens social stability in a way that ordinary betrayal does not.

By now the analogy with fraud should be obvious. Although a general prohibition of deceit would be overly demanding for a theory of political morality, libertarians can distinguish fraud from mere deceit by noting that deceit threatens to undermine the important institution of the market only in cases of fraud. Since markets are central social institutions for the libertarian, acts that threaten their stability may be prohibited, but since more pedestrian forms of deceit do not threaten markets, they are not prohibited.

Yet, this appeal to the negative effects of fraudulent transactions, is only one side of the argument. Although the market stability granted by a no-fraud clause is a necessary component of its justification, it is not sufficient to justify a libertarian prohibition on fraud. One can imagine many other actions that might also be prohibited by appealing to the positive social effects of the prohibition.⁹⁵ Indeed, such arguments might justify an intolerably large set of self-ownership limiting prohibitions. Without a second component in the justification, the problems brought by the inconsistency thesis resurface.

Libertarians might argue that the *severity* of the effect of fraud and the import of markets distinguishes it from related cases. While this diminishes the force of the inconsistency thesis, it does not exclude the possibility of state enforcement of significant personal sacrifices if their social effects are significant. Such situations are surely contrary to the libertarian ethos. Libertarians want a theory of political morality that grants only minimal powers to the state. Fraudulent transactions are problematic for the libertarian both because they undermine the effectiveness of markets and because they cannot be prohibited by appealing to existing libertarian principles. Libertarians must sacrifice either markets or their moral minimalism; of the two, the latter seems the best choice. However, in the *spirit* of minimalism, libertarians should endorse a justification for a no-fraud clause that departs the least from their original framework. A justification that appeals *only* to the severe effects of fraud will

⁹³Classic examples include Hobbes (1996), Narveson (1988), and Gauthier (1986).

⁹⁴Locke (1764, §92)

⁹⁵In fact, *prohibition* (of alcohol) is one that comes to mind.

not accomplish this task, something more must be added.

Libertarians can claim they are justified in prohibiting fraud not simply because it undermines a social institution, or even because it does so *severely*, but also because fraud involves a violation of general morality. That is, the wrongfulness of the deceit involved in fraud is also a necessary component of the justification for a no-fraud clause. Not all acts that undermine important social institutions should be prohibited by political morality, only those that are also wrongful according to general morality.⁹⁶ An appeal to acts prohibited by general morality allows the libertarian to avoid the inconsistency thesis. Not only does this restrict the scope of acts with negative social effects, it restricts the scope to acts that individuals have *independent* reasons (provided by general morality) to refrain from performing. Conversely, not all acts that violate principles of general morality should also be prohibited by a theory of political morality, only those wrongful acts that also severely threaten important social institutions should be prohibited. An appeal to the negative social effects of fraudulent transactions allows libertarians to avoid the demandingness challenge. They can avoid prohibiting deceit *tout court*, by appealing to the social effects of deceit involved in fraudulent transactions. Therefore, both components are necessary and jointly sufficient to justify the inclusion of a no-fraud clause within the libertarian theory of political morality. Taken together, they allow libertarians to prohibit fraud while avoiding both the demandingness challenge and the inconsistency thesis.⁹⁷ When RIGHTS is supplemented by this no-fraud clause, it can successfully handle cases like the lemon case.

I have argued that RIGHTS is preferable to CONTRIBUTION on practical grounds and to USE on normative grounds. Wertheimer's FMV principle lacks full justification since market prices seem to be a proxy for other notions of fairness that Wertheimer does not specify. More importantly, though, the existence of a micro-level principle like FMV implies that macro-level principles like rights cannot account for all cases of unfairness. Yet, as I argued in subsections 5.3.3 and 5.3.4, RIGHTS (when supplemented with a no-fraud clause) can account for all forms of unfairness within the domain of transactions between moral persons. Neither coercion nor fraud gives us reason to employ a micro-level principle like FMV. Further, since FMV presupposes a macro-level principle like RIGHTS, it acts only as a refinement on such principles. Thus, at this point, it seems appropriate to continue with RIGHTS as the working principle of fair transaction, supplementing it with FMV only if problems of insufficiency arise. For now, though, I will turn to the second face of exploitation: taking advantage.

5.4 Taking Advantage

Jon Elster argues that the definition of exploitation as 'taking unfair advantage' embodies

... a causal and a normative aspect. The normative aspect is that of unfairness, defined with reference to some theory of justice. . . The causal aspect of exploitation is that of taking advantage of someone: when I derive a benefit from another

⁹⁶This solution applies only to problem fraud presents for libertarianism as a theory of political morality. Some libertarians (such as Peter Vallentyne) advance the framework as a general moral theory. Those defending libertarianism as a general moral theory cannot claim that lying is wrongful according to general moral theories since their own framework is a general theory and does not prohibit lies.

⁹⁷Although writers like Milton Friedman might be receptive to this justification of a fraud prohibition, non-consequentialist libertarians, e.g. Nozick, may be more resistant. However, this resistance is purchased at the price of the plausibility of their theory. Since, as I have argued, fraud is unjust and undermines markets, a prohibition on fraud is essential for libertarians. And since existing principles do not prohibit fraud, this leaves these libertarians with a significant problem. I thank Mike Otsuka for pressing me on this point.

person being placed in such a situation that his best option is to act in a way that is to my benefit. Taking advantage of someone allows me to expand the opportunity set of the person of whom I am taking advantage, by making him an offer, but it does not allow me to restrict it by threats and similar means.⁹⁸

As I noted in chapter 3, Steiner and Roemer both provide systemic accounts in which the conditions that generate the normative aspect (maldistribution) are prior to the systemic or institutional elements that bring it about. For Goodin and Sample, the priority of causal and normative conditions is inverted. According to these duty-based accounts, the particular source of an individual's vulnerability is not normatively important; all vulnerability creates the opportunity for exploitation. What makes a particular transaction exploitative is the use of these vulnerabilities to gain benefit. For the duty-based accounts, injustice can create vulnerability, but it is not necessary for exploitation. I have already shown that this non-judicial approach to exploitation is mistaken: non-judicial accounts are either self-frustrating or incomplete. A transaction is an exploitation only if the vulnerabilities that are taken advantage of have arisen through injustice (unfairness). Elster's claim that exploitation requires an *unfair distribution* is correct and, in the previous section, I argued that RIGHTS is the best of many principles of fair transaction.

Just as there are many principles of fair transaction, there are many specifications of what it means to 'take advantage'. In the present section I consider three senses of 'taking advantage'. Elster's Pareto-based account, a vulnerability-based interpretation, and an awareness-based interpretation. Although each is appropriate in certain contexts, I argue that the advantage taking involved in exploitation is best captured by the awareness-based interpretation.

5.4.1 Advantage as (Strict) Pareto Improvement

According to the above quote from Elster, we take advantage of someone by "adding an option to his opportunity set that is better both for him and for [us] than the option preferred by him prior to the expansion."⁹⁹ Advantage taking in this sense is equivalent to *offering* someone a strictly Pareto-improving transaction. Strictly speaking this cannot be correct; the offer may be rejected, in which case, no advantage would accrue to either party. So it seems that we actually *take* advantage in Elster's sense when we offer a strictly Pareto-improving transaction and the transaction, in fact, occurs. That is, for Elster (with this repair), A takes advantage when the following two conditions are satisfied:

- (A1) A's preferences are better satisfied by transacting with B than they would be if he did not transact with B,
- (A2) B's preferences are better satisfied by transacting with A than they would be if she did not transact with A.

When these two conditions are both satisfied, the transaction is a strictly Pareto-improving transaction. A and B take advantage of each other's *preferences* when both *gain* from the transaction. That is, A and B each gain *because of* the other's preferences.

Elster's preference-based account of advantage taking coheres with *one* widely used sense of 'advantage taking'. Unfortunately, this sense adds nothing to the unfairness conditions

⁹⁸Elster (1982)

⁹⁹Elster (1982, p.364)

used in Steiner’s original account. Preference-based advantage taking is trivially satisfied anytime A and B engage in a strictly Pareto-improving transaction, fair or otherwise. According to Elster’s sense of ‘taking advantage’, the ‘taking advantage’ portion of the ordinary language definition of exploitation as ‘taking unfair advantage’ is redundant. When we employ Elster’s sense of ‘taking advantage’, exploitation is simply a strictly Pareto-improving transaction in which the RIGHTS principle is violated. As we have already seen in the analysis of Steiner’s account, the scope of this conception of exploitation is far wider than ordinary use. We require a more restrictive sense of taking advantage if this condition is to limit the scope of exploitation. Some further condition must be added to (A1) and (A2).

5.4.2 Vulnerability-Based Advantage

Elster claims that taking advantage is the causal component of exploitation. His own account is causal in the sense that A’s preference for B’s money over his car and B’s preference for A’s car over her money are both necessary conditions for an ordinary transaction. If either party’s preferences were reversed (i.e., A would rather keep his car), the transaction would not occur (at least not consensually). But we can also distinguish a stronger sense of taking advantage. It may be the case that transacting with B results in a gain for A *only because* B has been unjustly disadvantaged. Consider the following case:

C14: The unemployed.

Suppose that A runs a factory that produces widgets. The cost of widget production means that A cannot pay his workers more than \$10 per hour without making a loss. Suppose that B previously worked at a job she enjoyed for \$40 per hour, but because of a violation of her property rights, she has lost her job. Now that she is unemployed, B prefers to work for A for \$10 per hour over remaining unemployed.

Prior to the property rights violation, A and B would not have transacted, but now that B is unjustly disadvantaged—vulnerable—she transacts with A. That is:

- (A1) A’s preferences are better satisfied by transacting with B than they would be if he did not transact with B,
- (A2) B’s preferences are better satisfied by transacting with A than they would be if she did not transact with A, and
- (A3) The satisfaction of Steiner’s conditions (1)–(4) was necessary for A and B to transact.

Here A takes advantage of B’s unjust vulnerability (or disadvantage) in order to gain. According to this sense of advantage taking, B’s unjust vulnerability is *necessary* for the transaction to occur. The vulnerability-based sense of advantage taking is stronger than, and implies, the preference-based sense. When A works for B for \$10, (A1)–(A3) are satisfied, but obviously not all those transactions satisfying (A1) and (A2) also satisfy (A3). This sense of advantage taking also, of course, has the advantage of (at least partially) incorporating Sample’s and Goodin’s claims that exploitation involves the mistreatment of vulnerable people.

Despite its initial attractiveness, (A3) is *too* strong. Suppose that in addition to her \$40 per hour job, B wants to make some spending money on the side, working occasionally for A for \$10 per hour. However, before beginning to work for A, a property rights violation causes B to lose her higher-paying job. Knowing that she is desperate for money, A reduces his offer to \$5 per hour and B reluctantly accepts. It seems natural to say that A exploits B here,

and, indeed, that A takes advantage of B's unjust vulnerability; yet (A3) is not satisfied. In the absence of the rights violation, A and B would still transact (but not on the same terms). It seems that we require a sense of exploitation that is stronger than the conjunction of (A1) and (A2), but weaker than (A1)–(A3).

5.4.3 Advantage as Awareness

Consider again the second case of the unemployed. What strikes us as exploitative about the case is not only that B is forced to work for A for less because of a rights violation, but also that A *adjusts* his offer because of the advantage in bargaining power the rights violation gives him. Exploitation, it seems, does not merely depend on the conditions that cause B to enter into an unfair transaction. A's response to B's condition—as Goodin and Sample remind us—is also important in our ordinary judgement that a particular transaction is exploitative. The idea that exploitation requires an intentionality condition has appeared through the first half of the thesis. An intentionality condition provides one way in which A's attitudes towards B can be captured. Consider the following:

(A4) A intentionally engages in a transaction in which RIGHTS is violated.

Although promising, like (A3), (A4) is too strong. In general, we may claim that an agent is morally responsible for the side-effect of an intentional action without committing ourselves to the claim that the side-effect itself was an intentional action (indeed, Alfred Mele (2001) has argued that these side-effects of intentional actions are never brought about intentionally). This general claim seems true for advantage taking in possible cases of exploitation. Suppose A knows that B chooses to work for him for \$5 per hour only because of a past rights violation, but in hiring B, A does not *intend* to exploit, harm, transact unfairly, or act towards B in an impermissible manner. Rather, A intends to make a profit by producing widgets.

Yet, provided A is aware of B's plight, he does not passively or ignorantly take advantage of B in the sense characterised by (A1) and (A2). When A knows about B's situation, regardless of whether the *reason* on which he acts is tied to the fact that RIGHTS is violated, he actively takes advantage of B.¹⁰⁰ In some cases where (A4) is not satisfied, A still seems to take advantage of her unjust disadvantage in a stronger sense than that characterised by (A1) and (A2) when he hires her for \$5. We must replace the 'intentionally' that appears in (A4) with 'knowingly'.

However, this too is too strong. It does not matter whether A's belief that Steiner's conditions are satisfied is *justified*,¹⁰¹ but only that A believes they are satisfied and that it is true that they are satisfied. Suppose A truly believes transacting with B would be a violation of the RIGHTS principle, but his belief is based on an incredible backstory that in no way relates to the true reason that the transaction would violate RIGHTS. Nevertheless, RIGHTS is violated, the transaction is unfair, and A truly believes this (but given the information on which his belief is based he oughtn't to believe it). Does A still exploit B when he transacts with her in a way that violates RIGHTS? What is at stake here are two possible relationships

¹⁰⁰We may be sceptical about the claim that the intention behind A's action is not suspect. Indeed, the so-called Knobe effect, first outlined by Joshua Knobe, shows that there is an asymmetry in our willingness to ascribe intentionality: "people... seem considerably more willing to say that a side-effect was brought about intentionally when they regard that side-effect as bad than when they regard it as good." (Knobe, 2003, p.193)

¹⁰¹Of course, the characterisation of knowledge as 'justified, true belief' is problematic. See, for example, Gettier (1963) and the many papers that extend and generalise Gettier's problem. The point here is that what *does* matter is only that A has a belief that the conditions are satisfied and that this belief is true. That it is justified, or justified and satisfies additional robustness conditions, is not necessary for A to take advantage in the sense I am outlining.

between the unfairness condition and the advantage condition. Does exploitation require only that the conditions on both sides are satisfied, or does it require that the state of the world that makes it such that RIGHTS is violated is the same state that leads to A's true belief (A5) that it is violated? Or, to put the issue another way, in giving up the requirement that A's true belief is justified, have we weakened the advantage condition too much? I do not think so. His true belief seems sufficient to give him a reason to constrain his actions.

A's true belief about B's plight gives him a *reason* to modify his behaviour towards her. Even if A's *intention* in engaging in the transaction includes no reference to the violation of RIGHTS, when he ignores the reason to modify his behaviour generated by his awareness of B's situation, he understands that the kind of advantage he has over B is more than the sense of advantage characterised by (A1) and (A2). In other words, when A's true belief gives him a reason to not transact, and yet he transacts, he can be said to *actively* 'take' the advantage that B's disadvantage places before him. Thus, the sense of taking advantage relevant to the context of exploitation seems to be given by the following:

(A5) A truly believes that he engages in a transaction that is unfair according to RIGHTS.¹⁰² A actively takes advantage of B when A is aware¹⁰³—truly believes—he receives an advantage from B. Weakening the condition by removing the belief component leaves us with only Elster's sense of taking advantage; weakening the condition by removing the truth component leaves us with a case where A only *attempts* to take advantage of B, but does not in fact take advantage of B. On the other hand, strengthening the condition by requiring either that A knows or that A intends to engage in a transaction in which RIGHTS is violated is overly restrictive. It rules out the possibility of A's 'taking advantage' of B when A is aware that he receives a benefit from B's unjust disadvantage, but either lacks justification for this belief or focuses his attention on other reasons for transacting.

5.5 Chapter 5 Summary

In this chapter I argued, in section 5.2, that—contra Wood—exploitation is best captured by the ordinary language definition of 'taking unfair advantage' and is therefore a negatively moralised concept. In section 5.3 I claimed that the unfairness involved in exploitation is best captured by the RIGHTS principle of fair transaction.

Finally, in section 5.4 I claimed that the advantage taking component of exploitation is best captured by the 'awareness-based' sense of advantage taking, which is operationalised by (A5). When the claim in RIGHTS that a past rights violation "makes one party worse off than they would be" is operationalised using my modified conditions for Steinerian exploitation (1)–(4) and (A5) is added to these conditions, we obtain the following working account of exploitation:

A exploits B in the bilateral and voluntary transaction ϕ just in case:

- (1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that

¹⁰²There are two possible readings of A5. According to one, the content of A's belief must actually reference the RIGHTS principle. Not only does A believe that the transaction is one that we would ordinarily call unfair, he knows the wording of the RIGHTS condition and believes the transaction satisfies this condition. According to the second, weaker reading, A5 requires that A truly believes he engages in a transaction that is what we would ordinarily call unfair and this transaction also happens to satisfy the RIGHTS condition. This is the reading that applies to A5, not only in its present rendering, but in subsequent renderings, when the general RIGHTS is replaced with various operationalisations of the principle.

¹⁰³In my use, awareness is simply equivalent to true belief. Thanks to Julian Fink for the suggestion of 'awareness' in this condition.

- (2) makes at least one element of B's original¹⁰⁴ (and finite) feasible set, γ_B , unfeasible, leaving her with a new feasible set, δ_B (where $\delta_B \subsetneq \gamma_B$),
- (3) B prefers the unfeasible option in γ to her most preferred option in δ ,
- (4) A's egoistic preferences:
 - (a) are better satisfied in ϕ because γ is restricted to δ ,
 - (b) and if (1)–(3) also hold for A, then the positive change in the degree of satisfaction of A's preferences, given (1)–(3) hold for B, is strictly greater than any positive change in the degree of satisfaction of B's preferences, given (1)–(3) hold for A.¹⁰⁵
- (A5) A truly believes (is aware) that he engages in a transaction in which (1)–(4) are satisfied.¹⁰⁶

In the following chapter I discuss and assess this working characterisation of exploitation.

