Risk and Justice

An Account of Just Risk Imposition and Fair Risk Sharing

Name: Yunmeng Cai
College: Nuffield College, University of Oxford
Degree: DPhil in Politics

June 2023

Abstract

This thesis examines how a society should deal with risk justly. Many risks commonly faced by members of a society could be regarded as being imposed on them through the collective choice of social institutions. Since subjecting people to risk sets back their basic interest in autonomy, risk imposition of this kind requires justification. In this thesis, I aim to answer questions including under what conditions it is justifiable for a society to impose risk on people and how the burden of social risk materialisation should be fairly allocated among members of a society.

To answer these questions, I develop an account of just social risk imposition and fair risk sharing by employing the Scanlonian contractualist framework. I argue that social risk imposition should be regulated by not only principles concerning potential outcomes of the risk in question but also principles concerning the nature or the characteristics of the risk. Moreover, the burden of compensating the victims of risk materialisation should be shared by all those who benefit from the imposition of the risk and its fair allocation should also take personal responsibility into consideration. This account is then expanded to cover topics such as long-term risk imposition on people from future generations and risk sharing with people who are not fully capable of contributing to social cooperation (such as people with severe disabilities). Finally, the thesis considers the institutional and policy implications of the account. I argue that the account provides a fairness-based justification for social insurance and may help to address policy questions such as what risks should be covered by social insurance and how a social insurance scheme should be funded.

Acknowledgements

I am deeply grateful to my university supervisor, Prof. Stuart White, and college supervisor, Prof. David Miller, for their invaluable guidance, advice and support during my MPhil and DPhil study. I would like to thank Stuart for being such a supportive and patient supervisor, who has read every draft of my thesis and provided valuable comments and edits. I have also benefited significantly from the tutorials we have had over the years. I would like to thank David for always offering to read my drafts and for providing insightful comments as well as detailed edits for the thesis, without which it could never have taken its current shape. He has also been an extremely kind mentor and has provided me with valuable academic and life advice throughout time.

I would also like to thank Prof. Cécile Laborde for her academic guidance as well as her kind encouragement during the difficult times, and my confirmation examiners, Prof. Jonathan Wolff and Prof. Daniel Butt, for their thoughtful comments on the three chapters of this thesis. I would also like to express my gratitude to Prof. Jonathan Wolff and Prof. Alex Voorhoeve for being my VIVA examiners and for providing inspiring feedback for the thesis and for further research.

Chapters of this thesis have been presented in the Nuffield Political Theory Workshop, Oxford Graduate Conference in Political Theory, Harvard Graduate Conference in Political Theory, the MANCEPT Workshop at Manchester University and the Political Philosophy and the Future of Capitalism Conference at Waseda University. I am appreciative of the feedback received from the participants.

I am also grateful for receiving the Oxford-Adrian Swire Graduate Scholarship, which has supported me financially through my MPhil and DPhil study.

Finally, my deepest gratitude to my parents and to Dr. Tian Wei—I could not imagine completing this journey without their love, understanding and support.

Table of Contents

Chapter 1 — Introduction: Risk and Justice **1**

1.1 Risk and the Measurement of Wellbeing for Distributive Justice 3

1.2 Risk and Responsibility 8

1.3 Risk Imposition 15

1.4 The Contractualist Approach to Risk Imposition 24

1.5 Risk Sharing and Social Cooperation 27

1.6 The Concept of Risk and Thesis Plan 34

Chapter 2 — Risk and Autonomy **40**

2.1 Risk and Background Options 42

2.2 Can Risk Be Valuable for Autonomy? 52

2.3 Risky Choices and Paternalism 60

Chapter 3 — Justifying Social Risk Imposition: A Contractualist Approach **68**

3.1 The Imposition View of Common Risk 69

3.2 A Contractualist Account of Just Social Risk Imposition 72

3.3 The Compensation Principle and the Demand for Risk Sharing 85

3.4 Fair Social Risk Sharing 93

3.5 The Problem of Plural Risk Preferences 97

Chapter 4 — Responsibility and Fair Risk Sharing **104**

4.1 Scanlon's Account of Substantive Responsibility 105

4.2 Substantive Responsibility and Risk 115

4.3 The Reasonable Expectation Account 130

Chapter 5 — Risk Imposition on Future Generations **139**

5.1 The Multi-Generation Contractualist Framework 141

5.2 Imposing Risks on Future Generations 149

5.3 Risks Imposed on Future Generations—Anything Special? 155

5.4 Fair Burden Allocation Across Generations 165

Chapter 6 — Reciprocity, Risk Sharing and Social Cooperation **173**

6.1 Reciprocity as a Demand of Justice 175

6.2 Reciprocity as the Conditional Motivation for Justice 183
6.3 Reciprocity as a Disposition Toward Justice 195

Chapter 7 — Conclusion: Institutions and Policies for Fair Risk Sharing 201

7.1 Just Risk Imposition and Fair Risk Sharing—A Summary 201
7.2 Risk and Property-Owning Democracy 203
7.3 Social Insurance for Fair Risk Sharing 207

Bibliography 222

Chapter 1 — Introduction: Risk and Justice

People living in modern societies face various risks throughout their lives. They may fall ill, have accidents, experience temporary or permanent disabilities, lose jobs, or suffer losses in personal assets. Some of these risks seem to naturally exist in people's lives (such as the risk of natural disasters) while others are introduced by the design of modern societies (such as the risk of unemployment). In some countries the state may protect its citizens from some of the common risks by funding social security programmes or making insurance purchase mandatory; in other societies it may be believed that individuals should take personal responsibility in managing their own risk exposures.

Since risk represents a major threat to people's ability to live an autonomous and dignified life, it is a matter of justice how a society deals with risk. For example, under what conditions is it justifiable to impose a risk on people? Who should bear the burden in the event of its materialisation? Should individuals be responsible for managing their own risk exposure or should the state play a role in protecting its citizens against some common risks? How to handle long-term risks of which we lack full knowledge and whose consequences will only affect future generations?

How a society should deal with risk justly is the subject of this thesis. I aim to answer the above questions by developing an account of just social risk imposition and fair risk sharing. I will lay out the main arguments of this account in Chapters 2 to 4. I argue that many risks commonly faced by members of a society could be regarded as being imposed on them through the collective choice of social institutions and risk imposition of this kind requires justification because subjecting people to risk sets back their basic interest in autonomy. I then use the Scanlonian contractualist framework to examine what principles should regulate a scheme of social risk imposition. I suggest that its justification calls for not only principles concerning potential outcomes of the risk in question but also principles concerning the nature or the characteristics of the risk. Moreover, the burden of risk materialisation should be shared by all those who benefit from the imposition of the risk and its fair allocation should also take into consideration concerns of personal responsibility. I then expand this account to cover people from

future generations and people who are not fully capable of contributing to risk sharing (such as people with severe disabilities) in Chapter 5 and 6. Finally, in Chapter 7, I consider the institutional and policy implications of my account. I argue that fair risk sharing may serve as an additional justification for social insurance schemes. This fairness-based justification can also help to address policy questions such as what risks should be covered by social insurance and how such a scheme should be funded.

Before delving into the main arguments, I will provide an overview of how existing literatures in philosophy deal with risk in this first chapter. Section 1.1 looks at how risk may be incorporated into the measurement of wellbeing for a theory of distributive justice. I take Rawls's account of primary goods as an example to show that measures based on average or expected life prospects fail to properly account for the morally significant impact of risk. I then suggest that such impact may be captured by a modified version of autonomy as self-authorship as proposed by Raz. Section 1.2 examines how risk complicates the idea of responsibility in responsibility-sensitive theories of justice. In particular, I examine Dworkin's account of option and brute luck and argue that Dworkin's view of risk-related responsibility is merely assumed rather than justified. Moreover, his arguments ignore the possibility that whether a risk is taken on by an individual or it exists in the background of his life may impact our view of how the resulting responsibility with regard to the risk should be attributed.

Section 1.3 turns to the literature on the ethics of risk imposition. This topic is relevant because many common risks existing in the background of individuals' lives are in fact imposed on them through the design of social institutions and public policies. The justification for such common risks may have implications for the fair allocation of their resulting burdens. I suggest that both the consequentialist approach (such as Cost and Benefit Analysis) and the rights-based approach (in particular I will look at Nozick's account) fail to provide a satisfactory solution to the problem of risk imposition. I then turn to Scanlon's contractualist framework in Section 1.4 and suggest that it could be employed to identify principles of just social risk imposition. In particular, its justification may point toward a demand for all members of the society to fairly share the burden of social risk. Section 1.5 looks at the idea of risk sharing as a form of social

cooperation. I examine existing justifications for the state to intervene in voluntary cooperation in risk sharing (which in modern societies mainly takes the form of private insurance schemes) and suggest that just risk imposition may provide an additional reason for the state to enforce compulsory risk sharing among all members of society with a fair allocation of the accompanying burden. Finally, in Section 1.6, I clarify the concept of risk to be employed in this thesis and lay out the thesis plan.

1.1 Risk and the Measurement of Wellbeing for Distributive Justice

Some theories of distributive justice measure individual wellbeing using what I call an average life prospects approach. That is, they refer to the average or the expected level of wellbeing enjoyed by each individual as a metric for what needs to be distributed justly. By doing so, they tend to overlook the impact of risk on individuals, because potential fluctuations introduced by risk are evened out in such measurements. To illustrate, I will use Rawls's theory as an example, but other accounts of justice which take a similar approach of assessing wellbeing based on average (or expected) resources, opportunities or access to advantages over individuals' lifetime are all subject to similar concerns that I am going to outline.[1]

Rawls measures people's social and economic advantage (or disadvantage) based on their reasonable expectations of primary goods over a complete life.[2] Primary goods are social conditions and all-purpose means that enable individuals to carry out life projects and pursue their own conceptions of the good. They are goods that any rational person generally wants no matter what else they want.[3] Rawls's list of primary goods includes basic rights and liberties, opportunities, income and wealth, and the social bases of self-

---

[1] For such accounts, see, for example, R. Dworkin, "What is Equality? Part 2: Equality of Resources", *Philosophy and Public Affairs* 10 (1981), 283-345; R. Arneson, "Equality and Equality of Opportunity for Welfare", in *Equality: Selected Readings*, ed. L. Pojman and R. Westmoreland (New York: Oxford University Press, 1997); G. A. Cohen, "On the Currency of Egalitarian Justice", *Ethics* 99 (1989), 906-44.

[2] J. Rawls, *Justice as Fairness: A Restatement* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2001), p. 59.

[3] J. Rawls, *A Theory of Justice* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999), p. 54; Rawls, *Justice as Fairness*, p. 57.

respect.[4] By looking at the expected level of an index of primary goods over a person's life time, Rawls's theory does not seem to take account of the volatility in the level of the index that may be caused by the existence of risk.

Note that this does not mean that such measurement cannot capture the consequence of risk—if an economic risk results in a reduction in a person's wealth, it will be reflected in their index of primary goods being lowered. What it does not capture, however, is the deviations from the expected level of the index experienced by the individual throughout his life. Consider two individuals who have the same expected level of primary goods throughout their complete life. However, one of them is expecting a relatively stable level of primary goods over time while the other is expecting lots of potential ups and downs that are beyond his control. Rawls's theory of justice would consider them equally well-off as long as the expected levels of primary goods over their lifetime are the same, ignoring the more risks the second person is being exposed to.

Being exposed to risks can burden individuals in various morally significant ways that are not captured by an average life prospects approach. First, facing significant ups and downs in one's either physical or economic conditions, or even merely anticipating such instability, may cause people to suffer psychological burdens such as fear or anxiety.[5] These emotional experiences may then exhaust people's mental energy and prevent them from engaging in other important activities in life. Second, being at risk may also undermine a person's ability to make long-term plans and enter into long-term commitments. For example, a person facing a constant threat to his health conditions may find it difficult to get a good job; a person without a stable financial outlook may not be able to obtain a loan for starting a business. Moreover, as pointed out by Goodin, since a person's plans and life projects often constitutes many interlocking parts, and contributions to these parts are often made sequentially rather than simultaneously,

---

[4] Rawls, *A Theory of Justice,* p. 54. In *Justice as Fairness*, Rawls expands opportunities into i) the freedom of movement and free choice of occupation against a background diverse opportunities, and ii) powers and prerogatives of offices and positions of authority and responsibility. See *Justice as Fairness,* p. 58.

[5] M. Hayenhjelm and J. Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature", *European Journal of Philosophy* 20 (2012), 26-51, p. 31.

people may be unwilling to carry out long-term projects without a guarantee of a relatively stable future because they may fear that their plans will be interrupted and their current efforts will be wasted.[6] Third, exposure to risk may also lead people to take costly measures to reduce the likelihood or impact of a feared event, even though the risk in question may turn out never to eventuate. As noted by Wolff and De-Shalit, individuals without financial security are likely to be forced to take on many other risks in order to address the more urgent ones.[7] This may lead people whose material well-being expectations are already low to be even more burdened with their compounded risk exposure.[8]

How, then, can a theory of justice account for the value of security or stability for individuals' life prospects? One strategy is to adopt the index of primary goods proposed by Rawls and add security or stability as an additional primary good to the list. This is in fact not an unusual approach—similar proposals have been made to add, among others, care[9] and leisure[10] as additional primary goods so as to account for the moral importance of receiving care during dependency or enjoying discretionary time as part of a person's life, because such needs cannot be independently captured by existing primary goods on the list. In a similar spirit, then, it may be proposed that security could also be included as a primary good so that, all things being equal, a person who suffers more disruptions in their life could be considered worse off in terms of the level of primary goods they enjoy over lifetime compared to others.

However, this solution is problematic because security or stability is not just another social good unrelated to existing ones on the list. It in fact reflects the volatility or the expected volatility of the levels of existing primary goods such as opportunities, income and wealth that are enjoyed or are expected to be enjoyed by a person over their life

---

[6] R. Goodin, "Stabilizing Expectations: The Role of Earnings-Related Benefits in Social Welfare Policy", *Ethics* 100 (1990), 530-53, p. 548.

[7] J. Wolff and A. De-Shalit, *Disadvantage* (Oxford: Oxford University Press, 2007), pp. 66-70.

[8] Ibid., p.10.

[9] E. Kittay, *Love's Labor: Essays on Women, Equality and Dependency* (London: Routledge, 1999).

[10] R. Goodin, *Discretionary Time: A New Measure of Freedom* (Cambridge: Cambridge University Press, 2008).

course. Adding it as a primary good merely transforms the problem of how to account for risk in distributive justice into the problem of how to compile this new index of primary goods so that interpersonal comparisons could be made between, for example, a person who has a higher level of average income but suffers a higher level of volatility and a person with a lower level of both income and volatility.

Alternatively, it may be suggested that, instead of being added to the index as a primary good, security could be viewed as a good that contributes to an existing good in the index. For example, this is how Daniels deals with the case of health—he suggests that the moral importance of health and, relatedly, the social factors that promote health, lies in their contribution to securing the primary good of opportunities.[11] Health is important because impairments of normal functioning would reduce the range of exercisable opportunities from which individuals may pursue their life plans and goals.[12] Similarly, it may be said that a lower level of security would also negatively impact the range of opportunities available to individuals. Hence, a modified approach to risk may be adopted which treats security, or social arrangements that manage risk and alleviate disruptions, as contributing to the preservation of opportunities.

There are, however, also problems with this suggestion. Opportunities for Rawls mainly refer to people's possibilities to exercise freedom of occupational choices and freedom of movement, those that constitute the institutional background for them to form and execute long-term life plans.[13] They do not refer to the whole set of background options that are available to individuals during their life course, which is represented by the other "all-purpose" primary good—income and wealth. Since fluctuations in the level of income and wealth impact the level of security, security, in particular economic security, is not a pure exogenous factor to the index of primary goods. While physical security, on the other hand, may bear more resemblance to health (in fact, they may be said to be the same kind of primary good), economic security is in fact a derivative of, rather than a contributor to, the level of the index.

---

[11] N. Daniels, *Just Health: Meeting Health Needs Fairly* (Cambridge: Cambridge University Press, 2008), p. 21.

[12] Ibid., p. 35.

[13] See Rawls, *Justice as Fairness*, p. 38.

Moreover, security, especially economic security, also differs from health in an important aspect. Unlike health, of which a reasonable person always wants more rather than less (setting aside the concern of cost), it does not seem to be the case that more security and less risk is always preferred by all reasonable individuals at all times. An obvious example is personal investment—it seems perfectly reasonable that some individuals are willing to take on the risk of stock investment, thus sacrificing some financial stability, as part of their life projects. Since a primary good is by definition something that rational people always want more rather than less of,[14] security cannot be treated either as a primary good or a good which always positively contributes to the level of the index of primary goods. Therefore, in order to deal with risk in the context of justice, we need an alternative measure of its impact on individuals.

Interestingly, several philosophers who explicitly discuss how risk may affect individual interests have all appealed to the notion of autonomy as a tool of measurement. In particular, they mostly refer to the concept of autonomy as self-authorship formulated by Raz.[15] According to Raz, individuals have an interest in living an autonomous life, a life they could look back on and recognise themselves as its author. This choice of referring to autonomy as self-authorship is, upon some reflection, not surprising. The idea of autonomy seems very promising in capturing what is potentially at stake in subjecting individuals to risk. Intuitively, it demands the kind of life in which people are not subject to constant or severe disruptions so that they could focus on personal projects and carry out long-term plans. It also captures the subjective feeling that a life frequently shocked by unwanted risks or disrupted by the need of managing risks that are not voluntarily taken on feels less like a life authored by a person himself.

More formally, three preconditions for autonomy as self-authorship are laid out by Raz. For an individual to live an autonomous life, he needs independence from coercion and manipulation, sufficient internal capacity to make rational choices and an adequate range of valuable background options to choose from.[16] While Raz is not particularly concerned with the issue of risk, the last precondition regarding the set of life options

---

[14] Rawls, *A Theory of Justice*, p. 54.

[15] J. Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), ch. 14-15.

[16] Ibid., p. 372.

could be modified to properly accommodate its impact. The existence of risk may affect not only the range but also the robustness of background options available to individuals. Therefore, in this thesis, I will use an individual's prospect for autonomy as a desideratum so as to account for the impact of risk on justice.[17] Chapter 2 of this thesis will provide a detailed analysis on how risk may set back people's interest by impacting the preconditions for autonomy.

1.2 Risk and Responsibility

In the previous section I have looked at accounts which discuss how a risk may set back individual interest and suggests that the impact of risk could be captured by a person's interest in autonomy. In other theories of justice, risk factors in not because it may undermine autonomy but because it complicates the notion of personal responsibility, which is in some sense the other side of autonomy. The most notable efforts in addressing the relationship between risk and personal responsibility have been made by Ronald Dworkin in his theory of equality of resources.[18] His approach, labelled by Anderson as equality of fortune or luck egalitarianism (commonly referred to as the latter)[19], is also endorsed by other egalitarians including Arneson, Cohen, Roemer, Van Parijs and so on, though each takes a slightly different form.[20] In this section, I will mainly focus on Dworkin's account as his version is the best known.

Dworkin's concern about risk is derivative of his concern to incorporate the ideal of personal responsibility into an egalitarian theory of justice. His theory starts with a hypothetical auction of external resources to ensure that their initial distribution represents equal value to all community members. However, such an equal distribution

---

[17] It may be concerned that autonomy is a perfectionist ideal and is therefore not suitable for serving as a desideratum of distributive justice. However, my focus here is the preconditions of autonomy listed by Raz, which measure the extent to which the background conditions of a society are conducive to people's prospect of living a self-authored life. It is not to dictate that people should actually *achieve* an autonomous life.

[18] Dworkin, "What is Equality? Part 2: Equality of Resources".

[19] E. Anderson, "What is the Point of Equality?", *Ethics* 109 (1999), 287-337. Note that Dworkin himself does not accept these labels.

[20] Arneson, "Equality and Equality of Opportunity for Welfare"; Cohen, "On the Currency of Egalitarian Justice"; J. Roemer, "A Pragmatic Theory of Responsibility for the Egalitarian Planner", *Philosophy and Public Affairs* 22 (1993), 146-66; P. Van Parijs, *Real Freedom for All* (Oxford: Clarendon Press, 1995).

would not be preserved after the auction because people will make different use of their resources and are also subject to external disruptions that they have no control over. The question, then, is whether subsequent changes in people's holdings of resources are a concern for egalitarian justice, that is, whether such differences are subject to redistribution in order to maintain equality in resources.

When changes in a person's possessed resources are entirely due to his voluntary choices or factors he has full control over, it seems more appealing to insist that he should keep such changes himself. For example, if an increase in holdings is due to the person's labour effort, it seems fair that he now enjoys more resources than someone who does not put in such efforts. Insisting on equal distribution of the gain among all members of the community seems to represent a failure to respect the person as an autonomous being whose actions and choices could shape his own life course. On the other hand, if a reduction in resources is a result of his squandering them, it also seems fair to hold the person responsible for the outcome of his own autonomous choice rather than asking others to compensate him through redistribution.

However, when the change is at least partially due to factors a person has no control over (for example, due to the materialisation of a risk that he is exposed to), it becomes less clear if the person should also be held responsible for the outcome as a matter of fairness or justice. Dworkin regards these uncertain, disruptive factors as involving luck. He then distinguishes between option luck and brute luck—option luck, either good or bad, represents unpredictable consequences of people's conscious choices; brute luck, on the other hand, materialises from risks that individuals do not consciously take on.[21] According to Dworkin, gains or losses resulting from option luck should stay with individuals, as people should be held responsible for their own choices, while those resulting from brute luck represent no one's fault and should therefore be subject to redistribution in order to restore equality among individuals with regard to such fortunes.[22]

---

[21] Dworkin, "What is Equality? Part 2: Equality of Resources", p. 293.

[22] Ibid., pp. 293-96.

This distinction between option and brute luck, however, may turn out to be less clear than it initially appears to be. In particular, some risks are not actively taken on by individuals as their personal choices, though individuals could perceive or anticipate their existence. They could, therefore, have allocated some resources to buy insurance to cover their exposure if they prefer to do so. For Dworkin, then, insurance is a device whose existence converts risks which would otherwise eventuate as brute luck into a matter of option luck.[23] For example, the risk of falling ill is often not a matter of personal choice, thus seeming to represent brute bad luck. With the availability of health care insurance, however, an individual is granted a choice, according to Dworkin, between covering the risk of illness with insurance and leaving himself exposed to its potential cost. If this person chooses the latter and is then hit by illness, the fact that he would not receive an insurance payment to cover the cost of paying for health care is a matter of bad option luck resulting from his deliberate decision not to buy insurance.

It should be noted that Dworkin does not consider the choice of not covering a risk with insurance to be irrational or blameworthy, as he acknowledged that such a decision involves a balance between a person's risk preference and his other preferences in life.[24] He only insists that individuals making such a decision should be held responsible for whatever outcome that follows.

Dworkin then recognises that, while it is possible to convert some brute bad luck such as illness into option luck by buying insurance from the insurance market, this is not always the case. Misfortunes such as congenital disease (or genetic defects that are likely to cause illness later in life) and lack of natural talents are brute bad luck resulting from risks that have eventuated before a person is born and is able to make an insurance decision.[25] Acknowledging this fact, he proposes that such brute bad luck should be compensated with an amount equivalent to what a person could have received had he

---

[23] Ibid., p. 296.

[24] Ibid.

[25] It may be argued that parents could have bought insurance for their children. However, if they fail to do so, it is problematic that children would have to bear the consequence of their parents' actions. Dworkin explicitly rejected this approach because individuals should only bear the consequences of their own choices. See R. Dworkin, "Justice, Insurance, and Luck" in *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, Massachusetts: Harvard University Press, 2000), p. 339.

covered it in a hypothetical insurance market. Here, the hypothetical insurance market is constructed as a conceptual tool to measure the extent to which those born with bad luck should be compensated and to thereby justify the corresponding redistribution. The aim, according to Dworkin, is to mitigate brute bad luck to the degree that each person's actual insurance choices would have produced so that everyone is put in an equal position with respect to risk.[26]

In order to implement such redistribution in reality, however, since it is impossible to obtain reliable information regarding the level of coverage each person would rationally prefer in the hypothetical market, such a scheme has to be designed based on the preference of an average person. This could take the form of a compulsory social insurance scheme with a universal level of coverage. It may be objected that the unavailability of an actual insurance market does not in itself give rise to reasons for the state to enforce participation based on the hypothetical choice of the average person. Since Dworkin accepts that people may rationally choose not to buy insurance to cover a specific risk in the private market based on their risk attitudes and personal preferences, he should also accept that this is still the case in the hypothetical market he imagined.

Dworkin acknowledges that the averaging assumption only represents the second best in absence of more detailed information but he believes such a programme is in any case better than nothing.[27] By saying so, he essentially meant that social insurance is justified as an approximate way of realising the goal of equality of fortune when it cannot be served by the private market. This justification also implies that social insurance should only aim to cover risks like congenital problems which eventuate before a person could possibly buy insurance to cover himself. Failure to buy insurance for other risks, such as

[26] Dworkin, “Justice, Insurance, and Luck”, p. 341.

[27] Dworkin, “What is Equality? Part 2: Equality of Resources”, p. 298.

disease or disability acquired after a person has grown up, even if fatal, is a matter of personal responsibility which need not be a concern of the state.[28]

Dworkin's account is subject to various objections which criticise his readiness to abandon those who fail to insure themselves against calamitous consequences, his failure to respect those in need of assistance, and so on.[29] Among them some are particularly concerned with his treatment of risk and responsibility. When Dworkin insists that individuals should be held responsible for option luck and option luck includes all those risks that could have been anticipated and/or managed, implicit in his argument is the premise that justice requires individuals to always bear the full consequence of their risk decisions, which may include taking on a risk, avoiding a risk and, in cases where they are inevitably exposed to a risk, either buying insurance to cover the risk or leaving themselves uncovered (provided that the background distribution of resources is equal and effects of brute luck are neutralised).

However, whether individuals should be held responsible for the full consequences of these risk decisions is a question that needs to be addressed rather than the answer being simply assumed. In fact, it is doubtful that it could be addressed without taking into account the source and nature of the risk in question. For example, if a person faces a serious risk because his neighbour chooses to store explosive chemicals next door, it is implausible to insist that, should the chemicals explode, he needs to bear the full consequence because he chooses not to move away and not to cover himself with the relevant insurance. Intuitively, his neighbour should bear most if not all responsibility in this case by exposing him to such danger.

This example points to a more general issue in deciding the allocation of responsibility in a theory of egalitarian justice. As argued by Stemplowska, responsibility is only

---

[28] Dworkin accepted in his later work that, in the real world, due to unfair distribution of wealth and difficulty in obtaining full information, compulsory social insurance for other risks, such as health care, may also be justified based on what a well-informed individual could have bought under conditions of equal resources. See Dworkin, "Justice and the High Cost of Health" in *Sovereign Virtue*, p. 311. But so far as an ideal theory is concerned, his account implies that the hypothetical insurance device should only be employed for theoretically uninsurable risks.

[29] Anderson, "What is the Point of Equality?".

significant against certain background conditions.[30] While Dworkin seems to believe that an equal initial distribution of resources is a sufficient condition to hold individuals fully responsible for bearing the consequence of their subsequent choices, Stemplowska points out that duty fulfilment, rights violation and illegitimate exercise of available options may all impact the just allocation of responsibility. For example, in the context of risk, the allocation of consequences cannot be reasonably determined without taking into consideration the right not to have certain risks involuntarily imposed, the duty of performing due care in dealing with a risk, and so on. These rights and duties, if observed, would influence what risks an individual would be exposed to just by living in a society. They may also demand that, as a matter of justice, possible consequences of the resulting risk be allocated in a way that is different from how they naturally fall upon each individual. Luck egalitarianism, however, seems to be silent on such background issues.

Moreover, even when a risk is voluntarily taken on by an individual, what consequences would fall on him also depend heavily on the background conditions of the society he lives in. Suppose a motorcyclist drives on the motorway without wearing a helmet and he also fails to buy himself private accident insurance. If he crashes (due to no one's fault) and injures himself, the consequences he needs to bear may include paying for the cost of the ambulance and medical care. However, as pointed out by Olsaretti, how much those costs would be depends on factors such as how expensive medical treatment is in his society, whether passersby are ready to help him, and so on.[31] Perhaps more important for the purpose of my discussion, such background conditions also include what the health care system or social insurance system is like in his society. If the motorcyclist lives in a society like the UK where free health care is available, then the cost for his imprudent choice would be much lower than the case in which he lives in a society where health care is expensive and social insurance for medical treatment is not available.

---

[30] Z. Stemplowska, "Making Justice Sensitive to Responsibility", *Political Studies* 57 (2009), 237-59.

[31] S. Olsaretti, "Responsibility and the Consequences of Choice", *Proceedings of the Aristotelian Society* 109 (2009), 165-88, p. 172.

From a slightly different angle, Stemplowska also gives an example where a group of individuals choose between two schemes of car damage insurance—one requires equal contribution and offers free repair to every car; the other asks for less (also equal) contribution from everyone and offers free repair only to those cars well maintained by the owner.[32] The choice thus involves a trade-off between a higher level of security for all (the first option) and higher efficiency (the second option). These social arrangements will determine what kind of consequences fall on which individuals. And whether these arrangements are themselves just is a question that cannot be addressed by referring only to the equal distribution of resources.

The choice of car damage insurance is a relatively trivial example and may not be an issue of justice in itself. Other risk-related laws, policies and social arrangements, however, may involve questions of justice. In particular, they determine how common risks are dealt with in a society, which forms the background against which individuals exercise personal choices including risk-related ones, and what consequences of risk individuals should bear. A major task of this thesis is to examine these issues concerning the justice of institutions in dealing with common risks in a society.

Before moving on to the next section, it should be noted that, besides luck egalitarianism, there are also other views of risk and responsibility in egalitarian theories of justice. For example, Fleurbaey holds a more radical view that individuals should never be held responsible for their bad (or good) luck because luck does not lie within their control.[33] Its underlying premise is that people should only be held responsible for what they can control. According to Fleurbaey, it follows that individuals are responsible for their taste or preference for risk but not for being lucky or unlucky because the former represents personal choices while the latter results from pure random factor.[34] For example, suppose a person makes an investment with an expected return of -$1,000. Then, whether he ends up gaining $10,000 or losing

---

[32] Stemplowska, "Making Justice Sensitive to Responsibility", p. 248.

[33] M. Fleurbaey, *Fairness, Responsibility, and Welfare* (Oxford: Oxford University Press, 2008), p. 153. According to Fleurbaey, Arneson, Cohen and Roemer all hold such a view regarding luck and responsibility.

[34] Ibid., pp. 154-55. Fleurbaey takes the view that tastes and preferences are within individual control, which may be controversial but is not my concern here.

$20,000, he should always bear the expected consequence of his choice, i.e. -$1,000. The implication of this view is essentially that individuals taking on personal risks are required by justice always to use insurance to fully cover their risk exposure.[35] I believe this is too restrictive and is not required by a reasonable view of personal responsibility —I will come back to this and offer my account of risk and responsibility in Chapter 4.

1.3 Risk Imposition

Besides the literature on distributive justice, there is an increasing interest in the ethics of risk imposition in philosophy. This literature is relevant for my project because, when a risk is imposed on a person by others, in order to decide the just allocation of its burdens, it is important to understand what justifies its imposition in the first place.

In fact, in the previous two sections, the discussion often seems to assume that the risk in question either naturally exists in people's lives or is something voluntarily taken on by individuals. What it ignores is that many risks commonly faced by people living in a modern society are often imposed on them by fellow members of that society. A typical example is car driving. If a pedestrian is hit by a car, a luck egalitarian may treat him as voluntarily taking on the risk by choosing to walk on the side of the road. This overlooks that the risk of car accident is (at least partially) imposed on road users by those who drive a car or take a car ride. Imagine a society which prohibits car driving altogether—people living in that society would not be subject to the risk of car accidents by taking a walk on the road.

So why or under what conditions is it ever permissible to impose a risk on others which may eventuate into actual harm? The justification of risk imposition is a perplexing topic. Consider the paradigm case of driving again. On the one hand, it seems unreasonable to prohibit car driving, thus sacrificing the convenience made possible, simply because it introduces a possibility of harm or death. If all actions with possible lethal consequences for others are to be prohibited, it is likely that no action would ever be permissible. On the other hand, however, it is not clear why it is justifiable to subject people to deadly risks just for the benefit of convenience.

---

[35] The insurance should be funded by all those taking the same type of activities. See ibid., p. 154.

A natural response may be that, though the benefit of convenience brought about by the practice of car driving is small in amount for each individual, for the whole population it adds to a sum that is large enough to outweigh the burden it imposes. This is essentially the method of Cost Benefit Analysis (CBA) that is often employed by policy-makers in deciding what levels of risk are acceptable to be imposed on members of the society.[36] For example, CBA is widely used in the UK to assess workplace risks.[37] If improved safety in the workplace (calculated by, for example, the reduced likelihood of accident multiplied by the value of lives saved) outweighs the spending necessary for achieving this, then the level of risk imposed on workers should be reduced accordingly.

However, this simple idea of weighing cost and benefit faces various difficulties in application.[38] First, it is often disputable what estimation of probability should be used in the expected value calculation.[39] Probability numbers used in policy analysis are often based on statistical estimation. For example, to estimate the rate of car accidents given a specific speed limit, it needs to be decided what historical period and region of past data should be used. Another difficulty is to assign a value to human life when deadly risks are involved. It is notoriously difficult to find a robust and widely acceptable way to arrive at a monetary figure for life in order to calculate the expected cost of death (or the expected benefit of preventing fatality).[40]

Moreover, since the cost-benefit approach bases the justification of risk imposition on interpersonal aggregation, which is consequentialist in nature, it is also vulnerable to the standard objections laid against consequentialism. For example, allowing aggregated benefits to outweigh the cost suffered by a few, the approach seems to sanction the setback of the actual victims' interest in autonomy for the benefit of others, which

---

[36] See, for example, T. Lewens, "Introduction: Risk and Philosophy", in *Risk: Philosophical Perspectives*, ed. T. Lewens (London: Routledge, 2007), pp. 5-6.

[37] J. Wolff, "Fear, Blame, Shame, and the Regulation of Public Safety", *Economics and Philosophy* 22 (2006), 409-27, section 1. The government publication that explains the approach is *Reducing Risk, Protecting People* (Health and Safety Executive, UK, 2001).

[38] For a general discussion, see Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions", section 6.

[39] See a good discussion in S. Perry, "Risk, Harm, Interests, and Rights", in *Risk: Philosophical Perspectives*. I will also come back to this problem in Chapter 3.

[40] See, for example, J. Wolff, "What is the Value of Preventing a Fatality?", in *Risk: Philosophical Perspectives*.

violates what Rawls calls the separateness of persons[41]. It also cannot accommodate considerations of fairness in distributing benefits and burdens related to the risk among all affected individuals.[42] These problems should not be neglected simply because the approach can be conveniently used or can successfully explain the intuition people have about permitting certain forms of risk imposition.

These inherent problems of (aggregative) consequentialism may lead us to consider the theory's major rival, namely the rights-based approach, and see if the latter could offer better insights into the issue of risk imposition. However, as I will explain, it also faces serious difficulties in dealing with risk.

According to the rights-based approach, people should have a general right against risking (that is, against having risks imposed on them) which derives from their fundamental interest in, for example, liberty or autonomy.[43] Since risk exposure sets back their basic interest, individuals should have a general right against risking unless they consent to its imposition. Allowing this right against risking, however, would lead to the problem of paralysis.[44] Arguably, almost all actions involves potential risks that may eventuate into autonomy-undermining consequences for some individuals. Moreover, cases of risk imposition are usually such that it imposes a small possibility of a bad outcome on a large number of people, making consent-seeking impractical, if not impossible.[45] This would render most social and economic arrangements, indeed most human actions, impermissible.

To avoid the state of paralysis, it may be suggested that a threshold be introduced in determining the permissibility of risk imposition—only those actions with a small enough probability of harm may be allowed. Or, since the severity of harm also seems to matter, maybe the threshold should be based on the expected harm (total harm

---

[41] Rawls, *A Theory of Justice*, p. 166.

[42] Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions", p. 33.

[43] See, for example, D. McCarthy, "Rights, Explanation, and Risks", *Ethics* 107 (1997), 205-225; J. Oberdiek, *Imposing Risk: A Normative Framework* (Oxford: Oxford University Press, 2017).

[44] This term is used by Hayenhjelm and Wolff in "The Moral Problem of Risk Impositions", p. 26. For the same problem, also see McCarthy, "Rights, Explanation, and Risks", p. 208.

[45] McCarthy, "Rights, Explanation, and Risks", p. 216.

discounted by probability) it may cause. However, such suggestions are also problematic. Not only is the selection of a threshold bound to be arbitrary, calculating the probability or the expected harm of an action would expose the rights-based approach to the same problems of estimating probability and assigning appropriate values to human lives that were discussed earlier.[46]

The challenge for a rights-based approach raised by risk is well noticed by Nozick. Acknowledging that risk poses serious problems for a natural-rights position,[47] Nozick offers an interesting discussion about how, given that no natural-law theory is able to specify the content of a right against risking, an action involving a boundary-crossing risk for others should be dealt with in the state of nature. He outlines three possibilities: 1) the action is prohibited and punishable; 2) the action is permitted provided compensation is paid to those persons whose boundaries actually are crossed due to risk materialisation; 3) the action is permitted provided compensation is paid to all those persons exposed to the risk.[48]

It may be questioned why, if Nozick takes boundary-crossing seriously, he would consider Options 2 and 3 which allow the possibility of boundary-crossing without consent so long as compensation is paid afterwards. Surely he would not permit stealing or breaking one's leg just because compensation is paid thereafter. However, Nozick is aware of the problem of paralysis that may result if such prohibition extends too far. So he suggests that some border-crossing acts may permissibly be done without consent and provided compensation is paid if they fulfil the following criteria—consent-seeking is impossible or very costly; the act does not produce significant fear to others; the

---

[46] There is also a more theoretical objection against the threshold approach which is concerned with its effect on a combination of risk-imposing actions. Suppose a doctor has two patients and he gives them a drug with a 0.9% chance of causing death to each. If the acceptable threshold of 1%, the doctor is permitted to prescribe the drug to either patient but not to both as the chance of causing death in the latter case is roughly 1.79%, which exceeds the allowed risk level. In other words, one may be permitted to do A and be permitted to do B but not be permitted to do A and B. See discussion in F. Jackson and M. Smith, "Absolutist Moral Theories and Uncertainty", *The Journal of Philosophy* 103 (2006), 267-83; R. Abode, A. Borer and D. Enoch, "Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith", *The Journal of Philosophy* 105 (2008), 259-72.

[47] R. Nozick, *Anarchy, State, and Utopia* (Oxford: Basil Blackwell, 1974), p. 74.

[48] Ibid., pp. 75-76.

benefit of the act is worthwhile, and can pay its way (that is, it outweighs the corresponding burden).[49]

According to Nozick, some cases of risk imposition may fit into this description. If all risks are permitted (with compensation paid), the totality of fear they would cause is likely to be too large to bear. Hence some risks are valuable enough to permit, but others should be prohibited so that overall fear could be reduced to an acceptable level.[50] Among those that may be permitted, compensation should be paid for the possible or actual boundary crossing caused by the imposed risk, which leads to Options 2 and 3. Between them Nozick does not seem to have a clear preference. He suggests that Option 2 is more manageable due to far smaller transaction costs as well as the advantage that it fully compensates the actual victims in case of risk materialisation, but Option 3 makes more sense when the action creates fear to all those subject to the risk regardless of whether it eventuates into actual harm.[51]

Nozick is instead more concerned with a different question—with either Option 2 or 3, what if a person wishing to perform a risk-imposing action cannot afford the relevant compensation? Should he then be prohibited from performing the action which may in the end involve no cost at all to anyone else (suppose it is not particularly fear-arousing)?[52] Especially, there are some types of risk-imposing actions, such as driving, which are generally done in a society and play an important role in people's lives. For easy reference, let me call a risk imposed by such an action a common risk and the corresponding action a common action. Nozick argues that it would seriously disadvantage a person if he is prohibited from performing a common action simply because he cannot afford the relevant compensation (as required by either Option 2 or

---

[49] Ibid., p. 72.

[50] Ibid., p. 74.

[51] Ibid., pp. 76-78. Nozick does not attempt to make a choice probably also because he believes they are convertible. According to him, under the third option, people can choose to pool the compensation they receive to pay to the actual victims; under the second option, especially in the case of a deadly risk, people may want to sell the right to compensation after death to a private company which purchases such rights in exchange for a fixed payment immediately.

[52] Ibid., p. 78.

3).[53] Therefore, in the case of a common action, those who wish to forbid a person from performing it should compensate him for the disadvantage he suffers due to the prohibition. He calls this the Principle of Compensation.[54]

It is not clear, however, how this principle could fit into the general picture of Nozick's discussion. In particular, it seems unclear to me whether Nozick intends his Principle of Compensation to apply only to cases where the person who wishes to perform a common action is unable to afford the relevant compensation or he takes the principle to cover everyone engaging in common risk imposition. According to the first possible interpretation, a person who can afford compensation still has an obligation to pay those affected should he wish to perform the action. In other words, a person's right to impose the common risk is still conditional on the guarantee that he compensates accordingly. Meanwhile, others have no right to prohibit him (even if they are willing to pay compensation to stop him) from carrying out the action as long as he compensates properly. Only when the person cannot afford to pay compensation do others then acquire a right to stop him from doing it but this right is also conditional on them fulfilling the duty to compensate him for the resulting disadvantage.

However, this interpretation cannot reasonably answer the question of what should happen when a person subjected to an imposed common risk also cannot afford the required compensation for stopping the risk imposition. To illustrate, let me call this person the Potential Victim and the compensation he needs to pay the Pre-emption Payment. On the other hand, call the person who wishes to impose the common risk the Imposer and the compensation he needs to pay for imposing the risk the Imposition Payment. The problem is this—if the Potential Victim cannot afford the Pre-emption Payment, is the Imposer who also cannot afford the Imposition Payment now permitted to go ahead without paying? Is the Potential Victim not owed compensation (either ex ante or ex post) any more simply because he cannot afford to pay to stop the risk from being imposed? If Imposition Payment is owed in general to victims of risk imposition

---

[53] Ibid., pp. 81-82. Nozick does not intend to specify a theory of disadvantage but suggests that disadvantage is compared to the normal situation rather than being made worse off than one otherwise would be.

[54] Ibid., pp. 82-83.

for crossing their boundaries, then this fact is not changed by either the Imposer or the Potential Victim's ability to pay.

This may lead to a second way of understanding Nozick's principle. Under this second interpretation, when a risk-imposing action is considered as a common action, everyone now has a right to perform it and this right is not conditional on the performer's payment of compensation. In other words, those exposed to a common risk are not owed Imposition Payment any more, but they have a right to stop others from carrying out the action by making Pre-emption Payment instead. This interpretation is potentially more coherent, but it also cannot explain why compensation for boundary crossing is no longer required simply because the risk is a common risk. In some places, Nozick seems to suggest that it is because we all subject each other to common risks, so the overall burdens approximately cancel out.[55] While it may make sense for ex ante burdens such as fear, it is hardly true that the ex post burden of actual harm would cancel out even for the most commonplace risk materialisation like car accidents. When a person cannot afford compensation for stopping a risk from being imposed on him and that risk eventualises, wouldn't it also seriously disadvantage this person that he does not get any compensation to rectify the harm he suffers?

In fact, it seems that what motivates Nozick's introduction of the Principle of Compensation is the concern that some people wishing to engage in common risk-imposing actions may not be able to afford compensation owed to possible victims. Yet whether people should be allowed to perform such actions and whether they have to bear the full burden of paying for the attached compensation are potentially two separate questions. For example, instead of having the imposer of a particular risk pay full compensation, others engaging in the same type of common action may be required to share the burden together.[56] Treating the two questions separately may thereby

[55] See Nozick's reference to Fried in ibid., p. 76.

[56] Nozick actually mentions this possibility when discussing the risk of pollution, yet he believes that having those who actually benefit to pay, that is, each company paying for the risk it imposes, is a fairer option. See ibid., p. 80. This, however, may differ from the case of common risk imposition where many people other than the risk imposer may benefit from the practice. I will discuss this in detail in Chapter 3.

dissolve Nozick's concern that a person's need to engage in common practices and his ability to pay sometimes mismatch.

While Nozick's treatment of risk, especially his formulation of the Principle of Compensation, seems a bit unusual and deviates from how the issue of risk imposition is usually discussed, it does offer some valuable insights into the problem that are worth noting. First, Nozick introduces the idea of common risk (though he does not use the term), that is, risk imposed by common practices or arrangements in a society and that affects most if not all its members, and suggests that common risk should potentially be dealt with in a way that is different from how other individual risks are treated. This distinction is not often emphasised in other literatures on risk imposition, such as most of the contractualist literature that I am going to discuss shortly, yet it is an important one because, unlike personal risks, common risks in some sense affect the background conditions for individuals to exercise autonomy and take personal responsibility. As I have already suggested when discussing Dworkin, it is therefore a question of justice how institutions should deal with such risks. In Chapter 3 of this thesis, I will offer a detailed analysis of what risk might be treated as common risk.

Second, Nozick's discussion draws attention to the idea of compensation. It seems that quite often risk-related policy discussion chooses to use deadly risks as examples. In those scenarios, the issue of compensation becomes secondary because, for the victim himself, death is something that cannot be compensated for. But, in reality, many cases of risk imposition result in physical harm or material losses, which could be either fully or partially rectified by the payment of compensation. If compensation is required as a condition for legitimately imposing a risk in the first place, then payment to victims no longer needs to be justified by referring to, for example, the egalitarian demand for redistribution or the humanitarian need of providing aid. This is what I will defend in this thesis for the case of common risk—I will argue that the guarantee of compensation upon risk materialisation is an important condition for justifying a scheme of common risk imposition to members of a society. In doing so, I will also defend a compensation principle in Chapter 3, though I will use this term in a way that is different from Nozick's.

Lastly, as noticed by Hayenhjelm and Wolff, some of Nozick's discussion points toward a contractualist framework for dealing with the issue of risk imposition. Nozick suggests that the fairest way of deciding which risks to permit and which to forbid is to see whether the activity in question could "pay its way".[57] If it can pay its way and all exchanges occur on a truly voluntary basis, then supposedly it means that everyone affected ends up better off with the risk imposed. Then the activity is permissible in an intuitive sense—it is permissible because it is based on a voluntary contract agreed by everyone involved. However, it is doubtful that this could ever be achieved in a case of common risk imposition. In addition to the practical problem noticed by Nozick that it will involve huge transaction costs, a more serious concern is that, when a risk-imposing action impacts a large group of people, there will inevitably be some individuals who value security exceptionally highly,[58] which would ultimately make the price unacceptable for the risk imposer. It is therefore very likely that the idea of voluntary exchange will lead back to a world of paralysis where no scheme of common risk imposition could ever be agreed by all.

This concern may be addressed by endorsing Nozick's Principle of Compensation, which transfers the right to compensation from those subject to the risk to those prohibited from imposing the risk. But, probably more plausibly, the above discussion may point toward an alternative approach, namely the contractualist framework developed by Scanlon.[59] According to the Scanlonian contractualist approach, whether a case of risk imposition, or, more precisely, whether a general principle that regulates the type of risk-imposing activities, could be justified depends on whether it is a principle that cannot be reasonably rejected by appropriately motivated individuals in the contracting position.[60] Instead of asking what each would voluntarily agree to in order to advance their own interests, it seeks to identify principles that everyone could

---

[57] Ibid., pp. 76-77, p. 79.

[58] Either security in general or security against the specific risk in question.

[59] T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1998).

[60] Ibid., p. 5.

reasonably endorse should they be motivated by a desire for reasonable agreement.[61] It is therefore not vulnerable to the problem of paralysis faced by rights-based approaches. On the other hand, by emphasising that principles need to be justifiable to each individual in the contracting position, it supposedly also avoids problematic aggregation inherent in consequentialist approaches.

1.4 The Contractualist Approach to Risk Imposition

Scanlon himself has discussed the issue of risk imposition in various places in *What We Owe to Each Other* to demonstrate how his contractualism could justify intuitively permissible cases of risk imposition while ruling out other problematic cases. He suggests that, in order to evaluate whether a risk is justifiable to each individual, we could rely on the idea of "intrapersonal aggregation".[62] For example, in the case of car driving, instead of aggregating total benefit and cost across all affected individuals, the permissibility of the practice may be explained from each individual's perspective by aggregating the potential cost and benefit they themselves would experience throughout their life course. Even for a non-driver, the advantage of being able to benefit from the general practice of driving may outweigh the burden of the corresponding possibility of suffering automobile accidents. If this is the case, then everyone may reasonably prefer the social arrangement where activities like convenience driving are not prohibited.

Scanlonian contractualism has indeed been suggested by many philosophers as the most promising approach to the justification of risk imposition.[63] But, to be sure, evaluation based on interpersonal aggregation alone is not sufficient in deciding which cases of risk imposition are permissible. Some of the potential problems are recognised by Scanlon himself. For example, he considers a case in which a few individuals from the

---

[61] Ibid., p. 191; B. Barry, *Justice as Impartiality* (Oxford: Clarendon Press, 1995), p. 67. Also note that Scanlon chooses to use the term "cannot reasonably reject", see p. 189. However, I agree with Barry that there is no significant difference between the two terms. See Barry's explanation in *Justice as Impartiality*, p. 70.

[62] Scanlon, *What We Owe to Each Other,* pp. 237-38.

[63] See Hayenhjelm and Wolff, "The Moral Problem of Risk Impositions"; Oberdiek, *Imposing Risk*, ch. 5. For other works focusing the contractualist approach in justifying risk imposition, also see J. Lenman, "Contractualism and Risk Imposition", *Politics, Philosophy and Economics* 7 (2008), 99-122; R. Kumar, "Risking and Wronging", *Philosophy and Public Affairs* 43 (2015), 27-51; J. Frick, "Contractualism and Social Risk", *Philosophy and Public Affairs* 43 (2015), 175-223.

whole population are randomly selected to be killed for medical experiments.[64] However large an amount of net benefit this could bring for every member of this society (by comparing the health benefit with the expected cost of being chosen to die), the risk it imposes cannot be justifiable. What are the reasons or principles within his framework that can successfully distinguish normal risk-imposing cases like car driving from problematic cases like this medical experiment? Scanlon's own discussion is not entirely clear, but he seems to suggest that, in order for a risk to be justifiable, it has to be imposed with reasonable precaution. That is, the eventuated harm has to be a result that "occurs despite the fact that reasonable precautions have been taken".[65] In the medical experiment example, harm is deliberately inflicted on a few people rather than occurring "by accident", thus violating reasonable precaution.[66]

Scanlon's discussion has inspired a growing literature attempting to make sense of various acceptable and unacceptable cases of risk imposition using the contractualist framework. Lenman, for example, considers more scenarios of risk policy selection and proposes, in addition to the demand for reasonable precaution, the Aim Consistency Principle, which requires the aim of the risk-imposing action be consistent with the intention of precaution, and the Redistribution Principle, which requires the burden of risk be redistributed in a way that is justifiable to all.[67] On the other hand, Kumar suggests that what makes a case like medical experiment wrong is the violation of an independent principle which demands respect for people's bodily autonomy.[68] These discussions provide the starting point for me to identify principles that should guide the imposition of common risks on members of a society through its institutions and arrangements. I will discuss them in detail in Chapter 3.

It should be noted that, while I will base my account of common risk imposition on Scanlonian contractualism, Scanlon's original project (as well as many other studies of

---

[64] Scanlon, *What We Owe to Each Other,* pp. 208-9.