¹⁰⁴It is possible to interpret 'original' in condition (2) as original in either a temporal or modal sense. Here and in the remainder of the thesis, 'original' should be interpreted in the modal sense: exploitation requires that B's feasible set is reduced from the feasible set that would have been available to her had the rights violation not occurred.

¹⁰⁵Note that condition (4b) presupposes the ability to make interpersonal comparisons of preference satisfaction. Although we have little trouble making interpersonal comparisons in practice, whether our doing so is *justified* is a difficult theoretical issue (see Robbins (1932, 1938)). Two justifications for the theoretical legitimacy of interpersonal comparisons have been suggested by (Hausman, 1995) and (Harsanyi, 1955), however, Hausman's approach, as he acknowledges, is ethically indefensible and Harsanyi's approach, as some latter commentators have observed, is unsatisfactory.

In condition (4b) and its later variants (G4b, G4'b, G4''b) I do not favour any particular approach to the theoretical problem of interpersonal comparisons of utility over any other, I simply suppose that, though the theoretical problems are interesting, they do not significantly affect our ability to make interpersonal comparisons in practice. The import of pointing out the necessity of interpersonal comparisons is to show that the informational requirements of this account of exploitation are richer than certain other moral concepts; i.e. the Pareto principle.

¹⁰⁶For my purposes the reference in (A5) to A's beliefs is to his *actual* beliefs, not merely those that he might be expected to have. This rules out the possibility of the occurrence of exploitation in cases where A's mistaken beliefs are caused by negligence. I discuss the difference made by the reference to actual beliefs in more detail in 6.2.1.

Note that a more generalised account of exploitation can be obtained by leaving the interpretation of 'beliefs' unspecified.

Chapter 6

Refining the Working Conception

6.1 Chapter 6 Introduction

Chapter 5 closed with a working account of exploitation. On this account, exploitation's unfairness is captured by the RIGHTS principle of fair transaction—in particular, the operationalisation of rights given by my revised conditions of Steinerian exploitation. The advantage taking aspect of exploitation is captured by condition (A5): A takes advantage of B in an exploitative transaction when he is aware that the transaction he engages in is unfair, but nevertheless fails to constrain his actions.

In this chapter I assess the working characterisation in section 6.2. The first part of the section considers the features of the working characterisation. I consider the role of negligence in exploitation and the possibility of mutual exploitation and also show that the account solves the problems of moral responsibility and the pervasiveness issues we encountered in the analysis of Steiner's and Roemer's accounts. In the second part of the section I present a problem of obligation for the working characterisation. Sections 6.3 considers two unsuccessful solutions to this problem and 6.4 presents a successful solution that abandons the preference satisfaction-based fairness conditions. Here I argue that an account of exploitation that can function in deontically imperfect situations must make reference to a non-outcome-based non-exploitative baseline. I advocate a bargaining-based baseline and develop a second characterisation of exploitation that avoids the problems of the first while retaining its positive features. Section 6.5 concludes the chapter.

6.2 Assessing the Working Characterisation

Before considering the features of the working account, I will briefly comment on its characterising conditions. Recall from the previous chapter that A exploits B in the bilateral and voluntary transaction ϕ just in case:

- (F1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that
- (F2) makes at least one element of B's original (and finite) feasible set, γ_B , unfeasible, leaving her with a new feasible set, δ_B (where $\delta_B \subsetneq \gamma_B$),
- (F3) B prefers the unfeasible option in γ to her most preferred option in δ ,
- (G4) A's egoistic preferences:
 - (a) are better satisfied in ϕ because γ is restricted to δ ,

- (b) and if (F1)–(F3) also hold for A, then the positive change in the degree of satisfaction of A’s preferences, given (F1)–(F3) hold for B, is strictly greater than any positive change in the degree of satisfaction of B’s preferences, given (F1)–(F3) hold for A.

(A5) A truly believes (is aware) that he engages in a transaction in which (F1)–(G4) are satisfied.

In this presentation, the original revised Steinerian conditions have been re-labelled according to their role (fairness, gain, advantage). The *unfairness* of exploitation to B is captured by conditions (F1)–(F3). B is unfairly worse off when a past property rights violation (F1) causes her to lose options (F2) that she preferred to the options that are available after the rights violation (F3). While these conditions alone characterise a transaction that is unfair to B, they do not say anything about whether A *gains* from the unfairness.

Exploitation requires not only that B is worse off than she should be, but also that A *gains* from B being worse off. This intuition is captured by (G4), the gain condition. (G4a) requires that A gains from B’s being disadvantaged by (F1)–(F3), and that he gains *in the right way*. Schadenfreude is (arguably) not exploitation. If the only additional gain A receives from B’s situation satisfying (F1)–(F3) is the satisfaction of his preference to see B worse off when transacting with him, we would not normally describe A’s action as one of exploitation, but of spite. Although the concept could be extended to cover cases involving non-egoistic preferences, most of the time we think of the gain that A receives as being a form of *material* gain. Sometimes we extend this notion to include egoistic preference satisfaction, but rarely would we take it to include non-egoistic preference satisfaction.

In some cases, it is possible for conditions (F1)–(F3) to hold for both A and B. For example, suppose A and B both live in an apartment complex that has been the target of arson. The arson may make both worse off in a way that satisfies (F1)–(F3). Supposing the two then transact, if they are equally disadvantaged neither can exploit the other. Exploitation depends on the *relative* disadvantage of one of the transactors. Thus, in order for A to exploit B, the arson must have left B relatively worse off—according to (F1)–(F3)—than A. Condition (G4b) adds this restriction.

Finally, as we saw in chapter 5, in cases of exploitation, the advantage that A takes of B is best captured by A5. A takes advantage of B when he truly believes he engages in a transaction in which (F1)–(G4) hold. That is, he truly believes that in transacting with B, his preferences will be better satisfied than they would have been, and B’s preferences will be less well satisfied than they would have been, because of a past property rights violation.

This working characterisation of exploitation has many benefits. It solves the pervasive-ness issues plaguing Steiner’s and Roemer’s accounts, it solves the problem of moral responsibility presented by unintentional exploitation, and it incorporates the intuition shared by Goodin, Sample, and Wood that exploitation involves some form of active advantage taking while also avoiding the pitfalls of non-judicial exploitation. It also excludes the possibility of mutual exploitation and negligent exploitation. In the subsections that follow I address these benefits and exclusions, beginning with the latter.

6.2.1 Excluding Mutual Exploitation and Negligence

Mutual Exploitation. First consider the exclusion of mutual exploitation. Although paradigmatic cases of exploitation are asymmetric, some writers have also defended the

existence of mutual exploitation. Judith Tormey writes “we do, after all, commonly describe persons as standing in a relation of mutual exploitation (A exploits B relative to x ; B exploits A relative to y)... In mutual exploitation each gets more than a fair share of some burdens (or benefits) and less than a fair share of others.”¹ While Tormey’s claim that we often *speak* of mutual exploitation is correct, this does not entail that such talk is coherent. The existence of mutual exploitation requires the existence of two incommensurable domains of value, for although A may gain unfairly from B in domain x and B may gain unfairly from A in domain y , if there exists a domain z common to both x and y , the relative gains of A and B can be weighed on the common scale. If overall A gains more than B, we can claim that on the whole, A exploits B. We might also say that A x -exploits B and B y -exploits A, but this does not entail that they both all-things-considered exploit each other. Such forms of exploitation are possible even in a preference satisfaction framework; x and y may simply be labels for two different transactions. However, in a uni-dimensional value theory (or a value theory with multiple, commensurable dimensions), mutual exploitation in a *single* transaction is impossible, because on such a value theory, advantage is a zero-sum game. The value theory underpinning both Steiner’s account of exploitation and the above characterisation is that of subjective preference satisfaction. Therefore, it does not allow for mutual exploitation.

Nevertheless, the working conception of exploitation outlined above is not as starkly opposed to mutual exploitation as it may seem. If mutual exploitation is possible, then the structure of exploitation in each particular incommensurable dimension of value is surely similar to that characterised by (F1)–(A5). The preference-based value theory would be replaced by (or augmented with) another value (incommensurable with preference satisfaction) for the second dimension, but the basic conditions of unfairness and awareness would remain the same. Therefore, the basic structure of the working conception is compatible with many theories of value and can account for different kinds of intuitions about exploitation when combined with different theories of value.²

Negligence. The second remark concerns the existence of cases in which (A5) is not satisfied, but only because A is negligent. In such cases, although A *ought* to truly believe (F1)–(G4) hold, he does not. There are two responses we might give to such cases. First, we might simply add a negligence disjunction to (A5). The condition is satisfied when either A truly believes (F1)–(G4) hold, or when he *ought* to truly believe they hold. Alternatively, we might claim that because (A5) is not, in fact, satisfied, A does not exploit B. Rather, he is guilty of a negligent and unfair transaction. Much like murder can be distinguished from manslaughter by ‘malice aforethought’—the intention to kill or harm—we may claim that exploitation requires A to truly believe he is engaging in an unfair transaction.

The debate between these two approaches is more than merely terminological. Since A is culpable for his negligence, exploitation would remain wrongful with and without the addition of a negligence disjunct to (A5). However, the fact that exploitation seems to be closely related to intentional action in ordinary use provides one reason for excluding the negligence disjunct. But a stronger reason for excluding the negligence disjunct is provided by the fact that its inclusion would require an account of negligence. Whether A exploits B does not seem to depend on whether A *ought* to have known that a particular transaction satisfied (F1)–(G4). An account of exploitation should not depend on an account of negligence. By excluding the negligence disjunct, the independence of exploitation is maintained. Further,

¹Tormey (1976, p.212)

²Thanks to Sebastian Köehler for pointing this out to me.

as I noted above, excluding negligent and unfair transactions from the scope of exploitation does not exonerate the negligent transactor, it merely classifies the transactions as a different kind of normative transaction.

6.2.2 Solving Responsibility and Pervasiveness Problems

Pervasiveness. As we saw in chapters 2 and 3, both Steiner and Roemer provide accounts of exploitation that make its presence pervasive. The inclusion of (A5) in the working conception limits the scope of exploitation to those cases in which A is aware that the transaction is unfair to B. This circumscription excludes what we might normally consider ‘ordinary transactions’ from the scope of exploitation. We would not ordinarily call transactions like the purchase of a loaf of bread from the store, dog-sitting in exchange for a bottle of wine, etc., exploitative transactions. Yet, neither Roemer’s account, Steiner’s original account, nor my revised version of his account, provide a way to exclude these cases from the scope of exploitation. Of course, we may be unaware that there is actually something unfair about these seemingly ordinary transactions, but we would not normally call them exploitations. The addition of (A5) restricts exploitation to those unfair transactions in which A *is* aware that the transaction is out of the ordinary—in particular, to those situations in which he is aware that it is unfair. The addition of (A5) does not deny the normative significance of unfair transactions; indeed, the unfairness characterised by (F1)–(G4) supplies the core normative content of the conception I have outlined. What my account denies is that we should call transactions satisfying (F1)–(G4) *exploitation* rather than simply labelling them ‘unfair transactions’. This is, of course, merely a matter of convention. We may label transactions satisfying any number of properties ‘exploitation’. But if our labelling is to be useful, it should bear at least some resemblance to our ordinary use of the concept. The justification for calling (F1)–(A5) ‘exploitation’ is a topic to which I return in section 7.3.1.

Moral Responsibility. The addition of (A5) not only restricts the domain of exploitation, it also offers a solution to the problem of moral responsibility for Steiner’s theory of exploitation. Recall from chapter 2 that the problem had the following form:

The Problem of Moral Responsibility.

- P1 Steiner’s theory permits non-culpably ($\neg C$) ignorant ($\neg K$) exploitation (E).
- P2 On Steiner’s theory exploitation is a negatively moralised concept, which means if exploitation occurs, then it is wrong (W).
- P3 If an action is wrong, then there is at least one person who is morally responsible (R) for the action.
- P4 If a person is morally responsible for an action they performed, then they acted either intentionally or from culpable ignorance.
- P5 If a person acted intentionally, then they must know (K) they performed the action.

In my discussion in chapter 2, I argued that consistency may be restored to the five premises through what I called the ‘badness solution’. Contrary to the implication of P2, negatively moralised concepts are those that imply badness *or* wrongdoing. By modifying P2 to include a ‘badness’ disjunction ($E \rightarrow W \vee B$) in the consequent, the inconsistency is removed. However, in that section I also briefly outlined an ‘intentional condition solution’, noting that there are those that may disagree with the badness solution, not because the solution

itself fails, but rather because the problem of moral responsibility sets out in the wrong direction before P2. If we think that exploitation must be intentional, then the ‘obvious’ solution to the paradox will be to abandon the first premise. In my brief discussion of the ‘intentional condition solution’ I noted two drawbacks: it limits the scope of a theory of exploitation and, because it necessitates modifications to the minimal ethical framework that underlies Steiner’s original approach to exploitation, it is strongly revisionary.

The first of these concerns is not a problem; rather, because Steiner’s original approach was both diachronically and synchronically pervasive, the limited scope brought by the addition of (A5) is—as we have just seen—a virtue. The second concern is a problem only for those committed to classical liberalism’s moral minimalism. Steiner has shown that even within this framework classical liberals have a reason to be concerned with exploitation. If classical liberalism cannot incorporate the augmentations necessary to capture the trust funder case³ (section 2.2), then they must restrict their concern to a subset of what we might ordinarily call exploitation. This, I suggest, is a problem for the classical liberal, not for a conception of exploitation that draws on a more extensive moral framework.⁴

Of course, as the discussion of (A5) makes clear, intentionality is too strong as a necessary condition for exploitation. I have claimed that the weaker requirement that A has a true belief that B is disadvantaged is, when combined with the other four conditions, sufficient for the truth of the claim that A exploits B. This ‘awareness’ is sufficient to hold A morally responsible for his actions toward B. If A is aware that he is engaging in an unfair transaction with B, then, provided that justice is one of our moral rules,⁵ A has a moral reason to constrain his own advantage when transacting with B. When A truly believes B is unjustly disadvantaged, his failure to constrain his own unjust advantage engages with his judgment-sensitive attitudes and thus warrants a negative moral appraisal of A.⁶ Thus, we may say that A is attributively responsible for exploiting B. Whether (and to what extent) A may also be *substantively responsible* for the exploitation (e.g., should bear some of the cost of rectifying the unfairness) is a further question that I do not consider in this chapter.

6.2.3 The Problem of Obligation

The working conception of exploitation characterised by (F1)–(A5) excludes negligence and mutual exploitation from its scope, solves the problem of moral responsibility, and limits the concept to transactions that are more naturally described as exploitation. I have argued that these features are all, in one way or another, *virtues* of the working conception. However, it also faces a significant challenge. The obligations implied by the conditions (F2) and (F3) are unduly strong in cases of non-obligatory transactions in which the only non-exploitative trade lies above A’s reserve price. Ultimately, these overly demanding obligations undermine the plausibility of the fairness conditions contained in the working conception. The remaining parts of this section outline this ‘problem of obligation’ and a solution is explored in the next section.

³Recall that in the case B owns an investment that pays her a dividend of \$100,000 per annum, an amount that allows her to realise a fairly comfortable standard of living. B values her leisure time, but is not opposed to work. Given the dividend, B is willing to work only if she can secure a salary greater than \$40,000. However, no such jobs are available, so B is unemployed. One day C steals B’s investment, putting an end to her dividend. Having lost her income, B must now find work. The only job she can secure is to work for A. Although A usually pays employees in B’s position \$30,000, knowing that B needs a job and that the next best job B could obtain pays less than \$20,000, A pays B only \$20,000.

⁴Whether my full account of exploitation is compatible with a stark-budget liberalism is a topic that I return to, briefly, in chapter 7.

⁵See the discussion in §5.2.1, Steiner (1994, pp.222–23), and Kramer (et. al., pp.282–83).

⁶See Scanlon (1998, p.248) and subsection 2.3.2.

Consider the following case in which a transaction between A and C is non-obligatory, neither is under any obligation to transact with the other, and in which the non-exploitative price lies above A's reserve price.⁷

C15: (Upsetting) The apple cart.

Suppose that every Saturday C drives his apple cart to B's village and sells apples in the market. Every Saturday B buys 10 apples for \$1 each. One Saturday C's apple cart is stolen. When B arrives at the market her only alternative for apples is A, who charges \$2 per apple. Because A's apple selling operation is smaller than C's, she does not enjoy economies of scale. A's break even price for an apple is \$1.50. Finally, suppose that the most B would be willing to pay for each apple is \$2.50.

In order to avoid exploiting B, A has two options. Since the transaction is non-obligatory (by hypothesis), A may simply refrain from transacting with B. Alternatively, he may engage in a *fair* transaction with B. Since the transaction is *unfair* to B when she receives less than she would have because of an antecedent property rights violation, it must be *fair* to B only if she receives what she would have, had the property rights violation not occurred. The conditions that specify the sense in which B receives less than she would have are (F2) and (F3). According to these conditions, B receives less than she would have when a rights violation makes a previously feasible and more preferred alternative unfeasible. Therefore, in order for B to receive what she would have (and thus for the transaction to be fair), A must transact with B on terms that restore her feasible set to the original γ_B . That is, in order to ensure (F2) and (F3) do not hold, A must offer his apples to B for \$1 each. Faced with the options of either not transacting or making a loss, a rational and morally motivated A will choose to refrain from transacting with B since this is the morally permissible option that maximises his own welfare.

Here we have a problem. We know from the reserve prices listed in the apple cart case that, given a choice between selling B 10 apples for \$20 and not transacting, A would prefer to sell the apples. Additionally, given a choice between buying 10 apples for \$20 and not transacting, B would prefer to buy the apples. Unfortunately, the only morally permissible options, according to the working conception of exploitation, are either to not transact, or to exchange 10 apples for \$10. The moral constraints imposed by (F1)–(A5) are worse for both A and B and better for no one. If (at least a portion of) the wrongfulness of exploitation stems from the harm to B when compared to a non-exploitative baseline, then the greater harm of a *no transaction* situation must also be, if not morally wrong, morally *bad*. Moral obligations that encourage the badness of a 'no transaction' situation are surely unacceptable. Further, the problem represented by the apple cart case is not limited to special cases. In any non-obligatory transaction in which the only non-exploitative trade lies beyond A's reserve price, A (if morally motivated and rational) will choose to refrain from transacting. It seems clear that the conditional obligations implied by (F2) and (F3) are too strong. The following section outlines two unsuccessful attempts to solve the problem of obligation.

⁷We may also assume that A and C provide identical products. Of course, from an economic perspective, this means that A's continued existence in the market is puzzling. Perhaps a more plausible version of the case places A and C in different village markets and the rights violation does not involve the theft of C's apple cart, but rather a barrier along B's path to C's village. Since this (arguably) more plausible version is also more complex, I will not concern myself with a realistic explanation of the existence of different apple prices in the same village. We may simply take the prices as given.

6.3 Unsuccessful Solutions to the Problem of Obligation

In many ways the problem of obligation is similar to the extended problem that the dilemma in chapter 4 creates for Goodin's and Sample's accounts. Goodin and Sample can avoid the dilemma by prohibiting A from taking prudent risks, but not only is this solution too stringent, it is not well-motivated. The source of the problem is not A's risk taking, but the insensitivity to personal responsibility of the come-what-may duties that underpin the accounts. In the problem of obligation, A can avoid the overly demanding fairness obligations implied by (F2) and (F3) by simply refraining from transaction with B, but like the prohibition on prudent risks, this does not address the root cause of the problem.

This section outlines two possible solutions to the problem of obligation that the working conception of exploitation faces. The first argues for a 'piecemeal' account of fairness where different obligations apply in different contexts. The second argues that the problem presented by the apple cart case can be solved by restricting the scope of exploitation to exclude such cases. Unfortunately, as we will see, neither solution is acceptable.

6.3.1 Piecemeal Fairness

Initially, we might claim that exploitation is not always a *strongly* moralised concept. Although *most* cases of exploitation are wrongful, some are merely morally bad. In cases where the only non-exploitative trade lies beyond A's reserve price, A's *pro tanto* obligation to refrain from exploiting B is overridden by feasibility constraints and concerns about welfare. Thus, in most exploitative situations, the obligations implied by (F2) and (F3) apply, but in apple cart cases A may charge B any price.

This approach is unsatisfactory for two reasons. First, part of the motivation for including the advantage condition A5 is to capture the intuition that exploitation is not only morally bad, but also morally wrong. Second, and more importantly, this solution—like the claim that A can avoid the demanding obligation by not transacting—fails to address the root cause of the problem, namely, a defect in (F2) and (F3). Surely, it is not the case that anything goes if the only non-exploitative trade according to (F1)–(A5) lies beyond A's reserve price. Provided there is range of prices over which A and B may bargain, it seems that there must be some subset of this range that represents a fair price, weighing the rights violation, feasibility constraints, and welfare concerns all together. If this is the case, then perhaps a different piecemeal approach to (F2) and (F3) would provide a more comprehensive solution to the problem.

A more plausible piecemeal solution may claim that when A can—without making a loss—offer B a price that means the original (F2) and (F3), then A should offer this price. However, if the 'fair according to (F2) and (F3) price' lies *beyond*⁸ A's reserve price, then alternative conditions apply that allow A to offer B a price within the range of both transactors' reserve prices.

In the apple cart case, this range is \$1.50 to \$2.50 per apple. A number of options suggest themselves here. We may say that A should receive a price that is as close to the price she would have received had the rights violation not occurred. In the apple case, this is \$1.50. But since \$1.50 is A's reserve price, he will be indifferent between transacting at \$1.50 and not transacting. Requiring a price at which A is indifferent to transacting is only marginally better than requiring a price at which A prefers to not transact. In both cases the transaction

⁸Either above, or below, depending on whether A is the buyer or seller.

may fail to occur, which would leave both worse off.⁹ Perhaps, in order to ensure A engages in the transaction, the principle that applies in these cases should provide him with some small incentive to transact. We may say that in the apple case A should offer B a price of \$1.51 per apple (or not transact). Here both A and B are better off transacting.

And yet, judged from one perspective, this is *unfair* to A. Why should he give B a price that is less than he would charge another customer (\$2) simply because someone has stolen a third party's (C's) apple cart? A has not changed his prices in order to capitalise on B's disadvantage. In fact, if A is only permitted to charge B \$1.51 for an apple, it seems that A does worse than he would have because of a past rights violation. Why should A, rather than B, pay for the rights violator's wrongdoing? According to this line of reasoning, A should receive \$2 per apple.

However, judged from another perspective, \$2 is unfair. It is *not* true that A does worse than he would have had the rights violation not occurred, because, had the rights violation not occurred, B would not have transacted with A at all. Although \$1.51 is less than other customers are charged, it is more than A would have received from B and it is a price at which A makes a profit.

Yet again, \$1.51 is also dubious. The requirement that A gain (however little) from the transaction means that A profits from B's disadvantage. A's refusal to trade without profit, in a sense, holds B hostage to A's demands. It seems odd to say that A should receive some gain if he trades with B, but the fact that the transaction is non-obligatory means that without such gain A has no incentive to transact. Although the 'normal' obligations of (F2) and (F3) in non-apple cart cases seem acceptable, none of the salient options for obligations that apply in apple cart cases are unproblematic. Without incentivising A, he may opt to not transact, which is bad for both, but when we *do* incentivise A, he profits from B's unfair disadvantage. No matter the approach, the piecemeal solution is unsatisfactory.

6.3.2 A Change of Scope

At this point we may want to change tack. In the piecemeal solution, justifications for each of the two plausible secondary obligations undermine the other. Further, these issues do not even address the looming question of whether that particular solution is *ad hoc*. Rather than developing two-part obligations, we might want to avoid the problem of obligation by excluding the apple cart case (and similar cases) from the scope of the working conception. According to this approach, the apple cart case would be a kind of unfair transaction, much like Steinerian exploitation without the addition of the advantage condition (A5). Of course, as Wertheimer notes, "we would still have to ask what should be said about [these] transactions, whether and why they are wrong, and whether or why they should be prohibited."¹⁰ However, by restricting the scope of exploitation, these tasks would not form part of the analysis of exploitation.

The apple cart and similar cases could be excluded from the set of exploitative transactions by an additional advantage condition. However, this is not merely an exclusion of convenience; there are reasons to think that there is a moral difference between apple cart cases and other forms of transaction. There is a sense in which A 'could not have done otherwise' in the apple cart case that, we might argue, excuses him from moral responsibility for *wrongdoing*. Consider the following options open to A in the apple cart case: A

⁹Alternatively, we could assume that there is a small cost (ϵ) to transacting. In this case, A would not transact at \$1.50, since a price of $\$1.50 - \epsilon$ is a loss-making transaction for him.

¹⁰Wertheimer (1996, p.259)

could give B a non-exploitative price (\$1), not transact, or charge his normal price (\$2). Not transacting is the least preferred option by both A and B; both can agree that some form of transaction is better than not transacting. Suppose A gives B the non-exploitative price of \$1 per apple. In this case, although B is not harmed by the transaction (when compared to a no rights violation baseline), A *is!* A is worse off than he would have been had the rights violation not occurred because his break even price is \$1.50. A should not be obliged to shoulder the burden of the antecedent rights violation any more than B. Requiring that A give B the non-exploitative price burdens an innocent bystander.¹¹ But if A charged every other customer \$2, then charging B \$2 means that he does not take *special* advantage of B, even if he takes advantage of her in the sense of the original (A5). We might claim that exploitation requires an additional advantage condition. Not only must (A5) hold—A truly believes (F1)–(G4) obtain—but the following condition must also be satisfied:

(A6) A modifies the terms of his transaction *in response to* this true belief.¹²

When (A5) and (A6) are satisfied, A takes advantage of B in a stronger sense than he does when only (A5) holds. He modifies his price in order to maximise his own gain from B's disadvantage. (A6) gives us a reason to claim that A does not take *wrongful* advantage/exploit B when selling the apples to him for \$2. Note that regardless of whether (A6) represents a plausible additional constraint, it is clear that the weight of A's wrongdoing in transactions where (A6) is satisfied is stronger than it is in those cases where only the weaker (A5) holds. When A does not modify his price in response to B's disadvantage, we might claim that he acts permissibly because his other options—not transacting or offering a non-exploitative price—are respectively worse for B and non-obligatory. Recall that in the apple cart case, B would pay up to \$2.50 for an apple. She would therefore prefer paying \$2.50 to not transacting. Since not transacting is permissible, and (at least part of) the wrongfulness of exploitation stems from the harm to B, then if B is harmed less when paying \$2 than she would be when not transacting, it seems permissible for A to charge her the same price he would charge anyone else. He is not taking *special* advantage of B's unjust vulnerability.

Although this 'change of scope' solution is more plausible than the piecemeal approach, it too, is problematic. First, the 'could not have done otherwise' argument that attempts to justify the claim that A is not guilty of wrongdoing in transactions satisfying the weaker (A5) is flawed. While it is true that (A6) implies greater wrongdoing, the claim that A could not have done otherwise depends on a false trichotomy. Of course, if A's *only* options were \$1, no transaction, or \$2, then we might plausibly claim that he could not have done otherwise. But A's options are not so limited; he may also offer B \$1.51 (or a number of other prices). A *could* have done otherwise (while still profiting). However, determining precisely *what* he should have done—i.e., which price between \$1.50 and \$2 A should offer B—returns us to the problems of the piecemeal solution. None of the prices between \$1.50 and \$2 seem to be free from problems.

Second, while the fact that A does not alter his price in the apple cart case makes him seem innocent relative to transactions in which he does alter his price to take advantage of B's unjust disadvantage, in other cases A's failure to alter his price *in B's favour* seems to be a natural case of exploitation. Consider the following:

¹¹There are certain similarities between the present case and Judith Thomson's (1991) discussion of self-defence.

¹²Note that (A6) is similar to (but differs slightly from) the 'because of' contained in the vulnerability-based sense of advantage taking characterised by (A3).

C16: Stuck in the mud.

Certain ‘unmaintained’ dirt roads in Nebraska become irresistible to teenagers on ATVs when heavy rains turn the roads into a soupy mess. Farmers that live along these roads dislike the ATVs because they cut deep ruts in the roads they use to access their homes and fields. In an attempt at deterrence, many area farmers charge \$100 for extracting those ATVs that inevitably become stuck in the mud, even though the cost of rescue to the farmer is much less than \$100.

Suppose a past property rights violation forces B to travel on an ATV down A’s unmaintained road and she becomes stuck. Although B explains her situation to A (and so he truly believes (F1)–(G4) hold), he maintains his \$100 price for extracting her ATV.

Although A does not *raise* his price in response to B’s unjust disadvantage, his failure to constrain the terms he offers her in light of her situation seems to be a natural case of exploitation. Further, although nothing about the case indicates that A has any obligation to pull B out—there is no imminent threat to her safety—if he does help her, he should not charge \$100.

Not only is A *not* excused from wrongdoing in cases involving (A5), because he *could* do otherwise, but also, as the stuck in the mud case suggests, it seems natural to claim that A *exploits* B in these cases. Solving the problem of obligation by excluding the apple cart cases from the scope of exploitation is unsatisfactory. Such cases seem to be natural examples of exploitation. Although additional wrongdoing may be present in cases satisfying the stronger (A6), that condition is not necessary for exploitation.

The problems of the piecemeal solution suggests that no non-controversial and immediately obvious secondary principle can govern cases like the apple cart case. Of course, this does not mean that such a principle cannot be found, but it does mean that the selection and justification of a secondary principle is not a straightforward task. The failure of the change of scope solution suggests that strengthening the conditions in order to exclude cases like the apple cart case from the scope of exploitation will arbitrarily exclude natural cases from the scope of exploitation.

Despite their failures, the discussion of these two possible solutions suggests a range of acceptable prices in the apple cart case. We are looking for a principle that does not require A to pay B the *price* she would have received absent the rights violation, for, provided the transaction is non-obligatory, this will result in a no-transaction situation. On the other hand, as the stuck in the mud case shows, the principle must not entirely absolve A of responsibility to modify his price in light of B’s disadvantage. The solution to the problem of obligation presented in the following section attempts to strike just this balance by modifying the fairness conditions of the working conception. Ultimately, regardless of the merits of the piecemeal and change of scope solutions the following bargaining-based account provides a solution that is both unified and more explanatory than either.

6.4 Bargaining-Based Exploitation

In chapter 2 we saw that a generalisation of the fairness principles solved the problem for Steiner’s account presented by the trust funder case. Steiner claimed that B was exploited just in case she received less from transaction ϕ than she would have in another transaction ψ which was identical to ϕ , save for the absence of the antecedent rights violation. However,

the trust funder case showed this account to be insufficiently general since, in some cases, had the rights violation not occurred, there would be no transaction. Steiner’s operationalisation of the claim in RIGHTS that an antecedent rights violation ‘makes one party worse off than they would be’ needed to be broadened. (F2) and (F3) provided this broader account. By generalising the domain of the fairness conditions to (F2) and (F3), I shifted *what* B receives less of from ‘benefits from transacting’ to ‘preferences satisfied.’ B can’t receive less benefit from transacting if she doesn’t, in fact, transact. But in these cases her preferences may still be less well satisfied.

The apple cart case reveals that the degree to which B’s preferences are satisfied is also an inappropriate currency for grounding the counterfactual by which the fairness of a transaction is judged. In particular, the inadequacy of (F2) and (F3) stems from their insensitivity to changes in the economic environment. When a rights violation alters the economic environment in which B transacts, obliging those in the new environment to satisfy her preferences to the same degree that they would have been satisfied in the previous environment is too demanding. Yet, the demandingness of the obligation does not adversely affect A—provided the transaction is non-obligatory he can (and will in apple cart cases) choose to not transact rather than to fairly (by (F2) and (F3)) transact. In this case, the demandingness of the obligation adversely affects only B, the very person the obligation is designed to protect. What we require is a currency for the counterfactual that is sensitive to changes in the economic environment.

In the section that follows I propose a further generalisation of (F2) and (F3) that shifts the currency of the counterfactual from preferences to bargaining power. The first subsection (6.4.1) shows how this generalisation solves the problem of obligation and considers its merits. In 6.4.2 I fill in some of the details of the account, arguing that an approach to bargaining power derived from the status quo point is appropriate for understanding exploitation. Subsection 6.4.3 revisits the question from the previous chapter (subsection 5.3.2) about the need for micro-level principles of fairness like FMV. Here I present a number of bargaining solutions and argue that neither these solutions nor FMV is necessary for a transaction to be fair.

6.4.1 The Bargaining-Based Solution

The appropriate currency for capturing the apple cart case is bargaining power. While A may not be able to offer B a price that C would have offered her (\$1), he can treat her as if her bargaining power had not been compromised by the antecedent rights violation. By bargaining as if B’s bargaining power were as strong as it would have been absent the rights violation, B does no worse than she would have done *in a transaction with A* had the rights violation not occurred.

Suppose C’s apple cart was not stolen and he turned up at the market. However, for some reason,¹³ B decides to shop at A’s stall instead of C’s. In this scenario, B’s bargaining power is strong. In order to obtain a better price from A she may always threaten to shop at C’s booth instead. In this case, B would be able to bargain to a price of \$1.51, an outcome that is far better for her than a no-transaction scenario. At \$1.51 and above A gains by transacting with B. By appealing to bargaining power, we ground the counterfactual that determines the fairness of the transaction in a currency that is sensitive to changes in the economic

¹³Since we assume that A and C offer identical products, B’s choice to shop at A’s stall would be irrational, but this detail does not affect the point being made here.

constraints, thus making A's obligation (when he chooses to transact) less demanding. This generalisation grants B not the same outcome that she would have received had the rights violation not occurred, but the same *power* to influence the outcome.

In order to solve the problem of obligation, we must replace (F2) and (F3) with (F2') and (F3').¹⁴ With the revision in place, A exploits B in transaction ϕ just in case:

(F1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that

(F2') reduces B's bargaining power, and

(F3') the degree to which B's preferences are satisfied in ϕ is less than it would be had B's bargaining power not been reduced.

(G4') A's egoistic preferences:

(a) are better satisfied in ϕ because B's bargaining power is limited and

(b) if (F1)–(F3') also hold for A, then the reduction in A's bargaining power is less than the reduction in B's bargaining power, and finally

(A5') A truly believes that (F1)–(G4') hold for ϕ .

When A truly believes that (F1)–(G4') are true of ϕ , then in order to not exploit B, A must treat B *as if* (F2') and (F3') do not hold. This modified account has some attractive properties: it is a more general version of my previous account and also of Steiner's original account, it places a stronger rhetorical emphasis on the defect in A's attitude that is necessary for exploitation, and finally, it goes some way towards incorporating a notion of power in my account of exploitation. However, like the other solutions, the bargaining power-based solution faces problems.