[65] Ibid., p. 209.

[66] Though, as I will discuss in Chapter 3, the problem of the medical experiment example may be more properly described as involving unjust intention rather than violating the demand for reasonable precautions.

[67] Lenman, "Contractualism and Risk Imposition".

[68] Kumar, "Risking and Wronging".

risk imposition) has a broader scope than mine. Scanlon is concerned with the subject matter of morality, that is, judgements of right and wrong. He employs the contract device to identify principles which determine the rightness or wrongness of an action.[69] His theory therefore applies not only to social institutions (or what Rawls calls the basic structure) but also to individual actions.[70] In the context of risk, it aims to identify principles that guide both public and private risk-imposing or risk-managing activities. My employment of the contractualist approach, on the other hand, has a more limited scope. My project is mainly concerned with those risks individuals are exposed to by living under common social institutions and arrangements. Instead of assessing the rightness and wrongness of all kinds of risk-related actions, I use the contractualist approach to evaluate the justice of institutions and public policies in dealing with common risks.[71]

Adopting the Scanlonian contractualist framework in a theory of justice is not a new idea. Barry, for example, employs the Scanlonian contract device in his account of justice as impartiality.[72] By comparing it with the alternative contract approach—Rawls's original position—Barry argues that the Scanlonian construction is a "more effective realisation of Rawls's objectives [of deriving fair terms of cooperation that can reasonably be endorsed by all] than his own original position".[73] Barry's main rationale is that the Scanlonian contract is better positioned than the original position to offer consistent answers to the following two questions—What is the motive for behaving justly? (both the Rawlsian and the Scanlonian answer: a sense of fairness or justice); What is the criterion for a just set of rules? (the Rawlsian answer: rational choice behind the Veil of Ignorance; the Scanlonian answer: rules that could be reasonably endorsed

---

[69] Scanlon, *What We Owe to Each Other*, pp. 1-5.

[70] Ibid., pp. 6-7. However, as noted by Scanlon, the scope of his project is narrower than how terms of moral rightness and wrongness are commonly used. It only explains the part of morality that could be described as "what we owe to each other". It does not include, for example, the morality of certain forms of sexual conduct or conduct that leads to the destruction of animal species.

[71] For the contrast between rightness or wrongness of actions and the justice or injustice of institutions, see Barry, *Justice as Impartiality,* p. 68.

[72] For a more practical application of the Scanlonian contractualist framework in theories of justice, see Daniels, *Just Health.*

[73] Barry, *Justice as Impartiality,* p. 67.

by people on equal footing, which logically connects with the answer to the first question).[74] I agree with Barry here and will employ the Scanlonian contractualist framework to assess the justice of social risk imposition.

In particular, I will argue in Chapter 3 that the underlying logic which leads to the endorsement of the reasonable precaution principle in most accounts of legitimate risk imposition should also justify a compensation principle. Unlike Nozick's account, compensation is not owed because the victim's right is violated by the risk-imposing activity. Rather, the principle holds that compensation should be guaranteed, just like reasonable precautions should be performed, so as to reasonably reduce the negative impact on the potential victims of risk materialisation. Moreover, the burden of paying for compensation should be shared fairly among all those who benefit from the imposition of the risk. Since almost all members of a society benefit directly or indirectly from a scheme of common risk imposition, they are all required by reasons of fairness to contribute to the payment of such compensation. My argument then effectively justifies a demand for fair risk sharing by all.

1.5 Risk Sharing and Social Cooperation

Risk sharing represents a distinctive form of social cooperation but it has not received much attention in political philosophy. In fact, as pointed out by Heath, when political philosophers refer to the society as a scheme of social cooperation, few discuss in detail the mechanisms through which it generates cooperative benefits.[75] By looking at the way they describe the advantages of social cooperation[76], it seems that what philosophers have in mind are usually efficiency gains generated in production and trade through the division of labour and economies of scale. What is often neglected in their discussion is that there are other mechanisms of cooperation through which

---

[74] Ibid., pp. 46-51.

[75] J. Heath, "The Benefits of Cooperation", *Philosophy and Public Affairs* 34 (2006), 313-51.

[76] For example, Rawls referred to social cooperation as "mak[ing] possible a better life for all than any would have if each were to try to live solely by his own efforts." Gauthier described the benefits of social cooperation as "an explosion in the quantity of material goods and in the numbers of persons…an unprecedented rise in the average life span, and a previously unimaginable broadening of the range of occupations and activities effectively accessible to most individuals." See Rawls, *A Theory of Justice*, p. 109; D. Gauthier, *Morals by Agreement* (Oxford: Clarendon Press, 1986), p.18.

individuals may help each other to achieve objectives they could not reach on their own. These additional forms of cooperation may include, according to Heath, self-binding, information transmission and risk pooling.[77]

Risk pooling is a distinctive cooperative mechanism because it functions not by promoting social production but by reducing uncertainty. In its strict form, based on the law of large numbers[78], people who pool similar, independent risks together could collectively transform significant, unpredictable individual losses into small, pre-fixed costs for everyone, thus relieving each other from the burden caused by significant economic uncertainty. But risk pooling may generate cooperative benefits even when the law of large numbers does not hold in this strict sense, say when the number of participants is not sufficiently large or when the risks they are exposed to are not exactly equal or fixed. In such cases, though certainty in future payoffs cannot be guaranteed, pooling risks according to some prior agreement, explicit or implicit, may still help participants to smooth out negative shocks and to alleviate disastrous consequences. Mutual aid practices found in various preindustrial or even archaic communities often reflect this less rigid form of risk pooling.[79] Since the term risk pooling is sometimes reserved for practices involving strict application of the law of large numbers, I will instead use the term risk sharing to cover both types of case.

Risk sharing, moreover, may also refer to a different type of scenarios where the law of large numbers is not at work. That is when different people share both the potential upsides and the potential downsides of a collective venture, thus spreading possibly significant losses across individuals. The way Islamic finance interprets equity investment may provide an example of risk sharing in this sense.[80] A business of a certain scale may involve too big a risk for one individual to bear. By having multiple

---

[77] Heath, "The Benefits of Cooperation", p. 319.

[78] The law of large numbers is a theorem in probability theory which states that, as a sample size grows, its mean gets closer to the average of the whole population. Therefore, when the number of people who are independently exposed to the same risk gets larger, it is more likely that the number of people in this group who actually experience the materialisation of the risk will converge to the expected number.

[79] See, for example, P. Kropotkin, *Mutual Aid: A Factor of Evolution* (London: William Heinemann, 1902).

[80] H. Askari et al., *Risk Sharing in Finance: The Islamic Finance Alternative* (Singapore: John Wiley & Sons Asia Pte. Ltd., 2012), pp. 83-84.

business partners who pool their resources together, the risk involved in running the business is then shared by all the stakeholders and is therefore more manageable for everyone. Correspondingly, they also have a claim against the potential profit generated from this pool of assets that is proportionate to their investment (that is, the risk they bear). In fact, this understanding of equity finance is not an invention by modern Islamic scholars. Dating back to seventeenth-century Europe, risk-bearing business partners who have supplied capital to a joint venture were similarly considered by jurists to have a right to a share of the profit as the "price of the peril".[81]

While cooperative practices of risk sharing in various forms have had a long history, as pointed out by Heath, since an understanding of statistical probability and the law of large numbers is a relatively late development, risk sharing has not been formally recognised by social theorists as a mechanism of cooperation until the mid-nineteenth century, when Emile de Girardin made the first serious attempt to expound this view.[82] In his book *La politique universelle*, de Girardin argued that the primary function of the social contract was to found a system of "universal insurance".[83] But even since then, there has been a continuous conceptual bias toward focusing on market exchange and neglecting other forms of social cooperation such as risk sharing.

Heath believes that this bias has led to the dominance of what he calls a residual view of the welfare state.[84] According to this view, the central economic role of the state is either to intervene in the private market to promote efficiency-maximising transactions or to fulfil redistributive goals that are motivated by egalitarian reasons. However, if we recognise social cooperation may take other forms, then, in addition to the efficiency-based reason for the state to facilitate production and trade and the egalitarian reason for the state to facilitate redistribution, there may also be an efficiency-based reason for the state to promote, for example, the form of cooperation represented by risk sharing. Viewed in this light, the tension involved in justifying the welfare state, argued by

---

[81] L. Daston, *Classical Probability in the Enlightenment* (Princeton: Princeton University Press, 1988), pp. 19-20.

[82] Heath, "The Benefits of Cooperation", p. 324.

[83] E. de Girardin, *La Politique Universelle* (Paris: Librairie Nouvelle, 1855); quoted in Heath, "The Benefits of Cooperation", p. 324.

[84] Ibid., p. 345.

Heath, may not always be a conflict between efficiency and equality; it may instead be a conflict between two types of efficiency gains.[85] For instance, the debate over the privatisation of pension systems reflects the conflicting intuition of those who prioritise gains in economic transactions versus those who prioritise gains in risk sharing.[86] Private pension schemes, by allowing different individuals to choose between pension contributions and other forms of consumption or investment according to their own preferences, may improve market efficiency through the complementarities between those who have a greater interest in consuming now and those with a greater interest in consuming later.[87] Universal public pensions, on the other hand, may generate higher cooperative benefits in risk sharing from a more diverse risk pool.

While it may be overlooked in political philosophy, such an efficiency-based justification for the risk-sharing function of the welfare state as suggested by Heath is in fact the predominant discourse in welfare economics. In modern societies, voluntary participation in risk sharing is mostly facilitated by the private insurance market through the provision of private insurance policies. State intervention in the insurance market, in the form of either regulation or the provision of social insurance as a substitute for private insurance, is regarded as justifiable according to welfare economics when it improves overall market efficiency.[88]

Private insurance is known to be subject to the problem of adverse selection[89] and moral hazards[90], which threatens the efficiency of the market. Sometimes these problems could be effectively dealt with by private insurers by specifying the scope and terms of cooperation. For example, to lessen the problem of moral hazard, insurers may

[85] Ibid., p. 346.

[86] Ibid., pp. 346-47.

[87] Ibid., p. 342.

[88] I use efficiency here as a technical term in welfare economics. A market is efficient when it maximises the sum of economic welfare of all players, including consumers and producers.

[89] Adverse selection may lead to market failure when an insurer is unable to obtain sufficient information about buyers so as to exclude those carrying a greater risk than what has been assumed in setting the price of insurance policies. For the classic economic analysis of adverse selection, see G. A. Akerlof, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *The Quarterly Journal of Economics* 84 (1970), 488-500.

[90] Moral hazards refer to the potential market failure when individuals covered by insurance are incentivised to act in a way that increases the likelihood of the adverse consequence their insurance is aimed to cover.

introduce mechanisms of deductibles and coinsurance to have buyers share part of the cost incurred, thus reducing their incentive to act carelessly or to deliberately take additional risks. Insurers may also conduct inspections and deny payment for losses caused through actions prohibited in the insurance contract. Another option is to penalise those who have claimed for losses by increasing their subsequent premiums. This tactic is commonly employed by automobile insurers to encourage cautious driving.

With regard to adverse selection, insurers often introduce underwriting policies such as risk classification to charge differential premiums for people with different levels of risk, and they may even deny coverage for those carrying too high a risk. Take private medical insurance as an example. Since a person's health condition is often negatively correlated to age, according to statistics, insurers can use age as an indicator for assessing health risk and thereby adopt a price schedule with higher prices for older buyers. Some may also set a maximum age limit for new policyholders.[91] Insurers may also require potential buyers to provide personal medical records, based on which they could further exclude those already diagnosed with certain diseases.

While these measures have been effective in preventing market failure for some types of insurance, it is still difficult for private insurers to offer coverage for other risks. Unemployment insurance, for example, is subject to both the problem of high correlation[92] and the difficulty in preventing adverse selection and moral hazard, thus rarely being sold by private insurers.[93] Since market failure would result in many people's failing to obtain coverage even when they want to, this would create a reason for the state to intervene and provide social insurance to promote market efficiency.

---

[91] These measures do not necessarily imply that there is no private insurance coverage for old people. They are only required to start buying insurance before they reach certain ages.

[92] The law of large numbers only holds when the risk faced by each individual is not highly correlated, that is, it is not to eventuate simultaneously for a large group of people. Unemployment risk tends to strike many workers at the same time during economic recessions, so the law of large numbers breaks down for this type of risk.

[93] See detailed analysis in N. Barr, *The Welfare State as Piggy Bank: Information, Risk, Uncertainty, and the Role of the State* (Oxford: Oxford University Press, 2001), pp. 36-40.

Concerns for correcting market failure may arise not only on the supply side of the insurance market but also on the demand side. That is, individuals may also lack the incentive to participate in risk sharing. Barr points out that failure to buy insurance coverage may create various negative externalities such as the cost imposed on taxpayers to give social assistance to an uninsured person (or the cost of leaving him to starve or to die), the impact of non-insurance on his dependents, the consequence of health hazard if non-insurance results in no intervention for contagious diseases and so on.[94] In addition to these explicit costs, Goodin also suggests that, since a person's life plans are inevitably intertwined with others', a certain level of stability in income guaranteed by insurance contributes to other people's being able to enter into long-term commitments in their lives. This makes disruption in income a case involving negative externality that is similar to contagious diseases.[95]

Due to the existence of externalities, individuals are unlikely to have the rational incentive to buy insurance coverage at a level that is optimal from the collective perspective. In such cases, the state may, again, be justified to intervene in the market in order to maximise efficiency. But the concern to correct for externalities does not necessarily justify compulsory participation in social insurance. Providing tax allowances for buying private insurance schemes (for example, deducting contributions to such schemes from the total amount of taxable income) may be more effective in achieving the purpose.

Hence, according to welfare economics, when the state should intervene in the private insurance market and what form of state intervention it should take depend on whether the choice would improve efficiency either from the supply side or the demand side of the market. Recall that Heath argues that the justification for the welfare state in political philosophy tends to neglect efficiency-based reasons for the state to facilitate the cooperation of risk sharing. Here, on the other hand, the justification for state intervention in risk sharing offered by welfare economists seems to focus solely on efficiency. When a type of insurance can be efficiently offered by the private market, it

---

[94] Ibid., p. 42.

[95] Goodin, "Stabilizing Expectations", p. 550.

is not concerned that efficiency may be achieved through measures that undermine other values such as equality and fairness.[96] For example, when private insurers deny coverage for people they deem too risky (such as the old, the sick or the poor), such decisions restrict those people's access to a secure and stable life and may also threaten their self-respect and cause a sense of exclusion. When insurers charge some people higher premiums based on characteristics over which they have no control (such as age, race or inborn health issues), this effectively penalises them for carrying a higher involuntary risk and may also arouse a feeling of stigmatisation.

In fact, empirical evidence shows that private insurers' decisions about exclusion or risk classification have often been contaminated by discrimination and prejudice. As shown by Murray, their disadvantageous treatment of lower-income communities and ethnic and racial minorities is often based on stereotypes rather than objective, empirical data.[97] On the other hand, even when differential treatment is well supported by statistical evidence, statistics are based on past information and only demonstrate the existence of a correlation between certain individual traits and the degree of their riskiness. To exclude or to classify a potential buyer based on some statistically significant factors is effectively to judge him based on past information about a group of people with whom he happens to share some common traits, which could be morally arbitrary. Moreover, even if, as a matter of empirical fact, a certain trait does causally increase the risk faced by a person, holding him responsible for such a causal relationship still calls for justification. Suppose that belonging to a disadvantaged ethnic minority group has causally increased a person's health risk due to poor hygiene and

---

[96] In fact, welfare economists such as Barr do not deny that normative considerations of fairness, equality and freedom should be taken into consideration in making policy decisions. But Barr argues that, once the trade-off between different values is agreed upon at the objective-setting stage, it is then a pure efficiency question in selecting the appropriate method to reach that objective. So, in the case of insurance, it is a political question whether a certain type of insurance should be made available but, once this is decided, it is then an economic question whether it is the state or the free market that will provide the coverage more efficiently. See N. Barr, *Economics of the Welfare State* (Oxford: Oxford University Press, 2012), p. 4.

However, what I want to show is that the decision between leaving a risk to the private market and covering it through universal insurance also inevitably impacts values such as fairness and equality and is therefore not a pure economic question.

[97] W. Murray, "Homeowners Insurance Redlining: The Inadequacy of Federal Remedies and the Future of the Property Insurance War", *Connecticut Insurance Law Journal* 4 (1997), 735-62, p. 761.

nutrition in his community, it may nevertheless be reasonably argued that he should not be held responsible for such misfortune.

Recognising the moral arbitrariness of differential treatment in private insurance practices, some governments have responded by prohibiting classifying insurance buyers based on social categories such as race, sex, age and national origin.[98] Such laws effectively require that any additional risk associated with these social factors be shared equally among insurance buyers. However, if fairness or equality is what motivates such regulations, then it still seems arbitrary to limit the sharing of risks which members of a society are commonly exposed to only among those who actually decide to buy commercial insurance policies.

In fact, as I will argue in Chapter 3, since many risks individuals living in contemporary societies are subject to are collectively imposed on them through the design of social institutions and public policies, justice requires that those who benefit from the imposition or retainment of such risks share the corresponding burden. If my argument is successful, then individual reasons for opting out of private insurance schemes, such as a greater risk tolerance level or attaching a higher priority to other life projects, could not in themselves justify someone's refusal to participate in the social cooperation of risk sharing together with fellow citizens.

1.6 The Concept of Risk and Thesis Plan

The main goal of this thesis is to offer an account of just social risk imposition and fair social risk sharing. Before I list out my thesis plan, let me briefly clarify how I will use the term risk. It seems that people may use risk in at least two different senses—it may refer to the possibility of a bad outcome (such as the risk of car accident); or it may be taken to mean, more broadly, the uncertainty of an outcome, which may turn out to be either good or bad (such as the risk of investment).

In this thesis, I will take risk to mean the possibility of a bad outcome. Restricting the use of the term in this way helps to distinguish risk from the potential benefit it may

---

[98] D. Stone, "Beyond Moral Hazard: Insurance as Moral Opportunity", in *Embracing Risk: The Changing Culture of Insurance and Responsibility*, eds. T. Baker and J. Simon (Chicago: University of Chicago Press, 2002), p. 72.

bring about, which can be either guaranteed or merely possible. For example, when an instance has both potential upside and potential downside, I will treat it as involving a risk of loss accompanied by a possibility of gain. The distinction makes it easier for us to consider cases where taking or imposing a risk may be justifiable because of the benefit it makes possible.

When we refer to risk as the possibility of a bad outcome, this possibility may sometimes be a fixed and known probability. For example, if I play a dice-throwing game in which I will lose $5 if the dice lands on six, the probability of the risk is 1/6, which can be derived from the setting of the game. On the other hand, the possibility can also be a fixed but unknown probability. For example, instead of throwing a dice, I now need to draw a ball from a black box containing a mix of red and white balls and I will lose $5 if I get a red ball. The probability of losing money is fixed but I lack knowledge about its specific number.

Meanwhile, the possibility of bad outcome may not always be fixed. For example, the chance for me to have an accident by driving a car may vary depending on factors such as the speed I drive at, my level of concentration, the quality of the car, and so on. In such cases, since I have a certain degree of control over the eventualisation of the bad outcome, its probability is not a fixed number like that of throwing a dice. But, in this case, it is still possible to estimate the probability by looking at other similar cases. For example, by drawing on past data of car accidents based on my region, age group, level of experience and so on, we may have a statistical estimation about how likely it is for me to have a car accident.

For other risks we may not be able to accurately estimate its likelihood based on statistics, maybe because the event is rare and we lack sufficient data to make such an estimation (such as the risk of meteor strike) or maybe because there is a lot of room for people to exercise agency and influence the outcome (such as the probability of failure for me to start a business). Finally, there are also risks which we know very little about. In addition to unknown probability, we may also lack knowledge about what kind or scale of harm it may eventualise into, though we vaguely know that something bad

might happen. An example might be the risk of climate change understood by people a century ago.

According to a famous distinction between risk and uncertainty made by Knight,[99] some of the above cases may fall into the category of uncertainty rather than risk. For Knight, pure risk is an instance whose probability is fixed, so that we can either derive the number based on probability theory (in the case of dice throwing) or estimate it through repeated observation (in the case of ball drawing from the black box). Pure uncertainty, on the other hand, refers to the case where we completely lack knowledge about its outcome and probability (Knight's examples include scientific discovery and running a business). All other cases can be seen as lying somewhere along the continuum between pure risk and pure uncertainty. The distinction between risk and uncertainty can also been seen in decision theory, though, in this context, uncertainty is used in a more restricted way—it refers to the case where we know what are the possible outcomes but completely lack knowledge about their respective probability, whereas risk represents the standard case where we know both the range of outcomes and their probabilities.

The distinction between risk and uncertainty, however, is not very helpful for my discussion. This is because, first, what Knight takes to be uncertainty includes a mix of different cases. For example, some natural phenomenon may appear as pure uncertainty in the sense that we completely lack knowledge about it but its chance of happening may in fact be fixed. The uncertainty involved in a start-up business, on the other hand, is due to the various factors including the business decisions to be made by the owner, so its probability of failure is never a fixed one. Besides, the distinction between risk and uncertainty is also unnecessary for this project because, as I will show (especially in Chapter 5), my account of just risk imposition and fair risk sharing may apply to both types of cases. Therefore, when I refer to a case as involving risk, its possibility of bad outcome need not be fixed. I will, however, borrow Knight's term "pure risk" to refer to the dice-throwing kind of risk whose outcome is purely random and can therefore be described with a fixed probability.

---

[99] F. Knight, *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin, 1921).

With this clarification on how I will use the term risk, let me now list out my thesis plan. The rest of this thesis will be organised as follows. Chapter 2 starts by examining how risk impacts autonomy as self-authorship. I argue that the presence of risk undermines people's prospect for autonomy mainly through its negative impact on the background options available to individuals. The impact may not only be reflected in the smaller number of accessible options but also in the reduced robustness of their availability, the less optimal combination of long-term and short-term options, and so on. I then examine an opposite view that risk sometimes promotes rather than undermines autonomy. I argue risk itself does not have positive impact on autonomy, but risk or risky options may be retained as a by-product of social arrangements that may nevertheless enhance the background conditions for autonomy.

Chapter 3 then looks at the justification for common risk imposition. I employ the Scanlonian contractualist framework to generate principles for just risk imposition and argue that a scheme of common risk imposition should be governed not only by principles concerning its outcome (such as the presumptive worthiness principle and the fair distribution principle) but also principles concerning the nature or the characteristics of the risk in question—whether its potential harm is merely foreseen rather than intended and whether reasonable precautions are taken to prevent it from materialising. I then suggest that the rationale that justifies the principle of reasonable precaution would also justify a principle of compensation which requires that compensation be paid to victims of risk materialisation so as to reasonably reduce the burden for them. This duty of paying for compensation should be shared fairly among all those who benefit from the scheme of risk sharing. This includes almost all members of society in the case of social risk imposition.

Chapter 4 addresses the question whether, in allocating the burden of risk sharing, personal responsibility should also be taken into consideration. In particular, if individual actions play a role in causing a common risk to materialise or if people could have avoided or managed their own risk exposure, should they bear a larger share of the burden than what would otherwise be allocated to them? To answer this question, I start by discussing Scanlon's own account of substantive responsibility in cases without risk.

I then extend my conclusions to risk-involving choices and examine two supplementary accounts—the full consequence view and the expected return view—that may bridge the gap created by risk between a person's choice and its consequences. I argue that both views may apply to some risks but not others, and their justifiability depends on what we reasonably expect people to do in the specific situation, including whether we expect that they could reasonably avoid the risk exposure or reduce the chance of risk materialisation. Reasonable expectations of this kind should be reflected in (though not fully determine) the allocation of substantive responsibility.

Chapter 5 then looks at the case of intergenerational risk imposition and examines if imposing risks on future generations calls for more cautious principles than what I have justified in the contemporary context. I consider the various arguments for adopting more conservative principles (such as the Precautionary Principle) based on reasons such that we stand in a special relationship with future generations, that we do not know their risk preferences, and that the risks we impose on them exhibit special features such as potential catastrophic outcome, unknown probability and being long-term risks. I conclude that all these concerns can be successfully addressed by the principles I have developed in Chapter 3 and 4 as long as we have a dynamic understanding of their demands, so my account of just risk imposition can be extended to cover the intergenerational cases. I also reply to a few potential objections concerning the fair burden allocation across time.

Chapter 6 turns to a slightly different problem. Based on my account, a scheme of fair risk sharing should cover people with severe disabilities and people from future generations. However, it is often suggested that the scope of social cooperation cannot be extended to these two groups of people due to a lack of reciprocity between them and contemporary people with normal capabilities. I examine this concern by first clarifying that reciprocity can be understood either as a demand of justice or as the conditional motivation for sustaining cooperation based on principles of justice. I then argue that including these two groups of people does not violate reciprocity as a demand of justice as long as it is taken to be subordinate to impartiality. On the other hand, the establishment of risk sharing institutions may encourage a more generous form of

reciprocity as the motivation for justice and hence enhance the stability of social cooperation.

Finally, Chapter 7 examines the implications of my account for the design of risk sharing institutions. In particular, I will argue that my account of fair risk sharing may provide an additional justification for universal social insurance and also guide the fair design of such programmes. I will use a case study of the funding of pension schemes to illustrate how my account may be applied.

Chapter 2 — Risk and Autonomy

As I have mentioned in the introduction, in this thesis I will use the conception of autonomy to assess the impact of risk on individuals. This chapter examines the ways in which risk may affect a person's prospects for living an autonomous life and how such impact can be evaluated. To do so, I will adopt Raz's conception of autonomy as self-authorship.

According to Raz, people have a general interest in autonomy as self-authorship.[1] That is, we want to live a life that we could look back upon and recognise as our own making. Self-authorship, however, does not mean that a person could fully determine what happens in his life. Among the various factors that may condition a person's prospect of authoring his own life stories, the existence of risk introduces additional indeterminacy into people's life courses and hinder their ability to predict and to control the way their lives unfold. Unlike other accounts of autonomy which are concerned with assessing the level of autonomy exhibited in an agent's particular choices, Raz's account focuses on the living of an autonomous life. Those who are autonomous in this sense can "shape their life and determine its course", thereby becoming a "part author of his own life".[2]

Raz lists three preconditions for autonomy as self-authorship. Firstly, to be autonomous, a person should enjoy a sufficient degree of independence. That is, he should be free from coercion and manipulation when making decisions.[3] This condition of independence is also what many literatures focus on when assessing the level of autonomy exhibited by a person in making a certain choice.[4] But living an autonomous life requires more than the mere making of independent choices. For a person to make

---

[1] J. Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986).

[2] Ibid., p. 154.

[3] Ibid., pp. 377-78.

[4] See, for example, H. Frankfurt, "Freedom of the Will and the Concept of a Person" and G. Dworkin, "The Concept of Autonomy", in *The Inner Citadel: Essays on Individual Autonomy*, ed. J. Christman (Oxford: Oxford University Press, 1989), both of which focus on the psychological condition of autonomy, and G. Sher, *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997), which emphasises the importance of responsiveness to reason (and being free from manipulation) in forming desires.

choices that fit into his own life plans, he also need an appropriate set of options in the background to choose from. This is the second precondition for autonomy as self-authorship.[5] Thirdly, autonomy also requires certain internal capacities.[6] For a person to live a life that is shaped by his choices, he should be capable of forming intentions, setting goals, evaluating available options and making informed decisions.[7] If we accept that people have a general interest in autonomy, then institutions should be designed in a way that promotes the satisfaction of these preconditions, thereby securing a fair prospect for everyone to live a self-authored life. This is what I will take as a background principle of justice when addressing the main question of the thesis—how a society should deal with risk justly.

In this chapter, I will start by analysing the impact of risk on the background condition for autonomy. In Section 2.1, I focus on the precondition concerning the availability of options and examine how the range, compositional structure and robustness of background options may impact people's prospect for living a self-authored life. With this analysis, I then examine how risk may undermine such prospect through its negative impact on the available options. I argue that the impact is not only quantitative (in terms of the number and variety of options) but also qualitative (in terms of the structure of different types of options, the robustness of their availability, and so on).

In Section 2.2, I look at an opposite view which states that, instead of undermining autonomy, the existence of risks may in fact be valuable for autonomy. I consider two types of argument. First, risk may be valuable in the sense that it is the inevitable by-product of providing certain options that are valuable for autonomy. Second, some suggest that the existence of risk may itself feature as a precondition for autonomy as self-authorship. Here I focus on Tomasi's argument about the value of economic risk and argue that it at best shows the value of keeping certain risky options available in the background, not the value of being exposed to risks as such.

---

[5] Raz, *The Morality of Freedom*, pp. 373-77.

[6] Ibid., pp. 369-73.

[7] Ibid., p. 371.

I then turn to the precondition for autonomy concerning people's ability to make rational choices in Section 2.3. I argue that, when faced with risky options, deviation from what the classic rational choice theory recommends does not necessarily mean that a decision is irrational. However, at the same time, people do often exhibit cognitive biases when making risk-related choices. Since such biases may hinder people's ability to realise their plans or achieve their life goals, creating an autonomy-supporting environment could imply paternalistic intervention in people's risky options. I will argue that the demand of maintaining a background condition that is conducive to autonomy at best justifies a basic safety net or soft paternalistic policies known as nudging. The justification cannot be extended to the coercive removal of risky options in general or the enforcement of mandatory insurance programmes.

2.1 Risk and Background Options

To see how risk may set back people's interest in having a fair prospect of autonomy, I will start by analysing its negative impact on one of the preconditions listed by Raz—the availability of background options. According to Raz, to be autonomous a person must be provided with an adequate range of options.[8] The kind of option set available in the background would affect how likely a person will be able to shape his life through his choices. Raz gives a few examples to illustrate the relationship between the availability of background options and a person's prospect of autonomy. I will examine these considerations so that we have the conceptual framework for assessing the impact of risk on autonomy.

2.1.1 Availability of options as a precondition for autonomy

i) Number and variety of options. Intuitively, the larger the set of options available in the background, the more likely that a person could choose from it those that fit into his life projects and hence the greater prospect for him to realise a life of his own design. If, by contrast, the person is faced with a very restricted set of options, then his life course is more likely to be shaped by the limited possibilities than by his autonomous choices. But the number of available options is not the only consideration in evaluating the

[8] Ibid., p. 373.

background condition for autonomy. For the same number of available options, a set with greater variety is more likely to contribute to a person's prospect of autonomy than a set with largely identical options. For example, a person's career choice is more likely to be autonomous if the options available to him include being a teacher, a policeman, an office worker, and so on, compared to a situation where he is only given the choice to be an employee for one of many similar companies.

In this sense, the value of an option lies in its potential contribution to an individual's prospect in authoring his own life stories. A larger range (number and/or variety) of options often implies a greater chance that the person could choose among them those that are better suited to his plans and projects, but, beyond a certain point, the marginal improvement in his prospect of autonomy that could be brought about by one additional option may become insignificant. After all, not all options are relevant to his chosen projects and he could only, at any given point of time, choose one or a few among the available options. In fact, beyond a reasonably adequate range of options, the significance of having one additional option may easily be overweighed by considerations concerning other features of the option set, which will be discussed shortly.

Supposedly, the same reasoning also applies to the elimination of one option. It seems that, given that a person still has an adequate range and variety of options to choose from, the impact of eliminating one particular option is also likely to be limited. However, in this case, as noted by Raz, it makes a difference if the option to be removed is one that the person is already committed to as part of his life plans.[9] If this is the case, then the negative impact of its removal may be much more significant than if the closed option is only remotely relevant to his projects.

Besides the observation that choices are not equally valuable for people's prospect of self-authorship, it may be further suggested that sometimes one additional option can even have negative impact on the overall choice situation. Gerald Dworkin lists out a

[9] Ibid., p. 411.

few scenarios of this kind.[10] For example, my analysis assumes that having one more option is always better because people could simply not choose it if they do not need it. In reality, however, having a popular option available may create social pressure for those who do not want it to follow others' choices.[11] Or, having an additional option may undermine the symbolic value of existing options. Dworkin's example is that, if the option of selling blood is available, then the nature of the existing option of blood donation will be changed.[12] Among these considerations, one case considered by Dworkin is especially relevant for my discussion here. Dworkin points out that some options, once chosen, may harm people's interest but people can nevertheless be tempted to choose them (gambling may be such an example). If this is the case, then removing this kind of options may enhance rather than undermine a person's overall prospect of autonomy throughout his life. I will come back to this paternalistic consideration in the third section.

ii) Compositional structure: survival vs. non-survival options. Besides the number and variety of options, another important consideration is that the option set should not be dominated by survival decisions. A person is not autonomous if he has to devote all his physical and mental energy to making choices so as to maintain minimal conditions of life.[13] This is well illustrated by Raz's example of the Hounded Woman.[14] A woman finds herself on a desert island with wild animals constantly hunting for her. She has to use all her time and efforts to try to escape from the beasts with no chance of doing or thinking of anything else. In this case, while she does have various options to choose from at any given point of time, the woman hardly enjoys an autonomous life because, other than struggling to stay alive, there is no scope for her to carry out her own projects.

iii) Temporal structure: long-term vs. short-term options. Since a person's prospect for autonomy should be evaluated over the horizon of his life course, the temporal structure

---

[10] G. Dworkin, "Is More Choice Better Than Less?", *Midwest Studies in Philosophy* 7 (1982), 47-61.

[11] Ibid., pp. 51-52.

[12] Ibid., pp. 53-54.

[13] Raz, *The Morality of Freedom*, p. 155.

[14] Ibid., p. 374.

of the option set is also a relevant consideration. In particular, some options are long-term choices with pervasive consequences while others are short-term in their effects. An autonomous life requires the availability of both types of options and a balance between them[15]—on the one hand, we need options with impact lasting into the future so that we are able to make long-term commitments or pursue long-term projects; on the other hand, we should also be allowed to make more trivial and short-term decisions such as what meals to eat and what clothes to wear, which are equally important for us to live a meaningful life.

iv) Robustness of availability. The temporal dimension introduces, moreover, a dynamic aspect that is relevant in evaluating the background condition for autonomy. The range of options available to an individual at a given point of time would always be impacted by his previous decisions. However, if the background condition is such that his choice at one point of time may significantly restrict his option set at a future point of time, then it is a less favourable condition for autonomy compared to a situation where a meaningful range of options is always available. For example, imagine a society in which each individual could freely choose his occupation before he turns eighteen years old but, after the decision is made, could never change his mind again. At the point of making his career choice, it may be said that the person enjoys a satisfying range of options. However, his decision would then significantly restrict the range of options in his future life. People in this society have a lower prospect of living an autonomous life compared to a society in which people enjoy greater freedom of occupation through their lives.

The concern for the robustness of the availability of background options also implies that an autonomous life should have sufficient scope for a person to regret his decisions or revise his life plans.[16] That is, the availability of a minimally satisfying range of options should not depend too heavily on a person's previous decisions. Surely, choices would always influence the range and quality of options available to a person, otherwise

---

[15] Ibid.

[16] For discussion on the importance of being able to revise life plans, also see M. Fleurbaey, *Fairness, Responsibility, and Welfare* (Oxford: Oxford University Press, 2008), pp. 179-183; N. Daniels, *Just Health: Meeting Health Needs Fairly* (Cambridge: Cambridge University Press, 2008), p. 61.

he could hardly recognise his life as one that is shaped by his decisions and actions. However, this does not mean that the background condition has to be such that he will always be left to bear pervasive and even disastrous consequences for his decisions. Compare, for example, a society in which individuals are allowed to file for bankruptcy with one in which they do not have such a right. By granting people the chance to have a fresh start after making poor financial decisions, the former society offers greater protection for people's prospect of living an autonomous life by guaranteeing an acceptable range of options throughout their lives.[17] It offers, in other words, a more robust background condition for autonomy.

2.1.2 How risk may undermine autonomy

Based on the above analysis, we can now see how the existence of risk may undermine a person's prospect of autonomy by impacting the various features of the background options. Two kinds of autonomy-undermining consequences are often associated with risk. First, when a risk eventuates into actual harm, it undermines the victims' prospect of autonomy in a quite straightforward way. If the outcome involves physical harm, it may injure people's normal functioning and potentially disrupt their life projects. If the negative consequence is mainly economic, such as losing one's assets or steady source of income, it may also leave the person under great financial stress, thus closing off many life options or forcing him to give up long-term life plans and focus on short-term survival. Even when a single shock resulting from a risk is mild and does not have disastrous consequences, small but frequent harm or losses may similarly remove valuable long-term options, because the constant need of dealing with unstable life conditions may exhaust a person's energy and prevent him from focusing on more important life projects.

But these actual bad outcomes are not the only source of harm a risk may bring to people. In fact, if these are the only negative impacts, then there is nothing unique about risk in affecting people's prospect of autonomy—it compromises autonomy in the same

[17] Surely, such protection comes with a cost, often on the level of autonomy of other members of the society, so whether the policy best protects autonomy for all need to be assessed by weighing its impact on different individuals. Here, what I want to show is that, for a particular individual who wishes to file for bankruptcy, his interest in autonomy is better protected in a society where it is possible for him to do so.

way as other actions that cause direct damage to people (such as intentional infliction of physical harm), and it only does so when a risk eventuates into actual harm. However, a risk may undermine autonomy in a different way. A person's prospect of living an autonomous life may be undermined by the mere existence of risk because risk affects the range, structure and robustness of background life options available to individuals.

First, attaching risk to available options may restrict the range (number and/or variety) of the option set. In most cases, the mere existence of risk does not remove existing options completely in the sense that it is no longer feasible for people to choose those options at all, but it could render the otherwise valuable choices more costly and hence less likely to be desirable. The availability of an option always comes in degrees. When an option is said to be removed by, say, state regulation, instead of being eliminated, it is merely made prohibitive by the threat of coercion or financial penalty. Similarly, when a risk is attached to an option, although it does not make the option completely inaccessible, the possible harm could make it no longer an attractive or acceptable choice. For example, the option of taking a university degree (suppose that it is available) would be less attractive if it is accompanied by the financial risk of not being able to find a job yet having to pay back a large student loan after graduation.

Of course, the (partial) removal of an option is not always problematic in itself. As discussed, its impact should be assessed according to the extent to which it compromises individual's prospect of self-authorship. How much less valuable an option is made by the existence of a risk would depend not only on the possibility and magnitude of the risk but also on the overall range of options that is available to an individual. In the above example, if the prospect of getting a job and paying back the loan upon graduation is good enough, or if other similar educational opportunities are available with much less financial burden, then the financial risk would have a less negative impact on autonomy. The impact also varies among individuals. If the risk mainly involves economic shocks, a person with enough disposable assets to guard against the potential downside would be less impacted than someone who is already facing a precarious financial situation.

Second, when people are actually exposed to a risk, either because they cannot avoid it or because they take it on by choosing an option attached with risk, most of them would not be indifferent to their risk exposure and would divert their external resources as well as internal energies to deal with it. For example, if a person faces precarious employment prospects, he may be forced to set aside a significant proportion of income for rainy days instead of investing it for his main life projects. Even for a person with more secure income sources, he may need to make various risk management decisions, such as the purchase of insurance, in order to avoid being hit by unexpected shocks. Besides financial and cognitive burdens, some people may also feel anxious and even fearful when living with such economic risks, thus suffering additional emotional burdens.[18] These burdens in turn affect the range and structure of options available to them. The need for anticipating and dealing with risks is likely to close up longer-term opportunities because it makes people less capable of focusing on future planning and entering into long-term commitments. When the potential consequences of risk exposure are grave, it may also turn risk management decisions into survival decisions, which could consume most of a person's energy and place him in a situation that is similar to that of the Hounded Woman.

Third, risk exposure may also diminish the robustness of existing options. Sometimes risk exposure itself does not render certain options prohibitive, but the materialisation of bad outcomes may, in the future, remove options or restrict their accessibility. Such uncertainty undermines the prospects for autonomy because it prevents people from foreseeing the availability of options in the future and therefore hinders their ability to make and execute long-term plans. Especially, if the potential negative shock may impose great distress on people's life or close up most other options once the risk materialises, it may, at that time, leave people stuck with their previous choices and deny them the chance to revise their life plans or have a fresh start.

Note that, in this analysis, I have focused on how risk undermines autonomy through its impact on the background options. It may be questioned why such impact needs to be measured by appealing to the notion of autonomy. Ferretti, for example, argues that

---

[18] A. Baderin and L. Barnes, "Risk and Self-Respect", *British Journal of Political Science* 50 (2018), 1-19.

autonomy is not a useful concept in this context because it is too complex for just capturing the shrinking of the range of options.[19] Let me examine her objection as well as her proposed alternative in assessing the impact of risk. As I will argue, I use a person's prospect of autonomy as a measure because autonomy can account for not only the quantitative impact of risk on background options, which Ferrett's model can assess neatly, but also its qualitative impact, which her model cannot capture.

2.1.3 Why use autonomy as the measure

Ferretti proposes that the impact of risk may be assessed by employing a formula-based measure of overall freedom. This model evaluates the level of overall freedom enjoyed by an individual based solely on the range of options, which Ferretti believes is not only sufficient for capturing the impact of risk but also does so more sharply than the comprehensive concept of autonomy because it is not mixed up with other concerns such as mental capacities.[20]

Ferretti's account builds on a model that was proposed by Steiner and subsequently modified by Carter.[21] In Steiner's original account, a person's freedom is measured by the ratio between the number of options he is free to exercise and the total number of feasible options.[22] This can be represented by the formula F/(F+U), where F is the number of free options and U is the number of options that are feasible but the person is unfree to choose. Presumably, a person is unfree to exercise a feasible option because he faces humanly (as opposed to naturally) imposed impediments. This formula therefore captures the intuition that, the more humanly imposed constraints a person is subject to, the less free he is. Carter then makes two modifications to this model. First, Steiner's formula does not take into consideration the possible combinations of options. So, to

---

[19] M. Ferretti, "Risk Imposition and Freedom", *Politics, Philosophy and Economics* 15 (2016), 261-79, p. 266.

[20] Ibid., pp. 266-72.

[21] See H. Steiner, "How Free: Computing Personal Liberty", *Royal Institute of Philosophy Supplement* 15 (1983), 73-89; H. Steiner, *An Essay on Rights* (Oxford: Blackwell, 1994), ch. 2; I. Carter, *A Measure of Freedom* (Oxford: Oxford University Press, 1999), ch. 7.

[22] Steiner, "How Free", p. 74. Note that there may be different ways to interpret "feasible" here. It may be taken to mean technologically feasible, potentially technologically feasible, logically feasible, and so on. For a discussion, see Carter, *A Measure of Freedom*, pp. 172-73. (But the precise definition need not concern us here.)

capture the freedom in exercising options in conjunction, Carter revises the formula to be the ratio between the number of options and sets of options that a person is free to exercise and the number of all possible options and sets of options.[23] Second, Carter further suggests that impediments usually exist in the form of possibility rather than certainty. Such possibility can be represented by probability numbers and be integrated into the formula by discounting the availability of options and sets of options with their probability of being available.[24]

As Ferretti notes, this revised model can naturally account for risk as a factor in diminishing freedom—the negative impact of risk on an individual can be captured by the formula through the increased probability of impediments on some of the options or sets of options.[25] The loss of freedom caused by introducing a particular risk, then, can be calculated by comparing the level of freedom the person enjoys at time T in absence of the risk and the level of freedom he would have at time T with the risk imposed. This loss of freedom could form the basis for evaluating the legitimacy of risk imposition.[26]

In defence of using the concept of autonomy as self-authorship, I have two replies to Ferretti's suggestion. First, Steiner chooses to evaluate the level of freedom in a ratio form because he is more concerned with social freedom in its negative sense. That is, the relevant notion of freedom for him is being free from humanly imposed impediments. According to Steiner, while technological development and increased production have significantly enlarged the range of options that are available to us, we are not necessarily freer than people who live in ancient Athens because we may at the same time be subject to more humanly imposed constraints than Athenians.[27] If an Athenian is free to exercise a proportion of those options available at his time that is larger than the proportion that I am free to exercise today, then the Athenian is in fact

---

[23] Ibid., pp. 181-83. Mathematically, the formula can be written as the following—

$$\sum_{i=1}^{n} F_{r,i} / \left(\sum_{i=1}^{n} F_{r,i} + \sum_{i=1}^{n} U_{r,i} \right)$$

[24] Ibid., pp. 189-91.

[25] Ferretti, "Risk Imposition and Freedom", p. 271.

[26] Ibid.

[27] Steiner, *An Essay on Rights*, pp. 43-44.

freer than me. In Steiner's words, "freedom is a social relation, not a technological one."[28]

Whether or not we accept Steiner's definition of freedom, however, this ratio-based concept is not very useful for my discussion because it cannot capture an important consideration in assessing the impact of risk—that is, human beings may not only act to limit the range of available options by imposing constraints on others but also expand the range of options by overcoming natural obstacles and achieving technological advancement. Such actions, moreover, often introduce risk as a byproduct. Therefore, despite the negative impact of risk, there may be net enhancement in terms of people's prospect for self-authorship due to these activities (I will come back to this in the next section). Such improvement cannot be captured by the ratio-based formula because it increases both the numerator and the denominator of the ratio. Raz's account, on the other hand, can account for both the negative impact of risk and the positive impact generated by risk-introducing activities. It can therefore be used to evaluate whether a risky case has a net positive impact for affected individuals.

If the concern is that the ratio-based formula cannot correctly capture the net impact of risk, it may be suggested that we can simply use a formula that calculates the range of available options and combinations of options, discounted by their probability of availability (that is, only the numerator of Steiner and Carter's model). Such a formula can then capture the net impact of risk while providing a measure that is more straightforward than the notion of autonomy. This then leads to my second reply to Ferretti. Ferretti's proposal of using a formula-based measure, whether it takes the ratio form based on Steiner and Carter's model or this revised form, can only assess a person's level of freedom at a certain point of time. The impact of risk on a person, on the other hand, may be more complicated. For example, as discussed, risk may undermine the right balance between long-term and short-term options or threaten the continued availability of some options over a period of time. These qualitative aspects of a person's choice situation throughout his life time cannot be captured by the level of

[28] Ibid., p. 44.

freedom he enjoys at a given time point T or the simple aggregation of such a level throughout his life.

The concept of autonomy as self-authorship, on the other hand, looks at a person's lifetime and considers the optimal background condition for him to potentially live a life according to his own goals and plans. It is therefor a more comprehensive concept that can capture the quantitative as well as the qualitative impact of risk. Moreover, based on Raz's account, the implication of creating an autonomy-supporting environment for people is not limited to the securing of the right kind of option set in the background. It may further require enhancing people's ability to effectively exercise these options. I will turn to this consideration in the third section.

In this section, I have examined how risk may undermine a person's prospect for autonomy. If risk has only negative impact on autonomy as self-authorship, then the implication for securing an autonomy-conducive background may simply be that risks should be eliminated as far as possible. However, the relationship between risk and autonomy can be more complicated than this. For example, as I have indicated, some activities that introduce risk into people's lives may nevertheless increase the range of background options. In the next section, I examine whether risk may be valuable for autonomy.

2.2 Can Risk Be Valuable for Autonomy?

As explained in Chapter 1, in this thesis I take risk to mean the possibility of a bad outcome. Why would the possibility of having a bad outcome be valuable for autonomy? There seems to be at least two types of reason. First, risk may be valuable in the sense that it is a price that is worth paying for having some options available. Second, the opportunity to experience and manage risk may be valuable on its own. Let me examine them in turn.

According to the first view, risk is a burden to autonomy, but it may nevertheless be a burden that is worth bearing because of the benefit it brings about. To illustrate, consider an example introduced by Oberdiek when he explains how risk undermines

autonomy by reducing the range of options and combinations of options available.[29] In his example, a car manufacturer has produced some models with grave defect—they may fatally explode through no fault of the driver's. There is no way to detect the faulty models so they are now in the market. According to Oberdiek, the chance for a potential buyer to get such a faulty model is a risk introduced by the manufacturer and it worsens the choice situation for its buyers by reducing the possibility for a buyer to successfully exercise the combination of options of buying a model and staying alive. It is like, according to him, laying a trap on a passage which the person could potentially take.[30] This argument is similar to Ferretti's which I discussed just now.

However, what Oberdiek (as well as Ferretti) does not note is that there does not necessarily exist a state of affairs in which the risk does not exist while all currently available options remain open. For example, in the case of car production, what if, as a matter of fact, the car manufacturer could never fully rule out the possibility of producing an explosive model? If this is the case, then the only alternative that does not involve the fatal risk would be no car production at all. In this case, the option set which Oberdiek claims to be foreclosed by the manufacturer could not have existed in the first place. Comparing the range of options available to the potential car buyer to the range that he may enjoy in a no-car-production world, it may be the case that car production with a non-zero defect rate widens rather than narrows the overall range of options.

Therefore, trap-laying may not be the appropriate analogy for all risk-introducing activities. Sometimes risk may be the inevitable by-product of having some options available. Instead of being traps laid on existing passages, they are like potholes on newly built roads which the builders are unable to fix. Of course, it does not mean that introducing risk is justifiable as long as doing so may increase the range of background options, but this way of looking at risk-imposing activities can be a starting point for us to consider the justifiability of risk imposition, which I will turn to in the next chapter.

---

[29] J. Oberdiek, *Imposing Risk: A Normative Framework* (Oxford: Oxford University Press, 2017), pp. 107-8.

[30] J. Oberdiek, "Towards a Right against Risking", *Law and Philosophy* 28 (2009), 367-92, p. 375.

In the remaining of this section, I will focus on the second type of arguments, which state that risk may be valuable on its own. Before I delve into their reasoning, one clarification should be made. As I have mentioned, risk can be accompanied by either a certainty or a possibility of gain—car driving carries the risk of accidents but it offers the benefit of convenience; gambling involves a possibility of loss paired with a possibility of winning. While I distinguish the value of risk from the judgement that risk is a price that is worth paying, I also do not mean that the value of risk, if any, must be considered in isolation from the relevant upside. For example, compared to getting a certain level of benefit, or a payoff with possibilities of small gains and small losses, some people may prefer to have both a chance for more gain and a chance of more loss. For such people, risk is valuable in the sense that, among options with different payoffs, the so-called "higher-risk higher-return" payoff is more preferable. Therefore, arguments for the value of risk may be interpreted in either a stronger or a weaker way. The strong thesis states that being exposed to a possibility of a bad outcome as such may be valuable. The weak thesis states that it may be more valuable to have a "higher-risk higher-return" payoff. With this clarification, let me now turn to the explanations why risk may be valuable.

According to the first argument, risk may be valuable because experiencing risk can bring excitement to people's lives.[31] For example, if my friend proposes to play a simple poker game with me (assume the outcome depends purely on luck) and the loser will have to pay the winner $5, we will probably both be more excited than if she simply offers to give me $3 for no reason. Presumably this argument reflects the weak thesis, as it is hard to imagine a case where a person is excited just by being exposed to a possibility of a bad outcome. It may be suggested that sometimes people do participate in risky activities such as bungee jumping and riding a roller coaster, which seem to have no obvious upside other than the excitement of experiencing risk. Note, however, that people do not really take on any meaningful risk by doing those jumps or rides (other than the small risk of technical problems, which of course no one values). What these examples do help to prove, on the other hand, is that some people indeed value the

[31] For example, see D. Lupton and J. Tulloch, "Life Would Be Pretty Dull Without Risk: Voluntary Risk-Taking and Its Pleasures", *Health, Risk and Society* 4 (2002), 113-24; quoted in Baderin and Barnes, "Risk and Self-Respect", p. 1.

excitement involved in the experience of being in risk or losing control, even though the danger here is merely illusory.