Strengths

Generality. First, (F2') and (F3') are more general than counterfactuals based on either degree of preference satisfaction or transactions since these are simply special cases of (F2') and (F3'). Suppose that A's break even point was \$0.50 (rather than \$1.50) in the apple cart case. In this case, if A restored B's *bargaining power* to the pre-rights violation level, B would obtain an apple from A for \$1. Why? If A treated B as if B's bargaining power were not reduced by the rights violation, then he would treat her as if she could also transact with C. Since C charges \$1 for an apple, A would be forced to charge B no more than \$1 as well. Since there is no other seller offering apples for less than \$1, any threat B might make in the hope of further reducing A's price would not be credible. Thus the exchange would take place at \$1. In this case, B's preferences would be satisfied to the same degree that they would have been, had the rights violation not occurred.

Similarly, we might imagine a case where A normally charges \$1 per apple, but after someone steals B's wallet, A charges her \$2 (in the form of an IOU). This is a case in which Steiner's original counterfactual works: in order to not exploit B, the price A charges in the post-rights violation transaction should be equal to the price he would charge in the pre-rights violation transaction. And this, in turn, is equivalent to treating B as if her bargaining

¹⁴The wording of (G4) and (A5) has also changed to reflect the shift to bargaining power and the reference to the revised fairness conditions, so these must also be updated to (G4' and (A5'). Nevertheless, these conditions have not changed *substantively*.

power had not decreased due to the rights violation. The new conditions (F2') and (F3') can be applied to cases in which both Steiner's original account and the conditions (F2) and (F3) apply, but as the trust funder and apple cart cases show, the opposite does not hold. The new conditions are more general than either of the previous operationalisations of RIGHTS.

Attitudes. Second, the rhetorical flavour of (F2') and (F3') emphasises some of Sample's and Goodin's concerns about exploitation that are not as apparent in the 'degree of preference satisfaction' approach. Unlike Roemer and Steiner, Sample and Goodin place a strong emphasis on the fact that exploitation involves a mistreatment of B by A that is deeper than a mere failure to ensure a fair outcome. Exploitation for the latter two involves a defective *attitude*. Goodin writes, "Sleeping sentries are guilty of all sorts of moral delicts, but exploitation is not one of them. It is . . . the *flagrant* violation of the norm, rather than of a mere failure to discharge the responsibility it lays upon us. This is a fundamentally process-based notion."¹⁵ Similarly, Sample argues, "exploitation is interaction for the sake advantage *in a way* that fails to respect the value of our interactor—that is, in a way that degrades her."¹⁶ While we have already seen that it is a mistake to claim that the use of these attitudes to drive a hard bargain is *sufficient* for exploitation, they are, contrary to Steiner's and Roemer's accounts, necessary for exploitation. And while they are formally captured by the advantage condition (A5), the modified fairness conditions (F2') and (F3') better emphasise the way in which A's failure to constrain his unjustly strengthened bargaining power is a manifestation of these attitudes. Fairness in the sense of (F2') and (F3') requires A to relinquish or rein in a form of power he has over B. When A is aware (A5') of his unjust power advantage, but does not alter his behaviour, he shows an attitude of disrespect towards B and a willingness to 'push her to the limit'.¹⁷

Power. Finally, critics of Roemer's PR account (as well as Roemer himself) have argued that because Roemer's criterion "is neutral with respect to social features of states, appealing only to their material features", it is insufficient to capture our ordinary understanding of exploitation.¹⁸ As we saw, Roemer argues for the inclusion of a dominance condition that would require the coalition S' to be in a relation of dominance to S, but he leaves the condition undefined. Recently, Veneziani has argued that accounts of exploitation must "go beyond Roemer to incorporate a notion of power" and that "a notion of power, or dominance, is an essential part of exploitive relations."¹⁹ While he references Wood's (1995) emphasis on vulnerability and Paul Warren's (1997) claim that exploitation should focus on unequal power, Veneziani does not himself offer an account of how power might be incorporated into a theory of exploitation. Yet, the emphasis on bargaining power in (F2') and (F3') seems to do just this. Characterising the fairness of the transaction in terms of B's unjustly weak bargaining power, and conversely A's unjustly *strong* bargaining power, successfully incorporates a notion of power into an account of exploitation. Further, as we saw, (F2') and (F3') do not define a *particular* material outcome as the non-exploitative baseline; rather,

¹⁵Goodin (1987, pp.188–89, emphasis added)

¹⁶Sample (2003, p.84, emphasis added)

¹⁷The English 'exploitation' may be translated into two words in German, either as '*Ausnutzen*' or '*Ausbeuten*'. According to the Duden dictionary (Clark, 1999), the latter of these may mean "*systematisch nutzen, ausschöpfen*", that is, 'to use systematically, to the last drop'. Rendering exploitation as the failure of A to curtail his unjust power advantage in the sense of (F2') and (F3') captures this 'to the last drop' connotation nicely.

¹⁸Sensat (1984, p.24)

¹⁹Veneziani (2013)

they allow the non-exploitative baseline to vary along with the economic constraints the agents face in post-rights violation transactions.

Objections

Despite these three benefits, certain features of the bargaining power-based solution seem unsatisfactory.

Isn't B Still Wronged? The first objection is that, although weakening A's obligations allows B to avoid 'no transaction' scenarios, in cases like the apple cart case the weaker obligation means that she still does worse (her preferences are less well satisfied) than she would have because of an antecedent rights violation. Surely there is *something* morally amiss when, for example, an antecedent rights violation forces B to pay \$1.50, rather than \$1, for her apple.

The difference between the original (F2) and (F3) and the modified versions is that the modified versions offer an account of fairness in deontically imperfect situations, while the original versions operate in deontically perfect situations. Given that A and B find themselves in a deontically *imperfect* world, a fair transaction is one in which A ignores B's unjust disadvantage. However, as the objection indicates, B will not receive as much from A when he restores her bargaining power as she would if A faced the obligation implied by the original fairness conditions. Yet, as we saw, placing this obligation on A means that in some cases *he* does worse than he would have because of a past rights violation.²⁰ This too is unacceptable. If it is unfair for B to 'pay' for the antecedent rights violation, it must also be unfair for A to 'pay'. Thus, there are two kinds of fairness relevant to cases like the apple cart case. These senses coincide in 'ordinary' (non-apple cart) cases.

The first sense, denoted by the original (F2) and (F3), is a deontically perfect sense of fairness. The transaction is fair in a deontically perfect sense when B's welfare (judged by the degree to which her subjective preferences are satisfied) is as it would have been had the rights violation not occurred. The second sense, denoted by the modified (F2') and (F3'), is a deontically *imperfect* sense of fairness. The transaction is fair in this sense when A fulfils his obligation to treat B's bargaining power as if it had not been reduced by the rights violation. A exploits B only if the transaction is not fair in the deontically *imperfect* sense.

But why should the deontically *imperfect* sense of fairness be the sense that is necessary for exploitation? The answer is that if we are committed to exploitation as both a strongly moralised concept *and* a relational concept, then we must be committed to the claim that it is unfair in the deontically imperfect sense. Although some accounts of exploitation, such as Roemer's, are non-relational, defining only the predicates 'is exploited' and 'is an exploiter',²¹ ordinarily when we say that B is exploited, we mean it as a relation: B is exploited *by* a person or group. While the claim that A exploits B because the transaction is not fair in the deontically perfect sense preserves the relational property of exploitation, we cannot also claim that A acts wrongly when he does not offer B terms that are fair in the deontically perfect sense because, as we saw, A cannot be obliged to offer B these terms. Since I have argued that exploitation *is* a strongly moralised concept, the sense of fairness relevant to exploitation is the deontically imperfect sense.

²⁰This happens when absent an antecedent rights violation B would have received a price that is loss-making for A.

²¹Elster (1982, p.363)

There is something to the objection that B does worse than she ‘should’ even if she is not exploited according to the revised conditions. B remains entitled to the satisfaction of her preferences that would have obtained absent the rights violation, but it is not A who is obliged to ensure that B’s preferences are satisfied to the appropriate degree. Rather, I suggest, it is the original rights violator.²² Thus, like many of the other forms of transaction we have considered, transactions that are not exploitative according to the revised conditions, but that are still unfair in the deontically perfect sense, are normatively significant. They give us *pro tanto* reasons for action. But, I have argued, since we ordinarily consider exploitation to be strongly moralised and relational, these transactions should not be labelled exploitation.

Wasn’t \$1.51 Unfair? In my discussion of the piecemeal solution I argued that A’s offer of \$1.51 was unacceptable because requiring that A make a profit in order to transact with B means that A gains from B’s disadvantage. I claimed that A’s refusal to trade without profit holds B hostage to A’s demands. These claims still hold. It is not the *price* that is a problem for the piecemeal solution, but the justification. In the bargaining power solution it is B, not A, who dictates the \$1.51 price by using her restored bargaining power to bargain A down from the original \$2 asking price. Of course, she cannot push a rational A past his indifference point of \$1.50, but restoring bargaining power to B allows her a degree of autonomy not found in the piecemeal solution.

Additionally, it is important to note two differences between the bargaining power solution and the piecemeal solution. First, the bargaining power solution does not argue for a particular fair *price*. Rather, it focuses on B’s *capacity* to obtain a price, on her opportunity for advantage. Thus, the ‘fair price’ is not a fixed price or proportion, but depends on the transactors’ relative bargaining power before the rights violation. Second, unlike the two-part piecemeal approach, the bargaining power solution provides a unified benchmark for fair transactions that applies to both apple cart and non-apple cart cases.

Can A Strategically Avoid Exploitation? Because A is not obliged to ensure that B’s preferences are satisfied to the degree that they would have been satisfied absent a rights violation when doing so would harm him, A has an *incentive* to ensure that satisfying B’s preferences *would* harm him. That is, transactors in A’s position in the apple cart case have an incentive to ensure their cost of production is high if they want to avoid the charge of exploitation. Of course, in many cases there are also countervailing incentives to ensure their cost of production is *low*. In the apple cart case, the presence of the cheaper seller C in the market (pre-apple cart theft) provided A with a reason to try to cut the cost of production of his apples from \$1.50 to \$1. However, in more complex situations companies may have a number of inputs over which the cost of production is distributed: machinery, assembly line salaries, managerial salaries, utility expenses, etc. Firms may increase managerial salaries while claiming that they cannot afford to pay their line workers non-exploitative prices without making a loss. As with the question of negligent exploitation, we have two possible responses to such strategic allocation of costs. We may add a rider to the fairness conditions that explicitly includes such behaviour within the scope of exploitation, or

²²Of course, the original rights violator may be long dead. I am unsure what should be said about such cases where he has disappeared (one way or another). But, at any rate, I have argued that whether the transaction is fair in the deontically perfect sense has no bearing on whether it is exploitative. Therefore, determining who is obliged to ensure B is treated fairly in the deontically perfect sense, should the rights violator be dead, is beyond the scope of the present project. However, for a related analysis of the distribution of responsibility for injustice within a group, see Steiner (1994, pp.266-69).

we may exclude the behaviour. And as is the case with negligence, excluding the behaviour from the scope of exploitation does not entail that the behaviour is not wrongful, only that it should not be labelled exploitation. My response to this problem is therefore the same as the negligence case: when individuals attempt to avoid exploitation through the manipulation of their costs, we should allow that when they do manage to avoid satisfying (F2') and (F3') they avoid exploitation. But, importantly, they do not avoid wrongdoing.

What Is Bargaining Power? The final objection is more difficult to respond to than the first two. The primary problem with using bargaining power as the currency for the counterfactual in the fairness conditions is that it does not offer an appropriately informative explanation of exploitation. Absent a detailed account of what, exactly, bargaining power *is*, the claim that A exploits B when he, *inter alia*, takes advantage of her unjustly weak bargaining power is analogous to the claim that A exploits B when he takes advantage of her unjust *disadvantage*. But this much we already know. Indeed, Steiner prefaces his (1987) account with the claim that “[a]n exploitative exchange in the most general (and non-theory-relative) sense, is one which issues from an unfair distribution of bargaining power.”²³ He goes on to offer an account of what unfair bargaining power is within the framework of stark-budget liberalism: “An unfair distribution of bargaining power is a distribution of valuable possessions that contravenes a classical liberal set of property rights.”²⁴ And, expanding further, he offers his familiar transaction-based fairness conditions. B is exploited when selling a product in a bilateral and voluntary transaction only if (1) someone would have bid higher than A for B’s product, but (2) the reason he did not “is that someone’s rights had been previously violated.”²⁵

There are two points to take away from Steiner’s claims. The first is that a strict interpretation of the stark-budget liberalism framework within which Steiner is working is incompatible with the idea of power as an opportunity for advantage. Steiner claims that for stark-budget liberalism “power consists in valuable possessions and unfairness consists in rights-violations.”²⁶ The conceptions of exploitation based on the modified (F2') and (F3') do not interpret power as valuable possessions, but rather as an individual’s ability to secure preference-satisfying outcomes for herself. Thus, (F2' and F3' seem to be incompatible with a strict interpretation of power according to stark-budget liberalism. If these modified versions are required to explain the natural claim that A exploits B in the trust funder and apple cart cases, then it seems that an account of exploitation within the framework of stark-budget liberalism cannot cover many cases of exploitation. This is not a significant defeat for libertarianism in general though, for left-libertarian frameworks that extend the scope of property rights are, it would seem, compatible with even the bargaining power-based account of exploitation. Although the bargaining power solution may not conform with stark-budget libertarianism, it is not *illiberal*.

The second point is that in order to offer an informative account of exploitation, fairness conditions based on bargaining power must also provide an account of what bargaining power. In order to incorporate the anomalies of the trust funder and apple cart case, we find ourselves, rather curiously, back where Steiner began: bargaining power. We have no particular explanation of how, exactly, bargaining power should be understood. The need

²³Steiner (1987, p.133)

²⁴Steiner (1987, p.133)

²⁵Steiner (1987, p.133)

²⁶Steiner (1987, p.133)

for an account of bargaining power does not mean that the bargaining power-based solution fails. Rather, it means that without an account of bargaining power it is incomplete. The following section provides an account that can meet this objection and, ultimately, ground a bargaining-based account of exploitation.

6.4.2 Characterising Bargaining Power

Before considering how bargaining power is best characterised in the context of exploitation, we must first take a slight detour to discuss bargaining problems and their solutions. Below I outline two-person cooperative games, or bargaining problems, and one prominent solution concept, the Nash bargaining solution.

Bargaining Problems

Transactions between two individuals that are consensual and strictly Pareto-improving are a form of two-person cooperative game. Since both transactors gain from transacting when compared to a no-transaction baseline, each has an incentive to agree upon *some* distribution of the gains from trade, but, since both are also concerned to maximise their own welfare, each has an incentive to bargain for a distribution that favours their own interests. Bargaining problems are a form of two-person cooperative games. Bargaining problems between A and B can be characterised by each individual's status quo²⁷ (0), a set of Pareto-improving points S ²⁸ called the feasible set, each individual's utility function (which shapes the feasible set), and a solution concept, which is a function mapping the feasible set and non-agreement point to a (usually unique and Pareto optimal) point in the feasible set S .²⁹ Players begin with utilities at the status quo point and bargain to a Pareto superior point in the feasible set. The point that rational bargainers would end up at is the outcome, which is given by a particular solution concept.

So, in the figure below, A and B each have linear utility functions, the status quo is denoted by 0 , and the feasible set of Pareto-improving transactions is denoted by S . By acting jointly (trading), A and B can achieve any point in S , but, since points on the red frontier line jointly dominate³⁰ all points in the interior, if A and B are rational they will focus only on trades corresponding to points on the frontier.

Some game theorists, most notably Von Neumann and Morgenstern, believe that further restrictions cannot be placed on the outcome of bargaining problems. They argue that all we can say about rational bargainers is that they will choose a point on the frontier; *which* point is chosen depends upon the players' psychologies.³¹ However, Nash (1950), Zeuthen (1930), Harsanyi (1956), and Barry (1991) all argue for a particular solution to bargaining problems, specifically, the Nash bargaining solution. These writers argue that rational bargainers will end up at a point on the frontier characterised by the following two conditions (where : μ_A

²⁷Also called the non-agreement point.

²⁸A few definitions for the technically inclined reader. The set S is convex and compact. A set is convex (in Euclidian space) iff a line drawn between any two points in the set lies entirely within the set. A set is compact (again, in Euclidian space) iff it is closed and bounded. A set is closed iff it contains all of its limit points (it contains its border) and a set is bounded iff it has a border on all sides (and is therefore not infinite). In other words, the set S is like a 'normal' shape: it is closed on all sides by a border and the points on the borders are included in the set.

²⁹Kalai and Smorodinsky (1975, p.514)

³⁰A point (μ_A, μ_B) in S is 'jointly dominated' by another (μ'_A, μ'_B) if both $\mu'_A \geq \mu_A$ and $\mu'_B \geq \mu_B$ (Luce and Raiffa, 1989, p.117)

³¹Luce and Raiffa (1989, p.118)

is A's utility, $\mu_A N$ is A's utility at the Nash solution point, μ_B is B's utility, and $\mu_B N$ is B's utility at the Nash solution point).³²

(NC1) $\mu_A > 0$ and $\mu_B > 0$

(NC2) $\mu_A N \mu_B N \geq \mu_A \mu_B$, for all (μ_A, μ_B) in S , such that, $\mu_A \geq 0$ and $\mu_B \geq 0$

Here the status quo point is assigned a utility of 0. Condition (NC1) stipulates that the point that the bargainers agree upon will be a strict Pareto improvement—better for A ($\mu_A > 0$) and better for B ($\mu_B > 0$)—over the status quo point (0). Condition (NC2) stipulates that the Nash solution is the unique point in the set of all of these Pareto improving points that maximises the product of the utilities of A and B for a particular division. That is, the Nash solution is the division at the points $\mu_A N$ and $\mu_B N$ and the product of these two points ($\mu_A N * \mu_B N$) is greater than the product of any other two Pareto improving points in the feasible set.³³

As an example consider the three points x , n , and y in figure 6.1 below. At point x , A's utility gain (when compared to o) is 0.6 and B's is 0.4.³⁴ The product of these utilities is 0.24. The product of the utilities at point y (0.3, 0.7) is 0.21. The product of the utilities at the Nash solution, point n is 0.25 and is greater than the product of the utilities at x and y and indeed, any other point in the feasible set S . Another way of putting the same point is that the solution picks the point on the Pareto frontier that maximises the area of a rectangle drawn from that point to the origin. In figure 6.1, this point is, again, that denoted by n .

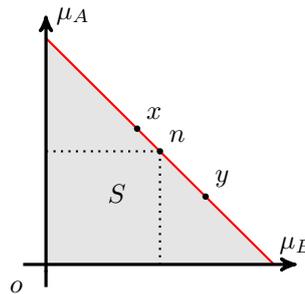


Figure 6.1: A basic bargaining problem.

There are three strong justifications for the Nash bargaining solution. Nash shows that his is the only solution concept that satisfies utility invariance, Pareto optimality, symmetry, and independence conditions³⁵ that, he argues, ideal rational agents would select, were they to select axioms to guide distribution in a bargaining problem. Since rational agents not involved in a bargain would select these conditions, they may be thought of as impartially rational. Thus, Nash argues that the Nash solution provides a 'fair' division that reflects the reasonable expectations of rational bargainers. Zeuthen and Harsanyi argue that if A and B each propose a distribution, the player whose relative loss is smaller should concede. Through a series of such alternating demands, Zeuthen and Harsanyi show that the players

³²Luce and Raiffa (1989, p.126)

³³Barry (1991, p.14)

³⁴Note that these points are not necessarily interpersonally comperable. The claims that A gains 0.6 and B gains 0.4 need not be made with reference to a shared scale of utility.

³⁵Nash argues that we want solutions to the bargaining problem that are invariant with respect to utility transformation (utility invariance), yield the same results when the game places A and B in completely symmetric situations (symmetry), and, if the feasible set is expanded in a way that leaves the status quo unchanged, then the bargaining solution is either unchanged or lies in the expanded part of the set (independence of irrelevant alternatives). Point n is the only point on the Pareto frontier (Pareto optimality) that satisfies these conditions.

will converge upon n . This justification emphasises Nash’s claim that the solution gives each individual as much as they can rationally expect to gain. Finally, Barry argues that if A and B hire a third party, or arbitrator, to estimate their respective utility functions, “each side will refuse to accept any arbitrator whose decision it expects to be less favourable than the outcome of direct bargaining ... [A]n arbitrator who is acceptable to both sides must be one whose decisions approximate the Nash solution.”³⁶ Successful arbitrators are those who, over time, give distributions equivalent to the Nash solution. Erratic arbitrators are a gamble that will be vetoed by both parties, and arbitrators using solutions other than the Nash solution will be vetoed by one of the parties. However, since arbitration takes time and resources, and rational agents are utility maximisers, they will (when possible) dispense with the arbitrator and agree upon the division recommended by the Nash solution.

Although the Nash solution is perhaps the most widely accepted bargaining solution, it is not the only solution. Many other solutions to bargaining problems exist; among the most prominent are the Kalai-Smorodinsky (1975) solution, which maintains the ratios of maximum gain for each agent, and the egalitarian bargaining solution, in which each receives the utility denoted by the maximum point of equal coordinates. The figure below shows the Nash (n), Kalai-Smorodinsky (k), and egalitarian (e) bargaining solutions.³⁷ In the following subsection I discuss each of these three bargaining solutions as possible micro-level principles of fair transaction. However, for now I will return to the task at hand, the characterisation of bargaining power.

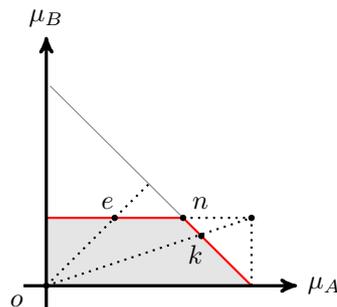


Figure 6.2: Three bargaining solutions.

Bargaining Power

Bargaining power, in a pre-theoretical sense, refers to an individual’s relative advantage and influence over the other when trading, bargaining, debating, etc. Consider a classic bargaining problem in which A and B must decide on a division of \$100 in an hour or neither receives anything. Intuitively, A has more bargaining power than B if A knows that B *needs* \$20 (to pay a debt, for example) even if she *wants* as much of the \$100 as she can get. When B’s utility for money is non-linear in this way and A knows this, then A has more bargaining power than B. Since B needs \$20 to pay her debt, she attaches a higher utility to the first \$20 than to the next \$20. A can push hard for a distribution close to (\$20, \$80),³⁸ knowing that B wants to ensure she receives at least \$20. Of course, non-linear utility is not always a disadvantage in bargaining. We may imagine that B needs \$80 to pay her debt. If receiving \$20 or \$50 is nearly as bad as receiving nothing, and A knows this about B, then B has the

³⁶Barry (1991, pp.26–27)

³⁷Note that in bargaining problems like that depicted in figure 6.1 these solutions coincide.

³⁸Payoffs are listed as (A, B).

upper hand. Therefore, bargaining power is shaped, at least in part, by the agents' utility functions.

Bargaining power is also shaped by the location of each individual's status quo point. Suppose again that A and B must decide on the division of \$100. As before, if they do not reach a decision in an hour A receives nothing, but this time B must *pay* (a third party) \$5. Since the consequences of non-agreement are costly for B, she will be more likely to accept less than half of the \$100. Her bargaining power is reduced. And, conversely, if B is guaranteed \$50 if no agreement is reached, then she will push for more than half of the \$100. In the latter case, provided A and B both have linear utilities for money, the Nash bargaining solution is (\$25, \$75). Graphically, shifts in the non-agreement point may be represented by moving the axis of one or both players (from o to o_2). When B's non-agreement point is shifted up, the Nash³⁹ bargaining solution shifts from n_1 to n_2 , a distribution that favours B, as shown in figure 6.3.

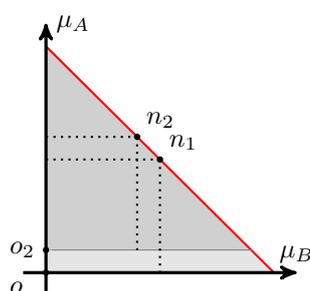


Figure 6.3: Shifting the status quo.

As we've seen, bargaining solutions are functions that map the non-agreement point and feasible set into a point in the feasible set (which, for our purposes, is unique and on the Pareto frontier). It is not surprising then that the outcome of a bargaining problem may be influenced by the non-agreement point and utility functions since the shape of S is determined by these variables. It remains to determine which of the two (or if both) offer the best characterisation of bargaining power in the context of exploitation.

Bargaining Power and the Status Quo

Let's consider what we are looking for. The bargaining-based versions of the fairness conditions, (F2') and (F3'), provide a particular operationalisation of the RIGHTS principle of fair transaction that is general enough to account for both the trust funder and apple cart cases. It is superior to both Steiner's original operationalisation in terms of counterfactual transactions and the degree of preference satisfaction-based ((F2) and (F3)) conditions. However, without an analysis of what, exactly, bargaining power *is*, the revised fairness conditions cannot provide an appropriately informative explanation of exploitation. Since the revised conditions are an operationalisation of the RIGHTS principle, the account of bargaining power that we are looking for must be appropriate to the rights principle. RIGHTS claims that a transaction is fair when there is no antecedent property rights violation that makes one party worse off than they would be (absent the violation) in a transaction.

In most cases,⁴⁰ a person's 'being worse off than they would be because of a property rights violation' involves a shift in their status quo point rather than a shift in their preferences.

³⁹And in this particular case, the other two solution concepts.

⁴⁰In chapter 2 footnote 16, I note that liberals may want to distinguish between an agent's 'true' preferences and her 'adaptive' or post-rights violation preferences, by for example taking 'meta-preferences'—preferences

Indeed, my example of preference change in the bargaining power discussion may be re-interpreted as a status quo alteration to bargaining power. It may be that B always has non-linear preferences for money when she is indebted to someone. In this case, we may say that her bargaining power is weak because her status quo is in a non-linear region of her utility function. If she had no debt, then her status quo point would have shifted upwards to a point at which she might be expected to have a linear utility for money. True cases of preference change that cannot be re-described in this way are, arguably, rare. We may operationalise RIGHTS by claiming that a transaction is unfair to B if and only if an antecedent property rights violation reduces her bargaining power by altering her status quo point in such a way that the proportion of the gains B receives from the transaction is lower than it would have been had the rights violation not occurred. This gives us the following revised (and final) characterisation of exploitation:

A exploits B in the bilateral, voluntary and Pareto-improving transaction ϕ just in case:

- (F1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that
- (F2'') reduces B's bargaining power *by altering her status quo point*, and
- (F3'') the degree to which B's preferences are satisfied in ϕ is less than it would be had B's bargaining power not been reduced (by the alteration of her status quo point).
- (G4'') A's egoistic preferences:
 - (a) are better satisfied in ϕ because B's bargaining power is limited by her altered status quo point and
 - (b) if (F1)–(F3'') also hold for A, then the reduction in A's bargaining power is less than the reduction in B's bargaining power, and finally
- (A5'') A truly believes that (F1)–(G4'') hold for ϕ .

In the apple cart case, when we require A to act as if B's bargaining power had not been compromised, we require him to restore her status quo point to a pre-property rights violation level (but, because the revised fairness conditions are in terms of bargaining power, we do not require A to extend to B the same offer that C would have). We may call this pre-rights violation level B's moral baseline (m_B). This shift in the actual status quo point to the moral baseline is shown on the left in figure 6.4 below.

Condition (G4''b) says that if *both* A and B are disadvantaged by a property rights violation (not necessarily the same one), then A exploits B only if (*inter alia*) the negative effect of a rights violation on B's status quo is greater than the negative effect of a rights violation on A's status quo. In the graph on the right in figure 6.4, both A and B are disadvantageously affected by a past property rights violation; however, A is *more* disadvantaged than B because the distance between his moral and actual baselines is greater. Thus in the situation depicted by that graph, when B does not constrain her unfair advantage, if the other conditions for exploitation are also met, she exploits A when she transacts with him.

But is the restoration of the moral status quo baseline all that is required for a transaction to be fair? Nash argued that his bargaining solution provided a 'fair' solution; in the figures in this chapter (6.2 excluded) both the egalitarian solution and the Kalai-Smorodinsky

over preference orderings—as given. Thus, in some cases a property rights violation might also impact an individual's preferences. However, this is a further 'epicycle' that may be applied to the basic case of unjustly diminished bargaining power which I believe applies most directly to the agent's status quo point.

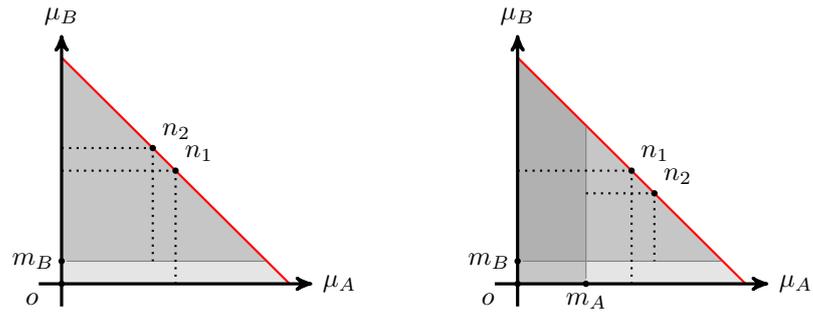


Figure 6.4: Adjusting for property rights violations.

solution coincide with the Nash solution. Yet the final conditions for bargaining-based exploitation do not require that transactions take place at these points. According to the final characterisation of exploitation, *any* point in the grey region of figure 6.5 is a fair (and thus non-exploitative) transaction. This *includes* point n_1 , the point at which rational agents would have coincided if no adjustments were made for unjustly diminished status quo points. It seems that the final characterisation of exploitation is too weak. Perhaps additional, micro-level principles of fairness are required.

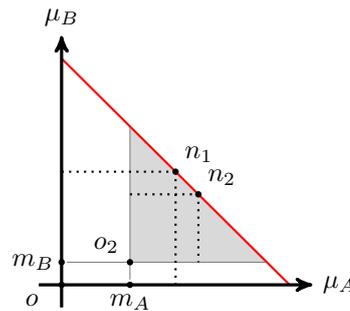


Figure 6.5: Many fair transactions.

6.4.3 Revisiting Micro-Level Fairness

In chapter 5, I considered Wertheimer’s FMV principle of fair transaction. FMV is a *micro*-level principle of fair transaction. As a micro-level principle, it operates against a background of institutional justice. If a situation is unjust on a macro-level, FMV provides neither necessary nor sufficient conditions for fair transaction. However, according to Wertheimer, when background conditions *are* fair, then a fair transaction is one that takes place at the “fair market value[.]... the price that an informed and unpressured seller would receive from an informed and unpressured buyer if the [good] were sold on the market.”⁴¹ In my discussion of FMV I offered two critiques. The first was that FMV does not itself embody any deep principles of fairness—or at least Wertheimer gives us no reason to believe that this is the case. Rather, FMV acts as a proxy for other more fundamental values that gives us a *practical* baseline for determining whether a transaction is fair. Since FMV is not itself a plausible moral principle, we should look to the deeper principles that it instantiates in order to find a plausible micro-level principle. Second, I argued that it is unclear that there is any need for micro-level principles of justice, since it is not straightforwardly clear that RIGHTS

⁴¹Wertheimer (1996, p.230)

(a *macro*-level principle) is insufficient to guarantee a transaction's fairness. The ability of RIGHTS to handle both fraud (with the addition of a no-fraud clause) and coercion provided further evidence for the cogency of the claim that it was sufficient.

However, at the end of subsection 6.4.2 we saw that the status quo approach to the bargaining power operationalisation of RIGHTS does not rule out some outcomes that might have occurred with no adjustment of the status quo. Agents may select the distribution denoted by n_1 in figure 6.5 both before and after adjustments for background unfairness (in the form of adjustments to the status quo point) are made. There is some reason to think that RIGHTS is insufficient to ensure that a transaction is fair. Further micro-level principles of fairness may be necessary in order to ensure the sufficiency of the RIGHTS principle.

Unfortunately for FMV, its lack of justification as a principle of fair transaction means that it is not a good candidate for a micro-level principle. However, above we saw three alternatives that might serve as a micro-level principle: the Nash, Kalai-Smorodinsky, and egalitarian bargaining solutions. Each of these solutions provides justification for its position, but the Nash solution in particular provides the kind of justification that FMV lacked: an axiomatic appeal to deeper principles of fairness. We may call macro-level principles of justice based on (any) rational bargaining solutions the BARGAINING principle of fairness. Wertheimer is aware of the appeal of BARGAINING and considers it a possible alternative to his own FMV account. However, he argues that it faces two problems: first, the multitude of bargaining solutions suggests that "bargaining theorists do not agree as to whether there is a determinate answer to the [division problem]" and, second, "even if there is a solution to what rational self-interest-maximi[s]ing bargainers would do," it is not clear that "this solution provides any guidance with respect to the criteria for a *fair* transaction."⁴²

We can allay the first of Wertheimer's concerns by considering bargaining problems in which the egalitarian, Kalai-Smorodinsky, and Nash solutions coincide, such as the one presented in figure 6.5. If the distribution denoted by n_2 (after the status quo point is moved) is not the only fair distributions, then no particular version of BARGAINING will be necessary for fair transaction. If n_2 is the only fair solution, then we can consider the merits of individual bargaining solutions. However, considering these additional bargaining problems will not be necessary since, I will argue, Wertheimer's second concern is correct. BARGAINING does not provide plausible principles of fair transaction for two reasons.

First, the alleged insufficiency of the status quo approach to the bargaining power operationalisation of RIGHTS is mistaken. The insufficiency charge stems from the fact that n_1 may be a fair distribution both before and after the status quo adjustments have been made. The thought is that if the adjustments make no difference in the outcome of the transaction, then further constraints on permissible outcomes are required. But this comparison of 'before adjustment' and 'after adjustment' outcomes involves a change not only in the status quo point, but also the bargaining strategies of the agents. If the agents select n_1 before adjustment, they act rationally according to the solution concepts. If they select n_1 after the adjustment, they act irrationally. When we fix the bargaining strategies by requiring the agents to be rational in both bargaining problems, the agents will choose n_1 before the restrictions and n_2 after the restrictions. The restrictions *do* make a difference in the distribution of gains. Similarly, we may consider the problem from the perspective of transactions that are fair according to (F2') and (F3'). Consider figure 6.5 and suppose that no antecedent rights violations have occurred. In this case, according to the conditions,

⁴²Wertheimer (1996, emphasis added)

any distribution within the feasible set is fair. Now suppose two rights violations occur that require shifting the status quo point to o_2 . Here the set of fair transactions is restricted to the grey region. Again, the restrictions *do* make a difference in the distribution of gains. The restrictions fail to make a difference only if the bargaining strategies of the agents are allowed to shift along with the shift in the status quo.