If risk is valuable because of the exciting experience it creates, then the implication for an autonomy-conducive background would be similar to that of other consumption goods. For example, since some people enjoy ice cream, its availability adds to the range of background options they may choose from and increases the chance for them to have their desire for ice cream satisfied. Similarly, for those who want the excitement of experiencing risk, they would prefer such risky options to exist in the background—the availability of risky options enhance the preconditions for autonomy by adding to the number and variety of background options.

To be sure, if a risky option involves a possibility of grave outcome, then its implication for autonomy may become more complicated—on the one hand, its availability adds to the range of options that may provide exciting experience to those who need it; on the other hand, for someone who is actually inclined to choose it, the possibility for the risk to materialise into significant harm may threaten his prospect of autonomy in the future. Some may argue that such risky options should be removed from the background for paternalistic reasons—I will come back to this issue in the next section.

Also note that, if people value risk for its excitement, then its implication for autonomy is relatively trivial. Experiencing risk is not the only way to create exciting experience in people's lives. For many people, the pure enjoyment of excitement is also not the most important or indispensable part of their lives. As long as there is an adequate range of consumption or recreational goods for a person to choose from, it seems that the unavailability of some risky options will not have a significant impact on his overall prospect for authoring a life he wants.

There are, however, arguments that attribute a more important role to risk in people's life. Tomasi, for example, suggests that the existence of risk may be valuable for people to develop the moral power of self-authorship.[32] According to him, living with various risks and having the chance to deal with them is an essential condition for an individual

[32] J. Tomasi, *Free Market Fairness* (Princeton: Princeton University Press, 2012), esp. pp. 79-82.

to feel that he has made a difference to his own life. If institutions are designed in a way that insulates people from risks (the example Tomasi is referring to is the European model of social democracy), then citizens are denied "opportunities to feel the special sense that they have done something genuinely important with their lives".[33] For similar views, Tomasi also quotes Murray who argues that "taking the trouble out of the stuff of life strips people…of major ways in which human beings look back on their lives and say, 'I made a difference'"[34], and poet Dean Alfange who writes "I do not wish to be a kept citizen, / Humbled and dulled, / By having the State look after me. / I want to take the calculated risk, / To dream and to build, / To fail and to succeed."[35]

Two points to note before I examine his arguments in more details. First, Tomasi's account is not explicitly about the value of risk for autonomy. His project is focused on the value of economic freedom—the freedom of making important economic decisions such as employment and long-term financial planning—for the ideal of self-authorship. But in various arguments, he does seem to equate the value of being able to make such decisions to the value of experiencing economic risk itself. It is these arguments that I will focus on.

Second, Tomasi seems to switch between two meanings of self-authorship in his argument and they are similar to but not exactly the same as how Raz uses the term autonomy as self-authorship. In most cases Tomasi appeals to self-authorship as the capacity or moral power individuals may have to develop and act upon a life plan,[36] which is similar to the precondition of internal capacities mentioned by Raz. On the other hand, when discussing the value of risk, Tomasi often talks about the sense of self-authorship, which is the subjective assessment people make about themselves when they look back at their own lives, a sense "of their worth as persons and to be able to

---

[33] Ibid., p. 80.

[34] C. Murray, *The Happiness of the People* (Washington, D.C.: AEI Press, 2009), p. 7, quoted in Tomasi, *Free Market Fairness*, p. 80.

[35] Ibid., p. xiii.

[36] Tomasi, p. 40.

advance their ends with self-confidence."[37] While a higher prospect of autonomy makes it more likely that a person will actually develop a sense of self-authorship, it is not guaranteed that everyone living with these preconditions would feel this way.

With these clarifications, let us come back to the argument that the existence of risk may enhance the sense of self-authorship. Why is it so? After all, Tomasi suggests that the sense of self-authorship requires a person to feel that he is "the central cause of the life" he is leading.[38] Why would it be more likely for a person to consider himself the central cause of his life when there are more, rather than less, external shocks to his life?

One possible answer is that the existence of risk provides people with the opportunity to exercise their agency. To illustrate, consider the example of risky sports. For those who enjoy mountain climbing or skydiving, the risk involved in the activities seems to be an important reason for them to find the experience valuable. Surely, it is also valuable that they can develop physical strength and probably enjoy excitement through these activities, but the sense of achievement would be significantly diminished if the potential risks are absent. In order to deal with the risks inherent in mountain climbing, participants need to prepare themselves both physically and mentally, to exercise relevant skills in different situations, to make judgments in unforeseeable circumstances and to overcome adversities. By exercising their agency in such ways, they can meaningfully reduce the possibility of accidents, hence achieving a sense of self-authorship.

Similar reasons may also apply to the handling of economic risk, which is the main focus of Tomasi's account, though more often such arguments take the form of the weak thesis and appeal to the possible downside as well as the possible upside—an economic risk may be valuable when it presents the challenge, thus also the opportunity, for people to avoid loss and potentially make a gain. For example, a person may take on significant financial risks by running a business, but if she can work hard, make the

---

[37] Ibid., p. 80. What Tomasi means by a sense of self-authorship can also be understood as a sense of self-respect. For this interpretation, see Baderin and Barnes, "Risk and Self-Respect", p. 5.

[38] Tomasi, *Free Market Fairness*, p. 83.

right judgment and learn from experience, she may be able to steer herself through the various obstacles and make the business successful. By dealing with such challenges, the person may indeed gain the sense that "I made a difference to my life".

It should be noted that, while it is possible for risk to contribute to a person's sense of self-authorship, not every risk is valuable in this way. In the case of risky sport and running a business, two factors seem to contribute to the value of risk. First, the outcomes of both risks are not purely random, so there is room for people to exercise their agency and potentially exert control over how they eventualise. Second, in both cases the risks are voluntarily taken on by individuals. Imagine a different case where a person is forced to play the Russian roulette. It is unlikely that he will gain a sense of self-authorship by facing this purely random chance of harm. The presumption that being subject to purely random risk has no positive value for self-respect can also be supported by empirical evidence. For example, Baderin and Barnes use the data in the Health Survey for England to examine the relationship between unemployment risk and workers's sense of self-worth. Contrary to the speculation that workers may gain self-confidence by dealing with precarious employment situations, the analysis shows that the risk introduced by adverse labour market condition, whose outcome workers have no control over, exhibit a consistently negative impact on people's sense of self-worth.[39]

Tomasi may object that, at least in the case of economic risk, it is not necessary for the two conditions (that the risk is not purely random and that it is taken on voluntarily) to hold in order for a risk to be valuable for self-authorship. This is because, even when a risk is imposed on a person and the person cannot prevent its potential bad outcome from eventualising, he can nevertheless exercise his agency in managing his risk exposure before the risk materialises. For example, he may buy insurance to cover the risk of accidents, buy a pension to cover the risk of outliving his savings after his retirement, and make financial investment to hedge against the risk of inflation. These economic decisions can be an important part of his life and he may derive a significant sense of self-authorship from such choices.

[39] Baderin and Barnes, "Risk and Self-Respect", p. 12.

Based on this line of reasoning, Tomasi can then argue that it is therefore essential for economic risks to exist in the background so that citizens can have the opportunity to develop their sense of self-authorship by managing such risks. This conclusion, however, does not seem to follow. Even if we agree about the value of dealing with economic risks, its implication for securing an autonomy-conducive background is only, similar to that of the excitement-based argument, the availability of the relevant options in the background, including those risky options people may want to take in order to achieve potential gains (such as personal investment) and those risk management options people can use to deal with existing risks (such as private insurance). What it does not require, on the other hand, is directly exposing people to economic risks. If people value economic decision-making as part of their life, they can always take on economic risks themselves by, for example, making personal investments or starting a personal business, as long as these risky options are available for them to choose. It is unlikely that a person would insist that the experience of managing risks existing in the background, such as the risk of unemployment, is so special that being denied the chance to manage them represents a significant loss to his prospect of autonomy.

Besides, economic risk is also not uniquely valuable in providing people with challenges and opportunities. After all, a life with less economic risks to deal with is not necessarily a life with fewer hurdles to cross or less chance for personal accomplishments. People may need to overcome personal shortcomings or other adversities in life. They can also proactively take up other challenges that are more relevant to their professional or personal plans. Other than those who treat investment as their main life project, most people will probably find addressing challenges in other areas to be a more meaningful way of exercising their agency than dealing with economic risks.

Finally, similar to what I have said about risky options that may bring people excitement, risky options that allow people to exercise agency and develop a sense of self-authorship may also involve potentially disastrous consequences, which, upon eventualisation, can significantly undermine a person's prospects for autonomy. If this is the case, then might the removal of such options have a net positive impact on

autonomy despite the opportunities they may provide? Let me turn to such paternalistic considerations in the next section.

2.3 Risky Choices and Paternalism

Raz argues that state intervention in people's choices may be justifiable when it enhances their overall condition of autonomy.[40] In particular, he focuses on paternalistic intervention with regard to the availability of background options. This is because, for example, if choosing an option makes it less likely for a person to successfully execute life plans and realise his aim, then removing this option could potentially have a net positive impact on his prospect of autonomy. However, Raz also notes that there is no simple, generic rule for determining whether paternalistic removal of options is justifiable based on considerations of autonomy.[41] To illustrate, he briefly considers the example of risky options. According to him, the availability of some risky options, such as unsafe consumer goods, only serves instrumental purposes. If a risky option has only instrumental value for everyone and removing it has a net positive impact on autonomy, then it is mostly likely that such paternalistic intervention would be justifiable. On the other hand, options such as risky sports are intrinsically valuable. In such cases, it would be trickier to assess the justifiability of removing them for the purpose of enhancing autonomy.[42]

Indeed, as I have analysed in the previous section, risky options may be valuable for autonomy in a few different ways. Some are intrinsically valuable for people to exercise their agency in overcoming obstacles and shaping their life stories. Others, while not intrinsically valuable in themselves, may nevertheless be the inevitable by-products of activities that are intrinsically valuable. Such value cannot simply be offset by the benefit of avoiding possible bad outcomes. However, among risky options that are valuable in this sense, some involve a possibility of eventualising into consequences that are not just bad but potentially disastrous and irreversible. Once materialised, the negative impact of such risks on autonomy could be significant and far-reaching—they

---

[40] Raz, pp. 412-21.

[41] Ibid., pp. 422-23.

[42] Ibid.

may threaten the availability of a minimally satisfying range of options for the rest of a person's life. In such cases, it is much more likely that the benefit of avoiding risk materialisation, hence enhancing the robustness of essential background options, will outweigh the negative effect on autonomy caused by removing such risky options despite the intrinsic value of their existence.

This is essentially the type of paternalism Gerald Dworkin has argued for. According to Dworkin, when potential consequences are dangerous, far-reaching and irreversible, paternalistic intervention can be justified to preserve people's ability in planning for and carrying out their autonomous lives.[43] Intervention justified in this way, however, should be limited to situations where the removal of an option is essential in preserving a minimally satisfying range of options throughout people's life. It may take the form of, for example, the collective enforcement of a basic safety net (that is, the removal of the option of not contributing to it). It may improve the overall preconditions for autonomy by helping to secure a minimally satisfying range of options throughout people's life. In fact, few question the justifiability of paternalism in this restrained version. The idea of a principled safety net is accepted even by Tomasi who champions the value of economic freedom.[44]

But paternalistic concerns about risky choices often go beyond disastrous options. They are often motivated by the observation that people exhibit various cognitive biases in decision-making and such biases are especially salient when risks and uncertainties are involved.[45] For example, according to behavioural studies, it is common for people to have problems in estimating probabilities—they tend to assign a higher probability to good outcomes based on unrealistic optimism, make judgments using only handily available or recent information or give special weight to information that confirms their antecedent beliefs.[46] In addition to the difficulties in correctly perceiving risks, people also make suboptimal risk management decisions because they tend to be biased toward

---

[43] G. Dworkin, "Paternalism", *Monist* 56 (1972), 64-84, pp. 80-83.

[44] See Tomasi, *Free Market Fairness*, p. 99.

[45] See, for example, C. Sunstein, *Why Nudge? The Politics of Libertarian Paternalism* (New Haven: Yale University Press, 2014), pp. 34-50; K. Ahlstrom-Vij, *Epistemic Paternalism: A Defence* (Basingstoke: Palgrave Macmillan, 2013), pp. 16-27.

[46] Sunstein, *Why Nudge?*, pp. 44-50.

the present and are not willing to incur a small certain cost now to cover future uncertain losses.[47] These biases may lead people to make sub-optimal decisions and prevent them from realising their own objectives.

This concern is related to another precondition for autonomy in Raz's account—the internal requirement that individuals possess the rational capacities to evaluate options and make informed decisions in order to pursue their chosen goals. For Raz, the institutional implication of the rationality condition is mainly the requirement to provide education for citizens. But behavioural failures are not always due to lack of knowledge in statistics and probability theory. Rather, it is often the result of people's tendency to rely on heuristics rather than formal analysis in making risk-related decisions and their unwillingness to engage in self-binding and self-correction to overcome such biases.[48] It may therefore be suggested that, if people tend to choose certain risky options because of cognitive biases, then removing those options may in fact improve their overall prospect for autonomy.

However, beyond the disastrous options, it is hardly obvious whether the removal of a risky option from the background could improve the overall condition for autonomy for everyone. First, it is doubtful whether the state can always correctly determine which options are valuable to keep. Depending on their personal preferences, financial conditions and life plans, an option that is unreasonable for one person to choose may be valuable for another. For example, suppose there is a gamble in which the participant has a 20% chance of winning £10,000 and a 80% chance of losing £5,000. Given its negative expected return, it may be suggested that it is not rational for people to take this gamble and removing the option to participate in it would not undermine the preconditions for autonomy. However, imagine a person who needs £10,000 to have a critical surgery. If he has no better alternative, the availability of this option may indeed be valuable for him.

More generally, people can also value risky options differently because of their different risk and time preferences. Classical rational choice theory evaluates a risky option by

---

[47] Ibid., pp. 34-38.

[48] Ibid., p. 34.

aggregating outcomes across possible states, each discounted by the probability of its occurrence. This expected return approach assumes that individuals have a risk-neutral attitude toward different outcomes that may result from the risk. For example, consider the following choice situation. An individual is asked to choose between (a) an amount of money which gives him 10 units of utility and (b) having a 50% chance of getting an amount of money that generates 21 units of utility for him and a 50% chance of getting nothing. Since the expected return of option (b) is 10.5 units of utility, which is greater than option (a), a rational person, according to classical rational choice theory, should prefer option (b). However, in reality, people often deviate from this risk-neutral attitude. For example, a person may reasonably choose option (a) if he prefers the certainty of winning at least something.[49]

In addition to risk attitudes, since a risk taken now may only materialise in the future, different time preferences could also affect a person's choice. For example, when an individual makes the long-term decision between consumption and investment, beyond the basic minimum, he may reasonably value current welfare more than welfare in the far future because of uncertainty about the length of his life. On the other hand, a person committed to achieving a long-term goal may want to invest most of his resources for the project instead of safeguarding against tolerable risks in the near term. The first person may save less than what is considered optimal into a pension scheme while the second may fail to buy a sufficient amount of health care insurance. Their decisions, however, can be considered reasonably fitting to their autonomously chosen life paths and the prospect for their autonomy would be undermined if they are simply deprived of such options.

Since risk and time preferences may affect how different people value a risky option, options should not be removed simply because they are deemed irrational to choose based on classical rational choice theory. On the other hand, however, treating people's risk decisions as reflecting their actual, considered risk and time preferences would only make sense if they actually do not regret when the risks they take on materialise later—

[49] Risk preference can be incorporated into decision theory using the Risk-Weighted Expected Utility model. See detailed discussion in L. Buchak, *Risk and Rationality* (Oxford: Oxford University Press, 2013), ch. 1 and ch. 2. By recognising such possibilities, however, the rational choice theory would no longer have clear action-guiding implications.

in reality they often do. As noted by Goodin, people's ex-ante risk preferences should be treated with a high degree of suspicion and what matters morally should be how people feel once they are hit by the negative shocks.[50]

In light of both considerations, it may be suggested that "nudging", a form of soft paternalism, could be employed to enhance the background condition for autonomy. Instead of removing certain risky options, nudging functions by presenting the same set of options in a different way so that people are more likely to choose what is considered good or effective for them to pursue their own ends—it influences people's choices by utilising their systematic behavioural patterns.[51] For example, the state may attempt to correct people's behavioural failure in purchasing insurance by enrolling citizens in state-mandated insurance programmes by default while allowing them to opt out if they declare that they understand the risk involved in doing so. Because individuals tend to stay with default options, such a policy nudges people to stick with a sufficient amount of insurance coverage. Meanwhile it does so without coercively removing options from the background—a person may still choose to opt out after going through the declaration process, which, being part of the nudging strategy, is designed to encourage him to carefully review his decision and to potentially overcome risk-related biases.

While nudging can be effective in helping individuals to select the appropriate means to pursue their autonomously chosen ends, some caution that it could be manipulative and therefore violate the independence condition.[52] According to Raz, manipulation "perverts the way that person reaches decisions, forms preferences or adopts goals."[53] In other words, it prevents people from responding to their own reasons in making decisions. This seems to be exactly what nudging aims to do—when a nudging policy

---

[50] R. Goodin, "Stabilizing Expectations: The Role of Earnings-Related Benefits in Social Welfare Policy", *Ethics* 100 (1990), 530-53, p. 551.

[51] For discussions about nudging, see, for example, R. Thaler and C. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (New Haven: Yale University Press, 2008); D. Hausman and B. Welch, "Debate: To Nudge or Not to Nudge", *The Journal of Political Philosophy* 18 (2010), 123-36; Sunstein, *Why Nudge?*; A. Schmidt, "Getting Real on Rationality—Behavioral Science, Nudging, and Public Policy", *Ethics* 129 (2019), 511-43.

[52] See, for example, T. M. Wilkinson, "Nudging and Manipulation", *Political Studies* 61 (2013), 341-55; T. Nys and B. Engelen, "Judging Nudging: Answering the Manipulation Objection", *Political Studies* 65 (2017), 199-214.

[53] Raz, *The Morality of Freedom*, p. 377.

steers a person toward a choice, it intends the person to choose what is considered to be the optimal option unconsciously without reflecting upon the reasons of his decision. If this is the case, then, even though nudging does not reduce the range of available options, it still compromises autonomy by violating the independence condition, which requires that individuals be free from coercion and manipulation.

However, it is too fast to conclude that nudging is manipulative due to its use of people's unconscious biases. First, it is not the case that, without the relevant nudge, people would no longer suffer from the behavioural biases nudging strategies intend to utilise. If the state does not enrol its citizens by default in insurance schemes, those who tend to stay with the default option rather than carefully consider their insurance choices would still stay with the default, which in this case is no insurance coverage. Nudging therefore does not necessarily worsen individuals' responsiveness to reasons, though it does not attempt to contribute to it either.

Second, manipulation is objectionable because of the kind of treatment it inflicts on people—it fails to respect them as ends in themselves and subjects them to the will of others.[54] Nudging could indeed be manipulative when it is designed to advance the interests of parties other than those affected. In this sense, nudges done by private businesses are often likely to be manipulative because they are designed to increase profit. On the other hand, public nudging, that is, nudging strategies adopted by the state through law, regulation and public policies, may be non-manipulative if they are designed properly and through a process that is transparent and adequately controlled by the affected individuals.[55] In fact, since people are often exposed to various nudges in the private market, public nudge may even act as a counter force which makes the overall choice environment less manipulative for citizens.

A further objection could be made—even when nudging is not manipulative, it is still a problematic form of paternalism because it treats people as if they were not capable of making right choices for their own lives, which is disrespectful because it implies a

---

[54] Ibid., p. 378; also see Sher, *Beyond Neutrality*, p. 47.

[55] Schmidt, "Getting Real on Rationality", pp. 532-35.

negative moral judgment about their rational capacity.[56] If empirical evidence shows that people are generally subject to cognitive problems in making risk-related decisions, then the respectful way for the state to intervene may include providing relevant information and reasoned arguments in order to persuade people to make the right choices.[57] To induce such decisions by appealing to their subconscious biases is to distrust their rational ability to understand facts or their will power to act accordingly.

What this objection fails to appreciate, however, is the cognitive burden people have to bear by living in a contemporary society. People already have to know so many things and make so many decisions and their cognitive capacity is not unlimited. Having to know everything and to perform active choosing in every situation could be a burden which distracts them from their major life projects. As pointed out by Sunstein, people do not always enjoy the freedom of choice, especially in complex and unfamiliar domains.[58] If we need to acquire knowledge for everything and make corresponding judgments, we would be overwhelmed and have no time or capacity left to work on our major projects.

In fact, it is partly due to the need to cope with such cognitive burden that people develop heuristics to assist decision-making. Instead of seeking to optimise each choice, they attempt to use decision-making procedures which, over time, allow them to reliably achieve their ends given their psychological features and choice environment. Schmidt defends such procedures as being ecologically rational.[59] If we accept this view of rationality, then employing public policy nudging would not imply treating agents as irrational. Rather, it merely aims to help ecologically rational agents to pursue their ends more reliably by adjusting the choice environment to better match their psychological makeup.[60] Therefore, considerations of creating a more favourable background

---

[56] See, for example, J. Waldron, "It's All for Your Own Good", *The New York Review of Books* (9 October 2014). For this type of objection on paternalism in general, also see J. Quong, *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011), p. 83. Note that Quong is not addressing the particular issue of nudging there.

[57] Waldron, "It's All for Your Own Good"; Quong, *Liberalism Without Perfection*, p. 95.

[58] Sunstein, *Why Nudge?*, p. 126.

[59] Schmidt, "Getting Real on Rationality", p. 522.

[60] Ibid., pp. 527-28.

condition for people to live an autonomous life may justify soft nudging policies which aim to correct people's behavioural biases in making risk-related decisions. Such paternalism-based justification does not, however, extend to hard intervention beyond the basic safety net, such as making medical insurance or pensions compulsory.

To conclude, then, in this chapter I have analysed how risk may undermine people's prospect of autonomy through its negative impact on the background options. Despite such impact, the availability of risky options may still be valuable under some circumstances, so creating an autonomy-conducive environment does not necessarily require the removal of all risky options. I have also considered the justification for state intervention in people's risk-related decisions for the sake of promoting autonomy and argued that paternalistic considerations only justify a basic safety net and nudging policies, not coercive measures such as mandatory insurance.

This does not mean, of course, that no considerations other than paternalistic ones may imply a demand for the state to intervene and deal with individuals' risk exposure in certain ways. In fact, the paternalist argument I have considered takes existing risks as given. That is, it treats people's background option set to be conditioned by those risks in a way that is similar to how its range is limited by natural obstacles, and then asks whether state intervention could be justified to modify that option set. However, most risks, or at least their potential impact on individuals, do not exist like natural obstacles. Rather, as I have already mentioned in this chapter, they are often by-products of chosen social arrangements—they could be seen as being introduced or imposed on society members through the collective choice of social institutions. If this is true, then, since risk sets back individuals' interest in autonomy, justifying its imposition may itself raise questions about how the potential burdens should be dealt with. I turn to this issue in the next chapter.

Chapter 3 — Justifying Social Risk Imposition: A Contractualist Approach

In the last chapter, I examined how risk may impact the preconditions for autonomy, especially by affecting the range, quality and robustness of life options available to individuals. I also mentioned that, while risk itself undermines the preconditions for autonomy, it is usually introduced or retained as by-products of certain actions or social arrangements that may on the other hand enhance these preconditions. In this chapter, I will explore if subjecting people to such risks for the corresponding benefits could be justified. In particular, I focus on common risks inherent in social institutions. Individuals have a general interest in autonomy, so institutions which influence what and how much common risk individuals are exposed to should be designed in a way that protects or promotes every individual's interest in having a fair prospect of living an autonomous life. This bears on the level and form of common risks that could be justly imposed on society members.

This chapter is organised as follows. In Section 3.1, I explain why many common risks could be regarded as being imposed on members of a society through their collective choice of social institutions. In particular, I argue that, although some risks appear to be natural risks, their impact on individuals is often determined by institutional design, so the moral implication of such choices is analogous to that of humanly imposed risks and is in similar need of justification.

Then, in Section 3.2, I use the Scanlonian contractualist framework to examine what principles would justify the imposition of social risk. I suggest that a scheme of social risk imposition resembles public goods provision in the sense that it imposes burdens on members of a society for some common benefits, so it may be guided by principles for supplying conventional public goods. To apply such principles, however, I need to address the difficulty of assessing risk-related burdens. The burdens introduced by a scheme of common risk imposition may be assessed based either on the ex-ante state before a risk materialises or on the ex-post state after the risk eventuates into actual harm, and disagreement about which state is relevant for justification leads to a debate between ex-ante contractualists and ex-post contractualists. I look at their arguments as well as Scanlon's solution to the problem and propose that contracting parties should

endorse a version of ex-ante contractualism which includes a just intention principle and a reasonable precaution principle.

I then suggest in Section 3.3 that the same logic which leads to the endorsement of the reasonable precaution principle would also justify a compensation principle. This principle requires that compensation be paid to victims of risk materialisation so as to reasonably alleviate their misfortunes. The burden of paying for compensation should be shared among those who have benefited from the imposition of the relevant risk. The next section considers the fair allocation of this burden. I propose a version of the benefit principle which requires that people who have benefited from a scheme of social risk imposition make a corresponding contribution so as to equalise the net benefit produced by the scheme for each individual. I suggest that such a principle is in contrast with the actuarial principle which guides the pricing of private insurance and this offers a strong case for a universal risk sharing scheme that embodies the benefit principle as a way of fairly sharing the burden of social risks among members of the society.

In the remainder of the chapter, I address a concern that the contractualist approach to risk imposition cannot settle on a single level of risk imposition that is justifiable to all contracting parties when they have diverse risk preferences. I appeal to Scanlon's grounds for reasonable rejection and argue that a single scheme of risk imposition could be justified when other alternatives impose greater burden on some society members in terms of the compromised prospects for their autonomy.

3.1 The Imposition View of Common Risk

To see how common risks are imposed through the choice of social institutions, take unemployment risk as an example. At first sight, it seems that this is an economic risk which naturally exists in most people's economic lives. However, considering an alternative arrangement of a planned economy in which all those able to work are guaranteed a job and a specific salary, it becomes clear that unemployment risk is in fact a risk imposed through the choice of an economic system. This risk is, moreover, retained for certain benefits. Economic theories as well as empirical evidence have shown that a planned economy not only sacrifices occupational freedom but is also

likely to end up with lower economic productivity. Therefore, the risk of unemployment could be seen as a common risk imposed on those living in a market economy for the benefit of better economic prospects and greater freedom in career choices, both of which may enlarge the range of background options available to individuals.

It may be objected that, by treating unemployment risk as a commonly imposed risk, this view ignores the role of personal causes for unemployment. For example, even though every employee is exposed to the risk of unemployment, it seems that a person who loses his job simply because he does not want to work and thus performs poorly should not be regarded as a pure victim hit by a commonly imposed risk. This legitimate worry points to a distinction between the personal and the social component of unemployment risk. In this example, the person should be seen as adding a personal risk of losing his job to the common unemployment risk imposed on everyone by the structure of the labour market. When I refer to unemployment risk as a common risk, I am only referring to its social component, as opposed to the personal component. I will also from now on use common risk and social risk interchangeably.

In reality, of course, it is often unclear how to distinguish between personal and social risk. When a person is not performing well in his current job, is it because his inborn talents and characters unfortunately do not fit the kind of job that is made available to him, or is it because he is simply unwilling to adapt himself to the job? It seems that a potential answer to this kind of question will not be an objective one, but I will later suggest that peoples' collective view of how this distinction should be made could be seen as reflected in rules of precaution that accompany a just risk imposition scheme in that society.

Unemployment risk is a relatively obvious example of social risk imposition because it is humanly imposed. Other examples of humanly imposed risk include the risk of workplace accidents, the risk of car accidents and the risk of nuclear disasters. In addition to human risks, there are also risks, such as natural disasters and some illnesses and disabilities, which are usually regarded as natural risks but could be similarly treated as cases of common risk imposition. Let me explain why they could be categorised in this way.

First, some risks that seem to be naturally caused in fact have a human cause. For example, some illnesses are caused by pollution or bad nutrition; some disabilities are a result of accidents due to lack of safety measures. Even some natural disasters are partially caused by human actions: flood may be caused by inappropriate use of farming lands; wildfires may be caused by bad forest management. Such risks could be regarded as being imposed on members of the society for the economic interests of some people or for the collective benefits of efficiency and could therefore be treated in the same way as humanly imposed risks.

Second, while some risks have a pure natural cause, the extent to which victims suffer bad consequences is a result of social arrangements. For example, while floods are sometimes purely caused by nature, how many people would be impacted and how much loss they suffer is partly a result of civil planning which determines the building of houses and flood control infrastructures in high-risk locations. Moreover, there are also cases where the initial bad consequences suffered by individuals are natural, but the extent to which such harm would then be alleviated is impacted by institutions. For example, while some physical disabilities are genetic, how they may hinder people's opportunity to move around freely and to take up certain jobs partly depends on the design of public spaces in their society. Similarly, for naturally caused illnesses, how much medical cost is needed to restore an ill person to a sufficient level of health and how soon he will be provided with the required medical service depends on the health care system of his society.

The recognition of these two types of cases, especially the second type, also blurs the distinction between risks that are actively imposed and risks that are retained because people choose not to remove them or not to neutralise their negative impact. Some may question why the latter could also be labelled as risk imposition. While it is true that such cases could be more generally described as cases of choosing among various schemes of social arrangements which carry higher or lower levels or impacts of social risk, I choose to label both types of cases as risk imposition because they are structurally similar. That is, when social institutions choose to retain a certain level of risk rather than eliminate it, the implication of such a choice is similar to cases of active

risk imposition. In both types of scenarios, the risk is allowed to exist for the sake of achieving some benefits. If a society chooses not to safeguard all its members against the risk of flooding by allowing some of them to live in potential flooding areas or by not spending all its resources in building flood infrastructure, it retains a level of flood risk for the benefit of more flexible or more efficient civil planning. Or, if a society chooses not to cover all citizens by providing free health care, it allows a level of health care risk for the possible benefit of higher efficiency or quality of health care services.

Because of this structural similarity, treating the latter as risk imposition also emphasises that social arrangements which deal with natural risks also need to be justified. That is, leaving people exposed to natural risks for the sake of efficiency, flexibility or productivity requires justification in the same way as does actively subjecting people to risks in order to achieve such benefits. It may be objected that a distinction should be drawn between actively imposing a risk and merely allowing a risk to exist, the former requiring stronger justification. While this may be true for interpersonal morality, the line is in fact blurred in institutional design and public policy making—whether to leave individuals exposed to the risk of health care or natural disasters is a decision that requires active policy selection just like the choice of permitting car driving. Besides, since the justification of active risk imposition is the focus of a growing literature in philosophy,[1] using the term risk imposition also points toward the possibility of drawing upon this literature to identify just social arrangements which involve common risks, whether or not they are actively imposed.

3.2 A Contractualist Account of Just Social Risk Imposition

I have explained in Chapter 1 that Scanlon's contractualism is the most promising approach to dealing with risk imposition. It is an approach which assess the permissibility of an action based on whether a principle which regulates the type of actions in question is one that cannot be reasonably rejected by each party in the contracting position. Unlike Rawls's original position, contracting parties in Scanlon's

[1] See, for example, J. Lenman, "Contractualism and Risk Imposition", *Politics, Philosophy and Economics* 7 (2008), 99-122; R. Kumar, "Risking and Wronging", *Philosophy and Public Affairs* 43 (2015), 27-51; J. Frick, "Contractualism and Social Risk", *Philosophy and Public Affairs* 43 (2015), 175-223; J. Oberdiek, *Imposing Risk: A Normative Framework* (Oxford: Oxford University Press, 2017).

contractualist framework are not behind a veil of ignorance. They know their personal characteristics, life goals, preferences, and so on, and they assess the acceptability or rejectability of a principle based on their own standpoints.[2]

To make sense why a principle that permits certain kinds of risk imposition (such as car driving) may be justifiable to every contracting party, all of whom know that they could end up being the victim when the risk materialises, Scanlon appeals to the idea of intrapersonal aggregation.[3] That is, for every affected individual, the benefit made possible by the imposition of the risk may aggregate throughout their life to an amount that outweighs the expected burden of risk materialisation. Based on the analysis of the previous chapter, this net benefit brought about by risk imposition may be described as an improved prospect for autonomy—if each individual enjoys better preconditions for autonomy as a result of a scheme of risk imposition, then it seems that a principle permitting this scheme may be acceptable to all.

The fact that a scheme brings net benefit to everyone, however, does not automatically justify the imposition of the corresponding burden on people. A person, especially if he is risk-averse, may reasonably wish to avoid the burden of risk exposure even though it generates net benefit. This problem is in some sense analogous to the general problem of public goods provision. Like a public good whose provision is costly, a social risk imposes an involuntary burden on society's members for some common benefits. The person who wishes to object to such an arrangement may quote the famous objection raised by Nozick about public goods that "one cannot…just act so as to give people benefits and then demand (or seize) payment."[4]

However, many people, unlike Nozick, do believe that the funding of at least some public goods can be justified. If so, then maybe similar reasons can also ground the imposition of common risks. Of course, the burden of being subject to a common risk differs from the burden of paying tax to fund public goods in an important way—instead of having every individual pay a predetermined amount of taxes, a common risk

---

[2] T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1998), esp. pp. 209-18.

[3] Ibid., pp. 237-38.

[4] R. Nozick, *Anarchy, State, and Utopia* (Oxford: Basil Blackwell, 1974), p. 95.

that every individual is subject to may eventualise into actual costs that are to be born by some but not others. With this difference in mind, however, I will look at the general case of public goods provision first and address the problem of assessing risk-specific burdens in later sections.

A public good is characterised by the feature of non-excludability.[5] That is, no one can effectively be excluded from its use. Because of this feature, when the production of a public good is left to the market, people will be tempted to free ride on others' efforts, leading to under-supply of the good. It is therefore generally agreed that it is justifiable for the state to intervene and to provide these goods with public funding.

If, once the good is provided by the state, individuals can choose whether or not they want to enjoy the good (even though the state is not able to exclude anyone that wants to do so), then those who actually choose to enjoy it should pay for its provision—the fact that a person has voluntarily accepted its benefit can by itself justify a duty for this person to make a corresponding contribution based on reasons of fairness. However, in addition to being non-excludable, some public goods are also non-avoidable. That is, people who are covered by the good's provision cannot avoid enjoying its benefit. This feature may undermine the fairness-based argument for contribution—a person who has enjoyed the good may claim that he would have wanted to reject the benefit had he been given the chance to do so.

To address this concern, it has been suggested that some public goods offer benefits that no reasonable person would prefer to avoid. For example, Klosko points out that there are some goods, such as clean air and national defence, whose benefits no one could claim in good faith that he wishes to avoid. Klosko labels such goods as presumptive goods and argues that, if a public good is presumptively beneficial for all, then mandatory contribution may be justified based on reasons of fairness even when it is not

---

[5] In economics, the definition of a public good further includes the feature of non-rivalry—one individual's enjoyment of it does not reduce its availability to others (as opposed to a common good, which is non-excludable but rivalrous). See J. Anomaly, "Public Goods and Government Action", *Politics, Philosophy and Economics* 14 (2015), 109-28, pp. 109-10. However, this definition deviates from how the term public good is used in common language and the feature of non-rivalry is less relevant for my discussion here. So I will refer to a public good as a non-excludable good.

avoidable, provided that, first, the goods are worth the recipients' effort in providing them and, second, the relevant burdens and benefits are distributed fairly.[6]

In a slightly different account, Miller and Taylor argue that some public goods are essential goods in the sense that their availability is required by justice—they may be necessary for social cooperation to exist (such as basic security), for correcting externalities that would otherwise interfere with individuals' justly acquired rights (such as a system of tort law) or for providing a social minimum that no one should fall below.[7] Similarly, fulfilling such demands of justice could be said to be essential for all and their provision may be justifiable on the condition that the benefits and burdens are allocated in a fair way.[8]

While the contractualism literature has not said much about the issue of public goods provision,[9] it seems that parties in the Scanlonian contracting position may reasonably accept principles which permit public goods provision of this kind. I will therefore assume, for the purpose of this thesis, that contracting parties could reasonably accept the following two principles for providing conventional public goods—*the presumptive worthiness principle* requires that, in order to justify mandatory contribution to the provision of a public good, its benefit must be presumptively beneficial or essential for all and sufficiently outweigh the corresponding burden for each individual;[10] *the fair distribution principle* requires that the relevant benefits and burdens be allocated in a

---

[6] G. Klosko, "Presumptive Benefit, Fairness, and Political Obligation", *Philosophy and Public Affairs* 16 (1987), 241-59, p. 246, p. 253.

[7] For essential public goods, see D. Miller and I. Taylor, Miller, "Public Goods", in *The Oxford Handbook of Distributive Justice*, ed. S. Olsaretti (Oxford: Oxford University Press, 2018), pp. 564-66.

[8] Both Klosko and Miller and Taylor also consider the further case of non-essential public goods, the mandatory provision of which is more difficult to justify. I will not attempt to address this question here.

[9] Except Rawls who briefly mentions that "assuming that the public good is to everyone's advantage, and one that all would agree to arrange for, the use of coercion is perfectly rational from each man's point of view". See J. Rawls, *A Theory of Justice* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999), p. 267.

[10] This principle is based on Klosko's account, though it is stronger than Klosko's version which merely requires that the benefit exceeds the cost for the society as a whole and so for the typical individual in the society. See Klosko, "Presumptive Benefit, Fairness, and Political Obligation", p. 252.

fair way.[11] The question is whether these principles could similarly justify a scheme of social risk imposition.

3.2.1 Presumptive benefits of social risk imposition

To treat a scheme of social risk imposition as a public good, it needs to be established that its benefit is non-excludable. A social risk is often imposed for benefits such as increased economic efficiency, greater freedom in occupational choices, and so on. These benefits are generally non-excludable, but they are not non-excludable in the sense that, theoretically, institutions cannot be designed in a way so that some people are completely excluded from the burden of risk exposure together with the enjoyment of these corresponding benefits. Rather, they are non-excludable in the sense that, as a matter of fact, social and economic institutions usually could not be organised in such an exclusive manner at a reasonable cost. There could hardly exist, for example, a parallel closed economy alongside the free market, which implements central planning and guarantees job allocation, within the same society.[12] Even when it is possible to design institutions or infrastructure in such a way that those unwilling to participate are excluded, they, by being members of the same society, still cannot be prevented from enjoying the direct and indirect benefits of an expanded range of goods, services and life possibilities.

Accepting this latter claim would further imply that the relevant benefits are also non-avoidable for society's members, including people who largely avoid direct participation in the specific scheme or arrangement. Again, it is not because people cannot, in theory, fully reject the fruits generated from such social arrangements but because it is not realistic for them to do so. Few people could, for example, choose to

---

[11] For the content of a fair distribution principle for public goods, see discussion in Miller and Taylor, "Public Goods", pp. 567-71. I will also come back to this issue later in the chapter when I discuss fair burden allocation with regard to social risk imposition.

[12] Olsaretti labels this type of goods as "deliberately socialised goods". See S. Olsaretti, "Children as Public Goods?", *Philosophy and Public Affairs* 41 (2013), 226-58, p. 248. It may be suggested that some degree of separation is possible and indeed exists in capitalist societies. For example, the public sector may provide much more job security than the private market. However, even in such a society, people's economic lives are still significantly intertwined and people who choose to work in the public sector still benefit from the wealth generated in the riskier private sector (because the public sector is partly funded by taxation on the private sector).

live in an area so isolated that they could avoid any benefit of a modern transportation system.

Since the benefit of social risk is non-avoidable, in order to justify a scheme which imposes mandatory burden (that is, unavoidable risk exposure) on everyone, it has to be established that its benefits are at the same time presumptively beneficial for all. The benefits generated by a scheme of social risk may be regarded as presumptively beneficial because it enlarges the range of valuable life options available to individuals. Even when a specific individual chooses not to exercise the additional options, the background condition for this person to live an autonomous life is still improved compared to a life with very limited exercisable options. If, for example, the person revises his life goals later, it is more likely that he can adapt his life to the new plan when he has a larger set of options to choose from. Therefore, if we agree that individuals have a general interest in securing a background condition that is conducive to autonomy, we could regard the benefit produced by a scheme of social risk imposition to be presumptively beneficial when its net impact on autonomy is positive.

Therefore, to justify imposing a social risk on members of a society, it should be required to satisfy the following two principles, similar to those that have been argued for public goods provision—

> *The presumptive worthiness principle*: the benefit for which a common risk is imposed should be presumptively beneficial in terms of its impact on autonomy for each impacted individual and sufficiently outweigh the burden introduced by the risk.
>
> *The fair distribution principle*: benefits and burdens resulting from a case of risk imposition should be allocated fairly among those subject to the risk.

The major challenge in applying these two principles, however, is to decide how the burden of risk imposition should be evaluated. Unlike a standard public good, there are two different states relevant to a case of risk imposition based on which its burden may be assessed—the ex-ante state when the risk is introduced but has not yet materialised and the ex-post state when the risk has produced bad consequences. Ex ante, every

affected individual faces a likelihood of a bad outcome whereas ex post a few of these individuals will have been hit by the risk and suffered the actual bad consequence of physical harm or material loss. How should such burdens be properly accounted for when assessing the permissibility of a case of risk imposition? In answering this question contractualists are divided into two camps. I will now turn to this debate between ex-ante and ex-post contractualists.

3.2.2 Ex-ante vs. ex-post contractualism

Ex-ante contractualism assesses the relevant benefit and burden of risk imposition based on the state when the risk is imposed but has not yet materialised, that is, when the identity of the victims is still unknown. Each contracting party calculates the burden by multiplying the level of potential harm with the probability that they will be hit by the risk throughout their life course. Ex-ante contractualism therefore captures our intuition that the lower the severity of harm and/or the probability of risk materialisation, the more likely that the expected burden will be outweighed by the corresponding benefit brought about by risk imposition and hence the more likely that the case will be justifiable based on the presumptive worthiness principle.

Ex-post contractualism, on the other hand, holds that a case of risk imposition can only be justified if the actual distribution of burdens after risk materialisation is justifiable to each individual including the victims.[13] The rationale underlying this position is that, when imposing a risk on a large population, it is almost statistically certain, due to the law of large numbers, that some unknown people will eventually suffer the bad consequence and such harm should be fully accounted for in assessing its justifiability. As stated by Moreau, "as long as we know that acceptance of a principle will affect someone in a certain way, we should assign that person a complaint that is based upon

[13] See S. R. Moreau, "Contractualism and Aggregation", *Ethics* 108 (1998), 296–311; M. Fleurbaey and A. Voorhoeve, "Decide As You Would with Full Information!", in *Inequalities in Health: Concepts, Measures, and Ethics*, eds. N. Eyal et al. (New York: Oxford University Press, 2013).

the full magnitude of the harm or benefit, even if we cannot identify the person in advance."[14]

However, both types of contractualism in the form stated above are problematic because they may lead to the endorsement of some intuitively unacceptable cases. The difficulties they face may be illustrated by comparing the following examples.[15]

> *Vaccination with Deadly Side Effect*: The state makes it compulsory for all society's members to take a vaccination, which is effective for 99.999% of the population in preventing an infectious disease with paralysing effect. However, there is a 0.001% chance that the vaccination will cause immediate death. Because of the law of large numbers, it is almost statistically certain that a number of unknown individuals (say X individuals in this society) will experience this bad outcome as a result.
>
> *Medical Experiment*: The state randomly selects X individuals in this society to be killed for medical experiments. Such experiments could achieve a level of health benefits for the whole population equivalent to the effects of vaccination in *Vaccination*.
>
> *Doing Nothing*: The state could also choose to do nothing. Let us suppose that, in this case, because the disease is highly infectious, ultimately the whole population (or a very high percentage of it) will be paralysed.

For ex-post contractualism, probability numbers in these cases do not matter. To assess their acceptability, all we need to compare is the ex-post distribution of outcomes. In both Vaccination with Deadly Side Effect and Medical Experiment, X individuals will suffer death and the rest of the population will be healthy. In Doing Nothing, all the population will be paralysed. An ex-post contractualist would therefore grant that the X individuals, whoever they turn out to be, could reasonably reject Vaccination and Medical experiment because the burden of death is much more significant than

---

[14] Moreau, "Contractualism and Aggregation", p. 304. Note that it does not mean that the ex-ante burden of risk exposure such as the psychological burden of fear and anxiety and/or the burden of managing the exposure is not to be accounted for—they could be added to the ex-post outcome so as to get the overall distribution. See M. Fleurbaey, *Fairness, Responsibility, and Welfare* (Oxford: Oxford University Press, 2008), p. 158 on this point.

[15] Similar examples are first discussed by Scanlon, *What We Owe to Each Other*, pp. 208-9.

paralysis. The rejection could be accepted as reasonable because their complaint is stronger than the complaint of becoming paralysed that might be issued by anyone in Doing Nothing. This criterion for reasonable rejection could be more formally stated as the following—it is unreasonable for contracting parties "to reject a principle because it imposes a burden on you when every alternative principle would impose much greater burdens on others." Frick labels it *the Greater Burden Principle*.[16] It is a principle that specifies when the complaint of a contracting party may be reasonably rejected, which I will appeal to again in my discussion later.

This conclusion is, however, counter-intuitive. Consider a person who is currently paralysed. There is some medical treatment which is 99.999% likely to completely cure her but has a 0.001% likelihood of causing immediate death. It seems entirely reasonable for this person to accept the treatment. If she does accept, we would not, after she realises that she is about to die and feels regretful, then deny the reasonableness of her prior choice. Why, then, is it not acceptable to apply this single-person choice to everyone in the society? It seems the problem ex-post contractualists are concerned with is the certainty of death the treatment causes when applied to a large population. However, as rightly pointed out by Frick, this kind of reply would imply a preoccupation with the overall shape or distribution of outcomes.[17] Being concerned with such patterns of distribution is not the right motivation for contractualists who should be focused on personal reasons in assessing the acceptability of principles.[18]

In fact, if the ex-post distribution of outcomes need to be reasonably acceptable to each contracting party, then no risk imposition with potentially disastrous consequences could ever be acceptable to actual victims[19] (unless there is no better alternative should

---

[16] Frick, "Contractualism and Social Risk", p. 177.

[17] Ibid., p. 197. Frick also considers another reply that may be made by ex-post contractualists —see ibid., Section V for a detailed discussion.

[18] It may be suggested that pattern of distribution matters when parties need to choose between, for example, saving one life and saving two lives (or preventing a hundred cases of paralysis). However, even in such cases, the reason for rejecting a principle which permits saving one life rather than two lives (or preventing a hundred cases of paralysis) need not be based on a concern about the overall distribution. Rather, it is based on the complaint that may be issued by any individual belonging to the latter group which states that the value of his life or his physical health is not properly counted in evaluating the cost of saving one life.

[19] Oberdiek, *Imposing Risk*, pp. 143-44; Frick, "Contractualism and Social Risk", p. 185.

the risk not be imposed—this is often the case in examples of vaccination and medical operations, on which most papers endorsing ex-post contractualism are focused). Requiring justifiability of this kind would prohibit most cases of common risk imposition like driving because few of them can completely rule out serious consequences ex post.

Ex-ante contractualism, on the other hand, could successfully explain our intuition that Vaccination with Deadly Side Effect is more justifiable than Doing Nothing. Even though the harm of death is significant, it is discounted by its small probability compared to the certainty of paralysis. However, ex-ante contractualism faces the different worry that its reasoning may also justify Medical Experiment. If Vaccination is acceptable because expected harm is outweighed by the corresponding health benefit for each individual, then the same could be said about Medical Experiment in which the expected harm of being selected to be killed can be outweighed by the expected benefit. This concern is what motivates Scanlon's discussion of risk imposition in *What We Owe to Each Other*. While he only intends the case as an illustration of how his contractualist framework might be applied and is quite vague in terms of the specific principles that he thinks could justify risk imposition in general, his brief discussion has inspired many subsequent papers on this topic and is therefore worth closer analysis.

3.2.3 Scanlon's account of risk imposition

On the basis of his argument for distinguishing between Vaccination and Medical Experiment, Scanlon seems to reject the employment of expected value calculation and has in mind a different form of ex-ante contractualism. According to him, when looking at a case where risk is involved, the probability of harm is morally relevant in assessing permissibility "not as a factor diminishing the 'complaint' of the affected parties" (discounting the amount of harm by the likelihood of their suffering it), "but rather as an indicator of the care that the agent has to take to avoid causing harm."[20] What he seems to be proposing here is an alternative principle for assessing the permissibility of risk imposition, namely the requirement of due care or reasonable precaution. Let me call this the reasonable precaution principle, which could be stated as follows—

[20] Scanlon, *What We Owe to Each Other*, p. 209.

The reasonable precaution principle: a risk must be imposed with a reasonable level of precaution. That is, its potential harm should not be a result of recklessness or lack of care.

He then goes on to explain the difference between cases like Vaccination with Deadly Side Effect and cases like Medical Experiment. He argues that Medical Experiment is unjustifiable because the harm is directly inflicted on particular people rather than "occurring by accident despite the fact that reasonable precautions have been taken".[21] This argument is a bit obscure because what it is primarily concerned with seems to be less about the taking of precaution. Rather, it is more related to the intention of risk imposition—justifiability of the two cases differ because the resulting harm is merely unavoidable in Vaccination but is intended in Medical Experiment. Let me call this the just intention principle. It could be stated as follows—

The just intention principle: a risk should not be imposed in such a way that the bad consequence is intended.[22]

While Scanlon does not further elaborate, it seems both principles play a role in his argument and the just intention principle is not meant to replace the reasonable precaution principle. To see why the latter has independent justificatory force, consider two cases of vaccination—both carry the same risk of side effect (that is, the same probability of causing potential harm of the same level) but one vaccination has gone through proper testing while no medical trial has been performed for the other. Here the side effect of either vaccination is not intended to harm people, so they both satisfy the just intention principle. The reasonable precaution principle, on the other hand, may work to approve the former case while rejecting the latter because, in the latter case, it fails to implement the risky scheme with a reasonable level of precaution.

---

[21] Ibid., p. 209.

[22] Lenman calls this the aim consistency principle. See Lenman, "Contractualism and Risk Imposition", pp. 104-5. He is concerned that referring to intention may cause confusion that the principle applies to the intention of the actor. This worry is worth clarifying. Lenman is right in arguing that, for example, it is justifiable to allow a person to work as an ambulance driver as long as he follows due care rules even though his real intention in taking up this risky job is to enjoy high-speed driving. It should therefore be noted that my just intention principle is concerned with the nature of the action and the harm it involves rather than the intention of the actor. In fact, what Lenman is referring to may be more properly counted as motivation rather than intention.