Second, although the various bargaining solutions provide an account of how agents ought to divide the gains from transacting if they are *rational*, they do not provide an account of how agents ought to divide the gains if they are *moral*. Consider the five distributions shown at points n , w , x , y , and z in figure 6.6. As before, distribution n is the distribution that all three bargaining solutions select as the ‘fair’ distribution. However, suppose that A and B end up not at n , but at w . Perhaps they agree to trade 10 apples for \$10 dollars and then they each burn half of their gains. If we also suppose that, whatever else they want, A and B want more money and apples, this behaviour is *irrational*. But it does not seem *unfair*.

The proponent of the BARGAINING principle may concede that this example shows that fairness does diverge from the maximising assumptions of rationality without granting the RIGHTS proponent’s claim that the entire shaded region of figure 6.6 is fair. If fairness is concerned with a distribution rather than a particular point, then all that matters is that the proportion of the distribution denoted by n is maintained. According to a refined version of BARGAINING, any distribution on the line segment \overline{on} is fair.

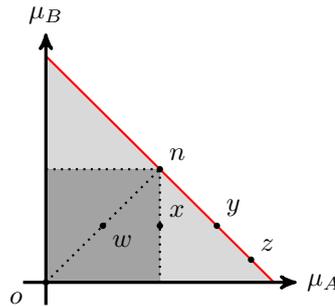


Figure 6.6: Fairness and Rationality.

But now suppose that A and B end up at the distribution denoted by x . Perhaps B burns half of her apples but A keeps all of his money. Is this *unfair*? If B acts irrationally, does fairness require that A do so as well? Perhaps if the two have agreed to both burn their gains, then A has a duty to keep his promise, but short of such an obligation, it does not seem that there is anything *unfair* in A’s keeping his money. The distribution is unequal, but it is not unfair. Suppose the proponent of RIGHTS concedes this point. She may make this further concession without extending fair distributions to the entire shaded region by arguing for a further weakening of BARGAINING according to which any distribution within the *dark* grey box is fair. If fairness does not require maximisation, then if either agent does not maximise, the transaction is still fair, provided the other agent does not take more than they would have received had the other player acted as a rational maximiser.

At this point the defender of BARGAINING has conceded half of the area to the defender of RIGHTS. As a micro-level supplement to RIGHTS, this weakened version of BARGAINING is far less restrictive than either the original BARGAINING or Wertheimer’s FMV. But the erosion of BARGAINING is not quite complete. Suppose now that B does not burn half of her share of the gain, but *gives*⁴³ it to A so that the distribution denoted by y obtains. Is this

⁴³The BARGAINING proponent may ask why B would give A a portion of the gain that she could have kept

distribution *unfair*? As is the case with x , it is *unequal*, but it does not seem unfair. And, since the location of y along the frontier depends on the amount that B gives to A, if y is fair, then the distribution denoted by z , and indeed the entire shaded area, must also be fair.

It would be odd for B to simply give A some of what she herself wanted. A far more likely case is one in which A and B are both equally proficient bargainers. Neither is perfectly rational, but both easily satisfy the conditions of moral agency (no one will be *giving away* what they want). However, A is a better bargainer and A and B end up at a distribution on the Pareto frontier very close to n , but that slightly favours A, perhaps (48, 52). Surely *this* is an unfair distribution.

Here the RIGHTS proponent will respond with a (by now) familiar trichotomy: B's having bargained poorly is either his own fault, someone else's fault, or no one's ('nature's') fault. If it is his own fault, and his ability to bargain has not been unfairly diminished, then there is nothing unfair about the distribution. If it is someone else's fault that his bargaining abilities have been diminished, then he should be compensated for his unjust disadvantage. The outcome is *unfair* but not because of the distribution, but rather because we have not adjusted B's bargaining power to take into account the past injustice. Finally, if it is nature's fault—a random fluke that A receives more—the RIGHTS proponent has a number of options available. He may claim that the existence of insurance schemes turns this fluke into something for which B can be held responsible, or he may claim that inequalities brought about by nature should be evenly distributed amongst everyone and, therefore, that the distribution is unfair. It is unclear what the *best* response to such a scenario is, but it is clear that the proponent of RIGHTS has at least one response. All differences in bargaining power and background conditions can be folded into the status quo point.⁴⁴

That a distribution of the gains from bargaining is fair is not determined by a particular bargaining solution (as in point n), nor by a range of outcomes that preserve the proportion of that distribution (as in line segment \overline{on}), nor even by a range of outcomes in which each individual receives no more than they would have if the other had acted fully rationally (as in the dark grey region). Fairness is a process-based notion. In the context of bargaining problems, provided each individual's bargaining power is just, any outcome is *fair*, even if it is not *rational*. Therefore, we do not need to supplement RIGHTS (and its bargaining power-based operationalisation) with a micro-level principle of fair transaction. Rather than providing us with criteria for judging the fairness of outcomes, bargaining solutions provide us with criteria for judging the rationality of outcomes.

6.5 Chapter 6 Summary

Chapter 5 concluded with a working conception of exploitation that added an 'awareness' condition, (A5) ((A5') and (A5'')) in later formulations), to my revised version of Steinerian

for herself. Surely this behaviour is *irrational*. This is true. When rational bargainers would have ended up at n , B's bargaining to z would be strongly irrational behaviour. It makes sense to ask why she would do such a thing. Yet, insofar as we can find a *reason* for her behaviour, we find that it is not irrational, but rather that the bargaining problem has been mis-characterised. Suppose B does not care at all about apples and this is why she gives a large portion of her apples to A. In this case, her utility function in the bargaining problem is mistaken and it does not, in fact, have the shape we see in figure 6.6. When the problem is re-described, we may find that B acted rationally.

⁴⁴Christian List has suggested that, at least under certain moral theories, the disadvantages of the would-be exploited could be caused not by an antecedent rights violation, but rather by (for example) Rawlsian unfairness. In my discussion of RIGHTS I have tried to provide some reasons for thinking this is not the case. However, it is also possible to replace my fairness condition (F1) with the following condition, which is more theory-neutral: (F1 _{n}) there exists background unfairness that... The phrase 'background unfairness' in (F1 _{n}) can be specified by a particular moral background theory.

exploitation. In section 6.2 I assessed this working characterisation. I defended its exclusion of both negligent and mutual exploitation. I also showed how it solves the pervasiveness and responsibility problems that plagued Steiner's original account. There I also introduced a problem for the working conception, the problem of obligation. In certain cases, the fairness conditions of the working conception of exploitation imply overly demanding obligations for A that lead to outcomes that are worse for both A and B (and better for no one). In section 6.3 I considered two unsuccessful solutions to the problem of obligation. I argued that the first—the piecemeal solution—fails to provide proper justification for a secondary obligation that applies in 'apple cart case' scenarios, and the second—a change of scope—is too restrictive since it introduces restrictions on the working characterisation that exclude cases that we would ordinarily describe as exploitative. In place of these solutions I proposed a bargaining-based solution in section 6.4. I argued that by generalising the working conception's fairness conditions to (F2') and (F3') the problem of obligation can be solved. Here I also argued that in the context of exploitation, bargaining power should be determined with reference to the status quo point of bargaining problems. The fairness conditions are updated to (F2'') and (F3'') to reflect this interpretation. Finally, I provided a counterargument to the claim that RIGHTS is insufficient to ensure that a transaction is fair.

At the end of chapter 6 we have a full and informative characterisation of exploitation that is grounded in the ordinary language account of exploitation as 'taking unfair advantage'. I have argued that a transaction is unfair to B when (F1) there exists an unrectified property rights violation at some point in a history of exchanges leading to the current transaction that (F2'') reduces B's bargaining power by altering her status quo point in such a way that (F3'') her preferences are less well satisfied than they would have been absent the rights violation. A exploits B when (G4'') his own egoistic preferences are better satisfied by transacting with B (to a greater degree than B's are) and (A5'') he truly believes this is the case. In the following chapter I argue that this conception of exploitation is the best solution to the paradox of exploitation.

Chapter 7

Conclusion: Open Questions

7.1 Chapter 7 Introduction

In the first half of the thesis I presented the paradox of exploitation and considered three kinds of accounts that offer a solution to the paradox. In the second half, I developed my own solution to this paradox, arguing for the conception of exploitation characterised by (F1)–(A5''). However, two questions remain open. One concerns the justification of my account; the other is concerned with the paradox of exploitation. I address these questions in the final section of the thesis, section 7.3. First, section 7.2 provides a summary of the claims I have made throughout.

7.2 Thesis Summary

In the first part of the thesis I argued against existing accounts of exploitation provided by Steiner (chapter 2), Roemer (chapter 3), and Sample and Goodin (chapter 4). In the second part, I developed my own account, first arguing for a working conception in chapter 5 and then a refinement in chapter 6. While this semi-dialectical presentation obviates the logic of discovery and builds rhetorical support for my account, it somewhat obscures the core structure of my arguments. In this summary I reconstruct the arguments in favour of my three main claims.

C1: Exploitation is a negatively moralised concept that involves taking unfair advantage. Most accounts of exploitation begin with the commonly accepted claim that exploitation is a negatively moralised concept; however, Allen Wood disagrees. In section 5.2 I consider Wood's arguments against the orthodox account. Wood claims that both Nietzsche's use and 'wrong because exploitative' claims provide two arguments against the orthodox approach, but I argue that neither of Wood's arguments are conclusive. Nietzsche, I claim, uses 'exploitation' in an inverted comma sense. Wood argues that 'wrong because exploitative' claims cannot be substantive if exploitation is negatively moralised, but since these claims are substantive, exploitation is not negatively moralised. I agree with Wood's first premise, but reject the second and therefore come to the opposite conclusion. Since Wood cannot provide any argument that speaks conclusively to his heterodox position, and there are good reasons to think the orthodox position is correct, I conclude that exploitation is a negatively moralised concept.

In chapter 4 I consider two accounts (provided by Sample and Goodin) that allow for non-juridical exploitation. In section 4.4 I show that such accounts are either self-frustrating

or incomplete. In order to avoid the dilemma, Goodin and Sample must either base their accounts of exploitation on responsibility-sensitive theories of justice or prohibit prudent risks (subsection 4.4.4). Since the latter is unacceptable, I argue that successful accounts of exploitation must involve unfairness and be based on responsibility-sensitive theories of justice.

Both Steiner and Roemer provide juridical accounts that are potentially responsibility-sensitive. However, both theories fail to incorporate strong advantage taking conditions and are therefore overextended. Steiner's account faces two pervasiveness problems, diachronic and synchronic pervasiveness (section 2.4), as well as the problem of moral responsibility (section 2.3). Both problems can be solved with the addition of an advantage condition. Roemer's basic account is also severely overextended, but many of his problems can be solved by the addition of the dependence condition and the restriction of his account to arm's-length economic transactions (3.4). The dominance condition promises to limit the account to an acceptable scope, but this condition is not fully formulated. A properly formulated dominance condition, I suggest, will restrict Roemer's account in the same way that the addition of an advantage condition would.

From these three arguments I conclude that exploitation is negatively moralised and involves transactions in which two necessary conditions are present: advantage taking and unfairness.

C2: RIGHTS provides the best account of exploitation's unfairness. Throughout the thesis, but primarily in chapter 5, I consider five accounts of the unfairness involved in exploitation including USE (section 4.2 and subsection 5.3.1), CONTRIBUTION (subsections 3.3.4 and 5.3.1), RIGHTS (throughout chapters 2, 5, and 6), FMV (subsections 5.3.2 and 6.4.3), and BARGAINING (subsection 6.4.3). I argue that the micro-level principles of BARGAINING and FMV are not necessary to ensure a transaction's fairness, and USE is only plausible when interpreted in such a way that it becomes equivalent to RIGHTS or CONTRIBUTION. I suggest that certain versions of the remaining two principles are coextensive, but the ease of use afforded by RIGHTS and its ability to capture difficult cases (5.3.3 and 5.3.4) makes this the preferred principle of fair transaction.

RIGHTS claims a transaction is fair iff there does not exist an unrectified property rights violation in the history of exchanges leading to a transfer ϕ that makes one party worse off than they would be. While the first portion of the principle is straightforward (and provides the first condition for exploitation), 'worse off than they should be' must be operationalised. I consider four interpretations. The first, in section 2.2, is Steiner's implicit claim that B is worse off than she would have been when there is a counterfactual transaction between A and B in which the violation is not present and B values this transaction over the exploitative one. However, I argue that in some cases there is no appropriate counterfactual transaction. Thus, RIGHTS should be reinterpreted in terms of preference satisfaction. In subsection 2.2.2 I claim that a transaction is unfair to B when: (F2) it makes at least one element of B's original (and finite) feasible set, γ_B , unfeasible, leaving her with a new feasible set, δ_B (where $\delta_B \subsetneq \gamma_B$) and (F3) B prefers the unfeasible option in γ to her most preferred option in δ .

In chapter 6 I revisit these fairness conditions and show that they are subject to the problem of obligation (6.2.3). In place of (F2) and (F3) I suggest the following fairness conditions: A transaction is unfair to B iff there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that: (F2') reduces B's bargaining

power, and (F3') the degree to which B's preferences are satisfied in ϕ is less than it would be had B's bargaining power not been reduced. Although this solution to the problem of obligation improves on those outlined in section 6.3, it leaves the concept of bargaining power unanalysed. In section 6.4.2 I provide an account of bargaining power based on the status quo point in bargaining problems and revise the fairness conditions for the final time. In their final form, they read as follows. A transaction is unfair to B iff there exists an unrectified property rights violation at some point in the history of exchanges leading to the transaction that:

(F2'') reduces B's bargaining power *by altering her status quo point*, and

(F3'') the degree to which B's preferences are satisfied in ϕ is less than it would be had B's bargaining power not been reduced (by the alteration of her status quo point).

C3: Conditions (A5'') and (G4'') together provide the best account of exploitation's advantage taking. Since the first characterisation of Steiner's position in 2.2, the working characterisation of exploitation has included a gain condition. A exploits B only if A gains from the transaction's unfairness and does so to a greater degree than B. Although this condition ensures that advantage taking in the sense outlined by Elster occurs (both A's and B's preferences are better satisfied by transacting), it does not restrict the scope of exploitation enough to deal with the pervasiveness issues that occur in Steiner's and Roemer's accounts. In order to handle these problems a stronger sense of taking advantage must be used. In section 5.4 I consider a number of advantage conditions before arguing in favour of (A5''), according to which, in the context of exploitation, A takes advantage of B just in case A truly believes that he engages in a transaction that is unfair according to RIGHTS. Elster's Pareto conditions are too weak, and stronger intentionality-based (A4) and opportunistic conditions (6.3.2) are too strong. Thus, I conclude that transaction ϕ is a case of exploitation just in case it is in the domain of bilateral voluntary and strictly Pareto-improving transactions and the following five conditions hold:

(F1) there exists an unrectified property rights violation at some point in the history of exchanges leading to ϕ that

(F2'') reduces B's bargaining power *by altering her status quo point*, and

(F3'') the degree to which B's preferences are satisfied in ϕ is less than it would be had B's bargaining power not been reduced (by the alteration of her status quo point).

(G4'') A's egoistic preferences:

(a) are better satisfied in ϕ because B's bargaining power is limited and

(b) if (F1)–(F3'') also hold for A, then the reduction in A's bargaining power is less than the reduction in B's bargaining power, and finally

(A5'') A truly believes that (F1)–(G4'') hold for ϕ .

7.3 Two Open Questions

While arguing for my own account over the alternatives presented in chapters 2, 3, and 4 I have created a hydra of sorts, introducing many types of normatively interesting transaction—unfair transactions, coercive exploitation, spiteful transaction, etc.—that are closely related

to the kind of transaction I have called ‘exploitation’. Of course, in the process of introducing each individual transaction, I defended its exclusion from the scope of exploitation. Yet, taken together, these various kinds of transaction seem to make a collective case against my own characterisation. Many of the normative transactions characterised by conditions that do not make the cut in my final characterisation of exploitation *are* involved in cases that we would ordinarily call exploitative. What, then, licences my exclusion of these cases from the scope of exploitation? I will take up this open question in subsection 7.3.1.

A second open question concerns the paradox of exploitation itself. Recall that the paradox claims that engaging in consensual and Pareto-improving transactions cannot be impermissible *and* that exploitative transactions that are Pareto-improving and consensual are impermissible. Although my shift to a bargaining power-based operationalisation of RIGHTS solved the original problem of obligation, a related issue remains. Suppose that in the apple cart case, contrary to the obligations implied by (F2'') and (F3''), A does not act as if B's bargaining power had not been reduced: he sells B apples for \$1.75. Suppose also that he knows this is an unfair transaction. Since the conditions (F1)–(A5'') are satisfied, A exploits B and, therefore, acts wrongly. Although it is clear that B has a *right* to a better price, is it really true—as the rejection of non-worseness seems to imply—that it would have been morally better for A and B to not transact at all? Even if *we* think so, A and B are likely to disagree. I revisit the paradox of exploitation in subsection 7.3.2.