By referring to principles like just intention and reasonable precaution, Scanlon's approach to risk imposition differs from both ex-ante and ex-post contractualism in the simplistic form I described earlier. Instead of assessing the permissibility of an action based solely on its outcome, either expected or actual, Scanlon's arguments focus on the nature or characteristics of the action. Two actions may produce the same outcome (or probabilities of outcomes) but differ in their justifiability because, for example, one action carries a risk due to limits in human capacity while the other carries a risk with the same probability and level of potential harm due to lack of care. Intuitively the former is more likely to be justifiable than the latter. This is a crucial aspect of risk imposition that needs to be taken into consideration by contractualists in assessing their justifiability. As rightly stated by James, "we judge what others have done to us based not on worldly outcomes per se but on what they knew or could have known would happen and the risks they took with our lives."[23]

This focus on the nature of actions does not necessitate the abandonment of the presumptive worthiness principle. Scanlon seems to be suggesting in some places that he intends principles like just intention and reasonable precaution to replace expected value calculation altogether in assessing the permissibility of risk imposition.[24] However, it seems highly implausible that expected benefits and costs do not factor in the assessment at all.[25] Intuitively, the smaller the probability of harm, the less burdensome it is to be exposed to the risk, thus making the imposition more likely to be worthwhile. Moreover, while the just intention and the reasonable precaution principles do rule out some unacceptable cases of risk imposition, they are not sufficient because, surely, not all imposed risks satisfying those demands are presumptively worthwhile for all. Imagine, for example, a high-speed automobile which, even when operated with reasonable precaution, causes accidents with a likelihood of 50%. This risk is imposed for convenience without intending any harm and is implemented with due care. The fact that the new technology is unlikely to be justifiable, while a normal type of car is, can

---

[23] A. James, *Fairness in Practice: A Social Contract for A Global Economy* (Oxford: Oxford University Press, 2012), p. 274.

[24] See, for example, Scanlon, *What We Owe to Each Other*, p. 209.

[25] See Note 24 in Kumar, "Risking and Wronging", pp. 23-24.

only be explained by referring to their different probabilities of causing accidents which act as the discounting factor in expected burden evaluation.

In fact, Scanlon later admits that rejecting ex-ante discounting was a mistake on his side.[26] He explains that his original intention was to rule out cases like Medical Experiment in which, at the time of decision-making, it is already certain that the bad outcome will occur, but, as I have argued, such cases could be covered and rejected by appealing to an additional principle like the just intention principle. The just intention principle is, moreover, not incompatible with probabilistic discounting. Kumar, for example, comes quite close to this approach of combining expected value calculation with the requirement of just intention in his contractualist account of risk imposition. He acknowledges the legitimacy of weighing discounted cost and benefit while insisting that risk cannot be imposed in a way that violates an individual's authority in using his own body and disrespects him as an autonomous agent, thus ruling out the case of Medical Experiment.[27] This latter requirement can largely be accounted for by the just intention principle—imposing such a risk fails to respect the victims' agency over his own body exactly because the bodily harm is intended rather than merely foreseen.

While it is easier to see how the just intention principle and the presumptive worthiness principle work together, it may be questioned why the reasonable precaution principle is separately needed when presumptive worthiness is already satisfied. To see why neither principle is redundant, imagine the following case. A patient was diagnosed with a disease which has a death rate of 50% in one year's time. He was informed that a certain operation could cure the disease with a likelihood of 99% but with a 1% chance of causing permanent disability. The patient agreed to the surgery because, based on expected cost and benefit, he considered the risk to be worth taking. After the surgery he

---

[26] T. M. Scanlon, "Reply to Zofia Stemplowska", *Journal of Moral Philosophy* 10 (2013), 508-14, pp. 510-11.

[27] Kumar, "Risking and Wronging", p. 37. Frick objects that Kumar's argument cannot explain the unacceptability of other cases of risk imposition where harm does not take the form of violating bodily integrity but reflects the failure to take a lower risk when it is available as an option. See his argument in Frick, "Contractualism and Social Risk", pp. 207-10. However, what Frick fails to recognise is that Kumar's principle concerning bodily autonomy is not intended as the sole principle for risk imposition and is to be combined with expected benefit and cost analysis. The example he is considering can therefore be successfully dealt with by the combination of these principles.

unfortunately suffered the consequence of becoming permanently disabled. Later, he learnt that, by the time the operation was performed, research had already discovered that the major cause of disability is genetic, and the risk could have been easily reduced to 0.1% by doing a pre-operation screening. In this example, supposing this piece of information was easily available, the doctor could be said to have failed to take reasonable precaution by ignoring it. If this is so, then the patient may reasonably regard the risk as being unjustly imposed even though the 1% risk is worth taking in terms of expected return.

To briefly summarise what I have argued so far, in order for a scheme of social risk imposition to be reasonably acceptable to contracting parties, it needs to satisfy, first of all, two principles concerning the nature or character of the risky action—it should not involve harm that is intentionally inflicted (the just intention principle) or harm resulting from lack of adequate precaution (the reasonable precaution principle). With these conditions met, its expected benefits and burdens are then to be assessed by the presumptive worthiness principle and the fair distribution principle. The scheme may be justifiable if it yields presumptive benefits that significantly outweigh the expected burden for each affected individual and the benefits and burdens are distributed in a way that is considered fair.

3.3 The Compensation Principle and the Demand for Risk Sharing

In the last section, I proposed two principles for determining the justifiability of risk imposition. While the case for the just intention principle is made quite clear by the comparison between Vaccination and Medical Experiment, the rationale underlying the reasonable precaution principle is worth closer examination. Why would contracting parties demand that a scheme of risk imposition should be regulated by the reasonable precaution principle? This relates to the fact that the same type of risk may be imposed in various ways with different levels of burden for the affected individuals.

In considering the justifiability of a typical case of public goods provision, it is implicitly assumed that the cost of provision is fixed. The presumptive benefit principle and the fair distribution principle then come into play to assess whether the fixed cost is

overweighed by the corresponding benefit and whether it is fairly distributed. When it comes to a scheme of risk imposition, however, the fact that its cost may vary depending on how the risk is imposed becomes centrally important. For the same level and type of benefit, if a risk could be imposed in a way that the probability or the magnitude of the potential harm is lower, then a higher level of risk would no longer represent the inevitable cost that individuals must bear in order to produce the corresponding benefit for all.

For example, consider reasonable precautions that may be required for allowing car driving. It is likely to include regulations such as a speed limit and a prohibition on drunk driving as well as the requirement for a regular quality check for automobiles and road maintenance. If drunk driving is not prohibited, then the practice of car driving could be reasonably rejected by contracting parties because the increased likelihood of accident brought about by drunk driving is not the inevitable cost of convenience and higher efficiency made possible by driving. The fact that an accident is not occurring, in Scanlon's words, "despite the fact that reasonable precautions have been taken" shows that it is disrespectful to those subject to the potential harm of the risk. Contracting parties could reasonably reject such a scheme because it imposes an unnecessary threat to their autonomy.

To be sure, these possible measures of precaution are themselves burdensome as they may take up resources or restrict individual behaviour. So it would not be reasonable to demand all possible precautions be taken so as to minimise the probability or magnitude of a bad outcome regardless of cost. In fact, doing so may end up undermining the preconditions for autonomy for all and introducing burdens that outweigh the benefit for which the risk is imposed. It seems that the justifiable level of precaution will need to be determined by actual members of a society, taking into consideration factors like human capacity, technological and economic development, and the social and cultural environment in that particular society. In the next chapter, I will come back to the content of the reasonable precaution principle and examine its relationship with the notion of responsibility.

For now, what I want to demonstrate from the discussion is the underlying rationale for the reasonable precaution principle—it is a demand that, in order for a scheme of risk imposition to be justifiable, its burden, including the probability and magnitude of its potential harm, should be reasonably reduced. When a similar benefit could be achieved with a lower level of risk, it is reasonably rejectable if the risk is not reduced to this level.

If this is the case, then the taking of precaution is not the only implication of this demand for reasonably reduced burden. It also points to an additional requirement which could similarly reduce the burden of risk imposition—the requirement of ex-post compensation for victims of risk materialisation. Compensation can effectively alleviate burdens in two ways. First, it helps to restore the condition of autonomy for victims of risk materialisation. In most cases, harm resulting from common risks, either physical or economic, can at least be partially rectified with external resources (including the payment of monetary compensation or the provision of services such as medical care). In such cases, compensation serves to reduce the magnitude of the burden to be borne by victims. Deadly risk may be an exception here because it cannot be compensated for. In fact, in discussing the justifiability of risk imposition, there is a tendency for philosophers to choose to focus on deadly risks. While it helps to provide the strongest case that even the imposition of a lethal risk may be permissible, this focus has probably led many to overlook the importance of compensation. When harm can be rectified, individuals may reasonably reject a scheme of risk imposition because it fails to rectify the harm, thus reducing the actual burden for victims, by providing compensation.

Second, even before a risk ripens, the guarantee of compensation would help to reduce fear and anxiety for all those who are subject to the risk in question. This also applies to deadly risks because compensation may serve to assure people that, in the case of risk materialisation, their families or loved ones will be provided with financial support. In addition, with the guarantee of compensation people are also relieved of the burden of taking risk management decisions for commonly imposed risks. Since social risks are not voluntarily taken on by individuals, having to assess and manage their exposure represents an additional burden for people and may distract them from their main life

projects. Hence, by reducing the burden of risk anticipation and management, the guarantee of compensation also makes the scheme more likely to be justifiable to contracting parties.

Since compensation, like the taking of precaution, could reduce the burden of risk imposition, it is justifiable for contracting parties to insist that the scheme has to guarantee that compensation will be paid if the risk materialises. This would be an additional principle for just risk imposition which I will call the compensation principle —

> *The compensation principle*: a risk must be imposed with the guarantee that victims of risk materialisation will be compensated in such a way that the harm they suffer is reasonably reduced.

Note that compensation justified in this way differs from the idea of compensation as a result of rights violation for imposing a risk or for causing a risk to materialise.[28] In the contractualist framework, contracting parties do not assume such an existing right against risk imposition. Rather, they consider principles of risk imposition that no one could reasonably reject and any right they may have concerning risk imposition is derived from such principles. The compensation principle stated here, for example, may be regarded as generating a right against imposing-risk-without-guaranteeing-compensation.

The implication of this distinction is two-fold. First, it breaks the otherwise tight link between the payment of compensation and the individuals who actually carry out the risk-related activities and/or cause a risk to eventuate. If compensation is required due to a rights violation, then the person who has violated the victims' right against being exposed to risk will be required to bear the full burden. This is what leads to Nozick's concern about the risk imposers' ability to pay as discussed in Chapter 1. Under my account, however, compensation, like the taking of precaution, is for the purpose of

---

[28] I have discussed Nozick's account in Chapter 1. For a different account on compensation and risk imposition, see McCarthy, "Rights, Explanation, and Risks". Instead of treating compensation as a result of rights violation, McCarthy distinguishes between rights violation and rights infringement and takes compensation for risk imposition as a result of the latter. This debate, however, does not need to concern us here as my account differs from both of them.

reasonably reducing the actual costs for potential victims so as to make the scheme justifiable to all ex ante. The payment of compensation is part of the overall burden of the scheme that is to be fairly allocated among all those impacted by risk imposition. It is a further question whether being engaged in the risky action or being responsible for causing the risk to materialise should factor in the fair allocation. I will address this question in the next chapter.

Setting this question aside, what could be said for now is that those who are directly involved in the causal chain leading to risk materialisation are, most likely, not the only ones that need to bear the burden of compensation in cases of common risk imposition. Since a common risk is imposed for presumptive benefits, similar to the fairness-based argument for public goods provision, those who have benefited from the imposition should also contribute to the payment. Indeed as acknowledged by Scanlon, it is unreasonable for contracting parties to reject a fairness principle which requires those who have benefitted from an arrangement or a scheme of cooperation to make a corresponding contribution.[29] Since common risk imposition is presumptively beneficial for almost all members of the society, the fairness principle implies that they should all make a contribution and bear a share in paying for compensation. This is effectively a demand for risk sharing by all.

Second, using compensation to reasonably reduce the burden of risk imposition also means that compensation need not always be full. If compensation is justified as a result of rights violation, then it is clear that full compensation should be paid or, where impossible, the situation should be rectified as far as possible. However, in my account, similar to the reasonable precaution principle, the compensation principle only requires that the burden for victims be reasonably reduced. This may imply full compensation when paying full compensation does not create unreasonable burden for those who need to bear a share. This seems often to be the case when the risk in question exhibits a low and stable probability of materialisation and is not highly correlated among individuals (that is, the risk is similar to those insurable in the private market). In such cases, because of the law of large numbers, the amount of compensation required at any given

[29] Scanlon, *What We Owe to Each Other*, p. 212.

time would be limited and predictable. If this amount is to be shared by all individuals subject to the duty of fairness, it is likely not to be significant for each person even when full compensation is required.

Moreover, since the total amount of contribution can be estimated beforehand, it can also be transformed into a series of small regular payments similar to the arrangement of private insurance schemes. While this commitment to regular contribution still restricts people's range of options at any given time, its impact on their overall condition of autonomy is likely to be much more limited than sudden, lump-sum payments and can usually be sufficiently outweighed, over their life course, by the benefit of avoiding disastrous consequences plus the opening up of valuable options made possible by risk imposition.

On the other hand, if the risk deviates from a typical insurable risk, that is, if it is difficult to estimate in terms of likelihood or is expected to affect a large group of individuals at the same time, then it seems that the demand for full compensation for all victims is more likely to impose significant burdens on duty bearers. Consider, for example, the risk of nuclear disaster or the risk of a great economic recession. In such cases, instead of demanding a one-off payment of full compensation for all victims, a less than full amount of compensation, or compensation to be paid over a longer time horizon after risk materialisation, may be optimal in smoothing disruptions without significantly restricting everyone's condition of autonomy. The trade-off may be made with the guidance of the Greater Burden Principle. Recall that this principle for reasonable rejection requires contracting parties to choose among principles the one that minimises the burden for the worst affected. Hence the compensation principle that would be reasonably endorsed by contracting parties would require that the burden for victims of risk materialisation be reduced unless further reduction would lead to an even greater burden for those who have to pay for the cost.

At this point, a question may be raised—by arguing that a risk is socially imposed and its victims should therefore be compensated by those who benefit from its existence, does my account implicitly assume a baseline of no risk? If so, it leads to a further concern—while referring to a state of no risk as the baseline seems natural for risks that

are plainly introduced into human society as a result of human decision (such as the risk of car accidents and the risk of nuclear disaster), should the hypothetical state of no health risk or no natural disaster also be treated as the baseline for justifying compensation to victims?

To answer this question, let me start from risks imposed through human actions. In such cases, a state of no risk may provide a natural starting point for assessing the justice of risk imposition, but it does not mean that any state without the existence of the concerned risk can serve as a baseline to justify fair risk sharing. The no-risk baseline has to be a state that is itself just. For example, imagine a society without car driving. This society may endorse justice as equal prospects of autonomy or some other conception of justice. When its institutions are arranged in a way that meets the demands of justice (call this State 0), this state can then be treated as the baseline for assessing the justice of introducing the risk of car driving. Because it is a no-risk state, we do not need an account of how to justly deal with the concerned risk in order to claim that it is a just baseline. So starting from this kind of case is especially useful when we are developing an account of just risk imposition.

On the other hand, to apply the account of social risk imposition, the state of no risk need not be the only acceptable baseline. To illustrate, let me continue with the above example. First, suppose this society wants to introduce a small risk of car accidents by allowing only emergency driving. In order to do so justly, it imposes the risk according to principles outlined by my account (call this State 1). If the society then considers imposing an even higher level of risk by allowing convenience driving in addition to emergency driving, this new case of risk imposition can be assessed either against the baseline of no risk (State 0) or the baseline where the lower level of risk has been *justly* imposed (State 1, which already demands a level of compensation for victims of the existing risk). Through either way of justification it will arrive at the same level of total compensation owed to victims.

However, things get more complicated when it comes to cases involving risk that naturally exists in human life, especially when it is the kind of risk that cannot be eliminated through human intervention. In such cases, applying the no-risk baseline no

longer seems appropriate. This is not because the no-risk state represents a hypothetical scenario but because it is infeasible (unlike, for example, the elimination of the unemployment risk, which is hypothetical but theoretically possible). Since we cannot achieve a state of no health risk no matter how we arrange our society, it does not make sense to consider how to justly move from this inaccessible state to a scenario in which we then choose to impose (that is, retain) a level of health risk.

Note that what I have said so far does not prevent us from dealing with natural risks from the perspective of social risk imposition because, as I have argued, institutions can be designed in ways that affect their likelihood and magnitude of impact, so retaining a higher rather than lower level of a natural risk may still be treated as a case of social risk imposition. The challenge is to choose an appropriate baseline in order to apply my account. It seems that, in such cases, what may serve as a baseline would be an arrangement where the relevant risk is reduced as much as possible with the constraint that other demands of justice that have priority over the elimination of the risk are being met. This baseline represents the scenario where best efforts have been made to eliminate the risk without violating other more urgent demands of justice, so a level of risk that is higher than this baseline can be treated as being socially imposed through collective choices.[30] Let me illustrate using the example of the current pandemic.

The coronavirus disease itself is naturally caused and seems impossible to eliminate at least in the short run, but we can imagine that the risk of contracting the virus may be reduced to the lowest possible level by enforcing a strict lockdown which prohibits anyone from leaving their homes. However, this cannot be a just state because in this hypothetical scenario all public services will have to stop and, as a result, the lives of many dependent or vulnerable individuals will fall below sufficiency, which violates the demands of most conceptions of justice. Hence, it seems that the baseline for assessing the justice of risk imposition is more likely to be a state where the society lowers the

[30] This is consistent with our intuition that we should adopt a lower baseline when the risk entails more severe consequences and when the risk is easier to reduce—in such cases keeping the existing level of the risk requires stronger justification (and more compensation). Also note that the baseline of the lowest possible level of risk compatible with other demands of justice can naturally incorporate the scenario of no risk for cases where the risk is introduced through human action, so it is consistent with my earlier argument that both types of cases are structurally similar.

risk as much as possible but still fulfils the basic demands of justice it endorses. This may be achieved by, for example, enforcing a reasonably strict lockdown while providing basic public services. Whether the society can justly move to a higher level of covid risk from this baseline can then be evaluated based on the principles of just risk imposition I have proposed, and the additional risk may be considered as being socially imposed and its resulting burden on victims calls for fair sharing by all.

Some may object that, in this example, there may be other reasons for the society to pay for the treatment of covid, whether or not we treat contracting the virus as a socially imposed risk. Indeed, in many cases of illnesses and diseases, due to the special moral importance of health, other considerations of justice may require a society to provide some level of universal healthcare, in which case there is no need to appeal to social risk imposition to justify the sharing of medical cost. However, treating the pandemic (as well as other diseases with a natural cause) at least partially as a case of social risk imposition may still be useful for two reasons. First, it provides a case for societies which do not otherwise recognise universal healthcare as a demand of justice to share part of the cost of medical treatment. Second, even for societies which already provide universal healthcare, recognising the risk as partly a socially imposed risk may additionally justify compensation for other burdens as a result of falling ill, such as the temporary loss of income.

Now, with a reasonable level of compensation justified, the final question is how the burden should be fairly allocated among those who bear the duty of compensation. I will turn to this question now. In particular, since I have argued that the justification for a scheme of social risk imposition resembles that of public goods provision, I will examine if existing literatures on the fair allocation of burdens in providing public goods may shed light on my account of fair risk sharing.

3.4 Fair Social Risk Sharing

In the general case of public goods provision, two different principles have often been proposed for fair burden allocation. One proposal is allocation based on the ability to pay. By demanding more from those who could afford to pay more, this principle is

deemed fair by some people because it implies the requirement of equal sacrifice. However, as pointed out by Murphy and Nagel, such reasoning treats the burden of funding public goods as if the scheme itself were merely a common disaster which is to be shared by all society members.[31] It ignores that what matters is not burden allocation per se but whether the scheme of public goods provision as a whole, which includes not only burden but also benefit, is overall fair and thus justifiable to all.

This leads to the other popular proposal—the proportional benefit principle, which requires that, against a background of just distribution, the funding of public goods be arranged according to some standard of proportionality to benefit.[32] Benefit could be measured by, for example, each individual's reserve price for the public good. Intuitively, this principle is likely to be justifiable to contracting parties because reserve price represents what people would voluntarily pay for the goods themselves.

One objection to this principle is concerned not with its spirit but with how it could be operationalised, especially how each person's reserve price could be measured. In the context of risk imposition, reserve price may in fact not be a good term because people gain from the scheme of social risk imposition mainly through the improved set of life options rather than the enjoyment of a particular good which they could in theory buy from a market. So, instead of referring to reserve price, the proportionate benefit principle may be stated in a way that makes burden proportional to each individual's enhanced life prospects.

Of course, this restatement does not in itself solve the problem—improvement in life prospects is still difficult to measure, especially when a particular risk is singled out for assessment. However, it seems that some rough evaluation could be made. For example, the institution of a market economy as a whole imposes various economic risks on individuals compared to the no-risk scenario (for example, a centrally planned economy with need-based pay, universal health care, etc). The extent to which each individual

---

[31] L. Murphy and T. Nagel, *The Myth of Ownership: Taxes and Justice* (Oxford: Oxford University Press, 2002), p. 25.

[32] See, for example, ibid., ch. 4, esp. p. 85; and B. Fried, "The Case for a Progressive Benefits Tax", in *Taxation: Philosophical Perspectives*, eds. M. O'Neill and S. Orr (Oxford: Oxford University Press, 2018).

benefits from this choice may very roughly be represented by the difference in income and wealth they would receive in the market economy and under central planning. If this is an acceptable proxy, then the proportionate benefit principle would require those who possess more income and/or wealth to bear a relatively larger burden in paying compensation to the victims of economic risks. In such cases, its implication may in fact be similar to that of the ability-to-pay principle in allocating the burden of economic risks.

Note, however, that there is a wrong way of specifying the proportionate benefit principle, which takes the principle to mean that those who benefit more, in the sense of greater improvement under the current state by comparison with the state of nature, should make a proportionally larger contribution. This understanding would lead to absurd conclusions such as that those who are in greater need of government support should bear more tax burden.[33] This interpretation is wrong because the proportionate benefit principle should apply to an otherwise just background distribution rather than to the state of nature. That is, what could serve as a benchmark for evaluating relative benefit should be a state of affairs that is itself justifiable to contracting parties. The equivalent of this common misunderstanding in the context of social risk is that the principle would require those who are subject to higher risk and are therefore more likely to receive compensation to make a proportionally higher contribution. This interpretation treats the state where every individual bears their natural risk exposure as a benchmark for the assessment of relative benefits. However, as I have discussed, it is not a just benchmark because leaving people exposed to existing risks is a state that calls for justification in the first place.

A different worry about the proportionate benefit principle is that it is not egalitarian in nature. If an equal overall distribution is what is reasonably acceptable to contracting parties, then the proportionate benefit principle is not necessarily going to preserve such equality. Miller gives an example for this objection: suppose the state could choose between two schemes, one generating 40 units of benefit for individual A and 20 units of benefit for individual B at a total cost of 30 units and the other generating 40 units of

[33] Murphy and Nagel, *The Myth of Ownership*, pp. 16-17.

benefit for each at a total cost of 50 units. The principle merely requires that the first scheme allocate 20 units of cost to A and 10 units to B and 25 units to each for the second scheme. It does not require the state to prefer the second scheme over the first. If the first scheme is chosen, A would enjoy a net benefit of 20 while B enjoys 10. If they start with an initial equal distribution, then such equality is undermined by the scheme of public goods provision.[34]

In fact, what the proportionate benefit principle attempts to replicate is the result of voluntary exchange in the free market. Therefore, unless the contracting parties regard any distributive result of a free market economy (or a market economy with, for example, a safety net) to be just, applying this principle will not produce or retain a just distribution. If what contracting parties endorse is a certain kind of egalitarian distribution, then it would only be justifiable that the scheme of public goods provision preserves such distribution as far as possible.[35] This leads to Miller's proposal of the equal net benefit principle, which requires that each individual's net benefit from the supply of the public goods be equal.[36] So, for my account, the equal net benefit principle implies that, starting from a condition of equal prospects for autonomy, a fair scheme of social risk sharing should allocate the burden in a way that preserves such equality as far as possible.

On the other hand, even if we cannot agree about the content of the benefit principle (whether it should require proportionate benefit or equal net benefit), as long as we agree that contracting parties would reasonably accept a version of the benefit principle, this would justify compulsory risk sharing by all members of the society in a form that is different from private actuarial insurance. This is because the benefit principle, regardless of its specific content, contrasts with the actuarial principle which allocates burden according to the size of the risk each insurance buyer carries and requires those with higher risk to pay a larger premium. This provides a strong case for fair risk

---

[34] Miller and Taylor, "Public Goods", pp. 568-69.

[35] If not possible, then something like Rawls' difference principle may be adopted which requires that we prioritise the worst-off and maximises their preconditions of autonomy. This should be reasonably acceptable to contracting parties in the Scanlonian framework based on the Greater Burden Principle.

[36] Ibid., p 569.

sharing through state mandated universal risk sharing schemes (for example, social insurance programmes—I will come back to this in Chapter 7) because the state could design the schemes in a way that deviates from actuarial insurance and fulfils the demand of the benefit principle. It should be noted, however, that the benefit principle only applies to the social component of a common risk, so such risk sharing schemes may often need to be supplemented by private insurance when some individuals need additional coverage for their higher risk exposures due to personal choices.

Moreover, my argument does not mean that existing social risks must be justifiable as long as they are accompanied by a scheme of fair risk sharing. It is perfectly possible that, given the stage of economic, technological and social development, members in a society decide that the benefit of, for example, a free labour market is outweighed by the burden of unemployment risk, so that they actually prefer an alternative economic arrangement in which, for example, people have guaranteed jobs with equal pay. What it does mean is that, when unemployment risk is actually imposed on members of the society, justice requires that the imposed risk be accompanied with a reasonable level of precaution in the form of labour law and a guarantee of compensation in the form of, for example, unemployment insurance, whether or not it further requires economic reform in the long-term to reduce or eliminate the risk.

Lastly, the fair risk sharing principle proposed here (which requires the burden of risk imposition be allocated based on its benefit for each individual) is only tentative because my discussion so far has not considered the role of personal responsibility. It may be suggested that, even for social risk, individuals are sometimes responsible for either causing the risk to materialise or failing to manage their own risk exposures and such responsibility should play a role in determining the fair burden allocation in a scheme of social risk imposition. I will turn to this issue of personal responsibility in the next chapter.

3.5 The Problem of Plural Risk Preferences

Before moving on to the next chapter, however, I will address one remaining concern about the viability of the contractualist approach in justifying a particular scheme of

common risk imposition. The concern is that, while my proposed principles may serve to rule out illegitimate cases of risk imposition, they cannot justify a specific level or form of common risk to contracting parties because they have different risk preferences.

The problem is often illustrated with the example of the Amish farmers. To justify imposing the risk of car accidents to all members of a society, it is stated that no one could reasonably reject the net benefit made possible by allowing car driving. However, this ignores that there may be people, such as the Amish farmers, whose life style barely benefits from the practice of car driving. They are, however, also exposed to the risk of car accidents because highways cut across their villages. It does not seem reasonable to insist that the risk is justifiable to them.

One way of interpreting this objection is that it is not justifiable to subject Amish farmers to the risk because they do not enjoy the corresponding benefit at all. However, first, while they may not benefit directly, it is likely that they still enjoy indirect benefits made possible by the use of cars, such as the production of certain goods or services that they need. Moreover, even setting aside such empirical possibilities, benefits that are relevant to the evaluation of the case should not be limited to that of car driving. As correctly pointed out by Oberdiek, according to the contractualist approach, discrete instances of risk imposition should be assessed as an instantiation of general principles which regulate types of risk-imposing activities.[37] Hence, in this case, while Amish farmers' possible objection to driving seems reasonable when viewed in isolation, it is less so if the objection is generalised to all activities with a similar profile to that of driving, including, for example, their own use of horse-drawn buggies, from which they have also benefited.

However, Oberdiek's reply only shows that it is not reasonable for people like Amish farmers to demand a total ban on activities similar to driving. It does not further explain why it is justifiable to require them to accept the specific form and level of risk imposition. Comparing the system which permits the use of high-speed automobiles to an alternative system which only allows less risky vehicles such as horse-drawn buggies, it seems obvious that Amish farmers would reasonably prefer the latter.

---

[37] Oberdiek, *Imposing Risk*, pp. 144-45; pp. 147-50.

This version of the objection could be generalised as the problem of justifying a particular level of risk to a population with plural risk preferences. In a Scanlonian contracting position, individuals are aware of their diverse risk attitudes and seek to reach a collective decision about a scheme of social risk imposition. However, given the plurality of risk preferences and the need to settle on a common social arrangement which embodies a certain level of risk, it seems that, whatever level is chosen, there would always exist contracting parties who regret the choice as being either too risky or too conservative. Some may therefore question if any specific scheme of risk imposition could be justified using the Scanlonian contractualist approach.

One possible response is simply that the contractualist approach is employed to identify a range of acceptable choices by ruling out those that violate principles of just risk imposition, but a decision within this range could not be settled within its framework and is something to be left to democratic decision-making. Instead of resorting to this solution right away, I would like to explore if Scanlonian contractualism has the resources to settle the disagreement at least to some extent, that is, whether it could offer people with different risk preferences a reason why they could not reasonably reject a certain level of common risk imposition.

Recall that earlier I have described how the Greater Burden Principle may help to settle this kind of disagreement. It states that it is unreasonable for contracting parties to reject a principle when every alternative principle would impose much greater burdens on others. The question, then, is how to evaluate the burden of being imposed with a level of risk that is either higher or lower than one's own risk preference.

In discussing physical risks, it is often suggested that a person's risk preference represents his desired balance between the liberty interest and security interest. For instance, allowing the risk of car driving threatens one's security interest in not being hit by a car, but it promotes his liberty interest in being able to move around more efficiently. A more risk-averse person, then, is someone who is more willing to sacrifice his liberty interest in return for more security. Presented as a trade-off between security and liberty interests, however, it seems difficult to weigh the complaints issued by a person who is more risk averse and another who is more risk tolerant—under what

condition would the burden of compromising liberty interest be worse than the burden of compromising security interest?

In fact, it may be noted that, what people call liberty interest in the case of risk imposition is mainly concerned with the number and variety of options available in the background; what they refer to as security interests, on the other hand, is concerned with the robustness of the availability of certain options. Based on what I discussed in Chapter 2, how an individual is impacted by a certain level of risk imposition can therefore be assessed based on how the person's prospects for autonomy have been affected. Even though the net impact of, for example, increased liberty and decreased security cannot be quantified straightforwardly using a single formula, in specific cases, we are often able to evaluate how a change in the balance between these two types of interests would affect the prospects for an individual to author his or her own life. Let me illustrate using the Amish farmer example.

An Amish farmer may regret that, because of his stronger interest in having a tranquil life and the much smaller benefit he enjoys from cars or planes, he suffers a greater burden than others by having highways cross his village or by living under flight paths. Therefore he would like to demand that highways or flight paths avoid his neighbourhood. He may even want to request an alternative transportation system which allows only his desired transport vehicles, namely horse-drawn buggies. The burden implied in such demands for other members of the society would vary depending on the actual society this Amish farmer lives in. If, for example, Amish farmers live in a concentrated area, then the cost for highways to avoid their villages may be much lower than the case where their habitats are more dispersed. If, in this society, planes and cars are used mainly as a form of entertainment, then the burden of prohibition on those who use them would be much lower in terms of their compromised prospects for autonomy compared to the case where they are an essential means of moving around that are integrated into most people's life projects.

Hence, with the reasonable requests of taking precaution and paying compensation being satisfied, at some point, further complaints of an Amish farmer about the risk of car accidents may be rejected based on the Greater Burden Principle as follows—the

remaining burden of the risk of being hit by a plane or a car is relatively small in terms of reducing the potential for you to realise your preferred life. Compared to this burden, the alternative arrangement of, for example, having all highways and flight paths avoid your living area would impose a much greater burden on others in terms of their prospects for living their desired life.

This example shows that risk preferences are not uniquely special compared to other considerations that are relevant for people's prospects to live an autonomous life. An individual may be said to enjoy a greater prospect of autonomy when he is better placed to author his preferred life story, whatever that may be, with his risk preference being just one of the considerations in that story. Therefore, instead of focusing on satisfying people's risk preferences, the design of social institutions should aim to secure the optimal background preconditions for autonomy for everyone, optimal in the sense that every individual enjoys the greatest potential for living a self-authored life that is compatible with others' enjoying the same.

In reality, there may be no arrangement that could achieve an equal promotion of prospects for autonomy for everyone. In such cases, contracting parties could apply the Greater Burden Principle to determine an optimal level of risk imposition that is justifiable to all contracting parties. However, this so-called "optimal level" is a simplified term which represents not just how high a risk is but the whole arrangement of its imposition, such as how the risk is imposed, what rules of reasonable precaution would accompany its imposition and how victims would be compensated. Therefore, in reality, it is likely that contracting parties will not be able to freely choose an optimal level from a continuous scale and their options may be limited to a few feasible alternatives. If, among these options of imposing a social risk, accepting one arrangement would satisfy the Greater Burden Principle, then it represents the most justifiable arrangement. If, on the other hand, no obvious choice could be justified according to the Greater Burden Principle because the burdens they impose on different individuals are of similar levels, then a choice among these alternatives could be made through fair democratic procedures.

Finally, a similar rationale could also be applied to settle the problem of different risk exposures. For certain risks, it is obvious even in the initial contracting position that they could not be imposed in a way that leaves all individuals equally exposed. Unemployment risk is such an example. In the Scanlonian contracting position, some of the contracting parties know that their talents would be in less demand in the labour market and they would therefore be subject to a higher risk of unemployment if a market economy (rather than a centralised economy) is chosen. It may be questioned why it would ever be reasonable to require such people to accept the imposition of the unemployment risk. This objection is slightly different from the Amish farmer objection because it is concerned with different risk exposures rather than different risk preferences, but it could be similarly addressed by appealing to the Greater Burden Principle. A market economy may undermine people's preconditions for autonomy because of the unemployment risk it introduces but it may also enhance prospects for autonomy if it increases economic efficiency and expands the range of life options. It is therefore possible that the overall scheme (with the satisfaction of the reasonable precaution principle and the compensation principle) cannot be reasonably rejected even by those with the worst risk exposure if every alternative arrangement would impose much greater net burden on others.

To recap, I have so far argued for the following principles for the just imposition of a social risk—

1. *The just intention principle.* It requires that a risk not be imposed in a way that the bad consequence is intended.

2. *The reasonable precaution principle.* It requires that a risk be imposed with a reasonable level of precaution. That is, its potential harm should not be a result of recklessness or lack of care.

3. *The compensation principle.* It requires that a risk must be imposed with the guarantee that victims of risk materialisation will be compensated in a way that the harm they suffer is reasonably reduced.

4. *The presumptive worthiness principle.* It requires that, with the satisfaction of the above principles, the expected burden of the imposed risk is sufficiently outweighed by the corresponding benefit for each contracting party in terms of their overall prospects of living a self-authored life.

5. *The fair risk sharing principle.* It requires, tentatively, that the burden of risk imposition be fairly allocated based on some version of the benefit principle.

While these principles do not by themselves determine what risks and what levels for each risk a society should choose to impose on its members, they do rule out some cases as being straightforwardly unjust. Among options which satisfy these principles, people may appeal to the Greater Burden Principle to deal with the problem of plural risk preferences or different risk exposures. Remaining disagreement could be settled using fair democratic procedures in order to reach a collective decision about social risk imposition.

Chapter 4 — Responsibility and Fair Risk Sharing

In the last chapter, I explained why we should treat many common risks as being imposed on members of a society by the choice of social and economic institutions and laid out principles of social risk imposition that are justifiable to contracting parties in the Scanlonian contracting position. Among these principles, the principle of fair risk sharing requires that compensation be paid to victims of social risk materialisation and that the burden of compensation be shared among all people who benefit from the imposition of the social risk. To specify the fair allocation of such burden, I tentatively proposed a benefit principle which states that those who benefit more should bear a larger share of the burden.

The proposal is tentative because it leaves an important concern unaddressed—many social risks are not imposed on people like a lottery. That is, the bad outcomes suffered by victims do not always happen by pure chance. There are often individuals who actively engage in the relevant activities, exercise their agency and manoeuvre themselves through the process which finally leads to the eventualisation of the potential harm (suffered either by themselves or by others). If we respect individuals as autonomous agents who author their own life stories, should we also let them pay for the outcome of their actions and choices, even when risk and uncertainty is involved in the process? That is, should we hold them responsible for their risk-involving activities by requiring them to bear a share of the burden of compensation that is larger than what the benefit principle would otherwise allocate to them?

It seems that we may have different intuitions about such questions in different contexts. Sometimes we tend to believe that it is sufficient to hold people responsible for the consequence of their risk choices, even though they have no control of the outcome at all. For example, if a person loses a voluntary gamble, most people's intuition is that he is fully responsible for bearing the loss. In other cases, the degree of control does seem to matter. We generally believe that a driver who hits a passer-by because the steering wheel suddenly stops working is in some sense less responsible for the accident than a driver who causes an accident simply because of recklessness and this may also affect our judgement about how much cost the person should pay for (for example, we may be

more willing to share the cost with the first person through insurance schemes than the second).

In this chapter, I will explore the relationship between risk and personal responsibility. I will start by looking at Scanlon's own account of responsibility in Section 4.1. In particular, I will examine his account of substantive responsibility, which specifies the conditions for holding people to bear the burden of their choices, and argue that the account needs to be supplemented by two additional conditions—the reasonably avoidable cost and the justifiable cost. I then extend this revised account to the case of risk in Section 4.2. Risk presents a challenge for applying the choice-based account of substantive responsibility because it creates a gap between choice and its outcome. To fill in this gap, I examine two candidate principles in the second section—I call them the full consequence view and the expected return view. I argue that their applicability is case-specific and depends on factors such as whether a risk could be reasonably avoided and whether the likelihood of the bad outcome could be reasonably reduced. I then propose the reasonable expectation account in Section 4.3. It outlines the expectations we may reasonably have for people to manage their risk exposure and/or the risk outcome. I argue that substantive responsibility should be allocated in a way that reflects these expectations.

4.1 Scanlon's Account of Substantive Responsibility

Since I employ the Scanlonian contractualist framework in this thesis, I will explore the issue of individual responsibility by starting from the account Scanlon proposes within his framework. Scanlon distinguishes between two notions of responsibility—responsibility as attributability (or what he later calls moral reaction responsibility) and substantive responsibility.[1] The former type of responsibility considers whether positive or negative moral appraisal can be appropriately attributed to a person as a result of his actions. On the other hand, claims about whether or not people have obligations to others or claims against others are judgements of substantive responsibility. The

[1] T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1998), pp. 248-49; T. M. Scanlon, "Forms and Conditions of Responsibility", in *The Nature of Moral Responsibility: New Essays*, eds. R. Clarke et al. (New York: Oxford University Press, 2005), pp. 89-90.

question that I need to address in this chapter—whether people should bear the consequences of their own risk-involving actions or whether they are entitled to having others share such burdens—is then, based on this distinction, a question of substantive responsibility.

According to Scanlon, to hold a person substantively responsible for the consequences of his action, it needs to be established that he has had adequate opportunities to choose otherwise.[2] To illustrate, Scanlon gives an example in which a city has to remove some hazardous waste found in a residential area. Officials of the city have taken reasonable precautions to reduce the risk by trying their best to inform everyone to stay indoors while they carry out the operation. However, several persons were still exposed and suffered health damage, among whom are the following—

> *The ignorant man*: the first person went out for his usual walk during the operation because, despite the government's best efforts to inform everyone, he did not get the warning.
>
> *The curious woman*: the second person got the warning but still chose to go out because she was curious to see what was being done.
>
> *The rational scientist*: the third person is similarly aware of the danger, but she needs to carry out a scientific project outdoors on that particular day. After carefully weighing the value of being able to advance her project against the health damage, she decided to go out.
>
> *The forgetful woman*: there is also a fourth person, who received the warning but went out during the operation because she simply forgot about it.[3]

According to Scanlon, in this example, the curious woman, the rational scientist and the forgetful woman have all been placed in a choice situation that is sufficiently good for

---

[2] Scanlon, *What We Owe to Each Other*, p. 249, p. 256.

[3] Ibid., pp. 256-57.

them to opt for the alternative of staying at home.[4] The fact they have instead chosen to go out then justifies having them bear the burden of their own choices. As Scanlon argues, "…the fact that a person could have avoided a certain burden by choosing appropriately make[s] it the case that he cannot reasonably reject a principle that makes him bear that burden."[5] The same cannot be said about the first man who does not receive the warning—since he is not aware of the danger and cannot therefore make an informed choice, he should not be held substantively responsible for his injury.[6]

Why does choice play such a decisive role in Scanlon's account of substantive responsibility? According to Scanlon, his account is grounded in the value of choice—because it is valuable to have choices and to have what happens to us depend on our choices, we should preserve such value by holding people responsible for the consequences of their choices.[7] Scanlon's argument is somewhat brief and calls for further elaboration. Let me look into his reasoning in more detail.

To start with, Scanlon lists the different types of value choice may have. The most obvious value of choice is instrumental. For example, by choosing a dish from a food menu, I make sure, or at least make it more likely, that what appears on my plate later will be what I prefer to eat. Choice may also have representative and symbolic value. The gift I carefully choose for my partner bears representative value that is otherwise lacking if I leave the choice to someone else. A vote cast by a citizen in a country where it does not have any influence on the outcome of an election may nevertheless have symbolic meaning for the (non-functioning) democracy. Scanlon believes that these values of choice provide reasons for contracting parties to endorse a principle of substantive responsibility which requires individuals to bear the burden resulting from their own choices, provided that they are well placed to choose otherwise. Why is it so?

---

[4] There are some doubts whether the curious woman is indeed well placed to choose otherwise. See, for example, A. Voorhoeve's discussion in "Scanlon on Substantive Responsibility", *Journal of Political Philosophy* 16 (2008), 184-200. But this issue need not concern us here.

[5] Ibid., p. 251.

[6] Ibid., p. 258.

[7] Ibid., p. 251.

It seems to me that there are at least three different ways to substantiate Scanlon's move from the value of choice to his account of substantive responsibility. According to the first possible account, the value of choice demands the assignment of substantive responsibility based on choice because the values people attach to their choices will be undermined if we do not also let them bear the relevant cost of those choices. Let me call this thesis *the intrinsic value of cost*. This thesis may be illustrated by a simple example in which a husband chooses to give his wife an expensive diamond ring as a birthday present. In this case, at least part of the representative value of his choice comes from the fact that he has actually paid a large amount of money for the ring. Such value will be discounted if it turns out that the ring was abandoned on the street and was picked up by him for free. Of course, the intrinsic value of cost is unlikely to apply to many other choices. Consider the case where I order a dish from a food menu. The value of my choice will not be impacted if the restaurant owner comes to tell me that I am the lucky customer who does not need to pay for the bill today.

But, instead of assessing specific choice scenarios, there may be a general case for the intrinsic value of cost. For example, it may be argued that, in order to respect individuals as authors of their own life stories, we should let them keep the achievements as well as the corresponding burdens resulting from their autonomous choices—a person who is protected from any burden or hardship throughout her journey may feel less an author of her own life even if she achieves what she aims for. This actually sounds a bit similar to Tomasi's argument about the value of risk which I discussed (and rejected) in Chapter 2. My objection to this kind of general argument for the intrinsic value of choice is this—there may be some burdens or costs that are intrinsic to the kind of life an individual chooses to live or to some particular life choice he makes, but it is unlikely that all burdens have such intrinsic value. An extreme sport athlete may consider the injuries she suffers to be an inevitable part of the sport which bears intrinsic value for the profession, but likely not the medical cost for treating such injuries. If the treatment is covered by the state or the sport team she belongs to or by the insurance company rather than herself, it certainly does not take away any meaning from her life as an extreme sport athlete. In fact, if a person has to navigate herself through many costly choices that are not part of her main projects, she is likely to feel

that her prospects for authoring her own life story is constantly being threatened rather than enhanced.

The second way of understanding Scanlon's move from the value of choice to substantive responsibility goes like this—if we value choice in the sense that we want what happens to us to depend on our choice, then, to preserve such value, we would also need a principle according to which we are not required to pay for the cost of other people's choices. It follows, then, that such a principle should hold people responsible for the consequence of their own choices because otherwise that cost will fall upon others who never choose the burden. I call this thesis *no unchosen cost*. The intuition behind no unchosen cost can be illustrated by the following case. A man threw a stone at my window and broke it. He should be held substantively responsible for fixing the window because, otherwise, I need to bear the cost for something that is not related to my choice or action. Allowing burdens to strike me like this would take away or undermine valuable choices that are otherwise available to me.

No unchosen cost, however, is not a defendable thesis. First, even without other people's actions and choices, various unchosen burdens constantly fall upon us. Consider, for example, natural illnesses and natural disasters. It is therefore unrealistic to require that we be protected from any burden we do not choose in order to preserve the value of our choices. One possible reply to this objection is that there may be a difference between "no unchosen cost" and "no cost chosen by others" and only the latter is required to explain why it is problematic for me to bear the cost of the broken window in the above example. However, even this more limited thesis is not generally justifiable. This can be shown by the fact that we can sometimes be under duties of justice to bear the burden of other people's choices—an example is the duty of rescue which may require a person to provide easy emergency assistance for someone who is in danger due to his own voluntary choices. But what the argument for no unchosen cost does remind us is that, to assess the justifiability of an account of substantive responsibility, we need to consider its impact not only on the choice maker himself but

also on others who are otherwise not involved[8] because what the choice maker need not be held responsible for will have to be borne by others.

Finally, there is the third and probably most promising interpretation of Scanlon's account. The justification involves two separate steps—firstly, the value of choice requires the provision of an adequate range of options so that people can make choices; secondly, if an individual is provided with such alternatives, then it makes sense to argue that, by choosing a particular option, he chooses not only what he aims to acquire but also the accompanying cost.[9] In other words, he is substantively responsible for his choice in such situations because the cost associated with his choice is also part of his choice—he could have avoided it by choosing otherwise. Call the second step *the chosen cost*.

This interpretation echoes with Scanlon's emphasis on the sufficiently good choice position or the existence of alternatives as the condition for substantive responsibility. Recall Scanlon's explanation of the hazardous waste removal case. Since those who are informed of the danger are placed in a choice situation where they understand the cost attached to the option of going out and could have chosen the other option (staying at home) to avoid the burden, it is reasonable, according to Scanlon, to argue that they have *chosen* to bear the burden of injury by choosing to go out.

There may be several objections to this third way of interpreting Scanlon's account. First, it may be questioned why the value of choice even applies to a case like the hazardous waste removal—what is the value of retaining the possibility of going out if it is an option that involves disastrous consequence? But as shown in the case of the rational scientist, the existence of this option may indeed be valuable for some people because the benefit they attach to the option outweighs the cost. In addition to the rational scientist, consider another individual whom Olsaretti adds to the case. This person is also aware of the danger but chooses to go out because he needs to go to

---

[8] Stemplowska makes a similar point. See Z. Stemplowska, "Harmful Choices: Scanlon and Voorhoeve on Substantive Responsibility", *Journal of Moral Philosophy* 10 (2013), 488-507.

[9] This is similar to Andrew Williams' interpretation. See A. Williams, "Liberty, Liability and Contractualism", in *Egalitarianism: New Essays on the Nature and Value of Equality*, eds. N. Holtug and K. Lippert-Rasmussen (Oxford: Oxford University Press, 2007).

hospital for an emergency dialysis.[10] For this needy man, it is also valuable to have the option of going out compared to a scenario where everyone is locked at home.

The case of the needy man, however, leads to a different objection, which is concerned with the second step—the chosen cost. Can it really be said that he is responsible for bearing the burden of injury because he has *chosen* so when the only alternative for him is dying from illness? Olsaretti argues that, to judge that a person has chosen the cost by making a choice, it needs to be established that it represents his truly voluntary choice.[11] This can also be illustrated by the classic example of life or money—we do not hold a person who hands over his money under the threat of a gun man responsible for his loss on the ground that he has chosen to bear the cost. According to Olsaretti, a choice is voluntary "if and only if it is not made because there is no *acceptable* alternative," where whether an alternative is acceptable depends on some objective standard of well-being.[12] If such a voluntariness requirement is accepted, then the chosen cost thesis needs to be restated as *the voluntarily chosen cost*—an individual should be substantively responsible for his truly voluntary choices because, by making such choices, he also voluntarily chooses to bear the corresponding cost.

But Scanlon does not want to introduce the idea of voluntariness, or similar conditions concerning the choice itself, into his account.[13] He contrasts his account with what he calls the forfeiture view, which grounds substantive responsibility in the actual choice people make. According to the forfeiture view, only by making an intentional, conscious choice can an individual then be said to have forfeited other alternatives, hence also the chance of avoiding the cost that now falls upon him.[14] Scanlon rejects the forfeiture view because, first, it cannot explain our intuition about cases like the forgetful woman —we tend to believe that she should be held substantively responsible for her action even though she did not *consciously* forfeit her chance of avoiding injury. In addition, Scanlon's more general worry is that invoking a concept like voluntariness will

---

[10] S. Olsaretti, "Scanlon on Responsibility and the Value of Choice", *Journal of Moral Philosophy* 10 (2013), 465-83, p. 475.

[11] Ibid.

[12] Ibid., p. 476. emphasis added.

[13] Scanlon, *What We Owe to Each Other*, pp. 260-61.

[14] Ibid., pp. 258-59.

introduce complicated issues such as determinism into his account which threaten its robustness.[15]

However, suppose we accept Scanlon's arguments against the forfeiture view[16]—even in this case, it seems that Scanlon should at least accept some additional condition in order to distinguish the needy man from the rational scientist. Instead of emphasising the voluntariness of the actual choice made, this can be done by incorporating the relevant requirement into what Scanlon calls the background conditions for choice. By stating that an individual should bear the burden of his choice when he is well placed to choose otherwise, "being well placed" should mean not only that that he has adequate alternatives to choose from but also that at least some of those alternatives do not involve unacceptable consequences. This can be reflected in a revised version of the chosen cost thesis—call it *the reasonably avoidable cost*.

> *The reasonably avoidable cost*: To hold an individual substantively responsible for his choice, the cost of his choice should be reasonably avoidable. That is, among the available options, there should at least exist some alternatives with acceptable cost.

Now, by specifying the right conditions for choice, the revised account can successfully distinguish between the needy man, who does not have acceptable alternatives and should therefore not be held responsible for his injury, and both the rational scientist and the forgetful woman, who are well placed to choose otherwise and are therefore substantively responsible for their choices. However, it still cannot address a different concern, which Olsaretti calls the question of stakes.[17] This third objection can be demonstrated by the following case.

> *Amy Cut Her Finger:* Amy lives in a society where there is no free public health care but private health care insurance is available at a reasonable price. Amy can afford such insurance but chooses not to buy it. One day she accidentally cuts her

---

[15] Ibid., pp. 261-62.

[16] Olsaretti questions whether the distinction between Scanlon's value of choice account and the forfeiture view can remain meaningful if we take the case of the needy man seriously. See Olsaretti, "Scanlon on Responsibility", pp. 479-82.

[17] Ibid., p. 478. Also see similar discussion in S. Olsaretti, "Responsibility and the Consequences of Choice", *Proceedings of the Aristotelian Society* 109 (2009), 165-88.

finger when preparing dinner, so she goes to a local hospital for emergency treatment. Now consider three possible worlds Amy may live in. In the first world, the cost of health care in her society is very affordable. Amy has an operation on her finger and paid the resulting bill. In the second world, health care is very expensive and Amy could not afford the cost of the necessary operation. However, in this society, hospitals are obliged to provide emergency treatment before billing the patient. As a result, Amy has the operation but becomes deep in debt. In the third world, not only is the cost of health care high in her society but hospitals also demand full payment before they carry out any operation. In this world, since Amy cannot afford the operation, she never gets treated and ends up with permanent damage to her finger.

In all the three possible worlds, the cost Amy ends up with can be said to be reasonably avoidable. She could have bought insurance to cover herself or have exercised more caution in cutting food or have avoided cutting food with a knife. Given these acceptable alternatives, we may agree that she should indeed be held responsible for at least some of the resulting burden. However, we probably do not want to further insist that she should bear whatever consequences that fall upon her. For example, we tend to think it is fair for Amy to pay the full bill in the first world, maybe less so in the second world, but there is certainly something wrong about the third world where Amy has to bear the consequence of being left untreated and becoming disabled. This shows that *the reasonably avoidable cost* alone cannot address a further question—what cost should be attached to each option to make it a just background against which substantive responsibility could be assigned based on choice. This question cannot be properly answered by only looking at the choice condition. It necessarily involves other considerations of justice such as the level of sufficiency that should be guaranteed to every individual, the just economic or health care system in the society, and so on. Only when all these requirements are met can we justifiably hold choice makers responsible for the *full* cost that falls upon them. Call this *the justifiable cost*.

The justifiable cost: To hold an individual substantively responsible for his choice, the cost attached to the available options should be justifiable based on considerations of justice.

It is clear now that both the requirements of *the reasonably avoidable cost* and of *the justifiable cost* cannot be derived solely from the value of choice. This point has been recognised by Scanlon in his later work. According to him,

> What individuals are (substantively) responsible for—i.e., cannot complain of—are the results of choices they make under conditions of the right kind—that is to say, under just institutions, in which the obligations of the state and various individuals have been fulfilled. So claims about substantive responsibility are conclusions of an argument about justice, not inputs into such an argument. The idea related to individual substantive responsibility that can serve as an input into arguments about what just institutions must be like is not an idea of responsibility but rather the value of choice: the instrumental and non-instrumental reasons that individuals have for wanting various things that happen to them to depend on how they respond when faced with the alternatives under the right conditions.[18]

What we can agree with the later Scanlon is then this—there are various principles, which contracting parties cannot reasonably reject, that are relevant for determining the just allocation of substantive responsibility in a particular context. These principles, or inputs, should work together to guide the design of background institutions. In addition to the value of choice, the other two principles identified in the discussion so far—the reasonable avoidability of the chosen burden and the justifiability of the attached cost—may also serve as this kind of input for an account of justice and (substantive) responsibility.

For the purpose of discussion, I will, in the rest of this chapter, assume that the conditions of *the value of choice* (there should be an adequate range of known, accessible options) and *the reasonably avoidable cost* (there are at least some alternatives with acceptable cost) are satisfied. I will also assume that *the justifiable cost*

---

[18] Scanlon, "Forms and Conditions of Responsibility", p. 109.

(the options are attached with justifiable cost) is satisfied with respect to considerations of justice that do not involve risk (as what the relevant considerations of justice are when risk is involved is the question I am going to address). I will then move on to my main task—examining how this revised account could be extended to cover the case of risk-involving choices.

4.2 Substantive Responsibility and Risk

The obvious challenge of applying the choice-based account of substantive responsibility to cases involving risk is that risk creates a gap between a choice an individual makes and its outcome. Instead of being accompanied by a fixed cost, a risky option involves various *possibilities* of cost. If a person chooses such an option (with all the conditions discussed above being met) and one of the relevant possibilities eventuates into actual burden, can the individual be said to have chosen the burden? If yes, why? If no, what should he be held substantively responsible for? It seems that some additional principles are needed to supplement the account of substantive responsibility in order to deal with risk.

In this section, I will examine two candidate principles concerning people's substantive responsibility for carrying out risky activities or making risk-involving choices. The first is what I call the full consequence view, which holds people responsible for the outcome of the risk they choose to take on; the second is the expected return view, according to which people should only be held responsible for the expected return of the chosen risk. By examining the potential arguments for these two principles, I will conclude that their justifiability seems to be context-dependent—different scenarios of risk may call for different ways of allocating substantive responsibility.

4.2.1 The full consequence view

According to the first candidate principle, what I call the full consequence view, if an outcome stems from a person's choice or is caused by his chosen action, then the person should be held substantively responsible for the full cost of that outcome, even though (at least part of) it is not directly chosen by him. To illustrate, consider the following simple example.

Bob Drove to Work: Bob drives to work everyday. On his way to work this Monday, Bob suddenly lost control of his car and the vehicle hit the guard-rail. As a result, both the guard-rail and the car bumper need fixing.

If the full consequence view applies, then Bob should be held responsible for bearing such cost, though the cost is not directly chosen by him as part of his decision to drive to work on Monday.[19] The question is—should we allocate substantive responsibility based on the full consequence view?

Before I delve into this question, it should first be noted that, in examining whether the full consequence view is justifiable, I will always assume that the full consequence satisfies the condition of the justifiable cost based on non-risk-related considerations of justice. That is, it would have been the justifiable cost to attach if the outcome is directly chosen by the agent. For example, if the road maintenance department charges Bob £1,000,000 for breaking the guard-rail, we may have non-risk related reasons to believe that the cost is not justifiable—maybe we think the charge is not fair because the market price for fixing the guard-rail is only £1000; maybe we are concerned that the charge would render Bob below a level of sufficiency. But suppose Bob is only charged with £1000. Also suppose we agree that the cost of £1000 would not be unjustifiable if it is Bob's deliberate choice to direct his car to the guard-rail with the somewhat stupid intention to test how strong his car is.[20] The question now is whether we also agree that Bob should be held responsible for paying £1000 when the damage is a result of bad luck instead of his deliberate choice.

To see what might justify holding people responsible for such bad luck, let me start by considering a real-world case that reflects the full consequence view of substantive responsibility—strict liability in tort law. Strict liability often applies to risk-imposing activities with potentially dangerous outcomes such as storing explosives, running nuclear power stations and, in some countries, car driving.[21] It requires that the person

---

[19] Whether or not Bob has covered himself with insurance so that the cost will be reimbursed by the insurance company is a separate question, to which I will come back shortly.

[20] Assume there is no safety concern here so we are not trying to punish Bob for his action.

[21] T. Honoré, "Responsibility and Luck: The Moral Basis of Strict Liability", in *Responsibility and Fault* (London: Hart Publishing, 1999), p. 23.

who carries out the relevant activity be held liable for the consequences resulting from his actions, regardless of whether the outcome is a matter of bad luck.[22]

What is the justification for strict liability? According to Honoré, strict liability may be justified based on a combination of consequential and non-consequential reasons. The consequential reason is, quite obviously, that it provides the incentive for those carrying out dangerous activities to be more cautious.[23] But incentive alone is not sufficient to justify strict liability because, to provide people with the right incentive, it does not always require them to bear the *full* consequence of such activities. So Honoré adds that strict liability is justifiable because it may, over the span of an individual's life time, produce results that are fair despite the presence of bad luck. That is, even though for a particular event of risk materialisation it is unfortunate that an individual suffering bad luck needs to bear its full consequence, it is not problematic if, by having individuals each bear the outcome of his or her own action, most members of the society (that is, those possessing a general capacity for deciding and acting when facing risks) would stand to win more than lose from the strict liability system. As long as the system is beneficial for each of them and applies impartially to all those who perform the same types of actions, the system could be considered fair and justifiable.[24]

However, it does not seem sufficient to justify the full consequence view simply to appeal to the fact that it delivers net benefits to everyone. Among the various ways of allocating benefits and burdens that would result in a positive net return for each participant, why is the particular allocation based on strict liability fair? Honoré then further argues, though briefly, that holding a person responsible for the outcome of his decisions and actions, either good or bad, is also important for respecting his status as a person who has a continuous history and character as a result of his actions.[25] Why is this the case?

---

[22] Ibid.

[23] Ibid., pp. 27-28.

[24] Ibid., pp. 26-27.

[25] Ibid., pp. 29-30.

One way of understanding this argument is to appeal to the *intrinsic value of cost* thesis I discussed earlier—taking away the cost of people's choices also takes away part of their life stories and therefore undermines their status as autonomous persons or self-authors. However, as I have argued, letting people bear the full cost of their choices does not always have the intrinsic value of respecting personhood or recognising their status as self-authors. It depends on factors such as whether the person willingly take on the risk and whether the cost attached to the outcome is justifiable. Some specific cases of risk taking may indeed satisfy these criteria, but the argument certainly cannot be generalised to all activities where strict liability applies.

On the other hand, Honoré's reference to personhood may be understood in a way that is similar to how libertarians appeal to the notion of self-ownership. This is what Perry takes to underlie the justification of strict liability. According to this view, strict liability is based on the libertarian idea that individuals own themselves (the self-ownership thesis) and should therefore own the consequences caused by their actions (the causation thesis).[26] Libertarianism holds that self-ownership implies that individuals are entitled to the full fruits of their labour. If people have a morally fundamental right in owning the good outcomes that causally flow from their actions, then, in a similarly fundamental sense, they should also own the bad outcome that flow from those actions.[27]

Note that this account does not provide a general justification for the full consequence view because its argument only applies to those outcomes resulting from individuals' active exercise of their agency.[28] It does not hold people substantively responsible for a bad outcome if, for example, it simply happens to him due to his choice of inaction. But the account, if successful, may indeed justify the particular case of strict liability because strict liability tends to cover those risky activities that someone actively carries out.

---

[26] S. Perry, "Libertarianism, Entitlement, and Responsibility", *Philosophy and Public Affairs* 26 (1997), 351-96, p. 353.

[27] Ibid., pp. 369-70.

[28] Ibid., p. 372.

Is the account successful then? It seems not even if we accept the premise of libertarianism. First, as noted by Perry, the libertarian account has difficulties in dealing with cases of human interaction because the causation thesis seems to assume a world where everyone lives an isolated life.[29] To argue that a person owns the outcomes of his actions, it must be established that those outcomes can be *uniquely* traced to this person's exercise of agency. However, in cases involving interaction between individuals, such unique attribution usually cannot be established. Consider a revised case of *Bob Drove to Work*. After Bob lost control of the car, instead of just hitting the guardrail, the car ran off the guardrail and hit a passer-by Chris. In this case, both Bob's action of driving and Chris's action of walking on the street seem to have contributed to the causal chain which finally leads to the accident. The simple requirement that each individual should internalise the cost caused by their own actions cannot provide an answer as to who should be treated as the unique cause of the outcome or what proportion of the outcome is uniquely caused, and should therefore be owned, by each individual.

Of course, in the real world, we tend to judge that strict liability, if it applies, should apply to Bob rather than Chris. To make sense of such judgment, some libertarians suggest that the active person, that is, the person who chooses to act upon or cause harm to the other should be regarded as "the sole originating source of the harm" suffered by the victim.[30] However, this cannot explain why our intuition about the case does not change whether Chris was standing still or walking when he was hit. In fact, as rightly argued by Perry, the active / passive distinction hardly makes sense as we, by living in the world, are all in a sense inevitably active beings.[31]

Or it may be suggested that, in the context of risk, the person who imposes risk on the other should be regarded as the originating source of harm. Since, in the revised case, Bob can be considered as imposing a risk on pedestrians by driving, he is then the one to assume strict liability for the outcome. However, similar to Perry's argument, by acting and interacting in the world, we also inevitably impose risk on each other. If the

---

[29] Ibid., p. 375, pp. 378-79.

[30] Ibid., pp. 380-81.

[31] Ibid., p. 386.

argument is that driving is a special type of risk-imposing activity that justifies strict liability, then what if, instead of hitting a passer-by, Bob's car hit another moving car? If we are to hold Bob responsible in such a case, it will have to be based on some additional criteria rather than a revised version of the causation thesis.

Besides the difficulty in dealing with cases where people interact, it is in fact doubtful if the causation thesis actually holds in the context of risk. If ownership of an outcome is derived from the exercise of a person's agency (the person owns himself, therefore his actions, therefore the outcome of those actions), why can such ownership be extended to cases where the outcome is a result of both the person's exercise of agency and pure luck? That is, why should a person own the part of the outcome, either good or bad, that is actually beyond the control of his agency? Here one may be tempted to appeal to some account, such as Raz's, that justifies responsibility despite lack of control. Raz argues that people may be held responsible for cases of negligence, non-intentional omission, and so on, based on a notion of "domains of secure competence".[32] We are responsible within such domains because it is the sphere in which we are competent in exercising our agency not only of acting but also of foreseeing, planning, intending, and so on, and may therefore be liable for failing to control.[33] However, it should be noted that what Raz is arguing for is a different notion of responsibility—he is considering whether an outcome can be ascribed to a person (which is sometimes called agent responsibility).[34] Being responsible in this sense does not directly entail property ownership of the outcome (that is, substantive responsibility). In fact, libertarians also appeal to a similar notion of moral sphere in some of their arguments, but they also do not further suggest that having such moral sphere implies having property ownership over outcomes eventualised within the sphere.

Having looked at legal theorists' attempts to justify the full consequence view in the context of strict liability and explained why they are not satisfactory, let me now turn to a different and more general case. It may be noted that luck egalitarianism in fact

---

[32] J. Raz, "Being in the World", *Ratio 23* (2010), 433-52, p. 444.

[33] Ibid., pp. 444-47.

[34] To be responsible in the sense Raz considers is when the actions can be appropriately relate to people's capacities of rational agency. Ibid., p. 434.

implicitly endorses a version of the full consequence view. When Dworkin argues that individuals should keep the outcome resulting from option luck, it effectively means that they should be held substantively responsible for the full consequence of their choices, including those stemming from bad option luck. What is then Dworkin's justification for the full consequence view?

Dworkin's argument goes like this, "people should pay the price of the life they have decided to lead…the price of a safer life…is precisely forgoing any chance of the gains whose prospect induces others to gamble…[for those who gamble] the possibility of loss was part of the life they chose—that it was the fair price of the possibility of gain."[35] It seems that this argument is effectively a revised version of *the chosen cost* discussed in Section 1. It can be restated as *the chosen possibility*—if a person has chosen to subject himself to a series of possible (foreseeable) outcomes, then he should be held responsible for any of those outcomes that eventualises into reality because all the possibilities, either good or bad, are part of his choice.

Due to their similar structure, *the chosen possibility* is subject to similar worries that have led to the proposal of the reasonably avoidable cost and the justifiable cost.[36] Let me consider them in the context of risk. The first worry is that most of our daily actions and choices involve a small likelihood of unacceptable outcomes. While it is reasonable to claim that, by taking a gamble, I have chosen both the possibility of gain and the possibility of loss, can it really be said that I have chosen the possibilities of unacceptable outcomes by carrying out daily activities such as cutting food, walking on a street or driving? This is not to say that, for a particular action that I perform, I cannot reasonably choose to do otherwise (for example, working at home today rather than driving to work). Rather, the concern is that, if I would like to live a life in which I am not subject to unacceptable possibilities, the choices left to me are likely to be very limited and inadequate for an autonomous life.

---

[35] R. Dworkin, "What is Equality? Part 2: Equality of Resources", *Philosophy and Public Affairs* 10 (1981), 283-345, p. 294.

[36] In fact, similar versions of those objections have also been raised against luck egalitarianism, though they are focused on the choice condition, not luck. See, for example, Z. Stemplowska, "Making Justice Sensitive to Responsibility", *Political Studies* 57 (2009), 237-59; Olsaretti, "Responsibility and the Consequences of Choice".

Dworkin in fact has a ready answer to address this objection—the option of insurance. Since insurance can convert the unacceptable possibilities into fixed costs, as long as insurance is available at reasonable price, the unacceptable possibilities are then reasonably avoidable. So Dworkin goes on to argue that, if insurance is readily available but a person chooses not to buy it, then he should bear the full consequence of his choice, which includes the full consequence of risk materialisation. This second part, however, leads to the other worry—is the full consequence the justifiable cost to attach to the option of carrying out risk-involving activities without buying insurance?

As I have argued, this question cannot be answered by looking at the choice condition alone. Instead, it needs to be addressed against background considerations of justice. For example, justice may require that people be guaranteed with a minimum level of welfare. If so, having a person bear the full consequence of risk materialisation may be unjustifiable when paying for the full cost means depriving him of sufficiency. This is indeed an objection that is often raised against luck egalitarianism.[37] However, in addition to the general worry about sufficiency, there is also a more specific concern for Dworkin—since Dworkin does have an account of background justice, which is the equal initial distribution of resources and the promise of neutralising the subsequent differences in holdings that cannot be traced to individual choices, is the full consequence view justifiable based on his own commitments?

According to Knight, it seems not.[38] Knight argues that, if luck egalitarians are to take the ideal of equalising fate (and not choice) seriously, then the element of luck involved in option luck should be neutralised in the same way as how brute luck is treated. To illustrate, Knight considers a case called the dominant gamble. A dominant gamble offers a chance of gain without any possibility of loss, so that individuals are rationally compelled to take it.[39] For example, imagine being offered with the following two options—1) directly receive $100; 2) enter a gamble with a 95% chance of winning

---

[37] Crisp suggests that this can be addressed by revising luck egalitarianism to include a sufficientarian requirement. See R. Crisp, "Equality, Priority, and Compassion", *Ethics* 113 (2003), 745-63.

[38] C. Knight, "An Argument for All-Luck Egalitarianism", *Philosophy and Public Affairs* 49 (2021), 350-78.

[39] Ibid., p. 354.

$200 and a 5% chance of getting $100. Since option 2) is dominantly better than option 1), it is not irrational for any individual to choose to enter the gamble. Now consider two persons who both choose option 2). One person is lucky and wins $200 while the other one is unlucky and ends up with $100. According to Dworkin, the difference in this outcome is a result of option luck and does not call for redistribution. Knight believes this violates the spirit of luck egalitarianism, which is committed to neutralising inequality that results from unchosen luck. In this case, both individuals make the same choice (which is also the dominantly better choice) yet face different outcomes. Such difference can only be traced back to luck. If luck egalitarians hold that brute luck should be neutralised because it does not reflect different individual choices, then the same rationale should also apply to the dominant gamble case, which would call for the neutralisation of option luck between the two participants.

This argument can then be extended to more general cases. What constitutes option luck is a combination of people's risk choice and the luck involved in the eventualisation of the final outcome. Hence, it should be contrasted with brute luck on the one hand and direct choice (that is, choice unmediated by risk) on the other.[40] What people should be held substantively responsible for is the risk choice they make, not the luck they experience. This leads to Knight's proposal of an alternative principle of substantive responsibility—the expected return view. Let me now turn to Knight's account.

4.2.2 The expected return view

To disentangle choice from luck in the case of option luck, Knight suggests that risk-involving choice could be represented by its expected return. Luck, then, is the outcome of a choice minus its expectation.[41] Substantive responsibility should be allocated based on the expected return of people's choices; the rest is due to luck and should be equally shared. Knight calls this account all-luck egalitarianism as opposed to brute-luck egalitarianism, which is the (standard) version that only aims to neutralise brute luck.

---

[40] Ibid., p. 351.

[41] Ibid., pp. 369-70.

What is the implication of the expected return view? To hold an individual responsible for the expected return of his choice means that the remaining burden of the risk outcome needs to be shared by others. So the expected return view effectively requires that individuals must cover all the risk they take by paying for insurance[42] or, to put it more generally, by participating in risk sharing with others—for an individual who needs to run a risk, to pay an insurance premium is to discharge the responsibility of paying for the expected return of the risk; should the risk materialise, the burden is then shared by others through insurance payment.

Is the expected return view then reasonably acceptable to contracting parties? While it addresses the problem of the full consequence view, it seems to be subject to two other worries—one concerning its applicability to cases where people willingly take gambles and the other concerning the legitimacy of using expected return to represent what people choose when there is room for individual agency to influence the outcome. Let me discuss them in turn.

First, the expected return view may create unreasonable burdens in scenarios where people are willing to bear the possibility of full loss in return for potential gain. Typical examples include those who invest in the stock market and those who participate in gambles. For investors, the principle of expected return means they need to forgo any gain beyond the expected return of the stock in return for the protection against downside risk. For gamblers, it effectively means the prohibition of real gambling[43] if the expected return of the game is theoretically zero or negative. For strict egalitarians, they may indeed welcome such implications.[44] However, it seems to me that many people would agree that allowing some risk taking of this kind (maybe excluding those with huge gains or losses) is a valuable option to keep in people's lives.

Lippert-Rasmussen has noted a similar distinction between what he calls "gambles proper" and "quasi-gambles". A gambler proper prefers facing the gamble to having its

[42] Note that here I am not referring to private insurance in the real world but insurance as an abstract, risk-pooling device.

[43] One can still participate in the game, maybe by using game tokens that do not have any real monetary value.

[44] Knight argues that the intuitions we may have about these two cases are non-egalitarian intuitions. See ibid., p. 375.

expected value; a quasi-gambler, on the other hand, prefers the expected value of the gamble to facing the gamble.[45] According to Lippert-Rasmussen, Dworkin should not lump these two types of risk-taking together as cases of option luck. Option luck should not be neutralised among gamblers proper because doing so would make both winners and losers worse off ex ante (as they prefer being able to keep the outcome of gambles proper), but redistribution among quasi-gamblers would actually make both winners and losers better off (as they get what they prefer from such redistribution—the expected return of the game).[46]

For Lippert-Rasmussen, what differentiates a quasi-gambler from a gambler proper is their attitude toward the game. It may be questioned if it is plausible to allow factors such as subjective preference to have such drastic implication on redistribution. However, based on what I have argued so far, the distinction between gambles proper and quasi-gambles can be defined in a more objective way. What Lippert-Rasmussen calls quasi-gambles are cases where people make choices or engage in activities that inevitably involve risk. They take on the risk not for the sake of risk-taking but as the inevitable cost they pay for their choices. Such cost is best measured by its expected return. Moreover, how these risks are dealt with by background institutions will affect what costs are attached to the background options available to individuals. Whether the cost should be fixed at its expected value for everyone or be allowed to strike some individuals but not others is therefore a matter of justice. Gambles proper, on the other hand, are like a consumption good which people choose for the sake of it. Other than paternalistic reasons,[47] it is not unreasonable to leave them to the choice of individuals and let them keep the consequences.

Besides the concern about the distinction itself, Knight has a different objection to Lippert-Rasmussen's account. He questions its coherency by considering a scenario in which losers are quasi-gamblers but winners are gamblers proper.[48] If, in a two-person

---

[45] K. Lippert-Rasmussen, "Egalitarianism, Option Luck, and Responsibility", *Ethics* 111 (2001), 548-79, pp. 554-55.

[46] Ibid.

[47] See discussion in Chapter 2.

[48] Knight, "An Argument for All-Luck Egalitarianism", pp. 365-67.

case, the losing quasi-gambler need to be compensated but the winning gambler proper is entitled to keep his gain—who is to pay for the compensation then? Knight believes this case shows that Lippert-Rasmussen's account contains two conflicting norms which make conflicting claims on resources.[49] However, it seems to me that Knight misunderstands how "compensation only for quasi-gambles" works. The distinction between quasi-gambles and gambles proper refers to different types of risk-involving activities, not different groups of people. A person may have a quasi-gamble by driving to a casino (and hence running the small risk of having a car accident on the way) and then a gamble proper by making a bet in the casino. Resources for compensating this person if a car accident happens to him would come from insurance premiums paid into an automobile insurance scheme. Depending on how automobile insurance is (or should be) funded, such premiums may be collected from all those who run a similar risk, all those who benefit from the running of the risk and so on. Such arrangements are independent of whether profits made in the casino should be taxed or not.

Knight has in fact considered a suggestion similar to my argument that, in practice, actions involving the so-called quasi-gambles are so prevalent in a large population that compensation for the losers of quasi-gambles can be funded by quasi-gamblers themselves; gamblers proper, on the other hand, can be left alone and keep both their good and bad luck.[50] Knight, however, dismisses this reasoning by arguing that an account like luck egalitarianism should not be dependent on empirical facts such as the existence of a large population and the account is also not supposed to apply only to large populations.

I disagree with Knight's arguments for two reasons. First, even though the suggestion he considers is phrased in a way that emphasises the large population, compensation for losers of quasi-gambles funded through risk sharing among quasi-gamblers is not really dependent on the existence of a large population. Even two persons can share certain risks between themselves (while at the same time making a bet on something in which one of them will win and the other lose). The limitation of a two-person or small-group

[49] Ibid., p. 366.

[50] Ibid. p. 367.

risk sharing arrangement is that it may be less effective in mitigating the shocks caused by risk materialisation. When the population size is sufficiently large, such a risk sharing scheme will function more like actuarial insurance—because of the law of large numbers, the burden to be shared by each participant will be more limited and predictable. In such cases, a principle that requires holding people responsible for the expected return of their quasi-gamble risk choices is more likely to be one that cannot be reasonably rejected by contracting parties because the burden it implies for anyone is likely to be smaller than what the most burdened need to bear under an alternative principle.[51]

It may be argued that this is exactly what Knight wants to reject—the reasonable acceptability of a principle of justice should not be dependent on empirical facts such as the existence of a large population. This leads to my second disagreement with Knight's reply. Based on what I conclude in the first section, an account of substantive responsibility has to be proposed against a background of just institutions which determine duties and rights and make available life options. It is therefore inevitably an account that applies specifically to a large group of people living together under a set of institutions. This does not seem to me to be problematic. It is in fact true that, if we are looking for an account of substantive responsibility for a group of ten people, we may end up with something different.[52]

The lesson to draw from what I have argued so far is then this—whether the principle of expected return, which holds people responsible for the expected return of their risk choice, is more justifiable than the principle of full consequence, which holds people responsible for the outcome of their risk choice, is likely to be case-specific. There are, on the one hand, cases like car driving, the risk of which people take on as an unwanted but inevitable part of their life and there are, on the other hand, cases like gambling, the risk of which people take on for the sake of risk-taking. The principle of expected return may be unduly restrictive for the latter type of cases but, as I discussed, seems more reasonably acceptable for the former. Moreover, among the former type of cases, the

---

[51] Recall that this is the Greater Burden Principle introduced in Chapter 3.

[52] Consider, for example, how a family of ten people would like to hold each other responsible for their choices.

justification for the principle of expected return would be stronger for risks run by a large number of people because the burden of risk sharing implied by the principle would be smaller. Presumably, this mostly applies to common risks as they are, by definition, commonly faced by members of a society.

Let me now move on to my second objection to Knight, which in fact applies to Lippert-Rasmussen as well. It seems that, by appealing to the term "gamble", both of them assume that all risks are gamble-like in the sense that the outcome is always purely random and there is no room for individual agency to exert any influence on how it may turn out. However, most cases of risk in real life are not pure risk of this kind. Consider *Bob Drove to Work* again. In this case, Bob certainly did not choose to direct his car to the guardrail, but is there nothing he could do to change the outcome? Maybe he could have checked the car more carefully before he drove off in the morning; maybe he could have applied the brakes more promptly when he felt something was wrong. Since there may be things Bob could have done to prevent the materialisation of the risk, it is not obvious to what extent the accident resulting from a series of his actions and choices could be represented by a gamble or its expected return.

This is not to say that expected return should not apply to any risk situation involving human actions. Consider the paradigm case of gamble—the throwing of a dice. We tend to treat dice-throwing performed by a human being (suppose the dice is a standard one) to be the same random event as machine dice-throwing because we believe it is beyond the capacity of a normal human being to influence how the dice lands on the table. Similarly, in the case of Bob's car accident, there is part of the event that is truly beyond normal human capacity to exert control over. It is reasonable to treat this dice-like element as random and to hold people who choose to run the risk responsible only for its expected value (that is, the resulting burden discounted by its likelihood). But sometimes there is also part of the event the eventualisation of which can be directed by human actions in a meaningful sense. In such cases, human intervention may effectively lower (or increase) the probability of the bad outcome. If Bob can reduce the chance of risk materialisation but fails to do so, we may think he does not do enough to prevent the accident from happening and we may also want to hold him substantively

responsible for more than what would otherwise be required. Therefore, while Knight is right to point out that, in the case of option luck, Dworkin fails to distinguish between what people actually choose and what they get as a result of good or bad luck, what he does not recognise is that it is not always clear where to draw the line between these two parts.

Moreover, even if we take a car accident to be the result of pure bad luck, there is still the remaining question of how to calculate its expected return because the probability of a car accident cannot be derived in the same way as how the probability of dice-throwing is calculated. The probability of getting a six by throwing a dice is an *a priori* probability that can be calculated based on mathematical principles. The same kind of reasoning cannot be applied to the chance of car accidents. The probability assigned to an event like a car accident would have to be estimated based on statistics, that is, the frequency of its occurrence in a certain set of empirical data.[53] The data set used—its time frame, geographical area, age group, gender group, car quality, speed range, and so on—will determine what is the probability of the outcome and hence the expected return of the choice.

For example, suppose a society sees 1000 car accidents per 1 million population over the last year. This data may be used to calculate the probability of having a car accident faced by Bob for every 1km he drives. However, if we break down the data and see that 800 out of the 1000 accidents occur at a speed of 120 km/h or above, we will then get two different probabilities—a smaller one for driving below 120km/h and a larger one for speed above that. We may further split the data within each speed range by age group and conclude that the rate of accidents is higher for the age group of 18-25 and for those above 60. By continuing with this reasoning, we would end up with a unique estimated probability for every individual. If everyone is running a slightly different risk of car driving, how similar do the probabilities need to be then for certain people to be considered to be running the risk of the same kind? The answer to this question needs to be justifiable to all the relevant parties because it is also a decision about who are

---

[53] This is called the relative-frequency theory of probability. See S. Perry, "Risk, Harm, Interests, and Rights", in *Risk: Philosophical Perspectives*, ed. T. Lewens (London: Routledge, 2007).

obliged to share the burden of such risks with each other but not with others. This is the reference class problem to which I will come back later.

Having discussed the potential justifications and objections for the two different views regarding risk and substantive responsibility, I reach the following tentative conclusions. First, both the full consequence view and the expected return view are subject to concerns that are particular to some kind of risk, so there is likely no single principle that can guide the allocation of substantive responsibility for all risk-taking events. Second, both how we expect a person to manage his risk exposure and to what extent we expect him to reduce the chance of risk materialisation may impact what he should be held substantively responsible for. In the next section, I will propose principles for allocating substantive responsibility that respond to these considerations.

4.3 The Reasonable Expectation Account

I call the account that I am going to propose the reasonable expectation account. It is meant to serve as additional "input" for the just allocation of substantive responsibility in the particular context of risk. For people who carry out risky activities or make risk-involving decisions, we may have two different kinds of expectations. On the one hand, we may expect them to manage their own risk exposure for some risks, including whether to take on the risk and whether to cover it with insurance. On the other hand, we may expect them to manage the outcome of some risks by exercising caution. I argue that substantive responsibility should be allocated in a way that reflects these expectations.

4.3.1 Reasonable expectations of managing risk exposure

As shown in the previous discussion, implicit in an account of substantive responsibility is the expectation of how risk exposure should be managed. The expected return view assumes that all risk exposure should be covered ex ante through risk sharing. If the risk eventuates into actual harm, the burden is to be covered by the funds collected from people who run similar risks. Call this *the collective risk management principle*. The full consequence view, on the other hand, assumes that it is up to those engaging in risky activities to manage their own risk exposures. Individuals are expected to fully

assess the situation, anticipate potential consequences, and decide whether or not to cover the risk exposure by buying insurance. Call this *the personal risk management principle*.

From the previous discussion, it seems that the more reasonably avoidable a risk is, the more likely the personal risk management principle is justifiable. By reasonably avoidable I mean the risk need not be taken on by most people living their everyday life.[54] For example, in a society where car driving is the main means of transportation, the risk of driving is less of a reasonably avoidable risk compared to a society where car driving is only for entertainment. If a risk is reasonably avoidable but a person chooses to take it on, it is likely that this risk is particular to his preferences or personal projects (such as dangerous sports or financial speculation). If this is true, then it is not unreasonable to expect this person to take efforts in understanding the risk and making an informed decision about it, including not only what level of risk to take but also whether to cover the exposure with insurance or using other risk management tools.[55] In such cases, the person's insurance decision is an integral part of his exercise of autonomy in executing his life plans.

On the other hand, there are also many risks people inevitably run as part of their everyday life. It seems that expecting people to carefully assess their everyday activities and to make insurance decisions for all the risks they are potentially exposed to would imply too high a cognitive burden for them. For an individual to do so, he has to identify all the potential risks involved in his daily actions, correctly understand the probability and magnitude of potential bad consequences (while avoiding the cognitive biases discussed in Chapter 2), assess his own financial situation and rationally decide about the level of coverage he needs. He then also needs to research about the insurance policies available in the market, understand their terms and conditions and try to judge if they are reasonably priced, and finally select the insurance policies based on all this information. Since expecting people to anticipate and manage all these common risks would consume a significant amount of their mental energy and divert them from their

[54] It does not mean people can avoid the risk exposure by taking insurance.

[55] An example of other risk management tools is hedging (for example, buying a put option to hedge a person's holding of stocks).

major life projects, a collective risk management principle that relieves people from the burden of managing common risk exposures is likely to be preferable.

The collective risk management principle may take a form that is different from the version embedded in Knight's expected return view. If everyone is to pay for the expected return of his own choice, as argued by Knight, then what the principle demands is effectively mandatory actuarial insurance. However, for reasons discussed in the previous chapter, we may reasonably endorse a version of the collective risk management principle that deviates from actuarial insurance. For example, we may allocate the burden of risk sharing based on the benefit people get from the risk-taking activities. If people who do not actively take on a risk benefit from others' risk-taking, then they will also be required to make a fair contribution to risk sharing. In this case, mandatory private insurance alone is unlikely to satisfy the demand of the collective risk management principle. The principle may require social insurance coverage or at least that private insurance be regulated with respect to its pricing and be supplemented by public funding.

It may be objected that, even for the case of common risk, the value of choice may be a valid reason for contracting parties to reject the collective risk management principle. The argument goes like this—even for a common risk, individuals may reasonably prefer different levels of insurance coverage (including no coverage). As such decisions are best made when taking personal circumstances into account, individuals may prefer to make their own choices by assessing their level of income and saving, level of risk tolerance, and other financial needs. Removing such choices by imposing a mandatory risk sharing scheme on everyone is therefore taking away valuable options from them. Moreover, if the concern about leaving risk management decisions to individuals is merely about the cognitive burden it implies, then this can be addressed by, for example, having the government issue insurance guidance which individuals could simply follow if they want to.

Two replies may be given to address this objection. First, it should be noted that endorsing the collective risk management principle does not necessarily remove the options of having no or different levels of insurance coverage. For example, the risk

sharing scheme may be designed in a way that allows individuals to opt out by paying a fee (which would be lower than the insurance premium and is for non-participants to discharge any remaining duty they may have in contributing to risk sharing[56]). This is in fact comparable to the case where the personal risk management principle applies—while individuals are by default uninsured, they still have the option of buying insurance from the private market.

Second, the real difference between these two principles is what option they set as the default. To endorse the collective risk management principle is to agree to relieve people of the burden of actively taking up insurance coverage. Applying it to common risks expresses the view that we do not expect people to assume personal responsibility in managing those risk exposures. We may reasonably adopt this principle because we acknowledge that we inevitably impose various risks on each other by living together and should naturally share some of our fate as a result. We also acknowledge that we have an interest in living as self-authors and in being able to carry out life projects that matter to us, and that our limited mental capacities are better devoted to these projects that are meaningful to ourselves rather than dealing with an infinite number of traps lying on our life paths. If someone decides that he, for whatever reason, prefers to take the riskier path by opting out of risk sharing, this is then similar to other risks he may deliberately take on as part of his personal projects. We would reasonably expect this person to manage such risk exposures and bear the relevant consequences (provided, of course, that the attached cost is justifiable).

4.3.2 Reasonable expectations of managing risk outcome

In addition to expectations relating to how people should manage their risk exposures, we may also expect those who carry out risky activities to put in reasonable efforts to reduce the likelihood of risk materialisation. We have such expectations for at least two types of cases. First, when the bad outcome may potentially fall upon people other than the risk taker himself, then, as discussed in Chapter 3, taking due care to reduce its

---

[56] This duty may result from the fact that they benefit from other people's risk-taking activities or that they are still entitled to a minimal level of assistance should the risk they take materialise. A real-world example of such fees is the tax penalties for not taking health insurance under Obamacare.

likelihood is what makes imposing the risk on others justifiable in the first place. Second, even when the consequence of risk materialisation does not fall directly on others, if the burden is to be shared by others, then those under such duties may reasonably demand that they only share the burden on the condition that due care has been taken or that the amount of burden they share depend on the level of precaution taken. Both reasons reflect the view that the risk in question is not a pure risk—there is room for human beings to exercise their agency and to meaningfully influence its outcome.

While there are no objective standards in determining, for example, which part of a car accident could have been prevented by taking precautions and which part represents pure randomness, we may decide on rules of due care that we reasonably expect people to follow. We expect drivers to have a sufficient level of driving skills by requiring them to get a driver's licence; we expect them to drive carefully by imposing a speed limit, prohibiting drunk driving and forbidding talking on the phone while driving. If a driver has followed all such rules, we may then exempt him from further expectations in preventing the risk from materialising. That is, we reasonably accept the remaining risk of car accidents as a pure risk, even though there may in fact still be things the driver could do to further reduce the chance of the bad outcome. We do not expect him to be able to do so because we recognise that human capacity is limited.

Reasonable expectations of this kind also apply to cases where the situation is complicated by human interaction. For example, we expect drivers to keep a safe distance from other vehicles, to give priority to pedestrians and to pay special attention when passing through school areas. We also expect pedestrians to keep off the motorway, to follow traffic lights and to watch out for oncoming traffic. When everyone has fulfilled such expectations yet a collision still occurs, we then consider it to be a case of pure risk materialisation.

Note that, in cases where risk materialisation impacts people other than the risk taker himself (that is, when they are cases of risk imposition), the due care rules I have been describing in fact represent part of the demands of the reasonable precaution principle I argued for in Chapter 3 for just risk imposition. But they represent only part of its

demands because the principle of reasonable precaution applies not only to individual risk imposers but also to institutions implementing the risk imposition scheme. For example, in order to legitimately subject citizens to the risk of car accidents, the state should fulfil the principle of reasonable precaution by working out the appropriate traffic rules and effectively enforcing such rules. Failing to do so would be a case of unjustifiable risk imposition by the state. On the other hand, these traffic rules implemented by the state also reflect the due care we expect individual drivers to take. When a person fails to follow the rules, it is a case of unjustifiable risk imposition by the individual and he may be subject to fines or other forms of legal punishment whether or not the risk eventuates into actual harm.

At the same time, when a risk does materialise into bad consequences, rules of due care may also serve to limit the scope of collective risk sharing (where it is required). As I have argued, the collective risk sharing principle should apply to common risks. While a common risk refers to the same kind of risk that is taken on or faced by many members of a society, there are in fact no two risks run by people that are completely identical.[57] Therefore, certain standards are needed to determine which risks could be regarded as belonging to "the same kind". It seems that the rules of due care could serve such a purpose. For example, suppose a speed limit of 120 km/h is imposed for driving on highway. The rule reflects the view that we do not generally expect drivers to further lower their speed for the sake of reducing the risk of car accidents. In this sense, the activity of car driving below 120 km/h (provided that other demands of due care are also fulfilled) can be regarded as involving the same kind of risk that we agree to pool together for the purpose of risk sharing. Hence, what we call the social risk of car driving is more specifically the risk of car driving within that speed limit. Meanwhile, a person who drives above 120 km/h would be regarded as taking on an additional risk which others are not obliged to share—he himself should be held substantively responsible for its consequences.

Of course, there are also cases where, even though we agree to share the risk, we want to recognise different levels of due care exercised in managing the risk outcome through

[57] Maybe except the gamble-type of risks.

the allocation of substantive responsibility. In private insurance, this is usually done through partial payment or future premium increase based on previous claims. For example, if a driver is found at fault for an accident (that is, he violates certain due care rules), then the insurer may only pay for part of the resulting cost; or it may pay the full amount but increase his premiums for subsequent years. These conditions can be similarly embedded into a non-actuarial risk sharing scheme. They serve both as incentives for risk takers to take precaution and as a way to allocate a fair share of burden to those who fail to meet reasonable expectations in taking due care.

Due to the demand of risk sharing, we may also extend our expectation of managing risk outcomes to other cases where those who run the risk only subject themselves to potential bad consequences. Consider, for example, the unemployment risk, which often does not involve issues of interpersonal risk imposition (except cases where the unemployed have dependents). Since it is a risk imposed on members of the society by its institutions, the demands of reasonable precaution should have mostly been directed toward the state rather than individuals. However, if others are to share the burden of unemployment, they may then expect the risk takers to reduce its likelihood by making reasonable efforts to find a job and fulfil job duties. If a person meets these expectations yet becomes unemployed (because of, for example, structural changes in the market), we consider it to be a consequence materialising from the social risk of unemployment the burden of which should be fairly shared. On the other hand, if an individual gets fired because he has never performed his job duties, then we may hold him (at least partly) substantively responsible for the resulting burden by, for example, withdrawing some of his unemployment benefits. Here, there is no issue of illegitimate risk imposition on others, but his failure to fulfil reasonable expectations may affect our judgement about his substantive responsibility for the materialisation of the unemployment risk.

Limiting social risk sharing to cases where due care rules are followed also addresses the reference class problem—it provides a reason why we should use the particular reference class (for example, all car accidents that has occurred without violating any of the relevant precautionary rules) to evaluate the expected return of a common risk, not a

class that is either narrower or wider. The particular reference class reflects what we reasonably expect others to do when carrying out a risky activity and what we do not reasonably expect them to be able to achieve. When people have met all the reasonable expectations, the remaining uncertainty can be regarded as a pure risk and it is in this sense that the risk can be measured by the expected return of the outcome calculated from a set of probability-based payoffs. It is also in this sense that the scope of risk sharing can be justified to all—since no reasonable expectation has been violated, it should be justifiable to all those who fall into this reference class to treat the risks they are running to be similar enough and to share the burden based on, for example, the benefit principle.

Finally, it may be noted that the kind of risks I have discussed so far, be it common risk or personal risk, are risks whose potential outcomes are well-known to people. That is, because there are enough observations about them, we know about the potential bad outcomes, their (frequency-based) probabilities as well as measures that may be effective in reducing their likelihoods. There are other risks, however, about which we know much less. This is either because they are very rare or because they take a long time, usually across decades or even centuries, to materialise. What we reasonably expect people to do with such risks may be quite different from the expectations I have discussed in this chapter. Such difference becomes especially relevant when it comes to the issue of risk imposition on future generations. I will turn to this kind of risk in the next chapter.

To conclude, in this chapter I have argued that parties in the Scanlonian contracting position may reasonably endorse a set of principles that guide the allocation of substantive responsibility. Some of these principles—the adequacy of available options (or what Scanlon calls the value of choice), the reasonable avoidability of cost, the justifiability of cost—apply to the choice condition in general. In the context of risk, additional principles are needed to supplement a general account of substantive responsibility. I propose two such principles—the reasonable expectation for managing personal risks and the reasonable expectation for taking due care. I argue that we may

relieve people of the burden of managing common risk exposure while expecting them to make careful decisions about those risks they take up as part of their personal projects. We may also relieve people of the burden of taking whatever measures that might serve to prevent risk materialisation while expecting them to follow commonly acknowledged due care rules. Against this background, we may then hold people substantively responsible for their risk-involving choices. Moreover, for cases of social risk imposition discussed in Chapter 3, this account of risk and substantive responsibility should supplement the benefit principle to fairly allocate the burden of social risk sharing.

Chapter 5 — Risk Imposition on Future Generations

Contemporaries impose various risks on future generations through their practices and policy decisions. If one generation adopts unsustainable monetary policies, the next generation may suffer hyper inflation or economic recession as a result. If the contemporaries pollute air and water, destroy forests and farmlands and deplete natural resources, they may significantly increase the risk of natural disasters as well as other health and economic hazards faced by the next few generations. Even the procreational choice of a generation may impose various risks on the people they create. For example, if a generation's collective preference or the social policies implemented in their society lead to a low fertility rate, the younger generations may face a higher risk of economic recession. On the other hand, if the population grows too fast, it may impose greater environmental risks on future generations.

Moreover, not all economic and environmental risks imposed by the current generation will materialise for the generations that come immediately after. Some policies which lead to persistent, structural economic problems (for example, a highly leveraged economy) or which fail to slow down global warming and delay the depletion of natural resources may only eventualise into actual harm over the long term and hit generations that do not overlap with those currently living. In this chapter, I will examine if the account of just risk imposition and fair risk sharing I have developed in previous chapters can be extended to cover these cases of risk imposition on future generations.

It has often been proposed that, for long-term risks like climate change, the so-called Precautionary Principle should apply. The specific demands of the Precautionary Principle are notoriously difficult to pin down. For example, in a weak form, the principle states that lack of full knowledge about a risk should not be a reason for failing to take precautions.[1] This interpretation, as pointed out by Sunstein, makes the principle quite trivial as it is obviously unreasonable to demand complete certainty

[1] C. Sunstein, *Laws of Fear: Beyond the Precautionary Principle* (Cambridge: Cambridge University Press, 2005), pp. 18-19; D. Steel, *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy* (Cambridge: Cambridge University Press, 2015), p. 3.

before justifying any action.[2] On the other hand, in its strong form, the principle is sometimes understood as the maximin strategy,[3] which represents an extremely risk-averse attitude. Instead of attempting to survey the different interpretations of the Precautionary Principle, I will take a different approach to engage with this literature in this chapter. Since I have argued for a set of principles for just risk imposition in the contemporary context, I will take those principles as the baseline. I then consider various arguments for more cautious principles, including some versions of the Precautionary Principle, on the ground that my baseline principles are insufficient, and examine whether they are successful. As I will argue, they are not justifiable and the motivation for adopting these more cautious alternatives can in fact be captured by a proper interpretation of the existing principles I have argued for.

The chapter is organised as follows. In Section 5.1, I extend the Scanlonian contractualist framework to include multiple generations by briefly addressing a few concerns about including past and future people in the contracting position. In so doing, I will also draw a lesson from the discussion which will be relevant for the rest of the chapter—for generic principles to apply across time, their specific demands for each generation need to be understood in a dynamic way.

I then consider two types of arguments for a more cautious principle in the second and the third section. Section 5.2 examines if the fact that we stand in a special relationship with future generations justifies that we act more cautiously—first, we need to be particularly careful because it involves moral hazards for us to impose risks on future generations where they have to bear the burden of risk materialisation while the benefit is enjoyed by us; second, we do not know the risk preferences of future generations and we should be more risk avoidant when imposing risks on people whose preferences are unknown to us. Section 5.3 then looks at the typical features of those risks that are imposed on future generations—their potential catastrophic outcome, their unknown probability and the fact that they may only materialise in the long-term—and examines if they justify a more cautious principle to regulate imposing risks of this kind. I will

---

[2] Sunstein, *Laws of Fear*, p. 24.

[3] Steel, *Philosophy and the Precautionary Principle*, p. 5. Also see S. Gardiner, "A Core Precautionary Principle", *Journal of Political Philosophy* 14 (2006), 33-60.

conclude that all these concerns can be addressed by the existing principles in my account of just risk imposition, provided that we adopt a dynamic reading of the principles.

Finally, in Section 5.4, I consider two objections concerning fair burden allocation across time. First, I argue that, while early generations' judgement in how to deal with long-term risks may turn out to be wrong, they do not wrong future generations as long as they have done their fair share and have done so based on the best available knowledge and in a way that keeps future options reasonably open. Second, I explain why requiring the current generation to devote resources to addressing long-term risks does not mean that they have to do so at the cost of leaving contemporaries to die.

5.1 The Multi-Generation Contractualist Framework

In my previous discussion of social risk imposition, I have assumed that the contracting position includes only contemporaries. This is a common assumption employed by philosophers as a starting point to develop an account of justice. Now, to deal with risk imposition that will affect those who live in the future, I need to modify this setup so that the interests of future generations are taken into account. To do so, I will continue with the Scanlonian framework and introduce people from different generations into the contracting position.

Scanlon himself does not have an issue with such a multi-generation contractualist framework. According to him, creatures that fall within the scope of his project include "those beings to whom we have good reason to want our actions to be justifiable".[4] Such beings should be conscious, capable of feeling pain and capable of holding judgment-sensitive attitudes, especially the kind of judgements involved in moral reasoning. They may include, as explicitly stated by Scanlon, "those who are already dead, or are not yet born, as well as…our contemporaries".[5] However, there are a few potential concerns about including different generations in the contracting position, especially when the framework is applied to justifying principles of justice. Let me

[4] T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1998), p. 179.

[5] Ibid., p. 179, p. 186.

briefly examine these worries in this section before moving on to the specific case of risk imposition.

5.1.1 The non-identity problem

The first worry about a multi-generation contract is whether it is subject to the non-identity problem. The non-identity problem arises due to the fact that individual actions and collective choices of those now living would affect the physical identities of future people.[6] That is, different people may come into existence if we make one decision instead of another or choose one social policy over another. If this is true, then the selection of principles of justice will also affect the identities of future people, because such principles will guide the design of institutions and therefore influence the choices made by individuals living under such institutions. In this case, it seems logically impossible to consider a multi-generation contracting position as what principles are chosen by the contracting parties would affect who the contracting parties are in the first place.

This worry, however, is based on an assumption that the physical or, more specifically, genetic identities of the contracting parties must be known or fixed so that we could determine what principles they could justify to each other. This, however, is not necessarily true. As pointed out by Kumar, even though the parties in the Scanlonian contracting position are, unlike those in Rawls's account, aware of their own identities, they cannot always refer to such personal information when giving reasons to each other in assessing the justifiability of principles, because the reasons they may appeal to need to fulfil the requirement of generality.[7]

---

[6] D. Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), pp. 351-55. Parfit points out that a person would never have existed if he had not been conceived within a month of the time when he was in fact conceived. This means that our action can easily affect the identities of future people. If we choose between two social policies, such a choice would affect people's daily lives and ultimately affect who meets who, who conceives with whom and when. After a century or two, everyone will owe their existence to our choice.

[7] R. Kumar, "Risking Future Generations", *Ethical Theory and Moral Practice* 21 (2018), 245-57, pp. 248-49.

According to Scanlon, a generic reason is "one that we can see people have in virtue of certain general characteristics; it is not attributed to specific individuals."[8] We can see, for example, reasons people may have in virtue of being more risk averse than others. When assessing a reason against risk imposition given by such a risk-averse person, his other characteristics, such as physical appearance or natural talent, would be irrelevant. As noted by Kumar, such reasons are reasons given from certain general standpoints.[9] In addition to standpoints of certain types of people, people may also have general standpoints in virtue of being in certain types of relationships with others. For example, a future person may have reasons for accepting a principle that requires parents to reasonably reduce the chance of giving birth to children with congenital defects. Even though such a person is yet to exist, he can give generic reasons in virtue of being a child of his parents.

Therefore, as long as we can access and weigh reasons given from the relevant standpoints, it does not matter that we do not know the specific identities of those who occupy those standpoints—we can still determine the justifiability of a principle based on such generic reasons. In this sense, the non-identity problem is not a concern for a multi-generation Scanlonian contract.

5.1.2 Issues concerning past generations

According to Scanlon, his contractualist framework may include not only people from the future but also people from the past. Some may wonder in what cases principles need to be justifiable to people who have already passed away. One example is that we may have a moral duty to treat a deceased person's body in a respectful way. This duty is at least partly owed to the deceased person himself. Or, consider another example—suppose there is a project that requires several persons to contribute in sequence. Person A, B, and C have agreed to cooperate for this project to achieve a common goal. If, after A has done her share, she unfortunately passes away, B and C may still wrong her if they fail to make their subsequent contributions and thereby waste A's efforts. If we

---

[8] Scanlon, *What We Owe to Each Other*, pp. 204-5.

[9] Kumar, "Risking Future Generations", pp. 248-49.

agree there are duties of this kind, then it is not problematic that the dead are included in the Scanlonian contract.