7.3.1 Normative Transactions

In a review of Sample's (2003) book, Wertheimer makes an important point, which I quote at length:

I am not particularly sympathetic to Sample's conceptual 'essentialism', to explaining what exploitation 'is'. Some argue that if a transaction is beneficial to B and if B consents, then it cannot be exploitative even if the transaction is unfair or degrading to B. Sample and I disagree. But does it matter who's right? For even if we were to concede the linguistic point, we would still have to ask what should be said about unfair or degrading mutually beneficial consensual transactions, whether and why they are wrong, and whether or why they should be prohibited. Those are the important questions. Sample suggests, for example, that to say that a transaction is exploitative when it is (merely) unfair does not account for the 'badness' of exploitation. . . If Sample wants to say we should not refer to unfair but not degrading transactions as exploitative, I don't have any strong objection, although I think she is wrong as a matter of common usage. The important task is to identify the way in which transactions can be wrong and what we should do about them, not what words we should use.¹

In my introduction to the first chapter I noted that exploitation is a 'family resemblance-like' concept that we use in various ways, in various contexts. If this is really true, it is unlikely that there is any one correct characterisation that captures all of ordinary use. The most important task is not the attempt to identify what, precisely, exploitation *is*, but rather, as Wertheimer argues, "to identify the way in which transactions can be wrong and what we should do about them."²

¹Wertheimer (2007, p.259)

²Wertheimer (2007, p.259)

This thesis has identified a number of ways in which transactions can be morally amiss (wrongful or morally bad). However, some of these normative transactions are not exploitation (again to quote Wertheimer) “as a matter of common usage”.³ Although restricting the label ‘exploitation’ to cases satisfying only (F1)–(A5’), excludes some cases that we might ordinarily call exploitation, it does so in order to ensure a consistency in our use. Suppose case α satisfies (F1)–(G4’’) and we ordinarily call it a case of exploitation, but β also satisfies (F1)–(G4’’) and we do not ordinarily call it exploitation. Here, our uses of ‘exploitation’ across these two cases are inconsistent with one another. Although fit with ordinary use is a strong desideratum, so is consistency. Any plausible analysis of exploitation must satisfy both. We can, as I have argued, allow these transactions to be normatively important without also labelling them exploitation; we can simply call them ‘unfair transactions’.

Conversely, if we adopt a more restrictive notion of exploitation, where the concept is satisfied by conditions (F1)–(A6)⁴, we exclude cases that we would ordinarily call exploitation. Suppose A accidentally happens upon someone with whom he can engage in a transactions satisfying (F1)–(A5’). We would not, I suggest, refrain from calling this exploitation simply because A did not *plan* to encounter B. That he did not constrain his actions towards her when he did transact with her is sufficient to say that he exploited her. Yet, refraining from labelling transactions satisfying (F1)–(A6) ‘exploitation’ does not mean that we must deny their normative importance. For example, premeditated exploitations may carry more moral weight than ‘accidental’ exploitations. So, just as (F1)–(G4’’) fail to ensure the consistency of ordinary use, (F1)–(A6) fail to assure a maximal fit with ordinary use. The conception characterised by (F1)–(A5’’) maximises fit with ordinary use while ensuring consistent use.

Interestingly, some paradigmatic cases of exploitation involve conditions that are far stronger than (A5’). In many sweatshop cases, multinational corporations in A’s position *actively pursue* unfair gain. The property which (perhaps) most exercises us about sweatshops is not merely that these groups gain unfairly from labourers, nor that they are aware that they do so, but that they actively and intentionally seek out those who are unjustly impoverished and from whom they stand to gain the most. Corporations like Nike or Wal-Mart do not accidentally hire impoverished workers in Nicaragua; they have a presence in these countries only because they know they can exploit. These cases of ‘wilful exploitation’ involve an additional ‘premeditation’ condition (A7) that makes the moral weight of their wrongdoing greater, but which is unnecessary for (mere) exploitation.

As I mentioned in the introduction to this section, my conception of exploitation must be defended not only against the competing conceptions provided in chapters 2, 3, and 4, but also against conceptions that are closely related to the one characterised by (F1)–(A5’). Table A.1 in appendix A summarises 16 conditions and 17 forms of normative transaction that have been discussed throughout the thesis and makes a case for the identification of a core set of characterisations that are candidates for exploitative transactions. There I argue that transactions characterised by (F1), (F2’), (F3’), (F4’), and (A5’’) are the best candidates for exploitative transactions.

7.3.2 Solving the Paradox

In chapter 1 I presented the following ‘paradox of exploitation’:

³Wertheimer (2007, p.259)

⁴Recall, (A6): A modifies the terms of his transaction *in response to* his true belief that (F1)–(G4’’) obtain.

The Paradox of Exploitation

1. In transaction, ϕ ,
 - (a) (Consent) to ϕ rather than to not ϕ is consensual and
 - (b) (Pareto) to ϕ rather than to not ϕ is a strict Pareto improvement.
2. (Impermissibility) ϕ -ing is impermissible.
3. (Non-worseness as permissibility) It cannot be impermissible to ϕ , provided:
 - (a) (Consent) to ϕ rather than to not ϕ is consensual and
 - (b) (Pareto) to ϕ rather than to not ϕ is a strict Pareto improvement.

The paradox may be solved by abandoning any of the three premises, but throughout the thesis I have really only focused on the third ‘non-worseness’ premise. I have assumed that the transactions we are considering are truly consensual and Pareto-improving and that they are also wrong. Thus the various solutions to the paradox have presented different ways in which the non-worseness premise might be false. They explain why transacting on certain terms can be said to be impermissible.

One way of reading the paradox is to see the non-worseness claim as a (supposedly) comprehensive list of properties that are necessary for the permissibility of transactions that is quite obviously far too short. The accounts of exploitation that I have considered have given us many other properties that may serve to make transactions impermissible. Transactions may be impermissible if they involve trade in non-commodifiable goods, contain terms that deviate too far from the cost of production (or the fair market value, or the labour value of the commodity), and if they involve gain from past rights violations. Clearly, as a list of impermissibility generating properties, the non-worseness claim is false. An account of the exploitativeness of a transaction may be based on any of these other impermissibility-generating properties; I have argued that the best solution to the paradox of exploitation, when the non-worseness premise is read in this way, is my RIGHTS-based account. Transactions may be impermissible—and indeed, wrongful—when they involve taking unfair advantage, as characterised by conditions (F1)–(A5’).

However, there is another, more nuanced, way that we can read the non-worseness premise. It may be seen as offering not a list of impermissibility-generating properties, but rather a moral *ordering* of these properties. Read in this way, the claim is not that transactions involve wrongdoing only if they are either non-consensual or harmful. Rather, the claim is that regardless of what *other* wrong-making properties are present in a transaction, if it is consensual and Pareto-improving, then it is morally better than not transacting *even if* it is also morally worse than another non-exploitative transaction. Read in this way, the non-worseness claim is true, or so I argue. Consider the following case (alluded to in this chapter’s introduction):

C17: The revised apple cart.

As before, suppose that every Saturday C drives his apple cart to B’s village and sells apples, and B usually buys 10 apples for \$1 each. One Saturday C’s apple cart is stolen, making A, who (lacking economies of scale) charges \$2 per apple, B’s only alternative for apples. Finally, also as before, A’s break even price for an apple is \$1.50, and the most B would be willing to pay for each apple is \$2.50. Now suppose that although A is aware of the circumstances of B’s situation, he sells B an apple for \$1.75.

We can ask three questions about this revised case:

Q1 Does A act wrongly?

Q2 Does A exploit B?

Q3 Would it have been *morally* better for A to not transact?

The answer to the first question should, by now, be obvious. A performs a wrongful action because he is aware that he gains from a transaction in which B's bargaining power is unjustly weakened from an antecedent rights violation. All conditions in my final characterisation of exploitation are satisfied, so the transaction is impermissible, wrongful and exploitative. In order for the transaction to be fair, and therefore non-exploitative, A should have acted as if B's bargaining power had not been reduced by the theft of C's cart. In this case, A should have offered B a price of \$1.51. That is to say, it would have been morally better for A to offer the lower price: $\$1.51 \succ_{\text{morally}} \1.75 .

The third question asks whether it would also be morally better for A to not transact with B than to transact with B for \$1.75. It is difficult to provide any reason this should be the case. Both A and B would prefer transacting at \$1.75 to not transacting, both consent to this transaction, both gain from this transaction, and no one else is harmed by this transaction. Further, in claiming that it is morally better for A to transact, we do not commit ourselves to the claim that in so doing A avoids wrongdoing. By transacting at \$1.75, A still acts wrongfully by exploiting B. When A transacts at \$1.75 he does something that is morally better than not transacting and morally worse than transacting at the lower price: $\$1.51 \succ_{\text{morally}} \$1.75 \succ_{\text{morally}} \neg \text{transacting}$.

If (on the second reading) the non-worseness condition is true, there is no paradox to resolve. We might say that ϕ -ing (transacting at \$1.75) is wrong because transacting at \$1.51 is morally better than transacting at \$1.75. And we can make this claim while still claiming that transacting at \$1.75 is morally better than not transacting. Thus, on the second reading of the non-worseness premise, the paradox is *dissolved*.

Unfortunately, this dissolution seems to present us with a new problem. If transacting at an unfair price is wrong *because* it is morally worse than transacting at a fair price, then it seems that not transacting must also be wrong *because* it is morally worse than transacting at either price. However, if this is true, then the conditional obligation that seems to be present in most cases of exploitation is false. In most cases—sweatshops, for example—we think A is obliged to provide B with a fair price *only when* A chooses to transact with B; A is not (except in special cases) *obliged* to transact with B. But if it is wrong to perform an action because it is morally worse than another, it is not only wrong for A to transact unfairly, it is wrong for A to fail to transact. In the apple cart case we may say if it is wrong to exploit B because it is morally worse than offering her a fair price, then it must also be wrong to fail to transact with her. But this seems clearly mistaken: A is not obliged to sell B an apple. She is not starving or in desperate need of help; she's just buying the week's groceries.

What we need is a solution that does not equate 'morally worse' with an action's being wrongful, so that we can explain why it is wrong to choose the morally worse option when A chooses to transact, but it is not wrong to choose the morally worse option when he does not choose to transact. This is just what my account of exploitation provides. A's transacting at an unfair price is not 'wrongful *because* morally worse', it is wrongful because he is *aware* that he gains unfairly from a past rights violation. This is also what makes it morally worse to transact at an unfair price. Thus, moral worseness does not cause wrongfulness, rather,

they have a common cause: A’s awareness of his unfair gain from a rights violation.

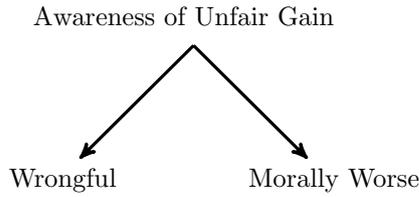


Figure 7.1: The wrongfulness of exploitation.

On the other hand, not transacting is morally worse than transacting at an unfair price because both A’s and B’s subjective preferences are less well satisfied. They would both prefer to transact—even at an unfair price—than to not transact. That is, not transacting is morally worse because there is a loss of welfare. But the fact that there is a loss of welfare from not transacting *does not* mean that A and B act wrongly (though they may act irrationally). There is no reason to think that A and B are morally required to perform actions (such as apple buying and selling) that provide them welfare, at least according to my account of exploitation.⁵



Figure 7.2: The permissibility of non-transaction.

Conclusion

Where does this leave us? On the first reading of the non-worseness claim, my account of exploitation provides a reason that transactions may be wrongful even when they are Pareto-improving and consensual. In this case, the non-worseness premise is false, and the paradox of exploitation is resolved. On the second reading, the non-worseness claim *and* the other premises are true. The paradox is not resolved, but *dissolved* since it is true both that not transacting is morally worse than transacting *and* that transacting at an exploitative price is wrong. However, in order to explain why transacting at an exploitative price is wrong when not transacting is permissible, we need an account of why the two differ. That is, we require an account of what *is* wrong with exploitation that is *not* wrong with not transacting. This is provided by my account of exploitation. So on one reading of the non-worseness premise, my account *resolves* the paradox of exploitation and on the other it *dissolves* it.

Although my account dissolves the *logical* inconsistency between the second reading of the non-worseness premise and the wrongfulness premise, it does not remove the residual *feeling* that there is something odd about exploitation. When we end up with a welfare-based moral ordering like $\$1.51 \succ_{\text{morally}} \$1.75 \succ_{\text{morally}} \neg \text{transacting}$, we still find ourselves in a situation in which, by moving from a non-transaction point to a transaction at \$1.75, we increase welfare for A and B and harm no one *and yet* we move from a morally permissible state to a morally impermissible state.

⁵Thanks to Sebastian Köehler for discussing these points with me.

In offering a bargaining power-based interpretation of the (relative) harm to B caused by a rights violation, my account blends consequentialist and non-consequentialist principles. Such blends often produce counterintuitive—or at least puzzling—moral obligations. Fortunately for my account, there is nothing very puzzling about transactions that take place at what I have argued are fair prices.

When A and B transact at \$1.51 in the apple cart case, B is not worse off (in terms of bargaining power) than she would be had the rights violation not occurred, and both A and B are better than they would have been had they not transacted. The satisfaction of the Pareto condition allays welfaristic concerns and A's restoration of B's bargaining power fulfils his non-consequentialist rectificatory duties. While B's preferences would be *even better* satisfied if A were to offer her the price she would have received absent a rights violation, A is not obliged to fully rectify the effects of C's injustice when doing so would harm him. In this case, A, not B, would be exploited: welfare constrains rights. Similarly, although A's preferences could be *even better* satisfied if he offered B a price of \$1.75, to do so would be to profit unfairly from her disadvantage: rights constrain welfare. The two overarching—and opposing—moral methods fit together to offer a clear account of a fair price.

There is something puzzling about transactions that take place at *exploitative* prices. My account of exploitation explains not only why transacting unfairly is worse than transacting fairly, but also why doing so is *wrong*. Consequentialist (welfare-based) considerations tell us why *not* transacting is morally worse than unfairly transacting. But if not transacting is morally worse than unfairly transacting and unfairly transacting is wrong, not transacting is morally worse than a wrongful action. And yet, common sense morality tells us that not transacting is not itself wrong.

Above I solved this problem by explaining that the wrongfulness of the exploitative transaction is caused by A's awareness of the fact that he unfairly transacts—a property that is not present in the non-transaction case. But this solution to the problem does not eliminate the perplexing fact that we can do wrong while also doing better for ourselves and others. As Joan Robinson famously wrote, “the misery of being exploited by capitalists is nothing compared to the misery of not being exploited at all.”⁶

⁶Robinson (1962, p.45)

Appendix A

Kinds of Normative Transaction

Table A.1: Normative transactions.

	P	C	T	F1	F2	F2''	F3	F3''	G4''a	G4''b	A5''	A4	A6	A7	O	Threat
Unjust Cases																
Theft (harmful)			1													
Systemic Unfairness	✓	✓		✓		✓		✓								
The Working Conception	✓	✓	✓	✓	✓		✓		✓	✓	✓					
Unfair Transactions																
Harmful Transactions			✓	✓				✓	✓	✓						
Non-Gainful Unfair Transactions	✓	✓	✓	✓				✓								
Attempted Exploitation	✓	✓	✓	✓				✓							2	
Mutually Bad Transactions	✓	✓	✓	✓				✓	✓							
Unfairly Gainful Transactions	✓	✓	✓	✓				✓	✓	✓						
Negligent Transactions	✓	✓	✓	✓				✓	✓	✓					3	
Core Transactions																
Harmful Exploitation			✓	✓				✓	✓	✓	✓					
Spiteful Transactions	✓	✓	✓	✓				✓	4	✓	✓					
Exploitative Transactions	✓	✓	✓	✓				✓	✓	✓	✓					
Mutual Exploitation	✓	✓	✓	✓		5		6	✓	✓	✓					
Intentional Exploitation	✓	✓	✓	✓		✓		✓	✓	✓	✓				✓	
Weighty Wrongdoing																
Opportunism	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓			
Wilful Exploitation	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		
Coercive Exploitation	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓

¹ Thefts are unilateral, not bilateral, transactions.

² Here A intends to exploit B, but fails to satisfy the gain conditions.

³ In negligent transactions, (A5) is not satisfied, and A is culpable for its non-satisfaction.

⁴ In spiteful transactions, part of A's gain is from the satisfaction of non-egoistic preferences.

⁵ In cases of mutual exploitation, F2' is satisfied in one dimension for A and another for B.

⁶ In cases of mutual exploitation, F3' is satisfied in one dimension for A and another for B.

Strict Pareto Improvement (P) Both A's and B's preferences are better satisfied by transacting.

Consent (C) Both A and B consent to the transaction.

Transaction (T) There is a transaction—an exchange of goods—between A and B.

Property Rights Violation (F1) There exists an unrectified property rights violation at some point in a history of exchanges leading to transaction ϕ .

Feasible Set Restriction (F2) The property rights violation makes at least one element of B's original (and finite) feasible set, γ_B , unfeasible, leaving her with a new feasible set, δ_B (where $\delta_B \subsetneq \gamma_B$).

Bargaining Power Restriction (F2'') The property rights violation reduces B's bargaining power by altering her status quo point.

Welfare Loss from Set Restriction (F3) B prefers the unfeasible option in set γ to her most preferred option in set δ .

Welfare Loss from Power Restriction (F3'') The degree to which B's preferences are satisfied in ϕ is less than it would be had B's bargaining power not been reduced.

Gain (G4''a) A's egoistic preferences are better satisfied in ϕ because B's bargaining power is limited.

Relative Disadvantage (G4''b) If A is also disadvantaged (according to either (F1), (F2), and (F3), or (F1), (F2''), and (F3'')), then A's disadvantage is less than B's disadvantage.

Intentionality (A4) A intentionally engages in a transaction in which (F1)–(G4'') hold.

Awareness (A5'') A is aware—truly believes—that (F1)–(G4'') hold for the transaction ϕ .

Opportunism (A6) A modifies the terms of his transaction in response to his true belief that (F1)–(G4'') hold for the transaction ϕ .

Premeditation (A7) A intentionally seeks out transactions in which (F1)–(G4'') hold.

Obligatory (O) A is obliged to engage in transaction ϕ with B.

Threat A proposes to make B worse off relative to the transaction that he is obliged engage in with B.