In fact, concerns about including past generations in the contracting position are less about whether contemporary people may wrong their ancestors than the other way around—since what people from the past have either done or not done cannot be changed, what is the point of coming up with principles with which the past generations may have, as a matter of fact, failed to comply? This worry is especially relevant when it comes to scenarios where past noncompliance has implications for what the current and the future generations are entitled to or are obliged to do. For example, if past generations have not passed down enough resources, it will impact not only what the current generation are entitled to consume but also how much they should reserve so as to discharge any duty they owe to future generations. It may be noted that this is one of the worries that motivates Rawls to rule out a multi-generation contract in *A Theory of Justice*. When discussing the issue of just saving, Rawls argues that "earlier generations will have saved or not" and that the contemporaries should save according to a principle of just saving that does not assume past compliance.[10]

While this worry is indeed relevant, it need not imply the exclusion of past generations from the contract because what parties in the contracting position justify to each other are generic principles. To have generic principles that apply across generations does not mean their force is conditional on the actual compliance of the past generations. This is because, instead of specifying what each generation should be doing assuming all past generations have done so, a generic principle may have conditional demands which specify what any generation should do based on the circumstances they are situated in—circumstances include, among other things, what the past generations have done to them or passed down to them. For example, a principle of just saving may specify a just rate of saving for each generation based on how much resources, technologies and knowledge they inherit from the previous generations. A generation that inherits more should also pass down more to their descendants. If one generation fails to meet this demand, it does not change the fact that the next generation still owes a duty of saving

[10] J. Rawls, *A Theory of Justice* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999), p. 255.

to future generations as required by the principle, but it does mean that how many resources the principle demands this generation to pass down may be less than if they had received more from the current generation.

Past noncompliance may not only reduce demands on later generations but also impose additional ones on them. Consider another example where a previous generation have unjustly planted a time bomb that may explode in a few centuries. This may cause all subsequent generations to acquire a duty to mitigate the risk based on a principle of just risk imposition. Here, the demand only kicks in because of past generations' unjust behaviour.

To conclude, it is not problematic to include past generations in the contracting position and justify principles with which they have failed to comply, though what such a principle implies for the current generation will be responsive to how they are impacted by past generations' behaviours. That a generic principle may be dynamic in terms of its specific demands is in fact an important observation for the discussion of this chapter, because it may address not only backward-looking but also forward-looking concerns. I will come back to this point again.

5.1.3 Issues of intergenerational distributive justice

Lastly, some concerns about having a multi-generation contract are not directed toward the Scanlonian framework in general but particularly toward my application of the framework to an account of distributive justice. These concerns can be further divided into two types. The first is about the possibility of justice across generations. The second is more specifically about egalitarian accounts of justice—it asks whether it makes sense to demand equal distribution among individuals living the past, the present and the future.

First, due to lack of cooperation or reciprocity between non-overlapping generations, it is often questioned whether issues of justice can even arise among people who do not live together temporally. Lack of reciprocity is not a problem for the Scanlonian framework itself—Scanlon has explicitly clarified that his project is concerned with the source of interpersonal morality and contracting parties need not be "those with whom it

is advantageous for us to enter into a system of mutual restraint and cooperation."[11] But since I apply his framework to an account of justice, this concern may become relevant. I will turn to this problem in the next chapter and I will argue that lack of reciprocity is not a problem for an account of justice intended to cover people who, for example, belong to far future generations or have severe disabilities.

The second type of worry, on the other hand, is more relevant for the discussion of this chapter. Recall that, in my discussion about the principle of fair risk sharing, I assume a background principle of equal prospects for autonomy for all members of a society—all meaning all contemporaries living together at a given time. Since the task of this chapter is to extend my account developed in the previous chapters across generations, let me briefly examine if it is plausible to continue assuming a background principle of equal distribution which applies to individuals from different generations.

The immediate response to an intergenerational egalitarian principle may be that it is simply impossible to achieve. First, since human beings may exist for several million years, there is going to be an extremely large number of people who live in the future. If existing resources are to be distributed equally between us and all of them, are we only entitled to an infinitely small share? This, however, need not be the case. Some resources are renewable and can therefore be passed down; some resources may no longer be needed or may be replaceable with others as technology develops. What an egalitarian principle requires is not that we leave most of the resources untouched for future use but that we use them in a sustainable way so future generations could potentially enjoy an equivalent amount of resources or the same welfare or prospects for autonomy, depending on which desideratum of distribution the egalitarian principle uses.

However, it may be objected that, even so, it is still quite impossible to achieve an equal distribution across time, whichever currency of equality we use. Compared to people living thousands of years ago, our lives are simply much better—our life expectancy is longer; we enjoy more welfare and more advanced technology; we have many more life opportunities. Although we do not know how history will unfold, we could well

[11] Scanlon, *What We Owe to Each Other*, p. 179.

imagine the possibility that people living in a few thousand years from now will also have a much better life than ours. On the other hand, should any global catastrophe happen, it is also not impossible that future people will live a much more miserable life than ours. Since we have very little control over how history develops into the far future, what, then, is the implication of an intergenerational egalitarian principle for us as the current generation?

Here, it is again relevant to note that a generic principle may issue demands that apply specifically to some people based on their circumstances. For a given generation, the demand of equal distribution may be that they pass down enough resources that can, based on their best knowledge, sustain the same prospects for autonomy for all those living in the future. This is quite similar to a principle proposed by Barry when he discusses intergenerational justice.[12] According to him, the current generation is required to ensure that some level of X that they enjoy (Barry suggests opportunity, which is similar in spirit to my measure of prospects for autonomy) can be expected to continue indefinitely into the future, provided that future generations are not themselves responsible for lowering that level.[13] Barry calls this the principle of sustainability, but in fact what underlies its demand is the egalitarian requirement that, for anyone who is currently living, he should count the interests of his own, his contemporaries and those living in the future equally by *aiming* to (at least) achieve equality with them.[14]

---

[12] B. Barry, "Sustainability and Intergenerational Justice", in *Fairness and Futurity: Essays on Environmental Sustainability and Social Justice*, ed. A. Dobson (Oxford: Oxford University Press, 1999).

[13] Ibid., pp. 109-11. Note that Barry assumes a constant population size. That is, his principle only requires each generation to maintain the expected indefinite continuation of the level of X based on the current population size. I will make no such assumption because the relevant policies of the current generation will impact the population size of the next generation. The current generation should therefore take this into consideration to decide how much they need to reserve for their descendants.

[14] Although it may be noted that Barry's principle is still different from a strict egalitarian principle that would require the "levelling down" of future people. For example, if the current generation foresees that the next generation may be able to achieve some technological breakthrough and could therefore enjoy a higher level of welfare, they would not try to forestall this in order to achieve equality with them. What matters, based on the principle of sustainability, is to ensure that future people will not be worse off compared to the current generation.

Compliance with this principle does not mean that there is going be an actual equal distribution across time because the level X may need to be redefined as time goes by.[15] For example, suppose individuals in Generation N enjoy a level of X and pass down resources that are enough for people from future generations to all enjoy X. Generation N+1, however, achieve a technological breakthrough and can now be expected to sustain a level of X', which is larger than X, indefinitely into the future. The principle of equality would then require Generation N+1 to pass down enough so that all subsequent generations can enjoy at least X' rather than X. The fact that Generation N+1 now enjoy more than Generation N is irrelevant—from the perspective of distribution, there is nothing Generation N+1 owe to those who have already passed away. What matters is that they count the interests of future people and of their own equally. To meet this demand, Generation N+1 should increase the level of X for all future generations to X' rather than keeping the gain to themselves. Based on this dynamic understanding, a principle of equal distribution will not achieve actual equality across generations, but it is egalitarian in spirit and can be reasonably accepted by all parties in a multi-generation Scanlonian contracting position.

I have now briefly addressed the concerns of having a multi-generation contract and have, in so doing, also made a point that will be important for the discussion of this chapter—that a generic principle which applies across generations may imply different demands on different generations based on their specific circumstances. I can now start to focus on the main topic of the chapter—what principles of just risk imposition and fair risk sharing should apply for risks that impact future generations.

Before I move on, there is one last point to note. In my discussion, I will often refer to a generation as a collective entity and talk about their obligations and entitlements. It may be questioned whether appealing to group agency or collective responsibility is compatible with the Scanlonian framework as each contracting party in the contracting position represents an individual rather than a whole generation. I believe there is no inconsistency because, for each individual in the contracting position, while they evaluate the justifiability of a principle from their own standpoint, they at the same time

[15] Ibid., pp. 109-11.

know that failure to comply with the demand of the justified principle in the intergenerational context is likely to be a matter of collective responsibility. Hence, it is not unreasonable for them to accept an account of group responsibility for decisions or policies that have to be made collectively by the generation they belong to.

5.2 Imposing Risks on Future Generations

Now that we have a contracting position with parties from different generations, we can deal with cases of intergenerational risk imposition by asking what principles the contracting parties would reasonably accept to regulate such activities. In particular, the question I want to explore is whether they would demand some more cautious principles than the ones I have argued for in Chapter 3 and 4 (in particular, the presumptive benefit principle and the reasonable precaution principle). I will start, in this section, by considering two reasons for contracting parties to endorse a more conservative approach to risk imposition in the intergenerational context, which are based solely on how the current generation stand in relation to people from the future, and I will argue that neither is successful. In the next section, I will then examine if special caution may be justified because of certain features of the risks that are typically imposed on future generations.

5.2.1 Potential moral hazards

According to the first argument, imposing risks intergenerationally is different from the standard case of social risk imposition considered in previous chapters because it involves potential moral hazards. To see why, first note that a case of risk imposition may involve three separate roles—the party that may benefit from the imposition of the risk, the party that may suffer the consequences of risk materialisation and the party that makes the decision about imposing the risk.[16] In the standard case of social risk imposition, risks are either imposed by citizens on each other (such as car driving) or imposed by a decision maker on people for their own benefit (such as mandatory vaccination). In both types of cases, those who benefit from the imposed risk and those

[16] J. Wolff, "Five Types of Risky Situation", *Law, Innovation and Technology* 2 (2010), 151-63. Also see H. Hermansson and S. O. Hansson, "A Three-Party Model Tool for Ethical Risk Analysis", *Risk Management* 9 (2007), 129-44.

who may become potential victims of risk materialisation largely overlap. This is what makes the activities potentially justifiable to all.

Risk imposition on future generations, on the other hand, is problematic because it is a typical case of moral hazard where "one party makes the decision and would gain the benefit while another would suffer any costs."[17] Take the ecological risks we are imposing on future generations as an example. Through excessive deforestation and overhunting, we significantly increase the risk of ecological disasters such as species extinction, resource depletion and climate change.[18] Such risks are, moreover, unlikely to materialise fully in the short term. While we enjoy the benefit of utilising what the nature has to offer, the burden of our actions will largely be borne by future generations who do not have decision making power over us. It may be argued that, because of this kind of moral hazard, a more cautious principle is needed to regulate intergenerational risk imposition.

However, if we take the problem of intergenerational risk imposition to be that future generations have to bear the burden of the risks imposed by us without enjoying any corresponding benefit, then we do not need a more cautious principle to conclude that it is not justifiable. The standard principle of presumptive benefit I have argued for would reject the case immediately—since the expected burden of environmental risks significantly outweighs its benefit (if any) for our descendants, imposing the risk on them does not satisfy the principle which requires risk imposition to be presumptively beneficial for all, all now including people from future generations.

In fact, based on this line of reasoning, it may even be suggested that, in the intergenerational context, the presumptive benefit principle seems too restrictive rather than too permissive in justifying risk imposition. Similar to the fact that we cannot avoid imposing various risks on those who coexist with us just by living our everyday life, many of our daily activities also inevitably impose various long-term risks (mostly

---

[17] J. Wolff, "The Precautionary Attitude: Asking Preliminary Questions", *Hastings Center Report* 44 (2014), p. S28.

[18] It may be pointed out that the occurrence of these events are more of certainties than of risks. However, for any given generation, it is still uncertain whether the risk will (partially) materialise during their lifetime and how bad the outcome may turn out to be.

environmental risks in the current context) on those who do not exist yet without generating much corresponding benefit. If the presumptive principle is to rule out this kind of risk imposition, then it may lead to a state of paralysis where we could hardly carry out most of the activities that make our lives meaningful.

This interpretation of the presumptive benefit principle, however, is unnecessarily restrictive. It ignores that a generic principle which prohibits an activity that imposes risks on future generations without generating much benefit for them is also burdensome for those future people because it means that they will also be restricted from carrying out similar activities. Therefore, what will be weighed by any given generation in order to determine the acceptability of a case of intergenerational risk imposition is the expected burden of such risks imposed on them by their ancestors and the benefit of being able to engage in such activities themselves (hence imposing similar risks on their descendants). Considered in this way, it seems that contracting parties would arrive at the same presumptive benefit principle as outlined in Chapter 3—a case of intergenerational risk imposition is presumptively beneficial for all and is therefore potentially justifiable in a multi-generation contracting position if the burden for everyone to be subject to the risk (imposed on them by their predecessors) can be significantly outweighed by the benefit they gain from being able to carry out similar activities.

Note that my response here does not deny the presence of moral hazard in cases of intergenerational risk imposition and the resulting challenges we need to deal with. Demanding a more cautious principle, however, would not help to solve the problem because the challenge is more of a practical one—since there is a strong incentive for the current generation to engage in activities that are not justifiable to future generations and those who do not exist yet have no power in deterring such behaviour, it is challenging to ensure actual compliance with principles of intergenerational justice (such as the presumptive benefit principle). This problem of compliance is relevant for an account of justice and I will come back to it (though not completely solve it) in the next chapter, but it should be distinguished from the question of what principles are justifiable in the multi-generation contracting position. Based on what I have argued,

the presumptive benefit principle remains justifiable in the intergenerational context and the fact that future generations are subject to the decisions of their predecessors does not in itself provide a reason to replace it with a more cautious principle.

5.2.2 Unknown risk preferences

A different argument for adopting a more conservative attitude when imposing risks on future generations is motivated by the observation that we do not know the risk preferences of those who live in the future. It is a relevant concern because I have argued in Chapter 3 that principles of just risk imposition can only serve to rule out illegitimate cases of social risk imposition and they cannot determine, in an abstract context, what level of risk is justifiable to all members of a society. This further question needs to be addressed by looking at the actual risk preferences of the citizens. In a society with a majority of risk-averse members, a less risky scheme can be justified based on the Greater Burden Principle.

This method can no longer apply when future generations are introduced into the contracting position because the risk preferences of future people are unknown to the current generation. That is, even though parties in the multi-generation contracting position are not put behind an artificial veil of ignorance, time constitutes a natural, partial veil among them. In order to determine the justifiable level of risk imposition on future people whose risk preferences are unknown to the decision maker (that is, the current generation), an additional principle may be needed.

Such a scenario has been considered by Buchak. She proposes that, when imposing risks on people whose risk preferences are unknown to the decision maker, the decision maker should adopt a more conservative risk attitude.[19] To illustrate, she considers the following case—imagine that you are helping an acquaintance to treat muscle pain and need to decide between heat treatment and ice treatment. Applying heat may lead to either complete relief or intense pain while applying ice may result in either reduced, mild pain or moderate pain. If you cannot ask the acquaintance about his own preference, the right choice seems to be applying ice. This is because, according to

[19] L. Buchak, "Taking Risks behind the Veil of Ignorance", *Ethics* 127 (2017), 610-44, pp. 630-32.

Buchak, intense pain is an outcome which we are not sure that the acquaintance would be willing to take. When the person's risk preference is unknown, we should "always err on the side of caution and choose the less risky option", so that no one could fault us for being too risky.[20] If this principle is endorsed in the multi-generation contracting position, then it may imply a demand that is similar to (some version of) the Precautionary Principle.

However, it is not obvious that parties in the Scanlonian contracting position would reasonably accept Buchak's proposal because, while people may fault others for being too risky, they could also fault others for being too cautious. Such a complaint may, moreover, apply to cases where the possibility of severe consequences is involved. For example, consider the following case where a doctor needs to decide whether to perform surgery on an unconscious patient. The patient is highly likely to suffer permanent brain damage if the surgery is not performed promptly, but the surgery will introduce a small but non-negligible risk of death. If the doctor refuses to perform the surgery on the ground that he prefers to err on the side of caution, it is at least not implausible for someone to reasonably fault him for being too unwilling to take risk.

Buchak may point out that, even though her conclusion in the heat or ice treatment example is to adopt the more risk avoidant attitude, she does not insist that people should always choose the option which implies the most risk avoidant attitude. Rather, her proposal is that the risk imposer should adopt "the most risk avoidant of the reasonable risk attitudes."[21] So a possible response to my example is that refusing to perform the surgery is not a reasonable risk avoidant attitude for the doctor to adopt.

However, this move would render Buchak's proposal without much action-guiding power, unless an additional principle is introduced to define what constitute reasonable risk attitudes. Instead of attempting to make such a definition, Buchak in fact suggests that the line between being reasonably and unreasonably risk-avoidant is bound to be arbitrary. If so, then the judgment of what counts as the most risk-avoidant of the reasonable risk attitudes is, by her definition, also arbitrary. Such an arbitrarily defined

[20] Ibid., p. 631.

[21] Ibid., p. 630.

risk attitude cannot provide a reason for contracting parties to endorse a particular choice of risk imposition.

In fact, looking at the two cases, it seems that we agree with Buchak about applying ice not because it is the more risk-avoidant choice but because it is consistent with our common preference in such a situation. Similarly, we disagree with the conservative doctor not because we have a clear line to define how risk-avoidant a doctor should be but because in this particular case most of us would prefer to take the small risk of death to prevent brain damage. As shown in these two examples, when choosing for others without knowing their risk preferences, we reasonably adopt what is to our best knowledge the most *common* rather than the more risk-avoidant attitude for the particular choice situation. If, on the other hand, we have evidence to believe that the person we are choosing for would prefer otherwise, we should then choose according to his own preference. It can be reasonably justified to those impacted because, by choosing in this way, we have tried our best to count their interests equally with the rest of us.

Coming back to the case of intergenerational risk imposition, then, while parties from the current generation do not know the risk preferences of their descendants, they can similarly justify their decision about a particular case of risk imposition to future generations based on what they could reasonably assume about future people's preferences. Without any counter evidence, the best guess should be that the same distribution of risk preferences in their own population would also apply to future generations. However, as time goes by, if, for example, the younger generations start to exhibit more risk-avoidant preferences than what is prevalent among the older people, then such information would require the current generation to adjust their assumption about the future generations and start to choose more conservatively for them. Without such evidence, the unknown risk preference of future generations should not in itself justify a more conservative principle for intergenerational risk imposition.

5.3 Risks Imposed on Future Generations—Anything Special?

In the previous section, I conclude that the mere fact that we are imposing risks on those who live in the future does not justify principles that are more risk-avoidant than the ones I have argued for in the contemporary context. But would special caution be warranted due to certain features of the risks we typically impose on future people? Arguments for the Precautionary Principle, for example, often appeal to reasons such that the probability of a risk is unknown and that a risk may involve a potentially catastrophic outcome. Such features are often seen in intergenerationally imposed risks. Can reasons of this kind justify a more cautious principle to deal with risk imposition on future generations? Let me explore this possibility in this section.

5.3.1 Potential catastrophe

The discussion of the Precautionary Principle as the maximin strategy sometimes restricts itself to focus on one type of risk—those with potentially catastrophic consequences. For example, Sunstein argues for an Anti-Catastrophe Principle which requires the adoption of the maximin strategy only for catastrophic risks.[22] Similarly, Gardiner justifies a Core Precautionary Principle (as the maximin strategy) based on the possibility of grave outcomes. Hence, even though he does not explicitly call it a Catastrophic Precautionary Principle, the justification essentially limits its scope of application to disastrous or catastrophic risks.[23]

Compared to an outcome that is merely bad, the term "catastrophe" is used to refer to situations where "many millions of people…suffer severely harmful outcomes."[24] If a risk involves a potentially catastrophic outcome, might it justify special caution in the Scanlonian contractualist framework? To those who are tempted to say yes, a concern may immediately be raised. It may be pointed out that reasons contracting parties could

---

[22] Sunstein, *Laws of Fear*, p. 61, pp. 109-15.

[23] Gardiner, "A Core Precautionary Principle", p. 47.

[24] L. Hartzell-Nichols, "Precaution and Solar Radiation Management", *Ethics, Policy and Environment* 15 (2012), 158-71, p. 160. It may be pointed out that the number (millions of people) in this definition should be contingent on the population size. For example, if there are only a hundred people left on earth, causing them to die would be catastrophic. I think this is correct, but the specific number need not concern us too much here as what I am going to argue does not depend on it.

legitimately appeal to in accepting or rejecting a principle according to Scanlon's framework need to be individual reasons.[25] If, for each individual, the worst possible outcome is death (or suffering from severe health conditions), then, from an individual perspective, how is a potential catastrophe worse than the worst outcome of, for example, being killed by a car? If a catastrophe is no worse than a common deadly risk for each individual, then what reasons do contracting parties have to demand a special principle to deal with catastrophic risks?

Nevertheless, contracting parties may indeed have individual reasons for assigning a higher value to the avoidance of catastrophe than those everyday deadly risks. First, it can be particularly devastating to live through a catastrophe, losing loved ones, waiting for one's own fate, watching millions of people suffer and seeing civilisation being destroyed. Moreover, in cases where a catastrophe may threaten the continued existence of mankind, many (though probably not all) people may have an even stronger interest in preventing the risk from materialising because they ground the value of their life projects in being part of human flourishing and in being able to pass down their history and culture to descendants.[26] So let us assume that contracting parties could legitimately assign a much higher value to the avoidance of catastrophic risks than that of everyday risks.

Given this, would they also demand a more cautious principle to deal with catastrophic risks? It seems not necessarily. In fact, if the concern is that people have a strong interest in avoiding catastrophe, then such a consideration can already be captured by the presumptive benefit principle through the expected value calculation. Since parties take catastrophe to be especially burdensome, they will assign a significant negative value to that possible outcome. It will then take a really small probability to make the expected burden low enough to be outweighed by the benefit of not eliminating it.

On the other hand, if the chance of risk materialisation is really slim, then it may not be reasonable to insist resources be devoted to fully eliminating a catastrophic risk. Recall

---

[25] Scanlon, *What We Owe to Each Other*, p. 229.

[26] For such an account, see S. Scheffler, "The Afterlife" (Part I and II), in *Death and the Afterlife*, ed. N. Kolodny (Oxford: Oxford University Press, 2016). Also see S. Scheffler, *Why Worry About Future Generations?* (Oxford: Oxford University Press, 2018).

that, in the standard context, the rationale for parties to endorse the presumptive benefit principle is that both the badness of an outcome and its likelihood matter. The same rationale still applies here—even among catastrophic risks, some would have a much higher likelihood of materialisation than others (consider, for example, the probability of 1% versus 0.000001%). Treating them the same and demanding whatever efforts are needed to remove the possibility of catastrophe does not seem reasonably acceptable to all.

While it is not reasonable to ignore the likelihood of a catastrophic risk, applying the presumptive benefit principle may also be problematic. First, it may be pointed out that a typical feature of catastrophic risks is that their probabilities of materialisation would increase over the long term, so the expected burden faced by the current generation may be much lower than what the future generations are subject to. But note that we are now considering a multi-generation contracting position and our actions need to be justified not only to our contemporaries but also to later generations who will be subject to the higher likelihood of catastrophe. So the justifiability of a case of risk imposition will be subject to the evaluation by the worst affected parties, who are usually those from future generations.

A different objection is concerned with the fact that, for long-term catastrophic risks, we often lack knowledge about their probability of materialisation. The presumptive benefit principle is not applicable for such risks because their expected burden cannot be calculated. In fact, it has often been suggested that it is the combination of unknown probability and unacceptable outcome that justifies a more cautious principle, such as the maximin strategy. I turn to this argument now.

5.3.2 Lack of probability

Adopting the maximin strategy means that one always chooses the option whose worst outcome is the least bad among available alternatives. Note that, while the Precautionary Principle is often understood as the maximin strategy applied to the context of catastrophe, maximin in its strict form demands not just the elimination of some especially bad outcome. If maximin is followed, then an option should not be

chosen, even when its worst outcome is not particularly bad, as long as there exists an alternative whose worst outcome is less bad. On the other hand, what most people intend the Precautionary Principle to demand is something like the following—I) where feasible, the chosen option should not involve a possibility of catastrophic outcomes; II) if I) is not feasible, choose the option whose possible catastrophe is the least bad. This is a partial maximin strategy. Let me borrow Sunstein's term and call the combination of I) and II) the anti-catastrophe principle.

The anti-catastrophe principle is less demanding than maximin, but it is still more cautious than my presumptive benefit principle because it does not allow the imposition of a risk no matter how much benefit it may bring about as long as it involves a possibility of catastrophe whose probability we do not know. So the question now is whether certain conditions may justify replacing the presumptive benefit principle with the maximin or the anti-catastrophe principle in the Scanlonian contractualist framework.

Let me try to answer this question in two ways. First, I will consider if unknown probability itself can justify the maximin strategy in general. If so, the justification would automatically cover the anti-catastrophe principle (and in fact demand revising it to the stricter form). If this argument does not work, I will then consider if unknown probability together with the possibility of catastrophe may provide a justification that is targeted more specifically to the anti-catastrophe principle.

Maximin in its general form may be understood as a particular attitude toward risk—when making risk-involving choices, the chooser adopts the maximin strategy if he cares about nothing other than how bad the worst possibility of each option is. If we understand the maximin strategy as such an attitude, then what seems to follow is this—in a given choice situation, if a person would not have adopted the maximin strategy when probabilities were known to him, then he should not be triggered to switch to the maximin strategy *simply because* probabilities are unknown. To show why this is the case, let me employ the Risk-weighted Expected Utility (REU) model developed by Buchak.[27]

---

[27] L. Buchak, *Risk and Rationality* (Oxford: Oxford University Press, 2013), ch. 2.

The REU model explains how people's risk attitudes may affect their preference for probabilistic payoffs. To illustrate, consider a game which offers a 20% chance of getting £10, 50% chance of getting £20 and 30% chance of getting £50. According to the REU model, this game can be understood in the following way—i) the player will get at least £10 (call this State 1); ii) there is a 80% chance that he will get an additional £10 (call this State 2, in which the player gets £20 in total); iii) there is a further 30% chance that, in addition to the £20 he gets, he will get £30 more (call this State 3, in which the player gets £50 in total). The risk-weighted expected utility of this game, for a given player, can be calculated by aggregating the utility value he assigns to i) State 1, ii) moving from State 1 to State 2 and iii) moving from State 2 to State 3, each discounted by the relevant probability.[28]

This model can then incorporate a person's risk attitude in the utility calculation in a rather straightforward way—if a person is risk neutral, he would assign utility value to i), ii) and iii) that is proportional to the incremental monetary value implied in those states; if he is risk averse, the values would be proportionally declining.[29] In the extreme case, if the person assigns zero value to ii) and iii) because he is concerned with nothing but the worst outcome (and uses the resulting REU to assess the game against other alternatives), then he is effectively adopting the maximin strategy.

Note that the REU model does not dictate which risk attitude a rational agent should adopt, so it has no recommendation for or against the maximin strategy. However, what the model does show is that, in a particular choice situation, if you would not adopt the maximin strategy when you know the probabilities, then being deprived of knowledge about the probability numbers would not *in itself* motivate you to switch to the maximin strategy. This is because choosing a non-maximin strategy in the context of known probabilities means that the person assigns positive utility values to ii) and iii). Such

---

[28] For example, if a person attaches 10 units of utility for getting 10, 10 units for getting the additional 10 and 30 units for getting the additional 30, then the REU of this game for him is 10 x 100% + 10 x 80% + 30 x 30% = 27.

[29] For example, a person is risk averse if he attaches 10 units for getting 10, 8 unites for getting the additional 10, and 18 units for getting the additional 30. If, on the other hand, a person attaches 10 units of utility for getting 10, 10 units for getting the additional 10 and 30 units for getting the additional 30, then he is risk neutral.

values should not change simply because probabilities are unknown—lack of probability numbers merely pose a challenge of properly discounting those values.

Since the first approach does not work, let me now turn to the second approach which focuses on the anti-catastrophe principle. Recall that this principle only aims to avoid unacceptable outcomes. We are, moreover, considering it in the specific context where we know nothing about the probability of the outcome.[30] In such cases, there may be a more limited justification for the principle, which does not aim to justify maximin in general. The rationale goes like this—If we know there is a possibility of catastrophe involved, naturally we would be particularly concerned about it. Suppose we do not know the probability at all, a reasonable assumption is that it may be high enough to warrant elimination based on the presumptive benefit principle. Since it involves catastrophe, contracting parties may prefer to err on the side of caution and assume that eliminating the risk is presumptively beneficial. This essentially gives us the anti-catastrophe principle (rather than strict maximin), which requires that catastrophic possibilities be eliminated or mitigated as far as possible.

However, the anti-catastrophe principle in this form seems to be subject to the objection that it will lead to a state of paralysis. Unlike a game in which each option only has a limited range of payoffs, in real life every choice we make may lead to an infinite number of possible outcomes, including a slight chance of causing catastrophe in the long-term. This objection, moreover, can no longer be addressed by appealing to a threshold of possibility because the problem here is exactly that we do not know the relevant probabilities. Therefore, instead of relying on a probability-based threshold, those who support the anti-catastrophe principle (or its similar forms) have suggested other criteria in the context of uncertainty.

Gardiner, for example, argues that the potentially catastrophic threats must satisfy some minimal threshold of plausibility. If they can be dismissed as unrealistic, then maximin should not be followed.[31] Steel elaborates on the idea of plausibility by pointing out that

---

[30] It may be worth noting that this is similar to the conditions Rawls appeals to in justifying the adoption of the maximin strategy behind of the veil of ignorance. See Rawls, *A Theory of Justice*, p. 134.

[31] Gardiner, "A Core Precautionary Principle", pp. 49-51.

the key is to distinguish between what is merely possible and what is scientifically plausible, only the latter requiring us to take action to eliminate the risk.[32] That is, for the anti-catastrophic principle to apply, there need to exist "a scientifically plausible mechanism…whereby an activity can lead to a catastrophe."[33] This requirement may serve to rule out far-fetched cases such that my eating an ice-cream today may lead to a hurricane that kills millions of people in one thousand years. With such a constraint, many of our daily activities may remain permissible and the problem of paralysis can be avoided.

However, while it makes intuitive sense to introduce the requirement of scientific plausibility, this condition seems to defeat the very justification for the anti-catastrophe principle. Recall that the principle is justified based on the rationale that we do not know anything at all about the probability of the catastrophe and we should in such cases err on the side of caution. This argument would only hold if we truly lack knowledge to make any reasonable estimation of the probability and therefore believe that any number, be it 0.01% or 50% or 99.99%, is possible. However, if the scientific mechanism that may lead to a catastrophe is, as required by Steel, well understood, then it could hardly be the case that, with such knowledge, we still completely lack the ability to make an estimation of the probability.

To be sure, probability estimation here is not referring to statistical estimation. For risks such as climate change or nuclear disasters, we would not have the statistical data from past events to calculate a frequency-based probability. But with the scientific knowledge, we should be able to indicate a rough range of probabilities such that the likelihood of catastrophe should lie somewhere between 0.1% to 10%. Despite being a wide range, it is still useful in giving people a sense of how likely the outcome is and whether it is a risk that can certainly be dismissed as not being presumptively beneficial to be imposed. Hence, being able to come up with such estimation means that the presumptive benefit principle may still be applicable.

---

[32] Steel, *Philosophy and the Precautionary* Principle, p. 40.

[33] Ibid. Note that Steel calls it the Catastrophic Precautionary Principle.

It may be pointed out that the estimated range can be wide enough to produce both positive and negative expected values, in which case contracting parties may reasonably disagree about whether it is presumptively beneficial to impose the risk or not. But note that the principle requires a case of risk imposition to be presumptively beneficial for *every* contracting party so that it may be justifiable. Since it is indicated that the probability is within a certain range, it is not unreasonable for at least some risk-avoidant individuals to prefer to err on the conservative side and appeal to the upper end of the range. If they then assess the risk to be not presumptively beneficial, the policy of imposing such a risk could be legitimately rejected as it is not presumptively beneficial for all.

Another concern about the approach is that non-statistical estimation is subjective and cannot, therefore, be used for interpersonal justification. However, even though it is not purely based on statistical estimation, this kind of probability is not purely subjective either. To arrive at such estimation, scientists may, based on the current knowledge of the scientific mechanism, break down the causal chain, analyse the physical dynamics, assume a range of scenarios, build simulation models, draw on available data of other events that involve similar processes or sub-processes, and so on.[34] Unlike a pure subjective probability, the validity of such a result can be discussed and potentially dismissed as unreasonable with scientific counter argument. It is, moreover, subject to review and adjustment as time goes by when new theory is discovered or new evidence is observed. This forms the basis for it to be used in interpersonal justification.

To summarise then, lack of probability does not in itself prompt the contracting parties to abandon the presumptive benefit principle in favour of the maximin or the anti-catastrophe principle. On the other hand, if we know there exist scientifically plausible mechanisms whereby a risk may eventualise into a catastrophe, then its probability should be estimated based on the best available knowledge and be used for applying the presumptive benefit principle. If people can at best estimate a range of probabilities, contracting parties may reasonably appeal to the more conservative end of the range to

[34] For a more formal account of this kind of probability, see, for example, W. Myrvold, *Beyond Chance and Credence: A Theory of Hybrid Probabilities* (Oxford: Oxford University Press, 2021). Myrvold calls this epistemic probability, which is a hybrid form of probability that applies objective physical dynamics to credences.

assess if keeping a catastrophic risk rather than trying to eliminate it is presumptively beneficial.

Before I conclude the discussion of this section, there is one more point to note—while we often assume that a catastrophic risk is a long-term risk, being potentially catastrophic and being long-term are two distinctive features of a risk. A catastrophic risk can also be imminent. If it is imminent, how to deal with the risk can only be decided based on what people currently know and what measures of prevention are currently available. On the other hand, if the risk is long-term, then there is potentially more time for people to acquire knowledge and develop relevant technologies to forestall its materialisation. Does the additional time also imply that more should be done? Let me turn to this point now.

5.3.3 Long-term risk

Since we are discussing risks imposed on future generations, the risks in question here are mostly long-term risks. Because they may only materialise in the future, time naturally introduces a wider range of possibilities for us to reduce or eliminate the likelihood of disastrous outcomes. For example, if a bomb in a residential building is about to explode in one minute and we do not know how to defuse it, what we can do is probably limited to evacuating as many residents as possible. If, on the other hand, we are talking about a time bomb that will take a hundred years to explode, it is hardly convincing to claim that what we are able to do at this moment is all that can possibly be done until the explosion happens.

Given the increased possibilities, future people may, correspondingly, reasonably expect more cautions to be taken if the risk is imposed on them. For example, for risks we still lack full understanding of, time now allows us to acquire more knowledge through scientific research and experiments. As our understanding about the mechanism of a risk improves, we may also search for effective measures or develop relevant technologies that may stop or slow down the processes that will potentially lead to catastrophe. If we expect certain measures to be effective, we may also start to test them on a small scale,

observe the results and adjust according to updated knowledge.[35] Because there is time for us to do these things, future people may reasonably expect us (and later generations who live before them) to make such *ongoing* efforts so that they could accept the remaining risk as being legitimately imposed on them, that any catastrophe they may suffer occurs "despite the fact that reasonable precautions have been taken." As I have discussed in previous chapters, such reasonable expectations can be captured by the reasonable precautionary principle (not to be confused with the Precautionary Principle) for just risk imposition.

In addition to demanding continued efforts in accumulating knowledge and searching for effective measures, future people may further expect us to deal with long-term risks in a way that keeps future options open. We should not, according to such expectation, implement precautionary measures which may themselves impose significant and irreversible risks on future people. An example is solar geoengineering technology (also known as Solar Radiation Management). Research has suggested that it may be used to slow down the process of global warming by reflecting more sunlight away from the earth. However, such attempts to mitigate the risk of climate change may itself create the so-called risk of termination shock—once deployed and then stopped suddenly, there would be a rapid increase in global temperature, which could be much more disastrous than global warming at its current pace.[36]

While the expectation of preserving flexibility for future people can be explained by appealing to the general egalitarian demand that we should ensure future generations enjoy an equal prospect of autonomy as we do, time also adds to the force of the demand. When dealing with imminent risks, it is not uncommon for us to face hard trade-offs—we may have to sacrifice some people in order to save others, or we may have no choice but to implement measures that are themselves harmful in order to prevent a worse outcome. Taking such actions may be justifiable based on what we are

---

[35] This is similar to an approach known as Robust Adaptive Planning (RAP). According to Steel, RAP recommends an experimentalist approach to dealing with uncertainties by requiring policy makers to take action to reduce uncertainty, to learn through the process and to constantly adjust to the improved understanding of the situation. See Steel, *Philosophy and the Precautionary* Principle, pp. 62-66.

[36] Hartzell-Nichols, "Precaution and Solar Radiation Management"; A. Parker and P. J. Irvine, "The Risk of Termination Shock From Solar Geoengineering", *Earth's Future* 6 (2018), 456-67.

capable of achieving within the short time limit. When a risk is still far away, however, it is less likely that our choices are limited to such hard trade-offs. For example, instead of implementing solar geoengineering right now, thus imposing the catastrophic termination risk, we could choose to focus on reducing carbon emissions instead while continuing to develop the technology alongside, testing and refining it in small-scale, just in case other precautionary measures fail. Future people may also reasonably expect us to research into other possible technologies that are safer to use than solar geoengineering.

These additional expectations for managing long-term risks are not motivated by the thought that we should be more risk-averse when risks are long term. Rather, they are based on what human beings may reasonably expect from their fellows—since time expands what generations of us, collectively, may be able to achieve, we should also be collectively responsible for devoting more efforts to mitigating long-term risks. This of course does not mean that the full burden should fall on the current generation. Since there is still time before the risk materialises, the task may be fairly allocated to different generations based on some intergenerational principle of equal distribution.[37]

5.4 Fair Burden Allocation Across Generations

I have now explained how my account could be extended to the case of risk imposition on future generations. For new risks which may be introduced now and materialise in the long-term, the principles discussed in previous chapters, including the presumptive benefit principle and the reasonable precautionary principle, may also guide us in assessing if their imposition is justifiable to future generations and what precautionary measures should be implemented. For existing risks that may strike future generations, be it natural risks or risks imposed by our predecessors or contemporaries, the principles can also work in a similar way to determine if resources should be devoted to actively eliminate or reduce the likelihood of disastrous outcomes. For both cases, the currently

[37] Such as the principle of equal prospects for autonomy I discussed in Section 5.1.3. Even if the contracting parties do not endorse an egalitarian principle of distribution over time, requiring the current generation to bear the full burden is unlikely to be justifiable because, based on the Greater Burden Principle, if all burdens need to be borne by one generation, people from that generation may reasonably complain that there exist some alternative arrangement where the worst impacted may bear less burden than what now falls upon them.

living generation(s) need to make decisions based on their best judgment, which are then subject to revision by subsequent generations if they acquire new knowledge or technology and can therefore update the relevant evaluation. The burden of taking precautionary measures should then be fairly allocated across generations.

This framework of intergenerational just risk imposition and fair risk sharing, however, may be subject to objections raised by two different camps. On the one hand, some may object that my account is not cautious enough because it allows people to delay taking actions. On the other hand, others may question if my account, which requires the current generation to spend a considerable amount of resources in reducing future risks, is against the common intuition that we should prioritise the saving of current lives. Let me address these concerns in turn.

5.4.1 What if it turns out to be too late?

According to the first objection, my approach to dealing with long-term risks is not cautious enough because sometimes we need to act early enough before it is too late. Climate change serves as a good example. While it is a long-term risk, it is not like a rolling dice which simply takes a few hundred years to settle into a position. As pointed out by Thoma, what we call the risk of climate change involves a series of instances of risk materialisation and the resulting harm accumulates until the consequences are irreversible.[38] This provides a strong reason for early generations to act promptly rather than expecting later generations to share the burden. It seems that my account fails to capture this kind of urgency in mitigating long-term risks.

I have two responses to this concern. First, if a given generation can be reasonably expected to understand the mechanism of the risk and hence the urgency of taking relevant measures to prevent the irreversible consequences, then they are already required by my fair risk sharing principle to contribute a corresponding level of effort, which is likely to be higher than the level implied by an equal distribution across generations. This is because failing to do so would violate the principle of equal prospects of autonomy (or other similar principles), even based on this generation's *own*

[38] J. Thoma, "Time for Caution", *Philosophy and Public Affairs* 50 (2021), 50-89, p. 65.

perspective—if they fail to spend such efforts, they cannot reasonably expect that those future people who will be subject to a high chance of catastrophe will enjoy the same prospects of autonomy as they themselves do.

Second, if the current generation lack sufficient knowledge to support the urgency of the issue, then it is not unreasonable for them to delay taking full-scale actions (other than doing research, making small-scale experiments and so on). This is because how much effort they should devote to mitigating one particular risk is also a matter of prioritisation. Because of limited resources, it makes sense for any generation to prioritise based on the best available information. Prioritising tasks without evidence usually means spending less resources and efforts on other projects that are better supported by existing knowledge.

In fact, if we reflect on our sense of urgency in mitigating the risk of climate change, it may be realised that addressing climate change is a priority not because it represents the most cautious choice when we know relatively little but because, and to the extent that, this is the most reasonable action to take based on our current level of knowledge, technology, resources, and so on.

Some may object that, in the particular case of climate change, even if generations preceding us actually did what they, based on the information available to them, should have done, retrospectively it may still turn out to be insufficient. Worse still, when we finally gain enough knowledge to realise this, it may already be too late because we may have missed the opportunity to intervene early enough so as to eliminate the risk or reverse the change. It is exactly this concern, it may be argued, that justifies a more cautious principle (such as the Anti-Catastrophe Principle) to deal with long-term risks.

However, it seems to me that our somewhat strong intuition about the case of climate change is mostly based on an ex-post perspective—it is motivated partly by the fact that climate change has actually turned out to be one of the most pressing issues we face globally and partly by the recognition that our predecessors have obviously failed to do enough even based on what they knew. So our intuition about this real-world case may not be the best guide for us to think about long-term risks in general.

When being put behind the natural veil of ignorance, would contracting parties demand others to take a level of precaution that exceeds what the best available knowledge demands of them? It seems they would not reasonably endorse such a principle. It is true that, as time goes by, what we now believe should be done based on our best judgment may prove to be insufficient, or what we decide to prioritise may turn out to be the less important issue. But even if it turns out that we have chosen the wrong path, we do not wrong future generations as long as we have done what they could reasonably expect from us. To illustrate, let us consider the following imaginary case.

> *Meteor strike:* We as the current generation will devote a certain amount of resources, as required by the relevant demands of intergenerational justice, to long-term projects such as mitigating climate change and preventing a potential meteor strike. Based on our current level of knowledge, we see the risk of climate change to be a more imminent threat to future generations than the possibility of a meteor strike, so we devote most of these resources to addressing the former. Through our efforts, we successfully set the right path to reduced carbon emission and a lower rate of resource depletion. However, 2000 years later, people living then realise that the earth is about to be struck by a giant meteor. Even though they start to devote a large share of their resources to prevent the strike, it is too late to develop the relevant technology. A few generations later, the meteor hits the ground and causes a large explosion, which does not lead to human extinction but significantly changes the climate of the earth. The resulting environmental deterioration makes any climate change that could have been caused by human activities comparatively trivial.

For people living in this post-meteor-strike era, they may well feel a bit unfortunate that their ancestors (that is, us) have chosen to prioritise addressing the issue of climate change over a meteor strike. However, by reading the history of our times, they are unlikely to conclude that they are *wronged* by our decision. We do not wrong future generations because it is indeed the most reasonable choice we could make based on the information and knowledge accessible to us. By choosing to prioritise what we

reasonably believe is more important for them, we have given a sufficient weight to the interests of our descendants and we do not therefore fail to respect them.

5.4.2 Shall we leave contemporaries to die?

Let me now turn to the other objection. The concern is that there are many deadly accidents and disasters happening right now. While it is true that devoting resources to reducing long-term risks could save future lives, it is also true that the same resources could be used to save contemporaries who are subject to immediate deadly threats. It simply sounds counterintuitive that, given limited resources, we should prioritise the saving of future lives over the prevention of immediate deaths. For example, we seem to have quite clear intuition for the following case introduced by Frick.[39]

> *Miners*: Gareth is trapped in a collapsed shaft and will die today if we do not rescue him. The cost of rescue is high and the fund can be used instead to improve the safety at the mine. Suppose the reduced risk of future accidents can save a hundred statistical lives. Shall we save Gareth or the one hundred persons whose identities we do not know yet?

Most people's intuition is that we should save Gareth. Frick believes that this intuition is consistent with what Scanlonian (ex-ante) contractualism would demand. His rationale is the following—for any given individual, they would prefer the avoidance of immediate death to the reduced possibility of future death, so contracting parties would reasonably endorse a principle that requires saving Gareth.[40] If this is correct, then it is in tension with (part of) what my account demands—the requirement to actively devote resources to reducing long-term risks is unlikely to take priority over the saving of contemporary people who are dying right now.

However, the way Frick formulates the principle that requires saving Gareth seems wrong to me. By stating the case as a choice between avoiding *immediate* death and reducing the possibility of *future* death for oneself, most people would reasonably

---

[39] J. Frick, "Contractualism and Social Risk", *Philosophy and Public Affairs* 43 (2015), 175-223, p. 212.

[40] Ibid., p. 215.

choose the former because it is presumptively valuable to live a bit longer (admittedly there may be some controversy about this claim, but let us accept it for the purpose of discussion here). However, even if it is indeed in everyone's interest to die later rather than now, this is not the accurate way of describing the situation here. If the interest at stake is living a bit longer, then the choice for everyone to make, when put in a generic way, is between 1) being killed at Gareth's age and 2) being exposed to a similar risk of workplace accidents throughout their working age, which yields a possibility of death either before or after Gareth's current age. It is not clear that everyone would reasonably prefer the latter.

If we agree that the explanation based on immediate versus future death fails, the question then is—is our strong intuition about saving Gareth in *Miners* actually misguided?[41] I think not. If we take a step back and ask ourselves what is the motivation behind our intuition, our answer is likely to be a simple one—we think we should definitely save Gareth because we do not actually believe in the assumption of future certainty. That is, even though the example claims that we will not be able to improve mine safety if we choose to save Gareth now, hence one hundred persons will certainly die in the future, we simply cannot take this assumption of certainty seriously. We think that, if there is still time before the risk materialises, there must be something that we could, in addition to saving Gareth, *also* do to reduce the risk in the long term and to save lives in the future. Maybe we can reallocate resources from other less urgent projects to improve mine safety; maybe some cheap technology will emerge in the near future which allows us to change how the mine operates; or maybe new regulations will force the mine to close before the next accident happens.

---

[41] It may be suggested that *Miners* is a rescue case, so our intuition can be explained by those reasons that are used to justify the duty of rescue, such as the proximity of the physical location, the determinacy of the situation, and so on. For such discussion, see, for example, F. Kamm, "Does Distance Matter Morally to the Duty to Rescue?", *Law and Philosophy* 19 (2000), 655-81; V. Igneski, "Distance, Determinacy and the Duty to Aid: A Reply to Kamm", *Law and Philosophy* 20 (2001), 605-16; D. Miller, "The Nature and Limits of the Duty of Rescue", *Journal of Moral Philosophy* 17 (2020), 320-41. It seems to me that there are indeed some similarity between *Miners* and the standard rescue case (such as saving a drowning kid from the pond) and the argument I am going to make is a bit similar to the determinacy account, but they not entirely the same. *Miners* can be formulated in a way that deviates from a rescue case (such that we need to travel a long way to save Gareth and the rescue is not easy), but our intuition should remain the same.

We are, moreover, not wrong in refusing to accept the assumption of future certainty. Because of the additional possibilities created by time, uncertainty is an inherent feature of those events that will only happen in the future. We therefore feel justified to discount *not the value but the urgency* of acting right now to avoid future deaths. If the current generation decide that it is less urgent to address a risk that is still far away, it does not necessarily mean that they fail to value future people's lives equally. In fact, those future people who are subject to a risk we choose not to mitigate immediately would also face similar choices between saving their contemporaries who are dying and addressing long-term risks that affect their descendants and are allowed to prioritise the former. They, like us, would not reasonably reject a principle that takes urgency into consideration in determining what should be done at present in the contracting position.

In *Miners*, the key factor determining urgency is time—the further away a potential disaster is, the more time and hence more possibilities for people to be able to do something before the risk materialises, the less urgent that the risk must be mitigated right now. But time is not the only relevant consideration. For example, as discussed earlier, risks like climate change are constantly evolving and materialising, so it may add to the urgency for us to act early enough before it is too late.

On the other hand, it also does not mean that projects like addressing long-term risks will never get priority over contemporary ones. This is because most projects that impact only current lives also involve uncertainty. In real life, many immediate deaths are not easily preventable as is assumed in *Miners*. Some severe illnesses, for example, cannot be cured right now no matter how much effort we spend. If the marginal expected effectiveness of devoting more resources to a contemporary project declines sharply, at some point it will be outweighed by the marginal expected effectiveness of committing them to long-term projects. Therefore, for any generation, it would be justifiable for them to allocate their resources between short-term and long-term projects. Moreover, requiring them to contribute to projects like mitigating long-term risks does not mean that they should do so at the cost of leaving contemporaries to die.

To conclude then, I have argued in this chapter that contracting parties in a multi-generation contracting position would reasonably endorse the same principles I have previously argued for—the presumptive benefit principle, the reasonable precautionary principle and the fair risk sharing principle—for risk imposition on future generations. However, due to the uncertainties introduced by time, the demands implied by these principles for each generation should be interpreted in a dynamic way. More specifically, every generation should always act based on the best knowledge available to them, including knowledge about the mechanism of the risk to be imposed, the effectiveness of different precautionary measures as well as future people's risk attitudes and what kind of life they may regard as valuable. They should, moreover, make constant efforts to improve their understanding and always be ready to adjust their actions in response to evolving situations and new information.

I have so far outlined an account of just risk imposition and fair risk sharing that includes both contemporaries and future people. In the next chapter, I turn to a slightly different problem. As shown in this chapter, the handling of a long-term risk is a project that requires contributions by different generations and its burden should be fairly allocated among them. However, it is often questioned if principles of justice, understood as fair rules of cooperation, can apply in the intergenerational context due to the lack of reciprocity between non-overlapping generations. A similar concern is sometimes also raised against the inclusion of people with severe disabilities in theories of distributive justice. Does the inclusion of these two groups of people represent a problem for my account of fair risk sharing? I will try to answer this question in the next chapter.

Chapter 6 — Reciprocity, Risk Sharing and Social Cooperation

I have so far outlined an account of just risk imposition which demands fair sharing of the burden of social risks. Among those covered by fair compensation are two groups of people who are often thought to fall outside the scope of justice—people with (severe) disabilities and people from (non-overlapping) future generations. My account requires all members of a society to share the burden of the disability risk because, as explained in Chapter 3, disability is at least partly a social risk whose negative consequences are to a greater or lesser extent determined by institutional designs (such as the planning of public spaces, the availability of medical care, the types of jobs available, and so on). On the other hand, as shown in Chapter 5, risk imposition by the contemporaries on far future generations through their policy choices also demands fair risk sharing across time.

However, it is often suggested that people with severe disabilities or people from future generations that do not overlap with the contemporaries cannot be dealt with within the framework of distributive justice. In explaining why including them is a problem, the reason often cited is that there is a lack of reciprocity between them and contemporary, contributing members of the society. For some, lack of reciprocity means no issues of justice arise at all. This view reflects a conception of justice that is often referred to as justice as reciprocity. Justice as reciprocity is different from the conception of justice that I endorse in this thesis (which, as I will explain, can be described as justice as impartiality[1]). So it may be suggested that, in order to address the potential objection to the inclusion of those unable to reciprocate in my account, I only need to explain why justice as reciprocity is not the right conception of justice to endorse.

Dismissing the objection in this way, however, does not seem to capture many philosophers' intuition that reciprocity is an indispensable idea for justice. For example, Barry, who rejects justice as reciprocity in favour of justice as impartiality, at the same time suggests that "any theory of justice that tries to eliminate reciprocity would be

[1] I adopt this term from Barry. See B. Barry, *Justice as Impartiality* (Oxford: Clarendon Press, 1995).

doomed from the start."[2] Nussbaum also notes that, by dropping the idea of reciprocity, a theory of justice would then need an account of "a more complex political psychology that makes it clear how a cooperation that is not understood in terms of mutual advantage alone will be sustained."[3] Based on the specific concerns expressed in the literature, I will in this chapter examine two reasons why reciprocity may still be relevant for an account of justice as impartiality—that is, 1) it may represent a demand of fairness and non-exploitation that is subordinate to impartiality and 2) it may refer to the conditional motivation people have for complying with principles of justice. I will also explain why, based on these ways of understanding reciprocity, including people with severe disabilities and people from future generations are not a problem for my account of fair risk sharing.

The chapter is organised as follows. In Section 6.1 I examine the notion of reciprocity as a demand of justice and distinguish it from mutual advantage and rational choice. With this clarification, I argue that a conception of justice founded solely on the demand of reciprocity is not defensible. On the other hand, my account of fair risk sharing does require people to make a fair contribution when they are able to, which is compatible with a form of fair reciprocity outlined by White.[4] Including people who are currently unable to reciprocate in the scheme does not violate this demand of reciprocity.

In Section 6.2, I turn to a different notion of reciprocity, reciprocity as a conditional motivation for justice, and explain how reciprocity in this sense may help to secure stable cooperation based on principles of justice. I also show that such stable compliance may be sustained when the scheme is extended to people with severe disabilities and people from future generations, provided that the cooperative scheme incorporates a properly designed risk-sharing mechanism. Lastly, I reconcile the treatment of reciprocity as a conditional cooperative strategy with it being an intrinsic disposition toward justice in Section 6.3. I do so by appealing to Ci Jiwei's account that

---

[2] B. Barry, "Justice as Reciprocity", in *Liberty and Justice: Essays in Political Theory 2* (Oxford: Clarendon Press, 1991), p. 235.

[3] M. Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2006), p. 153.

[4] S. White, *The Civic Minimum: On the Rights and Obligations of Economic Citizenship* (Oxford: Oxford University Press, 2003).

reciprocity as a motivation for justice can be both conditional and intrinsic at the same time.[5] Echoing his view that background institutions which secure the benefits of reciprocal compliance can help to foster an intrinsic disposition of reciprocity, I suggest that the establishment of risk sharing institutions may encourage a more generous form of reciprocity and hence enhance the stability of social cooperation.

6.1 Reciprocity as a Demand of Justice

According to a common view, reciprocity is an inherent feature of an account of distributive justice. That is, issues of distributive justice only arise in a scheme of social cooperation where participants engage with each other in a reciprocal relationship. Principles of justice, then, represent cooperative rules which regulate how the fruits of social cooperation should be divided in a fair way.[6] According to this view, only those capable of cooperating with other participants are owed considerations of justice. Hence, people with severe disabilities would be excluded if they cannot make any meaningful contribution to the scheme (though they may be owed humanitarian concern); people from non-overlapping future generations also cannot be dealt with under the framework of justice because they are not in a cooperative relationship with those now living.

Rawls's account is a typical example which excludes these two groups of people from the original contracting position.[7] In *A Theory of Justice*, Rawls states that parties in the original position know that their physical and psychological capacities are "within the normal range," that they "are full and active participants in society and directly or indirectly associate together over the whole span of their life."[8] This implies the exclusion of people with severe disabilities because their capacities are not within the normal range. Rawls also insists that parties know that they are contemporaries (either

---

[5] J. Ci, *The Two Faces of Justice* (Cambridge, Massachusetts: Harvard University Press, 2006).

[6] A. Buchanan, "Justice as Reciprocity versus Subject-Centered Justice", *Philosophy and Public Affairs* 19 (1990), 227-52, pp. 227-31.

[7] Though Rawls mentioned that these two groups of people should be dealt with by an account of justice, just at a later stage rather than in the original position. See, for example, J. Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), p. 20.

[8] J. Rawls, *A Theory of Justice* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999), pp. 83-4.

at the present time in *A Theory of Justice* or as belonging to an unknown generation in *Justice as Fairness*), thus excluding people from non-overlapping generations from the contracting position.[9]

In explaining why such exclusion is necessary, it is often suggested that the problem is there is no reciprocity between those who are unable to contribute to social cooperation and those who are contemporary, cooperating members of the society.[10] To see why reciprocity is deemed indispensable for some theorists of justice, let me first examine what its specific demand is.

6.1.1 What is the demand of reciprocity?

The demand of reciprocity is often associated with, or taken to be equivalent to, the demand of mutual advantage. However, depending on how mutual advantage is understood, the notion may be more or less demanding than reciprocity. Sometimes, the claim that justice only arises in the context of reciprocity is taken to simply mean that the relationship or the scheme should be mutually advantageous—mutually advantageous in the sense that every participant is better off compared to no social cooperation at all. Rawls, for example, often refers to society as a cooperative venture for mutual advantage and what he means is simply that "social cooperation makes possible a better life for all than any would have if each were to try to live solely by his own efforts."[11] If this is what reciprocity demands, then including people with disabilities in the scheme and sharing with them some fruits of social cooperation is unlikely to undermine the mutually advantageous result compared to the state of nature. Similarly, treating people from future generations as part of the scheme is also consistent with this notion of mutual advantage. But this is obviously not what people have in mind when they argue that including those unable to contribute to social cooperation violates the idea of reciprocity.

---

[9] Rawls, *A Theory of Justice*, pp. 109-10; J. Rawls, *Justice as Fairness: A Restatement* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2001), p. 160.

[10] See, for example, L. Becker, "Reciprocity, Justice, and Disability", *Ethics* 116 (2005), 9-39; H. McCormick, "Intergenerational Justice and the Non-reciprocity Problem", *Political Studies* 57 (2009), 451-58.

[11] Rawls, *A Theory of Justice*, p. 109.

On the other hand, the demand that principles of justice should be reciprocal may be taken to mean, more narrowly, that it should represent a rational choice for every individual to comply with the principles as rules of cooperation. That is, the benefit of complying with their demands should be guaranteed to outweigh the potential benefit of non-compliance so that people will be prompted by rationality to comply. If this is what reciprocity demands, then it may indeed imply the exclusion of those who suffer severe disabilities with no prospect to recover or those who live in the far future from the scheme—if justice requires cooperation with these two groups of people, defection rather than compliance is likely to be the rational choice for other participants.[12]

The interpretation of reciprocity as rational choice, however, is too demanding. In fact, few theories of justice which invoke the idea of reciprocity would meet this demand. It may be suggested that Rawls's principles of justice are justified as the rational choice of parities in the original position, but it is only a rational choice for hypothetical individuals who are put behind the veil of ignorance. It is obvious that, after the veil is lifted, the rational choice for an actual person who belongs to the better-off group would be to rescind the difference principle. What motivates such a person to comply is not rational self-interest but a sense of justice.[13] Some may also suggest that Gauthier's theory of rational contractarianism satisfies the demand of rational choice.[14] But even Gauthier requires bargainers in his contracting position to accept what he calls the principle of minimax relative concession.[15] By referring to the acceptance of equal

---

[12] It should be note that, even with this interpretation, reciprocity as rational choice does not automatically call for the exclusion of all those who are suffering disabilities or who live in the future. For example, it may not be irrational to include those temporarily suffering disabilities in a cooperative scheme of mutual aid. (See Barry, "Justice as Reciprocity", p. 220.) While those currently contributing to the provision of aid cannot claim direct repayment from the present beneficiaries, they may benefit from the scheme in the future when they need similar aid. But cooperation of this kind can only cover people who have a reasonable prospect of recovering. Including those with permanent disabilities in the scheme would no longer be a rational choice for existing participants. (Barry calls this "chronic bad risks". See ibid., p. 230.) On the other hand, it may also be rational for contemporaries to cooperate with overlapping future generations to smooth out long-term risks caused by economic cycles and demographic changes, but not with those from the far future.

[13] Rawls, *A Theory of Justice*, pp. 125-26.

[14] D. Gauthier, *Morals by Agreement* (Oxford: Clarendon Press, 1986).

[15] Ibid., p. 22.

proportional sacrifice as a kind of "practical rationality", he is obviously referring to it more as a common sense of fairness than as a rational choice in an egoistic sense.

With reciprocity being distinguished from narrowly egoistic rational choice on the one hand and mutual advantage (versus no cooperation) on the other, what, then, is its actual demand? It may be helpful to note that the term reciprocity is originally used to describe the practice of exchanging things with others for mutual benefit. In this original context of bilateral exchange, each party gives out something in return for something else. On the one hand, the exchange is reciprocal only if both parties receive something beneficial for themselves through the exchange. On the other hand, both parties have to contribute something—if the practice only involves one party giving something to the other, then it cannot be a case of reciprocity. A strong form of reciprocity may further require that the benefit received by each party be equivalent or proportional to their contribution, while a weak form merely requires *some* return be made. What neither version demands is that the exchange represents the rational choice (versus other potential options) for both parties in a narrowly egoistic sense.

Hence, when applied to social cooperation, the requirement that the scheme be reciprocal can also be broken down into two demands—1) every participant should receive some benefit from the scheme and 2) every participant should contribute something to the scheme. If we adopt the strong form, then it may be further demanded that 3) the benefit received by each be equivalent or proportional to their contribution. Note that, whether or not we endorse the third demand, it does not imply that including an individual who is a contributor to the scheme always represents a rational choice for other participants.[16] Also note that, when extended to multiple persons, reciprocity can take more complicated forms than simultaneous, bilateral exchange. For example, in a cooperative scheme of mutual aid, those who contribute in a given period do not necessarily expect to claim repayment from the present beneficiaries. Rather they expect to be reciprocated in the future when they need similar aid. Barry calls reciprocity of this kind "reciprocity as mutual aid".[17]

---

[16] Because if a participant has strong bargaining power, the rational choice for him may be demanding more than what is proportional.

[17] Barry, "Justice as Reciprocity", p. 220.

6.1.2 Justice as reciprocity vs. justice as impartiality

Theories which explicitly or implicitly appeal to the two demands of reciprocity represent a conception of justice which Buchanan labels as justice as reciprocity.[18] It is contrasted with a different conception of justice—what Buchanan calls subject-centred justice or what Barry calls justice as impartiality (I will use the latter label "justice as impartiality" from now on).[19] Justice as impartiality does not limit its scope to those who are capable of contributing to social cooperation. Rather, it is grounded in the idea of fundamental moral equality of persons and therefore applies to all members of a society.[20] Different accounts of justice as impartiality may differ in what fundamental equality implies—whether it merely requires equal respect and concern or it further demands equalisation of opportunities, advantages or some other desiderata of distributive justice. What such theories share in common, however, is the denial of the relevance of the ability to contribute to the scope of justice.[21]

Similar to Barry's own account of justice as impartiality, my account of fair risk sharing is also based on the Scanlonian contractualist framework which does not exclude people with severe disabilities or people from future generations from its scope.[22] By employing this framework, I effectively endorse the conception of justice as impartiality and have therefore seen no problem about extending considerations of justice to people who cannot reciprocate. The inclusion of these people is consistent with the intuition

---

[18] Buchanan, "Justice as Reciprocity versus Subject-Centered Justice", p. 229. Note that it is different from how Barry uses the term. Barry refers to justice as reciprocity as a hybrid idea of reciprocal compliance under justice as mutual advantage. I will come back to this in the next section when I discuss about reciprocity as a conditional motivation for justice.

[19] According to Buchanan, his notion of subject-centred justice differs from Barry's account of justice as impartiality in that the latter assumes that principles of justice are the choice or agreement among contracting parties, whereas his own version does not make such an assumption. See ibid., p. 230, note 6. However, the difference need not concern us here as my main purpose is to contrast this type of theories with theories which regard considerations of justice to be dependent on the ability to reciprocate.

[20] Ibid., p. 233.

[21] Ibid., p. 235.

[22] Though Scanlon does mention that people with severe disabilities who cannot reason in the contracting position should be represented by trustees. See T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1998), pp. 185-86. But this does not mean people with severe disabilities are excluded from the considerations of the other contracting parties.

that they can be treated justly or unjustly with regard to risk imposition and risk sharing (as well as other issues of distribution).

To be sure, the fact that justice as impartiality is consistent with (some) people's intuition does not in itself justify its superiority to justice as reciprocity. Critiques of justice as reciprocity often focus on the theory's internal inconsistency. For example, as pointed out by Buchanan, whether a person is a potential contributor and is therefore qualified to be owed justice depends not only on his own capacity but also on the cooperative framework chosen by the society he lives in. Different cooperative frameworks may recognise people with different types and levels of skills as contributors and may also affect the type and amount of cooperative benefits that will be available for distribution. Hence, the choice of cooperative framework is itself a question of justice, which cannot, however, be answered by a conception of justice as reciprocity.[23]

While reciprocity as a conception of justice is hard to defend, many people do share the strong intuition that reciprocity is an important idea for a theory of justice. There are two alternative ways of capturing this intuition. The first is to recognise it as a demand that is subordinate to justice as impartiality. The second is to appeal to reciprocity not as a demand of justice but as the motivation for people to comply with principles of justice. Let me examine the first argument now and turn to the second in the next section.

6.1.3 Reciprocity as a demand embedded in justice as impartiality

White, for example, endorses the first interpretation in formulating his account of fair-dues reciprocity, which requires that "where the institutions governing economic life are otherwise sufficiently just, those who claim the generous share of the social product available to them under these institutions…should make a decent productive contribution, *suitably proportioned and fitting to ability and circumstances*, to the community in return."[24] The focus is not on how much a person contributes to the

[23] Buchanan, "Justice as Reciprocity versus Subject-Centered Justice", pp. 236-38.

[24] White, *The Civic Minimum*, p. 59, emphasis added.

cooperative scheme per se, but whether he makes a reasonable effort to contribute when he is able to.

As White points out, such a demand of reciprocity is grounded in the ethos of mutual respect between equals.[25] If a person benefits from others' labour but refuses to do his bit when he has the capacity to do so, the concern is that he effectively treats others in an instrumental or exploitative way—he does not respect them as equals. This principle of reciprocity can therefore be naturally embedded in a conception of justice as impartiality. Reciprocity, in this sense, is a demand of justice that is subordinate to impartiality.[26]

If what is required is that some form of fair-dues reciprocity be recognised in an account of justice as impartiality, then my account of fair risk sharing does not violate this requirement by extending its scope to those who are currently not able to contribute. In fact, my account naturally incorporates a notion of reciprocity that is similar to what Barry describes as the reciprocity as mutual aid, but in formulating its principles, I did not explicitly specify the conditions with reference to people's ability to contribute. Let me briefly examine this issue now.

In the case of disabilities, I have explained that I treat the risk of suffering disabilities, whether congenital or acquired, (at least partly) as social risk. Therefore, those who lack the capacity to contribute to social cooperation due to disabilities can be seen as victims of social risk materialisation and are owed compensation by other members of the society. There is generally no concern if such a person has made some contribution before he loses the abilities to do so and gets compensated, but if a person is born with severe disabilities and has never been able to contribute, it may then be questioned whether the demand to compensate him violates some notion of reciprocity.

In fact, formally speaking, for those suffering congenital disabilities, their duty to contribute to risk sharing could be seen to be offset by the compensation owed to them.

---

[25] Ibid., pp. 62-63.

[26] It is a separate question whether reciprocity represents an independent demand under such a conception of justice or it is simply derived from some general egalitarian ideals. For discussion on this issue, see ibid., pp. 67-74.

Meanwhile, should they recover in the future and acquire the ability to contribute, they are then indeed required by my account of fair risk sharing to reciprocate others' contribution and bear an appropriate share of the burden going forward. But such interpretation can be equivalently stated in a more straightforward way as the following —the demand of fair risk sharing requires that every member of the society bear a fair share of the burden *when they are able to do so*. This is in line with the demand of fair-dues reciprocity proposed by White. Reciprocity is only expected when the lack of it implies instrumental or exploitative treatment of others, so it is not a concern (for an account of justice as impartiality) that people with congenital disabilities have never contributed because their failure to do so does not imply disrespect toward those who put in their efforts.

In the case of future generations, on the other hand, the concern about lack of reciprocity is actually a bit different. Few would hold the view that requiring those who exist at the present time to contribute to an intergenerational scheme to which future generations cannot currently contribute represents disrespectful or exploitative treatment of the current generation. As time goes by, people from the future will come into existence and it will become their turn to contribute based on what justice requires. It is only when they inherit what their ancestors made an effort to achieve but fail to do their share do they then violate the demand of reciprocity.

Instead, the concern about lack of reciprocity between non-overlapping generations is often based on the role of reciprocity in sustaining cooperation.[27] In this context, reciprocity is referred to not as a specific demand of justice but as a conditional motivation to cooperate with those who comply with the rules of cooperation and punish those who fail to do so. The worry with extending the scheme across generations is that, since there is no direct reciprocity between non-overlapping generations and the so-called cooperative benefits always flow in one way, earlier generations may not be motivated to contribute. Unlike the contemporary case where a person who otherwise

[27] See, for example, a survey of such concerns in McCormick, "Intergenerational Justice and the Non-reciprocity Problem", p. 452. Note that this worry about the lack of motivation is often commingled with the view that issues of justice only arise in reciprocal relationships. Since I have already rejected the latter (i.e. justice as reciprocity), I will focus here on the motivation question.

lacks the motivation to contribute may nevertheless choose not to defect knowing other fellow citizens could sanction him or condition their cooperation with him based on his behaviour, such a mechanism to secure cooperation does not exist in the intergenerational context. (In fact, a similar problem would also arise in the case of severe disabilities.) This understanding of reciprocity is consistent with how many sociologists and economists use the term when studying social cooperation—according to their analytical model, the motivation of reciprocity is what helps to secure stable cooperation among human beings.

For those who believe that reciprocity as a conditional motivation for cooperation is a relevant issue for theories of justice, lack of reciprocity between people from the current and the future generations, between people capable of contribution and people with severe disabilities may indeed be a problem. To address this concern, however, we need to first understand how reciprocity in this sense impacts justice and the stable compliance with its principles. I will turn to this motivational understanding of reciprocity in the next section.

6.2 Reciprocity as the Conditional Motivation for Justice

As pointed out by Ci, many theorists of distributive justice have appealed to the notion of reciprocity as the conditional motivation to be just or fair,[28] though they do not always do so in an explicit way. For example, according to Rawls, what motivates individuals to comply after the veil is lifted is a sense of justice.[29] When citizens with this sense of justice "believe that institutions or social practices are just or fair (as the conception specifies), they are ready and willing to do their part in these arrangements *provided that they have reasonable assurance that others will also do their part*."[30] Similarly, Barry argues in his account of justice as impartiality that individuals' motivation of compliance is the desire to behave fairly—"if I am motivated by a desire

---

[28] Ci, *The Two Faces of Justice,* p. 38. Ci refers to demands of justice as "broadly reasonable norms". However, since other literatures also refer to individuals' conditional compliance of these demands as the norm to reciprocate, I will avoid calling principles of justice as reasonable norms to avoid potential confusion.

[29] Rawls, *A Theory of Justice*, pp. 125-26.

[30] J. Rawls, *Political Liberalism*, p. 86, emphasis added by Ci (See Ci, *The Two Faces of Justice,* p. 24).

to behave fairly, I will want to do what the rules mandated by justice as impartiality require *so long as enough other people are doing the same*."[31] According to Ci, this conditional motivation to be just is based on reciprocity—that is, my sense of justice will motivate me to comply with principles of justice only if others reciprocate my efforts through their compliance and vice versa.[32]

Reciprocity as the conditional motivation to comply with principles of justice is to be distinguished from rational egoism on the one hand and altruism (or unconditional compliance) on the other. If human beings are, as a matter of fact, dispositioned in a way that they cannot commit themselves to reciprocal support of principles of justice and will always choose to defect whenever doing so accords with their self-interest, then no social cooperation based on predetermined terms could be stably maintained. On the other hand, if (most) individuals can be motivated in a way that they would follow fair terms of cooperation unconditionally regardless of other people's compliance, then the problem of stability would not arise in the first place.

Reciprocity in this sense is not part of the content of principles of justice. Hence, as Ci mentions, there could be, at least in theory, reciprocal compliance with justice as impartiality, with justice as mutual advantage or with some other principles of justice.[33] However, to say that the motivational question is separate from the content question does not mean that answers to these two questions can be completely independent. Ci suggests answers to the motivational question "seem to constrain…though not determine" answers to the content question.[34] The example he gives is that, if some rules are "blatantly prejudicial", then reciprocity will not be sufficient to sustain universal compliance.[35] On the other hand, Rawls also suggests that principles of justice should be suited to human moral psychology so that stable support can be expected—"however attractive a conception of justice might be on other grounds, it is seriously

---

[31] Barry, *Justice as Impartiality*, p. 51, emphasis added.

[32] Ci, *The Two Faces of Justice,* p. 5, p. 38.

[33] Ibid., p. 39.

[34] Ci, *The Two Faces of Justice*, p. 3.

[35] Ibid., p. 38.

defective if the principles of moral psychology are such that if fails to engender in human beings the requisite desire to act upon it."[36]

Rawls's concern with the motivational question may explain the tension in his theory when he tries to bring together the Humean idea of the circumstances of justice and the Kantian idea of impartiality. On the one hand, by introducing the veil of ignorance which prevents arbitrary distinctions of identities, natural assets and abilities from affecting the choice of cooperative terms, Rawls's construction almost points naturally toward the possibility of including people with disabilities and people from different generations. On the other hand, however, he also insists that parties in the original position know that they coexist temporally and geographically and that they should be roughly similar in their physical and mental abilities.[37] These conditions are part of the so-called circumstances of justice described by Hume which guarantees the contracting parties' prospect to reciprocate.[38] With these conditions met, the chosen principles by parties under the veil would reflect the conception of justice as reciprocity.

It has often been pointed out that it is potentially contradictory for Rawls to appeal to these two ideas simultaneously.[39] While I agree with this observation and I also believe that justice should not be limited to the Humean circumstances, it seems to me that the underlying rationale for Rawls to appeal to the circumstances of justice relates to his concern that principles of justice should be able to "generate [their] own support".[40] That is, after the veil is lifted, individuals need to have the motivation to comply with the chosen rules or to support institutions embodying the just principles. This is not to say that coercive measures cannot be employed to force them to comply. Rather the concern is that a system that needs to rely constantly on coercive enforcement is not going to be stable—it does not generate its own support.

---

[36] Rawls, *A Theory of Justice,* p. 398.

[37] Ibid., p. 99. pp. 109-10.

[38] Ibid., p. 109.

[39] See B. Barry, "Circumstances of Justice and Future Generations", in *Obligations to Future Generations*, eds. R. I. Sikora and B. Barry (Philadelphia: Temple University Press, 1978), p. 228; Nussbaum, *Frontiers of Justice*, p. 57.

[40] Rawls, *A Theory of Justice,* p. 230, p. 399.

But even if we accept the importance of stability, there is still a logical gap here—given that justice as reciprocity does not imply it is always a rational choice for every individual to comply with its principles, why does Rawls believe that the reciprocal nature of the derived principles can secure stable compliance? He seems to have bridged the gap with two related claims. First, people's sense of justice or their motivation to comply with principles of justice is reciprocal—that is, they can only be motivated to contribute to a scheme of cooperation when they know those who benefit from their contribution would reciprocate the effort. Second, only reciprocal principles are compatible with people's reciprocal motivation for compliance—that is, only when the justified principles are reciprocal in their demand can reciprocity as people's natural disposition be sufficient in generating stable compliance.[41]

While referring to the same term "reciprocity" makes the two claims appear naturally connected, when spelt out they may not be as intuitive as they appear to be. In the rest of this section, I am going to analyse the relationship between reciprocity and stable cooperation and I will draw the following conclusions. First, there is a case for reciprocity as the conditional motivation for compliance to help secure stable social cooperation. Second, it is possible to extend this case to cover people with severe disabilities and people from far future generations, but it requires that people exhibit a more generous sense of reciprocity and/or that the cooperative scheme be properly structured (in my case, with a fair risk sharing scheme). Third, while there is no guarantee that people can be motivated in this more generous manner, what my analysis shows is that properly designed institutions could provide the background condition to foster such a norm.

Before moving on to my analysis, I need to acknowledge here that not everyone shares Rawls's concern that people's motivation to comply should constrain the content of principles of justice.[42] But even for those who hold the view that the content question

---

[41] I have reconstructed Rawls's discussion about moral psychology and the problem of stability into these two claims. See his discussion in ibid., p. 433; pp. 434-40.

[42] For the debate on whether motivation is a feasibility constraint for justice, see, for example, D. Estlund, "Human Nature and the Limits (if any) of Political Philosophy", *Philosophy and Public Affairs* 39 (2011), 207-237; D. Wiens, "Motivational Limitations on the Demands of Justice", *European Journal of Political Theory* 15 (2016), 333-52; D. Estlund, "Reply to Wiens", *European Journal of Political Theory* 15 (2016), 353-62.

should be answered independent of the motivation question, what I am going to argue may still be relevant as long as they believe that it is relevant (for a normative theory) to understand what are the conditions for achieving stable compliance with principles of justice.

6.2.1 Reciprocal compliance and stable cooperation

In order to show how reciprocity as the conditional motivation for compliance may help to secure stable cooperation, I will employ a formal analysis based on a cooperative model because such a model can capture the key concepts in question here (that is, conditional motivation and stable cooperation). There may be concerns about the validity of formalising reciprocity into a conditional cooperative strategy. I will lay out my analysis first and address this important question in the next section.

It is common for economists and sociologists to model the scheme of social cooperation as a multi-player game of public goods provision.[43] In the public goods game, each player is required to cooperate by making a contribution to the provision of a public good. Assume that the good can only be provided when all players contribute. Once the good is provided, no one inside the scheme could be excluded from enjoying its benefit. Hence, this game resembles the structure of a multi-player prisoners' dilemma (PD)—collectively they would be better off if all choose to cooperate, yet the rational choice for each is to defect and free ride on others' efforts.[44]

This model is also consistent with how social cooperation is generally understood in political philosophy. In describing the stability problem, Rawls observes that complying with the principles of justice is collectively rational (compared to the absence of agreement) from the standpoint of the original position, but free-rider egoism would be

---

[43] See, for example, R. Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984); E. Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge: Cambridge University Press, 1990); S. Bowles and H. Gintis, "Homo Reciprocans", *Nature* 45 (2002), 125-28; T. Voss, "Game-Theoretical Perspectives on the Emergence of Social Norms", in *Social Norms*, ed. M. Hechter and K. D. Opp (New York: Russell Sage Foundation, 2001).

[44] For any given player, if a sufficient number of players choose to cooperate so that the good is provided, then the rational strategy for him would be to defect and free-ride; if there are not a sufficient number of players who choose to contribute, then the good will not be provided and the rational strategy for him would also be defection.

better from an individual perspective after the veil is lifted.[45] Moreover, his concern that the scheme of cooperation needs to be stable[46] could be translated into the requirement that universal cooperation should be an equilibrium of the public good game—every player chooses to cooperate and no player has the motivation to unilaterally deviate from this choice.

I further assume that this public goods game is going to be repeated indefinitely into the future, so players can potentially reciprocate others' behaviour in previous rounds. Note that, while I follow what this game is conventionally called, the goods players produce may be more appropriately described as social goods rather than public goods in its strict sense because, while we assume that a player cannot be excluded from enjoying its benefit if he has participated in that round of the game (that is, he can free-ride), the game does allow a player to be excluded from enjoying the good if he is banned from participating in a round altogether.

Now, to see how players' motivation for cooperation may affect the equilibrium strategy of the game, consider the following cases—

> *Rational egoism*. Assume all players are solely motivated by the pursuit of self-interest. In each stage of the game, the rational choice for every player is to defect. Collectively, universal defection is the unique Nash Equilibrium (NE) of the game.[47]
>
> *Rational egoism with third-party enforcement*. Assume all players are solely motivated by the pursuit of self-interest. An external party is introduced to the game who threatens to punish those who choose to defect. If the cost of being punished outweighs the benefit of defection, then, in each stage of the game, the rational choice for every player is to cooperate. Collectively, universal cooperation becomes the NE of the game.
>
> *Conditional cooperation*. Assume all players are motivated to act based on the following strategy—they will cooperate on the condition that others also cooperate

[45] Rawls, *A Theory of Justice*, p. 435.

[46] Ibid., pp. 434-35.

[47] Nash Equilibrium (NE) means that every player knows the equilibrium strategy of other players and no player has the incentive to deviate from the equilibrium strategy.

while punishing those who choose to defect by excluding them permanently from future rounds of the game.[48] Since the game is to repeat indefinitely into the future, the cost of foregoing future cooperative benefits would outweigh the benefit of defection in a given round, hence transforming universal (conditional) cooperation into a NE of the game.[49]

Both rational egoism with third-party enforcement and conditional cooperation could achieve an equilibrium of universal cooperation, but note that there is a problem with the introduction of external sanction. Third-party enforcement is not cost-free. The joint provision of such a system would itself introduce a new game of public goods provision that resembles the original dilemma. This also offers a game-theoretical perspective why coercion is not an ideal solution for the problem of stability.

Conditional cooperation, on the other hand, avoids this second-order public goods provision problem by assuming that the sanctioning system is internalised by each player. In this case, players can be described as acting on the motivation of reciprocity —they reciprocate cooperators with cooperation and defectors with (permanent) defection. The question, then, is whether this is a reasonable assumption to make.

The motivation of "cooperation for cooperation and defection for defection" is in fact quite similar to Rawls' description of reciprocity as a tendency to "answer in kind", which, according to Rawls, is a disposition rooted in human moral psychology.[50] The first part—"cooperation for cooperation"—is also a good reflection of how a sense of

---

[48] There can be a less harsh version of conditional cooperation, which is known as "tit-for-tat". According to this strategy, instead of permanently excluding defectors, players choose to cooperate with those who have cooperated and defect on those who have defected in the previous round. While the adoption of this strategy has been seen in lab setting, according to Voss, it is not a Nash Equilibrium strategy in this game due to its insufficient penalty for defection. See Voss, "Game-Theoretical Perspectives", p. 120. I will therefore not consider it here.

[49] In any round of the game, each player now faces the choice between the short-term benefit of defection in the current round and the long-term cost of foregoing the benefits of cooperation in all future rounds, which is the aggregation of each round's benefit discounted by the player's time preference and his personal conviction on the likelihood of the game to be repeated into the future. Unless his time preference is extremely biased toward the present and/or his conviction on the continued iteration of the game is very low, the opportunity cost of the future is likely to outweigh the present benefit of defection. In such a case, cooperation would become the preferred strategy for him provided that the other player cooperates. See Voss, "Game-Theoretical Perspectives", p. 116.

[50] Rawls, *A Theory of Justice*, p. 433.

justice is usually understood. That is, people are generally willing to comply with principles of justice on the condition that (most) others also comply. The second part—"defection for defection", on the other hand, is a bit dubious. Can individuals really be assumed to be motivated in this way? Moreover, even if they exhibit such willingness in the game setting, is it a good reflection of their motivation in the real world since "defection for defection" can be a cost-free action to take in the game but may become very costly when applied to a large group of anonymous individuals in real life?

Two replies might be given to address this worry. First, as proved by behavioural scientists through lab experiments, most players appear to be "altruistic punishers".[51] That is, they demonstrate willingness to punish defectors even when the game is structured in a way that it both incurs a cost and generates no future benefit for a player to punish.[52] This shows that, instead of a mere response to incentives based on cost-benefit analysis, players' sanctioning behaviour reflects their genuine willingness to punish as a way to express their negative moral judgement. This willingness to punish, together with their willingness to cooperate, form what Bowles and Gintis term as *homo reciprocans*. They believe that social cooperation can be sustained because, empirically, many human beings are *homo reciprocans*.[53]

Second, it should also be noted that conditional cooperation in real life often need not (or is not likely to) take the form of personal retaliation. Rather, as suggested by sociologists, such reciprocal strategies may be internalised by members of a society as a moral preference or disposition.[54] Once compliance becomes the social norm, deviating from it may yield sanction in the form of shame or guilt, which may achieve a similar effect of deterring defection without incurring much cost on punishers. This increases the plausibility to assume that upholding the conditional cooperation strategy actually yields net benefit for players.

---

[51] Bowles and Gintis,, "Homo Reciprocans"; E. Fehr and S. Gächter, "Altruistic Punishment in Humans", *Nature* 45 (2002), 137-40.

[52] Bowles and Gintis, "Homo Reciprocans", p. 127.

[53] Ibid.

[54] Voss, "Game-Theoretical Perspectives", p. 129.

Given these two points, conditional cooperation is a reasonable way to model how stable cooperation can be sustained by individuals' reciprocal motivation. The next question is whether people with (severe) disabilities and people from (non-overlapping) generations can be incorporated into this model without disrupting such cooperative equilibrium.

6.2.2 Indirect reciprocity and risk sharing

The common feature of the two groups of people is that they are currently unable to contribute to the scheme of public goods provision. Let me call them non-current-contributors of the public goods game. At first sight, it may be thought the potential problem of including non-current-contributors is that they will be punished as defectors. However, based on Bowles and Gintis' argument that people punish out of their negative emotional response to those who deliberately choose to defect, it seems quite unlikely that they would extend such negative judgment to people who are unable to contribute (either because they have disabilities which prevent them from doing so or because they do not even exist yet).

In fact, if we look at the literatures that describe how lack of reciprocity may create a problem for non-current-contributors,[55] the underlying worry is effectively this—since non-current-contributors do not contribute to the cooperative scheme, they also have no power to form effective threats of punishment on others (that is, the threat of withdrawing cooperation with those who choose to defect on them). Therefore, if, as a matter of fact, most players are not motivated to comply with a distributive rule that includes non-current-contributors as beneficiaries, then defection on non-current-contributors will be a NE from which no one has the incentive to deviate.[56] In such

---

[55] For example, Parfit points out that "earlier generations [cannot] be deterred by threats from later generations." See D. Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), pp. 524-25. Arrhenius similarly argues that "future generations cannot wield any threat against previous generation." See G. Arrhenius, "Mutual Advantage Contractarianism and Future Generations", *Theoria*, 65 (1999), 25-35, p. 25. Quoted in McCormick, "Intergenerational Justice and the Non-reciprocity Problem", p. 452.

[56] Defecting only on non-current-contributors could take the following form—among the public goods provided by the players, there is one good (say, the social security good) that will be distributed to non-current-contributors. Players may defect on non-current-contributors by choosing not to contribute to the provision of this particular good.

cases, reciprocity as the conditional motivation to comply can no longer act as the endogenous mechanism for sustaining stable cooperation.

However, this argument is based on the assumption that players can only be motivated to punish (or threaten to punish) defectors when they themselves are made worse off by the defective behaviours. That is, it is assumed that reciprocally motivated players would not be motivated to punish a defector if that person only chooses to defect on non-current-contributors. This premise is not necessarily true. First, players may be motivated by a more generous sense of justice and are therefore willing to punish defectors on behalf of non-current-contributors even though doing so yields no cooperative benefits and even incurs net costs to themselves. Second, if the game is structured in a way that punishing on behalf of non-current-contributors yields net benefit for all players, then they may also be motivated to adopt this cooperative strategy.[57] In either case, if players are motivated to punish a person for his defection on others, not just on themselves, then it may be said that they are motivated by a form of *indirect* reciprocity.[58] In a cooperative scheme that provides benefit to non-current-contributors, the adoption of this indirect reciprocal strategy may result in stable cooperation that is similar to what can be achieved by direct reciprocity in a conventional public goods game.

The question, then, is whether players can indeed exhibit the motivation of indirect reciprocity in cases where non-current-contributors are people with disabilities and people from future generations. In previous chapters, I have provided a justification why these two groups are owed duties of justice based on the demand of fair risk sharing. If people are indeed convinced by such an argument, they may also be motivated by their sense of justice to comply with rules reflecting this duty and to punish or exhibit negative moral judgment toward those who do not comply. To what extent people can actually be motivated in this way is, of course, an empirical question

---

[57] Or it can be a combination of both. For example, while players are motivated by a sense of justice to support the non-current-contributors, punishing defectors is also consistent with their self-interest because defection by others means they would need to bear a larger share of burden if they want to ensure the collective obligation to support the disabled group is discharged.

[58] J. Heath, "The Structure of Intergenerational Cooperation", *Philosophy and Public Affairs* 41 (2013), 31-66, p. 44.

and may vary from society to society. On the other hand, what I want to show here is that, when the game is properly structured, it may also provide players with the right incentives to uptake the strategy of indirect reciprocity, thus making universal conditional cooperation a NE of the game.

Consider, for example, a modified version of the public good game with risk introduced.[59] In each round of the game, all players in the game face a risk of becoming non-current-contributors. Moreover, it is crucial for their survival to the next round that non-current-contributors receive some benefit of the public goods. For any given player, knowing that he may also fall into the group of non-contributors in some round and thus face the possibility of "being killed" in those rounds if he does not receive support from the public goods provision, the potential benefit of punishing those defecting on non-current-contributors is then very likely to outweigh the corresponding cost. Therefore, in this public goods game with risk, the strategy of indirect reciprocity will be the preferred strategy for all players.

The structure of this game bears considerable similarity to the case where disabilities cause people to be unable to contribute. It is reasonable to assume that any individual faces a likelihood of becoming disabled in any stage of their lives. If they lose their source of income due to disabilities, they fall into the group of non-current-contributors and would need public support to live a dignified life. The public goods maintained by the players can be regarded as including a disability insurance fund, through which all players share the risk of disabilities. The game can, moreover, cover different cases of disabilities. If he finds himself in the disabled group for every round of the game, he then effectively suffers permanent disability. On the other hand, if a player only falls into this group in a few rounds, he experiences temporary disability. (But assume that at least some of the disabled players will recover in each round, so that the number of contributors remains relatively stable throughout the game.) Since able players in a given round know that they also face the risk, however slim, of becoming disabled in

[59] For a game with a similar structure, see an example called "the mutual-aid game" provided by R. Sugden in *The Economics of Rights, Co-operation and Welfare* (Oxford: Basil Blackwell, 1986), pp. 122-28.

the future, they may support a disability fund that covers all kinds of disabilities (and threaten to punish those who attempt to defect on the disabled).[60]

It is important to note that indirect reciprocity will only be the preferred cooperative strategy for all players when the game models the existence of the disability risk and also includes a risk sharing mechanism like the disability fund. Without the possibility of potentially benefiting from the fund in the future should they experience disability, it would not be sufficient to secure stable cooperation based on mutual advantage. A similar mechanism can also be introduced for the case of future generations, though the model of intergenerational cooperation and risk sharing is a bit more complicated (and less straightforward) to lay out. Heath has provided such a model and explained how indirect reciprocity may function in the case of intergenerational cooperation.[61] I will describe this model in an appendix to this chapter and also explain how intergenerational risk sharing may work to enhance the stability of the scheme in a way that is similar to what has been discussed so far.

The conclusion to draw from the formal analysis in this section, then, is this: if we believe that stable cooperation can be achieved through direct reciprocal compliance with principles of justice, then it can be similarly sustained when benefits of the scheme are extended to non-current-contributors if the game is designed in a way that makes indirect reciprocity the preferred strategy for all. In particular, risk sharing provides such a mechanism through which a form of indirect reciprocity can function to make universal (conditional) cooperation the equilibrium strategy of the game. So, with the establishment of risk sharing institutions that embody the principle of fair risk sharing,

---

[60] It may be pointed out that this game cannot cover the case of being born with permanent disabilities (that is, a player falls into the disabled group since the first round). Because other players already know such a possibility would not occur to themselves, they may want to exclude such players from the benefit of the fund if they are motivated solely by self-interest. This seems true based on this particular setting of the game, but note that my argument here is only that a properly structured game can provide an *additional* reason for players to adopt the more generous strategy of indirect reciprocity. In reality, people may have other reasons to extend coverage to such people, based either on self-interest (for example, the concern that they may give birth to children with disabilities) or on a sense of justice (for example, they feel obliged to discharge the duty to support those born with disabilities but want to have this burden fairly shared with other people).

[61] Heath, "The Structure of Intergenerational Cooperation", pp. 42-48.

it is possible that covering people with disabilities and people from future generations will not undermine the stability of a cooperative scheme.

6.3 Reciprocity as a Disposition Toward Justice

My analysis in the previous section may give rise to the following concern. On the one hand, I suggest that reciprocity is a sense of justice, or a disposition toward justice, that motivates people to comply with principles of justice or to support institutions guided by such principles. On the other hand, however, I model the motivation of reciprocity (either direct or indirect) as a conditional cooperative strategy to reciprocate cooperators with cooperation and defectors with defection. I then examine the conditions under which the adoption of such a strategy will make universal cooperation a Nash Equilibrium of the game. By appealing to game theory in this way, it seems that I treat reciprocity as a motivation toward the achievement of net benefit, which is in some sense indistinguishable from rational egoism. If we do recognise reciprocity as a disposition toward justice, is it a mistake, then, to model it using game theory? I would like to address this important concern in this section.

Indeed, as noted by Ci, reciprocity as a sense of justice or fairness represents an intrinsic disposition that should be distinguished from the conditional motivation to cooperate for potential benefits. Reciprocity as an intrinsic disposition should be oriented toward past actions rather than future benefits.[62] If a person has treated me well, intrinsic reciprocity should motivate me to return his kindness even if he has no power from now on to punish me or further benefit me. Similarly, if a person has benefited from others' cooperation in complying with principles of justice (or supporting institutions embodying such principles), then intrinsic reciprocity should motivate him to return the efforts without considering the possibility of sanction or future cooperative benefit.

However, while Ci distinguishes between intrinsic reciprocity (what he calls reciprocity as mutual good will) and instrumental reciprocity (the conditional motivation for future benefits), he very interestingly suggests that reciprocity as a disposition can be both

[62] Ci, *The Two Faces of Justice*, pp. 136-41.

intrinsic and conditional at the same time.[63] To illustrate how this is possible, Ci appeals to an example of gift exchange among friends.[64] Gift exchange is a practice of reciprocity. Unlike market exchange between strangers, people do not engage in gift giving for the sake of receiving future benefit. Indeed, the absence of conscious expectation of return is essential for the practice to retain its meaning. Yet, unlike charity, gift exchange among friends is not based on the motivation of unconditional benevolence—it will collapse if gift giving only goes one way. In fact, what sustains the practice is exactly the implicit knowledge about the predictability of reciprocation. Hence, Ci argues that, in a scheme of reciprocity, only when it can be taken for granted that actions will be answered in kind can people rise above the conscious thought of return.[65]

To be sure, in the context of intimate relationships such as friends, family and lovers, the expectation of being answered in kind does not always need to be translated into mutual advantage. For example, Becker notes that, in these more personal relationships, what people focus on in expectation of reciprocation is often appropriate efforts rather than net benefits.[66] Parents would happily accept a thank you card from their kids in return for their love; people who take care of their injured family members or friends would also readily accept a kiss, a gift or some not-in-kind return for their efforts. However, when extended to a more impersonal context, what "answering in kind" requires is more often the assurance of mutual advantage—it is this implicit knowledge of net cooperative benefit that makes it unnecessary for people to consciously think about potential return when they carry out their daily lives.

According to Becker, these different conceptions of reciprocity, some more generous than others, would compete in a society and turn the idea of reciprocity into more specific norms and standards. So his suggestion is to persuade people from different political backgrounds to accept the more generous form of reciprocity (the version that does not require people with disabilities to reciprocate when they lack the capacity to do

---

[63] Ibid., pp. 141-46.

[64] Ibid., p. 145.

[65] Ibid., pp. 144-45.

[66] Becker, "Reciprocity, Justice, and Disability", pp. 26-27.

so) by making them aware of its potential of generating mutual advantage.[67] While it sounds similar to Ci's argument, his emphasis on persuading people to accept a certain norm of reciprocity based on mutual advantage may be less accurate a description of how social norms emerge in a society and may also make it difficult to explain how reciprocity so achieved can at the same time be an intrinsic disposition toward justice. Ci, on the other hand, suggests that under certain social conditions reciprocity can progress from a conditional cooperation strategy for the sake of mutual advantage to an intrinsic disposition toward justice. That is, when social practices or institutions are established in such a way that the benefits of reciprocal compliance can be taken for granted, it is then possible for people to rise above the conscious thought of return and reciprocate out of good will. Such progress is "not so much an individual achievement as the product of a successful social practice".[68]

My analysis in the previous section, then, may give support to Ci's account. What the model shows is that, by structuring the scheme of cooperation in a proper way, a more generous form of reciprocity can become the equilibrium cooperative strategy for all players. This provides the institutional background Ci calls for which secures the prospects of mutual advantage for reciprocal compliance. While it does not guarantee that everyone living under such institutions would actually act out of a sense of justice or an intrinsic desire to reciprocate, it makes it possible for such a disposition to emerge among members of the society.

To conclude, this chapter seeks to reconcile two points—on the one hand, according to the account I have outlined in Chapter 2 to 5, justice requires that we share the burden of social risks with people with disabilities and people from future generations; on the other hand, it is often suggested that these two groups of people fall outside the scope of justice due to their lack of capacity to reciprocate. In this chapter, I examined the various reasons why lack of reciprocity may be a problem for an account of justice. I suggested that the worry may emerge from two different angles. Some take reciprocity

---

[67] Ibid., pp. 16-18.

[68] Ci, *The Two Faces of Justice*, p. 145.

to be an inherent demand of justice. While I rejected a conception of justice that is based solely on reciprocity, I recognised a demand of fair-dues reciprocity that is subordinate to impartiality and argued that my account of fair risk sharing is compatible with this demand. But lack of reciprocity may still be a concern for some people who are concerned with the problem of stability. I acknowledged the importance of reciprocity as the conditional motivation in securing stable cooperation, but argued that risk sharing can in fact create a mechanism of indirect reciprocity between contemporary, cooperating members of a society and people with severe disabilities and people from far future generations. Therefore the establishment of properly designed risk sharing institutions may foster a more generous norm of reciprocity among members of the society and enhance the stability of social cooperation.

Appendix — The Intergenerational Game

Heath constructs the following game to show that there exists a form of indirect reciprocity between different generations which could sustain cooperation across time.[69]

Imagine eight players sitting in a circle. In each round of the game, every player has two options—he can either choose to receive $2 from the game facilitator or he can let the game facilitator give $5 to the player sitting to his left. This game is a variation of the standard PD game because the rational choice for each player is to keep $2 to themselves (similar to the strategy of defection) but collectively they will be better off if all choose to give $5 to their neighbour (similar to the strategy of cooperation). Because of this structural similarity, in the repeated version of this game, players can adopt a similar conditional cooperation strategy in order to transform universal cooperation into a NE. This strategy states that a player will start by cooperating, that is, by letting the player on his left receive $5, and will continue to cooperate provided that the player on his left also chooses the same rather than getting $2 for himself. This is again a strategy of indirect reciprocity—for any player, he has no power to sanction the player to his right should that player choose to defect and not to give him the $5, but such behaviour will be reciprocated by the player sitting to the right of the defector. This form of indirect reciprocity can be sufficient to sustain a cooperative equilibrium even though the benefits of cooperation only flow one way.

Now consider some slight modification to the game so that it resembles the structure of intergenerational cooperation. Players will now sit in one line instead of a circle. Those sitting to the left represent earlier generations while those to the right represent later generations. In each round, every player is asked to move over one chair to the left, which mimics the process of ageing. The person at the head of the line will be asked to leave the game (i.e. death) and a new player will join to fill the empty chair at the right end (i.e. birth). It seems that the same cooperative strategy can also sustain the cooperative equilibrium of this game. The only difference is that the player who expects to die in the next round will have no one to give the $5 to so he will choose to receive $2 himself instead. But this does not affect the game as doing so yields no negative

[69] Heath, "The Structure of Intergenerational Cooperation", pp. 42-48.

consequence for anyone. He will not be sanctioned as he will no longer be in the game in the next round. Neither will this trigger defection of other players who are still being deterred by the threat of losing future benefits of cooperation.

It may be questioned why it can be assumed that cooperation with the next generation would always yield net benefit in the sense that each player gets $5 instead of $2 by cooperating. This can partly be explained by the fact that each generation experiences a period of dependency toward the end of their lives and needs the production and service provided from the next generation during that period. Such needs used to be covered by family members in pre-modern societies and are more often institutionalised nowadays in the form of (either private or public) pension schemes and care services for the elderly. When these schemes are publicly funded in the so-called pay-as-you-go style, generations who are currently working pay for the generations that have retired, which resembles the structure of the game in which each player hands over money to the "older" player. Furthermore, the net benefit of intergenerational cooperation can also be explained by sharing long-term risks such as economic recession and demographic changes with the subsequent generation. The negative consequences of these long-term risks are impossible to be alleviated through risk sharing among members of the same generation. While cooperating with the next generation is not going to achieve the effect of insurance, extending out cooperation temporally can at least smooth out such shocks to a certain extent, making intergenerational cooperation more likely to be a mutually beneficial arrangement.

Unlike the case of disability, in the intergenerational case, indirect reciprocity is not made possible by the existence of a formal risk sharing mechanism—without sharing long-term risks intergenerationally, other forms of cooperation such as exchange (pay to get old age care service) may also produce reciprocity. But, being a cooperative mechanism which yields additional cooperative benefit, risk sharing does provide a stronger reason for individuals to believe that intergenerational cooperation is mutually beneficial and could be sustained through indirect reciprocity.

Chapter 7 — Conclusion: Institutions and Policies for Fair Risk Sharing

In this thesis, I have developed an account of just social risk imposition and fair risk sharing. As I have suggested, this account should be part of a general theory of distributive justice and should therefore have implications for the design of background institutions to justly deal with social risks. In this concluding chapter, I will briefly examine the implications of my account for normative political economy and the design of public policies.

The chapter is organised as follows. I will first summarise what I have argued from Chapter 2 to 6 in Section 7.1. Then in Section 7.2 I use property-owning democracy as an example to show how considerations of social risk may impact the kind of economic institutions a theory of justice justifies. I argue that wide dispersal of capital ownership alone cannot realise the ideal of equal liberty or non-domination because risk may threaten both capital and the income it generates. A property-owning democracy therefore needs institutions to deal with risk justly. Section 7.3 then looks at a particular institution that may serve the purpose of fair risk sharing—social insurance. I argue that my account may provide not only an additional justification for social insurance but also guidance on what risks social insurance should cover and how the scheme should be funded. I will do a case study on pension schemes and argue that pensions should be funded in the so-called pay-as-you-go style because this funding structure may fulfil the demand of fair intergenerational risk sharing.

7.1 Just Risk Imposition and Fair Risk Sharing—A Summary

The starting point of my account concerning the justice of social risk is that individuals have an interest in autonomy as self-authorship and being exposed to risk may set back this interest. To author an autonomous life, individuals must be able to set and revise goals, form and execute plans and make both short-term and long-term decisions. In order to do these things, they would need, among others, a range of background options to choose from. The range, quality and robustness of these life options would affect the prospect for individuals to lead a life that can be considered as authored by themselves. Risk may affect such a prospect by removing valuable options from the background,

forcing people to focus on survival choices rather than life projects, and undermining the robustness of certain options' availability.