A.1 Core Transactions

In table A.1 I identify a 'core' of five normative transactions that correspond most directly to the ordinary use of 'exploitation'. Each of the transactions in the core satisfy the five conditions (F1)–(A5), but only the form of transaction I label 'exploitation' satisfies *exactly* these conditions, with the exception of 'harmful exploitation', which rejects the domain conditions. Throughout the thesis I have assumed that exploitation is in the domain of Pareto-improving consensual transactions, but as the debate between Sample and Wertheimer shows, not everyone shares this assumption. And indeed, unlike the inclusion of (A5'') or the fairness conditions (F2'') and (F3''), little of importance hangs in the balance. The main reason for the inclusion of the Pareto and consent conditions is that it is (relatively) easy to explain what has gone wrong when someone is *harmed*.

Spiteful exploitation excludes the restriction of A's gain in (G4'') to egoistic preferences. As far as *use* is concerned, we ordinarily think of the gains from exploitation as material—but not even psychological—but normatively speaking, when A transacts unfairly with B because he wants to see her preferences less well satisfied, he acts just as wrongfully as he does when he transacts in order to advance his own material gain.

The existence of mutual exploitation depends on the truth of the claim that there exist two *incommensurable* domains of value. While I tend to think that this thesis is false, Sample disagrees. If incommensurable values do exist, then a greater number of transactions qualify as exploitative than we (or I) thought. Here something important does hang in the balance, because determining the wrongfulness of mutually exploitative transactions is complicated. For example, if A happens to care more about values in dimension α and B cares more about values in dimension β , then if A is exploited in β and B in α , neither may care much, provided they gain sufficiently in the dimensions of value that they *do* care about. And, if neither is particularly bothered by the 'exploitation' involved in mutual exploitation, then it is difficult to explain how either is harmed without appealing to objective values. But as we saw in subsection 4.3.3 the defence of particular lists of objective values is controversial. Nevertheless, 'mutual exploitation' does satisfy the five conditions (F1)–(A5'').

Finally, intentional exploitation and a weaker alternative, knowing exploitation, adopt a stronger form of constraint than the awareness condition (A5''). These conditions require not only that A is aware (F1)–(G4'') hold, but also that A either intend to engage in or knowingly engage in a transaction in which (F1)–(G4'') are satisfied. The difference in scope between exploitative and intentionally exploitative transactions is not likely to be great. Those counterexamples that show the intentional condition (or the knowledge condition) to be too strong are fairly far-fetched and have—for the most part—only *theoretical* traction. Thus, in *practice* it makes little difference whether we adopt an ‘awareness’ or ‘intentionality’ condition since nearly all real-life cases will satisfy the stronger (A4) condition.

A.2 Other Normative Transactions

Transactions falling into the category of unjust transactions do not satisfy any of the advantage taking conditions (A4), (A5''), (A6), or (A7). Since these transactions do not involve ‘taking advantage’ in a sense that goes beyond mere Pareto improvement (found in (A1) and (A2)) they are not in the set of core transactions.¹ On the other side, transactions falling into the category of weighty wrongdoing are transactions that are not only exploitative, but also involve additional forms of wrongdoing, such as coercion (as in the drowning woman case), or premeditated and repeated exploitation (as in most sweatshop cases). Although some of these transactions are exploitative, they include extra conditions that are not necessary for exploitation.² Finally, cases falling into the ‘unjust’ category are not linked by a common theme and do not even necessarily involve bilateral transactions. Here I have placed theft, systemic unfairness of the kind characterised by Roemer’s (PR1) and (PR2) conditions³, and the original working conception of exploitation.

Thus, although I have argued that Pareto-improving, consensual and bilateral transactions characterised by the conditions (F1)–(A5'') are the best candidates for exploitative transactions, it is possible to slightly weaken (by removing the non-egoistic rider from (G4'')) or abandoning the Pareto and consent domain conditions), or strengthen (by including the stronger (A4) or allowing multiple domains of value) the conception without seriously altering the conception I have provided. Alterations that go beyond these changes either depart too significantly from ordinary use by overly restricting the concept or by labelling transactions that are merely unfair ‘exploitation’.

¹Harmful transactions are bilateral transactions that are much like harmful exploitation, but unlike harmful exploitation, in the case of harmful transactions, A is not aware that he harms B. Non-gainful unfair transactions are Pareto-improving, consensual transactions in which a past rights violation means that B’s preferences are less well satisfied than they would have been in the transaction, but A does not gain from B’s being worse off (even though he may still gain from the transaction). Transactions that are cases of attempted exploitation are similar to non-gainful unfair transactions, except that A also *intends* to exploit B, but since A fails to gain from the transaction, the exploitation attempt fails. Mutually bad transactions are the somewhat hard-to-imagine cases in which both A and B do better than they would have because conditions (F1)–(F3'') are satisfied, but neither does better relative to the other. Unfairly gainful transactions are transactions that are similar to those Steiner calls exploitation (but here I represent them with the revised fairness conditions). They are identical to exploitation except they lack condition (A5''). Finally negligent transactions are Pareto-improving consensual transactions in which A ought to truly believe that (F1)–(G4'') obtain, but he does not.

²Opportunism involves cases of exploitation in which, in addition to the normal conditions of exploitation, A also intends to exploit B, and he modifies the terms of his transaction in response to his true belief that B is unjustly disadvantaged. Willful exploitation involves cases in which, in addition to the normal conditions of exploitation, A intentionally seeks out exploitative transactions. Finally, coercive exploitation involves cases in which A both exploits and coerces B.

³Roemer’s account is not strictly characterised by what I have called ‘systemic unfairness’, but it is compatible with it when the withdrawal conditions are defined by (F2'') and (F3'').

Bibliography

- Arneson, R. (1992). Commodification and commercial surrogacy. *Philosophy & Public Affairs*, 21(2): 132–64.
- Barry, B. (1991). *Theories of justice*. Berkeley: University of California Press.
- Bronfenbrenner, M. (1973). Samuelson, Marx, and their latest critics. *Journal of Economic Literature*, 11(1): 58–63.
- Cartwright, N. (2009). If no capacities then no credible worlds. *Erkenntnis*, 70(1): 45–58.
- Child, J. (1994). Can libertarianism sustain a fraud standard? *Ethics*, 104(4): 722–38.
- Clark, M. et. al., eds. (1999). *Oxford-Duden German Dictionary*. Oxford: Oxford University Press.
- Cohen, G. (1979). The labour theory of value and the concept of exploitation. *Philosophy & Public Affairs*, 8(4): 338–60.
- Cohen, G. (1988). *History, labour, and freedom: Themes from Marx*. Oxford: Oxford University Press.
- Cohen, G. (1997). Where the action is: On the site of distributive justice. *Philosophy & Public Affairs*, 26(1): 3–30.
- Davidson, D. (2001). *Essays on actions and events*. 2nd edition. Oxford: Oxford University Press.
- Dworkin, R. (1981a). What is equality? Part 1: Equality of welfare. *Philosophy & Public Affairs*, 10(3): 185–246.
- Dworkin, R. (1981b). What is equality? Part 2: Equality of resources. *Philosophy & Public Affairs*, 10(4): 283–345.
- Elster, J. (1982). Roemer versus Roemer: A comment on “New directions in the Marxian theory of exploitation and class”. *Politics and Society*, 11(3): 363–73.
- Elster, J. (1986). *Karl Marx: A reader*. Cambridge: Cambridge University Press.
- Feinberg, J. (1988). *The moral limits of the criminal law: Harmless wrongdoing*. Oxford: Oxford University Press.
- Gauthier, D. (1986). *Morals by agreement*. Oxford: Oxford University Press.
- Geras, N. (1985). The controversy about Marx and justice. *New Left Review*, 150(3): 37–69.
- Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23(6): 121–23.

- Goodin, R. (1986). *Protecting the vulnerable: A re-analysis of our social responsibilities*. Chicago: University of Chicago Press.
- Goodin, R. (1987). Exploiting a situation and exploiting a person. In *Modern theories of exploitation*. ed. A. Reeve: 166–97. London: Sage.
- Hall, N. (2004). Two concepts of causation. In *Causation and counterfactuals*. eds. J. Collins, N. Hall, and L. Paul: 225–76. Cambridge: MIT Press.
- Hare, R. (1952). *The language of morals*. Oxford: Clarendon Press.
- Harsanyi, J. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4): 309–21.
- Harsanyi, J. (1956). Approaches to the bargaining problem before and after the theory of games: A critical discussion of Zeuthen's, Hick's, and Nash's theories. *Econometrica*, 24(2): 144–57.
- Hobbes, T. (1996). *Leviathan*, ed. R. Tuck. Cambridge: Cambridge University Press.
- Hausman, D. (1995). The impossibility of interpersonal utility comparisons. *Mind*, 104(415): 473–90.
- Kalai, E. and Smorodinsky, M. (1975). Other solutions to Nash's bargaining problem. *Econometrica*, 43(3): 513–18.
- Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, 63(279): 190–93.
- Kramer, M., Simmonds, N., and Steiner, H. (2000). *A debate over rights*. Oxford: Oxford University Press.
- Kristof, N. (2009). Where sweatshops are a dream. *New York Times*, January 14, 2009: 14D.
- Kymlicka, W. (1990). *Contemporary political philosophy: an introduction*. Oxford: Clarendon Press.
- Lewis, D. (1973). Causation. *The Journal of Philosophy*, 70(17): 556–67.
- Locke, J. (1764). *Two treatises of government*, ed. T. Hollis. London: A. Millar & Co.
- Lowe, E. (2010). Action theory and ontology. In *A companion to the philosophy of action*. eds. T. O'Connor and C. Sandis: 3–9. Oxford: Blackwell.
- Luce, D. and Raiffa, H. (1989). *Games and decisions*. New York: Dover.
- Maitland, I. (1996). The great non-debate over international sweatshops. In *Ethical theory and business*, eds. T. Beauchamp and N. Bowie: 593–605. New York: Prentice Hall.
- Marx, K. (1887). *Capital: A critique of political economy*. Moscow: Progress Publishers.
- Mele, A. (2001). Acting intentionally: Probing folk notions. In *Intentions and intentionality: Foundations of social cognition*, eds. B. Malle, L. Moses, and D. Baldwin: 27–43. Cambridge, MA: The MIT Press.
- Miller, D. (1987). Exploitation in the market. In *Modern theories of exploitation*. ed. A. Reeve: 149–65. London: Sage.

- Morris, S. (2009). Husband with chronic sleep disorder strangled wife during nightmare. *Guardian*, November 17, 2009. <http://www.guardian.co.uk/uk/2009/nov/17/husband-strangled-wife-during-nightmare>.
- Murphy, L. (1998). Institutions and the demands of justice. *Philosophy & Public Affairs*, 27(4): 251–91.
- Narveson, J. (1988). *The libertarian idea*. Philadelphia: Temple University Press.
- Nash, J. (1950). The bargaining problem. *Econometrica*, 18(2): 155–62.
- Nozick, R. (1969). Coercion. In *Philosophy, science, and method: Essays in honor of Ernest Nagel*. eds. S. Morgenbesser, P. Suppes, and M. White: 440–72. New York: St. Martin's Press.
- Nozick, R. (1974). *Anarchy, State, and Utopia*. Oxford: Blackwell.
- Nussbaum, M. (2003). Capabilities as fundamental entitlements: Sen and social justice. *Feminist Economics*, 9(2-3): 33–59.
- Olsaretti, S. (2009). Responsibility and the consequences of choice. *Proceedings of the Aristotelian Society*, 109(1): 165–88.
- Otsuka, M. (2003). *Libertarianism without inequality*. Oxford: Oxford University Press.
- Pigou, A. (1920). *The economics of welfare*. London: Macmillan.
- Ramsey, F. (1928). A mathematical theory of saving. *The Economic Journal*, 38(152): 543–59.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
- Reiff, M. (2013). *Exploitation and economic justice in the liberal capitalist state*. Oxford: Oxford University Press.
- Robbins, L. (1932). *An essay on the nature and significance of economic science*. London: Macmillan.
- Robbins, L. (1938). Interpersonal comparisons of utility: A comment. *Economic Journal*, 48(192): 635–41.
- Robinson, J. (1933). *The economics of imperfect competition*. London: Macmillan.
- Robinson, J. (1933). *Economic philosophy*. Bungay: Penguin.
- Roemer, J. (1982a). *A general theory of exploitation and class*. Cambridge, MA: Harvard University Press.
- Roemer, J. (1982b). New directions in the Marxian theory of exploitation and class. *Politics and Society*, 11(3): 253–87.
- Roemer, J. (1982c). Property relations vs. surplus value in Marxian exploitation. *Philosophy & Public Affairs*, 11(4): 281–313.
- Roemer, J. (1986). *Value, exploitation and class*. London: Routledge.
- Roemer, J. (1994). *Egalitarian perspectives*. Cambridge: Cambridge University Press.

- Ross, R. (2004). *Slaves to fashion: Poverty and abuse in the new sweatshops*. Ann Arbor, MI: University of Michigan Press.
- Sample, R. (2003). *Exploitation: What it is and why it's wrong*. Boulder, CO: Rowman and Littlefield.
- Samuelson, P. (1971). Understanding the Marxian notion of exploitation: A summary of the so-called transformation problem between Marxian values and competitive prices. *Journal of Economic Literature*, 9(2): 399–431.
- Sandel, M. (1998). What money can't buy: The moral limits of markets. *The Tanner Lectures on Human Values*. Oxford University: May 11–12, 1998.
- Satz, D. (2010). *Why some things should not be for sale: The moral limits of markets*. Oxford: Oxford University Press.
- Scanlon, T. (1998). *What we owe to each other*. Cambridge, MA: Harvard University Press.
- Schweickart, D. (1997). A democratic theory of economic exploitation dialectically developed. In *Exploitation*, eds. K. Neilson and R. Ware: 49–68. New York: Humanities Press.
- Sen, A. (1992). *Inequality reexamined*. Oxford: Clarendon Press.
- Sen, A. (1999). *Development as freedom*. Oxford: Oxford University Press.
- Sensat, J. (1984). Exploitation. *Nous*, 18(1): 21–38.
- Smilansky, S. (2007). *10 moral paradoxes*. Blackwell: Oxford.
- Steiner, H. (1974). The natural right to equal freedom. *Mind*, 83(330): 194–210.
- Steiner, H. (1977). The structure of a set of compossible rights. *Journal of Philosophy*, 74(12): 767–75.
- Steiner, H. (1984). A liberal theory of exploitation. *Ethics*, 94(2): 225–41.
- Steiner, H. (1987). Exploitation: A liberal theory amended, defended and extended. In *Modern theories of exploitation*. ed. A. Reeve: 132–48. London: Sage.
- Steiner, H. (1994). *An essay on rights*. Oxford: Blackwell.
- Steiner, H. (1999). Silver spoons and golden genes: Talent differentials and distributive justice. In *1998 Oxford amnesty lectures: the genetic revolution and human rights*. ed. J. Burley: 133–50. Oxford: Oxford University Press.
- Steiner, H. (2010). Exploitation takes time. In *Economic theory and economic thought: Essays in honour of Ian Steedman*. eds. J. Vint et. al.: 20–29. London: Routledge.
- Thomson, J. (1991). Self-defense. *Philosophy & Public Affairs*, 20(4): 283–310.
- Tormey, J. (1976). Exploitation, oppression, and self-sacrifice. In *Women and philosophy: Toward a theory of liberation*. eds. C. Gould and M. Wartofsky: 206–21. New York: Putnam.

- U.S. Attorney's Office: Northern District of Illinois (2012). William Mastro and two other executives of former Mastro Auctions indicted for allegedly defrauding bidders in online and live auctions of sports memorabilia and other collectibles. *Press Release*: July 25, 2012.
- Vallentyne, P. and Steiner, H. (2000). *Left-libertarianism and its critics: The contemporary debate*. Basingstoke: Palgrave.
- Veneziani, R. (2007). Exploitation and time. *Journal of Economic Theory*, 132(1): 189–207.
- Veneziani, R. (2013). Exploitation, inequality, and power. *Journal of Theoretical Politics*, (forthcoming).
- Veneziani, R. and Yoshihara, N. (2011). What is exploitation? *MANCEPT: Exploitation: Conceptual, normative and political perspectives*. Lecture. Manchester University: September 1, 2011.
- Vrousalis, N. (2013). Exploitation, vulnerability, and social domination *Philosophy & Public Affairs*, (forthcoming).
- Walt, S. (1984). Comment on Steiner's "Liberal theory of exploitation". *Ethics*, 94(2): 242–47.
- Warren, P. (1997). Should Marxists be liberal egalitarians? *Journal of Political Philosophy*, 5(1): 47–68.
- Wertheimer, A. (1987). *Coercion*. Princeton: Princeton University Press.
- Wertheimer, A. (1996). *Exploitation*. Princeton: Princeton University Press.
- Wertheimer, A. (2007). Review: Ruth J. Sample, "Exploitation: What it is and why it's wrong". *Utilitas*, 19(2): 259–61.
- Williams, B. (1985). *Ethics and the limits of philosophy*. London: Fontana Press.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.
- Wolff, J. (1999). Marx and exploitation. *The Journal of Ethics*, 3(2): 105–20.
- Wood, A. (1981). Marx and equality. In *Issues in Marxist philosophy, Vol. 4*. eds. J. Mepham and D. Hillel-Ruben: 195–221. Sussex: Harvester Press.
- Wood, A. (1995). Exploitation. *Social Philosophy and Policy*, 12(2): 135–58.
- Zeuthen, F. (1930). *Problems of monopoly and economic welfare*. London: Routledge.
- Zwolinski, M. (2007). Sweatshops, choice, and exploitation. *Business Ethics Quarterly*, 17(4): 689–727.