While risk sets back people's interest in autonomy as self-authorship, it does not imply that autonomy requires the elimination of all physical and economic risks (or as many as possible). This is because risk is often the by-product of choices or arrangements that may nevertheless enhance the prospects of autonomy. At the individual level, a person may choose to take on a physical risk by engaging in extreme sport or an economic risk by investing in the stock market. In these cases, without taking on risks, people will not be able to enjoy the potential gains they seek from these activities which form part of their life projects. At the institutional level, we often choose to retain or even impose various social and economic risks through our collective decisions about the design of institutions or social arrangements.

Because of the burden of risk exposure, collective risk imposition of this kind requires justification. That is, the type and level of risks imposed on members of a society through the design of institutions or the choice of social arrangements should be justifiable to all. A case of social risk imposition may be justifiable for reasons that are similar to those supporting the provision of public goods—when a certain kind of burden is imposed on citizens for benefits that are presumptively beneficial for all and significantly outweigh the corresponding costs and when the costs and benefits are fairly distributed, the arrangement may be regarded as being reasonably acceptable to all.

When the burden of providing for such presumptive benefits takes the form of risk imposition rather than a fixed amount of payment, its size (as well as the corresponding benefit) could be evaluated based on its impact on people's prospects on autonomy ex ante. Moreover, unlike a typical case of public goods provision in which the burden is assumed to be fixed, different ways of imposing a risk may imply different levels of burden on affected individuals. Hence, to justify a scheme of social risk imposition to all, it needs to be further demonstrated that its implied burden is reasonably reduced. This implies two demands—first, precautions should be taken to reasonably reduce the likelihood of risk materialisation and the severity of its outcome (*the reasonable*

precaution principle); second, compensation should be guaranteed to potential victims so as to reasonably reduce the expected burden of risk materialisation (*the compensation principle*).

The compensation principle needs to be supplemented with an account of how to distribute the burden of paying compensation fairly. I argue that, presumably, those who benefit from the imposition of a social risk should bear a share of the burden based on reasons of fairness. Since most social risks are imposed for benefits to be enjoyed by all members of a society, it effectively means that everyone should pay for a share of the required compensation (*the fair risk sharing principle*). In addition, burden allocation should also reflect considerations of personal responsibility. In the particular context of risk, we may expect people to manage the outcome of certain risks by following commonly acknowledged due care rules; we may also expect people to manage their own exposure to risks they deliberately take on as part of their personal projects by taking private insurance. Those who fail to meet such expectations may be allocated with a share of the burden that is larger than they otherwise need to bear.

7.2 Risk and Property-Owning Democracy

The inclusion of this account of just social risk imposition and fair risk sharing in a theory of justice has implications for the design of social and economic institutions. Theorists of justice who take the average life prospects approach would often, correspondingly, insufficiently address the issue of risk when discussing normative political economy. For example, after outlining his two principles of justice, Rawls suggests that a property-owning democracy (POD) may realise his ideal of justice as fairness.[1] A property-owning democracy is an economy that ensures the wide distribution of capital among members of the society. For Rawls, wide dispersal of capital helps to secure substantive equality of liberties and opportunities and limits economic inequalities, thus making the effective realisation of his two principles of

[1] J. Rawls, *A Theory of Justice* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999), p. 242; J. Rawls, *Justice as Fairness: A Restatement* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2001), pp. 135-40. Note that, according to Rawls, market socialism may also be compatible with his account of justice, though he does not elaborate on how it should be organised.

justice possible.[2] However, he does not elaborate on how property ownership can protect substantive equality given the fact that properties are constantly subject to risk. Physical properties such as land, buildings and machines are vulnerable to the risk of deterioration and destruction due to natural disasters like earthquake and floods and human-caused hazards like fire, misuse or simply lack of care. Non-physical capitals such as financial assets, on the other hand, are also subject to the risk of devaluation. So the mere fact that a person owns capital at a certain point of time does not guarantee continued ownership throughout his life.

There are also proponents of POD who take the negative impact of risk into consideration but suggest that capital ownership can guard individuals against (economic) risk. Thomas, for example, argue that capital ownership can secure self-sufficiency and non-domination because it can protect individuals from economic uncertainties such as a temporary loss of income.[3] However, his reasoning seems to rely on the assumption that capital can always generate a steady stream of income, which then ensures that individuals can meet their basic needs throughout their lives as well as take up meaningful projects (including getting education or professional training). Again, this assumption does not take adequate account of the fact that capital cannot generate income without being put at risk and that there are also uncertainties in the rate of return capital can yield. To generate income from physical capital, it requires either the owners' labouring efforts or the possibility of renting it out in exchange for yields, each of which brings its own risk. If the owner loses his labouring capacity temporarily or permanently, he will at the same time lose the ability to simply rely on the ownership of capital for income. Or he may choose to lend out the capital to earn rent or invest the capital in other people's businesses. In such cases, the owner is then subject to either credit risk or equity risk. The same also applies to non-physical capital—financial assets also cannot generate the cashflow that is believed to provide the source of financial security against economic uncertainties without themselves being put at risk.

---

[2] Rawls, *Justice as Fairness*, p. 139.

[3] A. Thomas, *Republic of Equals: Predistribution and Property-Owning Democracy* (New York: Oxford University Press, 2017).

Therefore, even in a property-owning democracy where everyone is entitled to a certain amount of capital, it further requires the existence of some mechanisms to deal with the risks capital owners are exposed to in order to effectively realise the ideal of limited inequality, self-sufficiency or non-domination that are valued by defendants of POD. To be sure, both Rawls and Thomas do incorporate some form of risk sharing in their proposal. Thomas, for example, suggests that individuals should be allowed to pool their risks by combining capital in stakeholder schemes.[4] Such schemes, according to him, may be adopted for efficiency reasons as well as solidaristic or egalitarian concerns.[5] Rawls, on the other hand, mentions in *Justice as Fairness* that widespread ownership of capital should be ensured at the beginning of each period.[6] While he does not explain what "each period" means, if we take it to mean frequent periodic redistribution, then in some sense such regular adjustment in capital ownership does imply a form of risk sharing among capital owners.

However, if the motivation behind risk sharing is merely to limit significant inequality of distribution, then it seems to blur the distinction between POD, which its defendants claim to be a form of predistribution rather than redistribution, and conventional welfare state capitalism. Moreover, in both proposals, it remains an open question what rules should guide the design of such risk sharing schemes. In fact, if we agree that at least some of the risks just described are social risks, then my account of just social risk imposition may help to (partially) supply such an account. Its demand of fair risk sharing implies that a property-owning democracy needs a universal risk-sharing scheme so as to fairly allocate the burden of capital and yield losses due to social risk materialisation—a just property-owning democracy should be one in which everyone lives not as independent capitalists but as risk-sharing capitalists.

It is not my purpose here to either reject property-owning democracy or defend some of its variations. The point I would like to make with the example of POD is that whether or not a theory of justice includes an account of risk affects the kind of institutional arrangements it justifies. If my account of just risk imposition and fair risk sharing is

---

[4] Ibid., p. 48.

[5] Ibid., p. 48, pp. 204-5.

[6] Rawls, *Justice as Fairness*, p. 139.

accepted as part of a theory of justice, then it may offer justification for certain institutions that serve to fulfil its demands, including the demand of taking reasonable precautions and the demand of paying compensation to victims of social risk materialisation. The latter demand is of particular importance for normative political economy because it implies the need to establish risk-sharing institutions in order to fairly distribute the burden of paying for compensation among members of the society. Such risk-sharing institutions may differ from other transfer branches of an economy that focus on, for example, income- or wealth-based redistribution.

There are various institutions that can perform the function of risk sharing, but one obvious candidate for this purpose, namely social insurance, is worth special attention. As noted by Rachel Friedman, social insurance has become one of the central economic institutions in liberal democracies and it has exhibited a considerable degree of political resilience over the last century.[7] Yet, on the other hand, political philosophers differ in their opinions over the grounds on which social insurance may be justified and what principles may guide its design. Some treat social insurance as an instrument for providing the social minimum.[8] Compared to the alternative of means-tested benefits, social insurance is believed to be a more dignified institution to provide for the social minimum because its benefit payments are based on predefined risk events. Without subjecting those hit by risk to the mercy of bureaucrats, social insurance could provide a guaranteed minimum without creating a sense of stigma or inferiority, thus preserving equal status among all members of the society.

The minimum-based justification works well for one type of social insurance—those schemes that pay a fixed and relatively low level of benefits to all.[9] But many social

---

[7] R. Friedman, *Probable Justice: Risk, Insurance, and the Welfare State* (Chicago: University of Chicago Press, 2020), pp. 1-4.

[8] See, for example, R. Goodin, *Reasons for Welfare: The Political Theory of the Welfare State* (Princeton: Princeton University Press, 1988); J. Waldron, "John Rawls and the Social Minimum", in *Liberal Rights: Collected Papers 1981-1991* (Cambridge: Cambridge University Press, 1993).

[9] This type of social insurance is sometimes referred to as the Beveridge style as it was first outlined by the British economist William Beveridge in his 1942 report *Social Insurance and Allied Services.* See E. Anderson, "Thomas Paine's 'Agrarian Justice' and the Origins of Social Insurance", in *Ten Neglected Classics of Philosophy*, ed. E. Schliesser (Oxford: Oxford University Press, 2016), p. 80.

insurance programs in contemporary societies have moved away from the flat-rate style and instead pay out benefits that are linked to individuals' contributions.[10] Arguments for social insurance that pays beyond the minimum often treat it as state intervention in the private insurance market to correct either people's behavioural failures in buying a sufficient amount of insurance or the private market's failure in providing certain types of insurance policies. However, as I have argued in Chapter 2, concerns of cognitive biases on the buyers' side may justify at best soft nudging strategies but not mandatory contribution to universal programmes. On the other hand, correction of potential market failures would only justify social insurance for risks that cannot be efficiently covered by private insurance (such as unemployment risk). This justificatory gap makes it difficult for political philosophy to address normative questions concerning the design of social insurance programmes, such as what risks social insurance should aim to cover, to what extent benefits should be linked to prior contributions and how the additional burden caused by population ageing should be shared between generations.

My account of just risk imposition and fair risk sharing, on the other hand, may offer an additional justification for social insurance that goes beyond the social minimum. Moreover, by treating it as the institution to fairly allocate the burden of socially imposed risks, the account may also provide some guidance to the just design of social insurance programmes. Let me turn to this topic in the next section.

7.3 Social Insurance for Fair Risk Sharing

Before I address the specific policy questions, let me first clarify how I use the term social insurance. Social insurance can be used both narrowly or broadly. In its narrow sense, it refers to state-sponsored contributory programmes which resemble private insurance. Citizens make regular contributions into the scheme which serve as insurance premiums and receive compensation when predefined events occur. Social insurance in this sense is often regarded as part of the welfare state. But I intend to use social insurance more broadly—I will take it to mean state-mandated schemes which use risk-sharing mechanisms to pay for the materialisation of common risks. Social insurance in

[10] This type of social insurance is known as the Bismarck model as it follows Bismarck's design of contributory social insurance programmes for the German empire in the late nineteenth century. See ibid.

this sense can be compatible with various forms of political economy including welfare state capitalism, property owning democracy and market socialism.

Justifying social insurance through the lens of just social risk imposition contrasts with other existing accounts or popular views of social insurance in the following respects. First, the imposition view of social risk implies that the coverage of social insurance should be extended to include risks that are not covered by conventional social insurance schemes. Conventional social insurance usually protects citizens against the risk of work accident, illness, unemployment, and dependency in old age. In addition to these risks, people living in a modern society may also be exposed to other risks that can be regarded as being (at least partly) collectively imposed. This may justify social insurance coverage for such risks even when everyone is not equally exposed or when the relevant insurance policies can be offered by the private market. For example, it may be argued from the perspective of social risk imposition that part of the burden of the risk of natural disasters (such as flooding, hurricane, and forest fires) should be fairly shared among all members of a society rather than only those living in high-risk areas. Another example is the risk of climate change. While it is impossible to be covered by insurance in its strict, risk-pooling sense, justice may require risk-sharing institutions be designed in a way that spreads its burden across generations, including those generations that are not immediately exposed to the risk.

Second, the benefit principle embodied in the demand for fair risk sharing also provides a reason why social insurance is more likely to realise the justice of social risk imposition than mandatory participation in the (regulated) private insurance market. Recall that, to allocate the burdens of a scheme of social risk imposition, fairness requires that those who benefit more from the imposition of the risk also bear a larger share of its cost. The specific demand of the benefit principle will depend on the background principles of distributive justice the society subscribes to. For example, it may require equal net benefit if the society endorses an egalitarian view of justice. But, as long as we endorse a version of the benefit principle, this would justify compulsory risk sharing by all members of the society in a form that is different from private actuarial insurance because the pricing of the latter needs to be guided by the actuarial

principle, which allocates burden according to the size of the risk each insurance buyer carries and requires those with higher risk to pay a larger premium.

This provides a strong case for fair risk sharing through universal social insurance programmes because the state could design such programmes in a way that deviates from actuarial insurance and allocate the burden of compensation according to the benefit principle. It may be suggested that sometimes fair risk sharing could also be achieved by mandatory participation in the private insurance market. The state could regulate the pricing of private insurance so as to prohibit discrimination based on gender, age, genetic information, and so on. However, regulations are often not sufficient for the private market to replicate the benefit principle because, even when risk classifications are prohibited, it is still difficult to require private insurers to price their policies based on equal net benefit or proportionate benefit generated by the scheme of social risk imposition.

Another reason why state-mandated social insurance is more likely to fulfil the demand for fair risk sharing is that many cases of social risk imposition involve intergenerational impact and hence demand fair risk sharing across generations. Because, as discussed in Chapter 5, contemporaries' collective actions and decisions will affect what kind of risks later generations are subject to, the resulting burden should be fairly allocated across generations. Moreover, if some generations fail to fulfil the demand to take reasonable precautions, they should also be held responsible for a share of the resulting burden that is larger than what they otherwise need to bear. The private market is usually unable to achieve, especially over the long-term, such cross-generational transfers based on a principle of fairness. The state, on the other hand, can design the funding structure of social insurance programmes in a way that dynamically adjusts to the demand of fair risk sharing and it is also more likely to have the capacity to sustain such institutions over a long-time horizon.

To illustrate how my account may shed light on the just design of social insurance schemes, I will use a case study on pension schemes. Before I move on to the case, there is one final point to note. While a type of social insurance is usually designed to cover one particular social risk (for example, a pension is to cover the risk of outliving

one's saving after retirement), how it should fairly allocate its burden may or may not take into account considerations other than the just imposition of this particular risk. The difference here mirrors the contrast that Caney draws between the "Method of Isolation" and the "Method of Integration" when he discusses how to justly allocate climate responsibilities.[11] According to the Method of Isolation, different considerations of justice should be dealt with in isolation.[12] Pension schemes, according to this view, should be concerned solely with how the risk of dependency in old age should be fairly shared. On the other hand, the Method of Integration implies that we may determine the allocation of the burden of paying for old age in conjunction with other considerations of justice.

It is not uncommon to invoke the integration view when considering the justice of insurance pricing. For example, in the case of life insurance, it may be suggested that, while it is prima facie unjust to price premiums based on biological sex, it seems less problematic in the real world because the fact that women are charged less due to their longer life expectancy may act to counterbalance the other disadvantages they systematically suffer in the society. Or, more generally, it is sometimes suggested that progressive pricing for social insurance schemes[13] is not motivated by concerns of fair risk sharing but serves to realise the goal of redistributing from the rich to the poor. According to this view, social insurance is used as a tool to integrate other redistributive goals of justice, probably because it is more politically acceptable than progressive taxation.

What I will adopt in the case study, on the other hand, is a different version of the integrated method. For some cases of social risk imposition (such as the risk of climate change), there does not exist a straightforward mechanism to fairly allocate the relevant burdens. I will suggest that such burdens may be fairly allocated across generations through properly designed social insurance schemes.

---

[11] S. Caney, "Just Emissions", *Philosophy and Public Affairs* 40 (2012), 255-300.

[12] Ibid., p. 259.

[13] That is, schemes that adopt a progressive benefit formula which provides higher returns to the first dollars of a person's contribution and lower returns to the last dollars. See C. E. Steuerle, A. Carasso, and L. Cohen, "How Progressive is Social Security and Why?", *Straight Talk on Social Security and Retirement Policy, Urban Institute* 37 (2004).

A Case Study: The Just Design of Pension Schemes

Many countries have a public pension scheme that pays out retirement benefits either universally to all citizens who have reached the retirement age or to those retirees who have made the required contributions during their working years. A public pension may be justified for different reasons. For example, if justice requires the provision of a decent social minimum to all members of the society, then this demand may be partially fulfilled through a universal pension which pays an amount of benefit that is sufficient for retirees to sustain a minimally decent life (though a public pension is not the only way of discharging this duty—the minimum could also be provided through, for example, a universal basic income). On the other hand, paternalistic reasons may justify contribution-based pension schemes. It may be argued that individuals often have cognitive biases when it comes to saving for retirement or adequately covering the risk of outliving their saving. Such potential behavioural failures may be corrected through mandatory participation or default participation with an opt-out option in a public pension. In such cases, pension benefits would be linked to the level of contributions made by each participant.

In addition to these arguments, the demand of fair risk sharing implied by my account may also justify a universal pension that goes beyond the minimum. The risk involved in having to provide for one's own retirement has a natural cause—most people living to old age will experience a stage of dependency where they are no longer able to work and to support themselves, in which case they need some form of resource and service provision to live a sufficient and decent life. But the retirement risk in its current form is at least partly imposed through the choice of economic systems. Consider an alternative social arrangement in which there is universal old age care or universal income with affordable and high-quality care services. Or consider a society without a defined retirement age, giving older people a variety of options ranging from retirement to doing meaningful work that is fitting to their needs and conditions. In such a society, the risk of not having saved adequately for retirement during one's income-earning years would have much less significant impact on people's lives.

Moreover, under the current social arrangement, the level of the risk involved in providing for one's old age is also impacted by various forms of intergenerational risk imposition. For example, monetary and fiscal policies may affect the economic conditions faced by several generations. If an unlucky generation experiences economic recession during their working years and inflation during their retirement due to policy decisions made by other generations, they need to deal with an increased risk of not having enough saving for old age compared to a situation where they live in a relatively stable economy. Since the retirement risk is a result of various cases of collective risk imposition that call for compensation, it may require a risk-sharing mechanism across generations such as a universal pension to fairly allocate the burden of paying for such compensation.

The demand of fair risk sharing may not only provide a justification for universal public pensions but also shed light on how such a pension should be funded, including who should pay for retirement benefits and how much benefit a retiree should be entitled to. Generally speaking, there are two approaches to funding a public pension.[14] According to the first approach, a public pension scheme can be fully funded by a generation itself. That is, individuals belonging to the same birth cohort save part of their income during the working years, which would accumulate in value over time due to investment and be paid out later as retirement benefits to themselves. On the other hand, a public pension scheme can adopt the so-called pay-as-you-go (PAYG) approach. In a PAYG pension, the working population make contributions to the pension which are then immediately paid out to the retired generations. Both types of pensions include a mechanism of risk sharing—it pools the risk of outliving one's saving by transferring benefits from those who live shorter than average to those who live longer. But a PAYG pension differs from the fully funded type in that it can further accommodate risk sharing across generations. In determining what proportion of the working generations' income should be used to pay for retirement benefits, a PAYG pension may be designed in a way that dynamically adjusts how additional burdens of pension payments are shared between adjacent generations.

---

[14] For a detailed explanation, see N. Barr and P. Diamond, *Reforming Pensions: Principles and Policy Choices* (Oxford: Oxford University Press, 2008), ch. 3.

Before I illustrate how different pension schemes imply different ways of allocating the burden of risk sharing, it should be noted that how much retirement benefits a retiree should receive is not entirely a question about fair risk sharing. As argued by Daniels, what proportion of income should be saved or be used to pay for retirement is, first of all, a problem of distribution among different age groups.[15] This age-group problem should be distinguished from fair distribution between generations as birth cohorts—while a person only belongs to a certain birth cohort, he will grow old over time and live through each age group (until he dies).[16] Therefore, the question in what way should resources be distributed between different age groups can be transformed into the intrapersonal question how a person should reasonably allocate resources between life stages. Daniels argues that a reasonably prudent allocation for an average individual should ensure that he always has a basic share of goods at each stage of his life and that he can always keep his options open. Such a prudent plan of resource allocation, which can be reasonably justified to everyone, can then guide the amount of saving required by a public pension.

I will assume that we accept Daniels' account for the sake of argument[17] because the question I am concerned with here is a different one—suppose we agree about the percentage of lifetime production or income that should be used for retirement, is there any reason for us to prefer a fully-funded pension, which only facilitates risk sharing within the same birth cohort, or a PAYG pension, which may further incorporate a mechanism for intergenerational risk sharing, to achieve this allocation?

It seems that, even setting aside the concern of just (intergenerational) risk imposition, there is a prima facie reason to prefer the PAYG arrangement due to the existence of economic cycles. In a stable economy, a fully funded pension would work well to allow a generation to collectively implement their desired age-group distribution, but an

---

[15] N. Daniels, *Am I My Parents' Keeper? An Essay on Justice Between the Young and the Old* (Oxford: Oxford University Press, 1988), p. 11.

[16] Ibid., p. 40, p. 117.

[17] For a critique of Daniels' account, see J. Bidadanure, *Justice Across Ages: Treating Young and Old as Equals* (Oxford: Oxford University Press, 2021), ch. 2. For example, Bidadanure suggests that, in addition to the concern of securing sufficiency over one's life time, there may also be reasons of efficiency to allocate more resources to earlier life stages because doing so may maximise a person's chance of living a good whole life.

economy that remains stable over a cohort's lifetime may be impossible to achieve because, even if a generation are not subject to risk imposition by other generations, it is very likely that they still cannot have full control of their (macro)economic conditions and are at risk of becoming an unlucky generation (when, for example, they experience recession during their working years and subsequently high inflation during their retirement). Due to such possibilities, it would be beneficial for a generation to be able to smooth out the potential impact of economic cycles by sharing risks with other generations.[18] This kind of intergenerational risk sharing can be facilitated by a PAYG pension whose rates of contributions and benefits are adjustable according to some predetermined rules. For example, during a period of high inflation, the working generations may be required to contribute more to the scheme, which then allows the pension to increase the retirement benefits it pays out so as to partly cover the increased living costs for the retirees.

However, the benefit of smoothing out economic cycles does not seem decisive in justifying a PAYG pension over a fully funded one. A generation may prefer a funded pension so as to have better control over their own economic fate. It is also possible for a generation with a fully funded pension to hedge against the risk of inflation by investing their pension contributions properly.[19] Moreover, while a PAYG pension has the potential of spreading out economic shocks, it is nevertheless vulnerable to a different problem. Many developed countries are seeing the phenomenon of population ageing. That is, due to longer life expectancy and a lower birth rate, there is an increase in the ratio of the retired population to the working population. Since the trend is expected to continue into the near future,[20] adopting a PAYG pension means that there will be an increasing burden per person for retirement benefit provision over time.

Upon whom this additional burden will fall would depend on the specific funding rules of the PAYG pension. For example, if the scheme is pay-as-you-go with a fixed

---

[18] E. Schokkaert and P. Van Parijs, "Social Justice and the Reform of Europe's Pension Systems", *Journal of European Social Policy* 13 (2003), 245-79, p. 249.

[19] For example, investing the contributions in a well-diversified equity portfolio may help to guard against the risk of inflation over the long-term.

[20] J. Myles, "A New Social Contract for the Elderly?", in *Why We Need a New Welfare State*, ed. G. Esping-Andersen (Oxford: Oxford University Press, 2002), p. 131.

contribution rate, then the additional burden caused by population ageing will be entirely borne by the retired generations in the form of reduced retirement benefits. Or, if it guarantees the rate of benefits, then the additional burden of contribution will be fully borne by the younger generation. Alternatively, generations may want to spread this risk in a way that is similar to how they share the impact of economic cycles, in which case the additional burden will not fall entirely on one particular cohort. This solution may appeal especially to those generations experiencing a transition from high to low population growth.

These alternatives then bring out an additional question if a PAYG pension instead of a fully-funded one is adopted—what rule should a PAYG pension follow to allocate the increasing burden caused by population ageing? It may be suggested that, without additional concerns of justice, it seems fair to have adjacent generations share the additional burden (compared to the amount implied from the endorsed age-group distribution) equally. To achieve such an equal allocation, a PAYG pension could adopt the so-called Musgrave rule. Musgrave proposes to maintain a constant ratio between the per capita income of retirees to the per capita net earnings of the active working population.[21] In so doing, the increased burden of pension funding would be equally allocated between overlapping generations.

However, so far the analysis has not taken into consideration the issue of just intergenerational risk imposition. As I have argued in Chapter 5, both economic policies and procreational choices collectively made by a generation may impose various risks on other generations, in which case the demand of fair risk sharing may justify the PAYG pension as a means to realise fair risk sharing between overlapping birth cohorts. PAYG funding justified in this way may require a rule of burden allocation that deviates from the Musgrave rule as it would account for other factors such as who have benefited from the imposition of the risk and whether reasonable precautions have been taken.

To illustrate, let me first consider economic policies. Oftentimes extreme macroeconomic events such as hyperinflation and great recessions are not random

---

[21] R. Musgrave, *Public Finance in a Democratic Society - Volume II* (Hemel Hempstead: Wheatsheaf Books, 1986), p. 116.

events. Rather, they are caused by a generation or a few generations' choice of monetary and fiscal policies.[22] If a generation suffers the negative impact of policies which they are not responsible for making, it may be said that the economic risk is imposed on them by other generations and that they should be compensated should the risk materialise. While there is no particular intergenerational risk sharing scheme to cover the risk of economic cycles, compensation may be paid by one generation to another by adjusting the contribution or payment rules of a PAYG pension. For example, if the younger generations experience a financial crisis as a result of risk imposition by the older generations, they may be compensated in the form of a lower rate of pension contributions (compared to the rate of equal sharing), which would correspondingly reduce pension benefits paid out to the older generations.

But the argument that the younger generations should be compensated does not mean that the burden of paying for such compensation should be fully borne by the generation(s) responsible for imposing the risk. Based on my account of social risk imposition, how to fairly allocate the burden of risk materialisation depends on whether the relevant risk is justly imposed as well as who have benefited from its imposition. First, the reasonable precaution principle requires a generation to make sound economic policies when such policies impose economic risks on other generations (as well as on themselves). For example, monetary policies should be designed with the goal of avoiding high inflation; regulations should be made to promote the stability and integrity of the financial market. If a financial crisis is caused by a generation's failure to regulate the market properly, then they may be considered to have violated the demand of reasonable precaution when imposing the risk and they should, as a result, be allocated a larger share of the burden than what they otherwise need to bear. On the other hand, if a generation has benefited from the imprudent economic policies made by another generation, then, even though they have not violated the principle of reasonable precaution themselves, they should also bear a share of the burden for compensating the victims based on reasons of fairness.

---

[22] Note that my analysis assumes a closed economy. In reality such events may also be caused by international affairs that are beyond their responsibility. While I will not consider the case of international risk imposition in details, it presumably implies a demand of fair international risk sharing. I will touch on this point again in the concluding remarks of this chapter.

Besides the design of economic policies, demographic changes may also impose risks on future generations. However, unlike economic cycles, it is more ambiguous whether a baby boom or population ageing may be regarded as cases of intergenerational risk imposition and, if so, what their implication would be for fair pension funding. To illustrate, let us consider the case of population ageing. Population ageing is a result of several factors and not all of them involve intergenerational risk imposition. For example, one cause of population ageing is increased life expectancy as a result of medical and scientific advance. The fact that people are now living longer does not seem to raise issues of just risk imposition. In fact, how to deal with the increased burden of providing for old age due to longer life expectancy is a question that can be transformed into an intrapersonal question that is similar to Daniels' account of age-group distribution. For individuals in a given generation, if they expect themselves to live longer than their predecessors, they would reasonably adjust their allocation of resources for different stages of life. Not only are they likely to save a larger proportion of their income for old age compared to their predecessors, but they may also reasonably decide to work until a later age in order to save more for retirement. To apply such adjustment to a public pension (either funded or PAYG), it would mean covering the additional burden introduced by greater longevity by raising the retirement age[23] and/or increasing the percentage of income as pension contributions.

However, population ageing is not just a result of greater longevity. Its other cause—a lower fertility rate—may have more ambiguous implications for just distribution between generations as birth cohorts (rather than age groups). On the one hand, if a generation collectively have fewer children, then they will create a greater (per capita) burden for their children's generation to provide for pension benefits if a PAYG pension with fixed rates of benefit is adopted. In fact, one may suggest that, since such a generation has saved on child rearing by having fewer children, they should also bear the associated cost by receiving less pension benefits.[24] Or, some may question the legitimacy of holding a generation collectively responsible for their procreational

[23] Schokkaert and Van Parijs, "Social Justice and the Reform of Europe's Pension Systems", p. 252.

[24] H. W. Sinn, "Why a Funded Pension System is Useful and Why It is Not Useful", *International Tax and Public Finance* 7 (2000), 383-410, p. 405.

choices but suggest that childless individuals and families should be required to compensate others through more pension contributions.[25] Both views are pro-natalist because they focus on the negative impact of reduced fertility on the sustainability of pension schemes (or, more generally, the welfare state) that are funded on the PAYG basis.[26]

On the other hand, however, a smaller population also means that the younger generations would have a larger amount of available resources per person. Moreover, as mentioned in Chapter 5, lower population growth may also reduce environmental risks for future generations. According to some, population growth is an important source of damage to the climate and may also accelerate natural resource depletion. If this is the case, then failing to achieve a fertility rate that is low enough for environmental sustainability would effectively mean that the current generations are imposing (additional) environmental risks on future generations. It may be justifiable to impose this risk if maintaining a higher fertility rate is presumptively beneficial for both the current and the future generations, probably because it is necessary for securing a sufficient level of economic production or for being able to establish and pass down just institutions, or because human beings have a fundamental interest in having (many) children, so that the value of retaining this option in their lives outweighs the value of those options that involve a safe and well-preserved environment. But such arguments, even if true, should be constrained by the reasonable precaution principle. For example, if the necessary level of economic production can be achieved with a reduced environment risk on future generations (for example, through automation rather than having a larger population), then maintaining a higher level of fertility rate would no longer represent a case of just risk imposition.

The design of pension schemes may also reflect these considerations about the justice of environmental risk imposition on future generations. This is not to say that pension schemes *must* serve to facilitate the fair sharing of other risks. But it may be a good idea

---

[25] Myles, "A New Social Contract for the Elderly?", pp. 265-66.

[26] Again, my analysis assumes a closed society. In an open society, the age structure of its population could also be affected by its immigration policies, for which the current generations are responsible. My argument may similarly apply to this case, but note that many other considerations of (global) justice may affect rights and duties relating to immigration.

to integrate the concern of environmental risks with the design of pension schemes (as well as other PAYG institutions such as old age care) for several reasons. First, while we have other institutions to implement policies such as the preservation of the natural environment and the restriction on greenhouse gas emissions, we do not seem to have reliable mechanisms to hold a generation accountable should they fail to fulfil the demands of the reasonable precaution principle. PAYG schemes may be designed in a way that respond to such failures—for example, it may specify rules to adjust the ratio between the per capita income of retirees to the per capita net earnings of the active working population according to some quantifiable measures that assess each generation's contribution to environmental sustainability. It is more promising to use a PAYG scheme to (partially) achieve such intergenerational redistribution compared to direct resource transfers because a pension is a reciprocal institution and, as I have argued in Chapter 6, is therefore possible to sustain into the future.[27]

Second, as I have explained, there is a tension between the economic and ecological impact of our procreational policies, so we need to reflect on the urgency of both issues in order to specify what count as reasonable precautions in dealing with intergenerational risks. Schokkaert and Van Parijs suggest that the complication in such cases may justify the Musgrave rule for equal sharing as a good proximation to resolve the indeterminacy.[28] However, if we, based on theoretical reflections as well as empirical findings from disciplines such as demography and ecology, decide to weigh our concerns about economic and environmental sustainability in a certain way, then we may reasonably deviate from the Musgrave rule in order to reflect our endorsed balance between these goals. By institutionalising the trade-off through the funding rules of a PAYG pension scheme, we may in turn exert influence on people's procreational decisions. In this sense, we can also, through our choice of a pension scheme and its funding structure, encourage the alignment of individual choices with ideals of intergenerational justice.

---

[27] This is not to deny that, in the real world, it is still politically challenging to incorporate such principles into a public pension scheme.

[28] Schokkaert and Van Parijs, "Social Justice and the Reform of Europe's Pension Systems", p. 252.

Before I conclude this thesis, let me list out a few topics that may be worth exploring in future work. First, in this thesis I have argued that the imposition of social risk implies a demand of fair risk sharing among not only contemporaries but also people from different generations. I have, however, limited my discussion within a society and have not considered the case of international risk imposition and risk sharing. If members of a society have, through their policies or collective actions, imposed risks on people living in other countries, then presumably the demand of fair risk sharing requires them to also share the burden of compensating victims abroad. International risk sharing, however, can be difficult to implement in reality due to the lack of a global state or other institutions that can mandate a transnational scheme of social insurance (or other risk sharing programmes). On the other hand, such a scheme is also not unimaginable. Schokkaert and Van Parijs have, for example, proposed joint provision of a basic pension among European Union member states.[29] There are also potential benefits for different countries to cooperate in risk sharing—by diversifying its covered region or population, such a scheme could pool contagious or rare risks more effectively. If this kind of transnational institution could be established, then it may also provide the opportunity to implement fair risk sharing internationally.

Second, the application of my account of just risk imposition may go beyond those conventional risks considered in this thesis. For example, there is a growing literature on the ethics of algorithms. This literature tends to focus on the injustice of algorithmic systems that are currently implemented, which often reflect rather than correct existing biases in the data that are fed into the model. But there is also the concern that a reasonably optimised algorithmic system (optimised in the sense that it is trained with data that is refined to correct existing biases) will still produce errors because it is ultimately based on statistical inference which may be unfair to people belonging to certain groups.[30] For example, young black people may be subject to a higher risk of being investigated due to the higher crime rate of their group. This is an interesting question that may be looked at from the perspective of just risk imposition. Based on

[29] Ibid., pp. 258-59.

[30] M. Di Bello and C. O'Neil, "Profile Evidence, Fairness, and the Risks of Mistaken Convictions", *Ethics* 130 (2020), 147-78.

my account, it seems that imposing the risk of being misidentified by algorithm may be justifiable if measures of reasonable precaution (such as procedures for requesting human review, constant refinement of the model to reduce errors, transparency in the principles used to design the system) are taken and the burden of higher risk exposure suffered by disadvantaged groups could be fairly shared by other members of the society in certain ways. On the other hand, the availability of algorithm-based decision support systems may also impact what is required by the reasonable precaution principle for dealing with other social risks. For example, if employing the so-called "predictive policing" technology can, as a matter of empirical fact, reduce the crime rate, does the reasonable precaution principle imply a demand to use such technology to reduce the risk of crime that citizens are subject to? It seems that the answer would depend on the potential of the technology in improving accuracy and preventing biases compared to human-conducted predictive policing, which has always existed.

Finally, it may also be illuminating to examine the relationship between my account and the politics of social insurance. On the one hand, democratic deliberation about the just design of social insurance schemes could potentially shape social norms (for example, what counts as reasonable precaution, what level of risk is justifiable to all). On the other hand, empirical surveys about people's attitude toward different social insurance schemes may help to identify sources of solidarity in sustaining cooperation in risk sharing and therefore shed light on the feasibility of applying my account to the design of real world policies. These issues could be examined in further research.

Bibliography

Abode, R., A. Borer and D. Enoch, "Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith", *The Journal of Philosophy* 105 (2008), 259-72.

Ahlstrom-Vij, K., *Epistemic Paternalism: A Defence* (Basingstoke: Palgrave Macmillan, 2013).

Akerlof, G. A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *The Quarterly Journal of Economics* 84 (1970), 488-500.

Anderson, E., "What is the Point of Equality?", *Ethics* 109 (1999), 287-337.

Anderson, E., "Thomas Paine's 'Agrarian Justice' and the Origins of Social Insurance", in *Ten Neglected Classics of Philosophy*, ed. E. Schliesser (Oxford: Oxford University Press, 2016).

Anomaly, J., "Public Goods and Government Action", *Politics, Philosophy and Economics* 14 (2015), 109-28.

Arneson, R., "Equality and Equality of Opportunity for Welfare", in *Equality: Selected Readings*, ed. L. Pojman and R. Westmoreland (New York: Oxford University Press, 1997).

Arrhenius, G., "Mutual Advantage Contractarianism and Future Generations", *Theoria*, 65 (1999), 25-35.

Askari, H. Z. Iqbal and N. Krichene, *Risk Sharing in Finance: The Islamic Finance Alternative* (Singapore: John Wiley & Sons Asia Pte. Ltd., 2012).

Axelrod, R., *The Evolution of Cooperation* (New York: Basic Books, 1984).

Baderin, A. and L. Barnes, "Risk and Self-Respect", *British Journal of Political Science* 50 (2018), 1-19.

Barr, N., *The Welfare State as Piggy Bank: Information, Risk, Uncertainty, and the Role of the State* (Oxford: Oxford University Press, 2001).

Barr, N., *Economics of the Welfare State* (Oxford: Oxford University Press, 2012).

Barr, N. and P. Diamond, *Reforming Pensions: Principles and Policy Choices* (Oxford: Oxford University Press, 2008).

Barry, B., "Circumstances of Justice and Future Generations", in *Obligations to Future Generations*, eds. R. I. Sikora and B. Barry (Philadelphia: Temple University Press, 1978).

Barry, B., "Justice as Reciprocity", in *Liberty and Justice: Essays in Political Theory* 2 (Oxford: Clarendon Press, 1991).

Barry, B., *Justice as Impartiality* (Oxford: Clarendon Press, 1995).

Barry, B., "Sustainability and Intergenerational Justice", in *Fairness and Futurity: Essays on Environmental Sustainability and Social Justice*, ed. A. Dobson (Oxford: Oxford University Press, 1999).

Becker, L., "Reciprocity, Justice, and Disability", *Ethics* 116 (2005).

Bidadanure, J., *Justice Across Ages: Treating Young and Old as Equals* (Oxford: Oxford University Press, 2021).

Bowles, S. and H. Gintis, "Homo Reciprocans", *Nature* 45 (2002), 125-28.

Buchak, L., *Risk and Rationality* (Oxford: Oxford University Press, 2013).

Buchak, L., "Taking Risks behind the Veil of Ignorance", *Ethics* 127 (2017), 610-44.

Buchanan, A., "Justice as Reciprocity versus Subject-Centered Justice", *Philosophy and Public Affairs* 19 (1990), 227-52.

Caney, S., "Just Emissions", *Philosophy and Public Affairs* 40 (2012), 255-300.

Carter, I., *A Measure of Freedom* (Oxford: Oxford University Press, 1999).

Ci, J., *The Two Faces of Justice* (Cambridge, Massachusetts: Harvard University Press, 2006).

Cohen, G. A., "On the Currency of Egalitarian Justice", *Ethics* 99 (1989), 906-44.

Crisp, R., "Equality, Priority, and Compassion", *Ethics* 113 (2003), 745-63.

Daniels, N., *Am I My Parents' Keeper? An Essay on Justice Between the Young and the Old* (Oxford: Oxford University Press, 1988).

Daniels, N., *Just Health: Meeting Health Needs Fairly* (Cambridge: Cambridge University Press, 2008).

Daston, L., *Classical Probability in the Enlightenment* (Princeton: Princeton University Press, 1988).

Di Bello, M. and C. O'Neil, "Profile Evidence, Fairness, and the Risks of Mistaken Convictions", *Ethics* 130 (2020), 147-78.

Dworkin, G., "Paternalism", *Monist* 56 (1972), 64-84.

Dworkin, G., "Is More Choice Better Than Less?", *Midwest Studies in Philosophy* 7 (1982), 47-61.

Dworkin, G., "The Concept of Autonomy", in *The Inner Citadel: Essays on Individual Autonomy*, ed. J. Christman (Oxford: Oxford University Press, 1989).

Dworkin, R., “What is Equality? Part 2: Equality of Resources”, *Philosophy and Public Affairs* 10 (1981), 283-345.

Dworkin, R., *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, Massachusetts: Harvard University Press, 2000).

Estlund, D., “Human Nature and the Limits (if any) of Political Philosophy”, *Philosophy and Public Affairs* 39 (2011), 207-237.

Estlund, D., “Reply to Wiens”, *European Journal of Political Theory* 15 (2016), 353-362.

Fehr, E. and S. Gächter, “Altruistic Punishment in Humans”, *Nature* 45 (2002), 137-40.

Ferretti, M., “Risk Imposition and Freedom”, *Politics, Philosophy and Economics* 15 (2016), 261-79.

Fleurbaey, M., *Fairness, Responsibility, and Welfare* (Oxford: Oxford University Press, 2008).

Fleurbaey, M. and A. Voorhoeve, “Decide As You Would with Full Information!”, in *Inequalities in Health: Concepts, Measures, and Ethics*, eds. N. Eyal et al. (New York: Oxford University Press, 2013).

Frankfurt, H., “Freedom of the Will and the Concept of a Person”, in *The Inner Citadel: Essays on Individual Autonomy*, ed. J. Christman (Oxford: Oxford University Press, 1989).

Frick, J., “Contractualism and Social Risk”, *Philosophy and Public Affairs* 43 (2015), 175-223.

Fried, B., “The Case for a Progressive Benefits Tax”, in *Taxation: Philosophical Perspectives*, eds. M. O’Neill and S. Orr (Oxford: Oxford University Press, 2018).

Friedman, R., *Probable Justice: Risk, Insurance, and the Welfare State* (Chicago: University of Chicago Press, 2020).

Gardiner, S., “A Core Precautionary Principle”, *Journal of Political Philosophy* 14 (2006), 33-60.

Gauthier, D., *Morals by Agreement* (Oxford: Clarendon Press, 1986).

Goodin, R., *Reasons for Welfare: The Political Theory of the Welfare State* (Princeton: Princeton University Press, 1988).

Goodin, R., “Stabilizing Expectations: The Role of Earnings-Related Benefits in Social Welfare Policy”, *Ethics* 100 (1990), 530-53.

Goodin, R., *Discretionary Time: A New Measure of Freedom* (Cambridge: Cambridge University Press, 2008).

Hartzell-Nichols, L., "Precaution and Solar Radiation Management", *Ethics, Policy and Environment* 15 (2012), 158-71.

Hausman, D. and B. Welch, "Debate: To Nudge or Not to Nudge", *The Journal of Political Philosophy* 18 (2010), 123-36.

Hayenhjelm, M. and J. Wolff, "The Moral Problem of Risk Impositions: A Survey of the Literature", *European Journal of Philosophy* 20 (2012), 26-51.

Heath, J., "The Benefits of Cooperation", *Philosophy and Public Affairs* 34 (2006), 313-51.

Heath, J., "The Structure of Intergenerational Cooperation", *Philosophy and Public Affairs* 41 (2013), 31-66

Hermansson, H. and S. O. Hansson, "A Three-Party Model Tool for Ethical Risk Analysis", *Risk Management* 9 (2007), 129-44.

Honoré, T., "Responsibility and Luck: The Moral Basis of Strict Liability", in *Responsibility and Fault* (London: Hart Publishing, 1999).

Igneski, V., "Distance, Determinacy and the Duty to Aid: A Reply to Kamm", *Law and Philosophy* 20 (2001), 605-16.

Jackson, F. and M. Smith, "Absolutist Moral Theories and Uncertainty", *The Journal of Philosophy* 103 (2006), 267-83.

James, A., *Fairness in Practice: A Social Contract for A Global Economy* (Oxford: Oxford University Press, 2012).

Kamm, F., "Does Distance Matter Morally to the Duty to Rescue?", *Law and Philosophy* 19 (2000), 655-81.

Kittay, E., *Love's Labor: Essays on Women, Equality and Dependency* (London: Routledge, 1999).

Klosko, G., "Presumptive Benefit, Fairness, and Political Obligation", *Philosophy and Public Affairs* 16 (1987), 241-59.

Knight, C., "An Argument for All-Luck Egalitarianism", *Philosophy and Public Affairs* 49 (2021), 350-78.

Knight, F., *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin, 1921).

Kropotkin, P., *Mutual Aid: A Factor of Evolution* (London: William Heinemann, 1902).

Kumar, R., "Risking and Wronging", *Philosophy and Public Affairs* 43 (2015), 27-51.

Kumar, R., "Risking Future Generations", *Ethical Theory and Moral Practice* 21 (2018), 245-57.

Lenman, J., “Contractualism and Risk Imposition”, *Politics, Philosophy and Economics* 7 (2008), 99-122.

Lewens, T., “Introduction: Risk and Philosophy”, in *Risk: Philosophical Perspectives*, ed. T. Lewens (London: Routledge, 2007).

Lippert-Rasmussen, K., “Egalitarianism, Option Luck, and Responsibility”, *Ethics* 111 (2001), 548-79.

Lupton, D. and J. Tulloch, “Life Would Be Pretty Dull Without Risk: Voluntary Risk-Taking and Its Pleasures”, *Health, Risk and Society* 4 (2002), 113-24.

McCarthy, D., “Rights, Explanation, and Risks”, *Ethics* 107 (1997), 205-225.

McCormick, H., “Intergenerational Justice and the Non-reciprocity Problem”, *Political Studies* 57 (2009).

Miller, D., “The Nature and Limits of the Duty of Rescue”, *Journal of Moral Philosophy* 17 (2020), 320-41.

Miller, D. and I. Taylor, Miller, “Public Goods”, in *The Oxford Handbook of Distributive Justice*, ed. S. Olsaretti (Oxford: Oxford University Press, 2018).

Moreau, S. R., “Contractualism and Aggregation”, *Ethics* 108 (1998), 296–311.

Murphy, L. and T. Nagel, *The Myth of Ownership: Taxes and Justice* (Oxford: Oxford University Press, 2002).

Murray, C., *The Happiness of the People* (Washington, D.C.: AEI Press, 2009).

Murray, W., “Homeowners Insurance Redlining: The Inadequacy of Federal Remedies and the Future of the Property Insurance War”, *Connecticut Insurance Law Journal* 4 (1997), 735-62.

Musgrave, R., *Public Finance in a Democratic Society - Volume II* (Hemel Hempstead: Wheatsheaf Books, 1986).

Myles, J., “A New Social Contract for the Elderly?”, in *Why We Need a New Welfare State*, ed. G. Esping-Andersen (Oxford: Oxford University Press, 2002).

Myrvold, W., *Beyond Chance and Credence: A Theory of Hybrid Probabilities* (Oxford: Oxford University Press, 2021).

Nozick, R., *Anarchy, State, and Utopia* (Oxford: Basil Blackwell, 1974).

Nussbaum, M., *Frontiers of Justice: Disability, Nationality, Species Membership* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2006).

Nys, T. and B. Engelen, “Judging Nudging: Answering the Manipulation Objection”, *Political Studies* 65 (2017), 199-214.

Oberdiek, J., “Towards a Right against Risking”, *Law and Philosophy* 28 (2009), 367-92.

Oberdiek, J., *Imposing Risk: A Normative Framework* (Oxford: Oxford University Press, 2017).

Olsaretti, S., “Responsibility and the Consequences of Choice”, *Proceedings of the Aristotelian Society* 109 (2009), 165-88.

Olsaretti, S., "Children as Public Goods?”, *Philosophy and Public Affairs* 41 (2013), 226-58.

Olsaretti, S., “Scanlon on Responsibility and the Value of Choice”, *Journal of Moral Philosophy* 10 (2013), 465-83.

Ostrom, E., *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge: Cambridge University Press, 1990).

Parfit, D., *Reasons and Persons* (Oxford: Clarendon Press, 1984).

Parker, A. and P. J. Irvine, “The Risk of Termination Shock From Solar Geoengineering”, *Earth's Future* 6 (2018), 456-67.

Perry, S., “Libertarianism, Entitlement, and Responsibility”, *Philosophy and Public Affairs* 26 (1997), 351-96.

Perry, S., “Risk, Harm, Interests, and Rights”, in *Risk: Philosophical Perspectives*, ed. T. Lewens (London: Routledge, 2007).

Quong, J., *Liberalism Without Perfection* (Oxford: Oxford University Press, 2011).

Rawls, J., *Political Liberalism* (New York: Columbia University Press, 1993).

Rawls, J., *A Theory of Justice (Revised Edition)* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999).

Rawls, J., *Justice as Fairness: A Restatement* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 2001).

Raz, J., *The Morality of Freedom* (Oxford: Clarendon Press, 1986).

Raz, J., “Being in the World”, *Ratio* 23 (2010), 433-52.

Roemer, J., “A Pragmatic Theory of Responsibility for the Egalitarian Planner”, *Philosophy and Public Affairs* 22 (1993), 146-66.

Scanlon, T. M., *What We Owe to Each Other* (Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1998).

Scanlon, T. M., “Forms and Conditions of Responsibility”, in *The Nature of Moral Responsibility: New Essays*, eds. R. Clarke et al. (New York: Oxford University Press, 2005).

Scanlon, T. M., “Reply to Zofia Stemplowska”, *Journal of Moral Philosophy* 10 (2013), 508-14.

Scheffler, S., “The Afterlife” (Part I and II), in *Death and the Afterlife*, ed. N. Kolodny (Oxford: Oxford University Press, 2016).

Scheffler, S., *Why Worry About Future Generations?* (Oxford: Oxford University Press, 2018).

Schmidt, A., “Getting Real on Rationality—Behavioral Science, Nudging, and Public Policy”, *Ethics* 129 (2019), 511-43.

Schokkaert, E. and P. Van Parijs, “Social Justice and the Reform of Europe’s Pension Systems”, *Journal of European Social Policy* 13 (2003), 245-79.

Sher, G., *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997).

Sinn, H. W., “Why a Funded Pension System is Useful and Why It is Not Useful”, *International Tax and Public Finance* 7 (2000), 383-410.

Steel, D., *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy* (Cambridge: Cambridge University Press, 2015).

Steiner, H., “How Free: Computing Personal Liberty”, *Royal Institute of Philosophy Supplement* 15 (1983), 73-89.

Steiner, H., *An Essay on Rights* (Oxford: Blackwell, 1994).

Stemplowska, Z., “Making Justice Sensitive to Responsibility”, *Political Studies* 57 (2009), 237-59.

Stemplowska, Z., “Harmful Choices: Scanlon and Voorhoeve on Substantive Responsibility”, *Journal of Moral Philosophy* 10 (2013), 488-507.

Steuerle, C. E., A. Carasso, and L. Cohen, “How Progressive is Social Security and Why?”, *Straight Talk on Social Security and Retirement Policy, Urban Institute* 37 (2004).

Stone, D., “Beyond Moral Hazard: Insurance as Moral Opportunity”, in *Embracing Risk: The Changing Culture of Insurance and Responsibility*, eds. T. Baker and J. Simon (Chicago: University of Chicago Press, 2002).

Sugden, R., *The Economics of Rights, Co-operation and Welfare* (Oxford: Basil Blackwell, 1986).

Sunstein, C., *Laws of Fear: Beyond the Precautionary Principle* (Cambridge: Cambridge University Press, 2005).

Sunstein, C., *Why Nudge? The Politics of Libertarian Paternalism* (New Haven: Yale University Press, 2014).

Thaler, R. and C. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (New Haven: Yale University Press, 2008).

Thoma, J., "Time for Caution", *Philosophy and Public Affairs* 50 (2021), 50-89.

Thomas, A., *Republic of Equals: Predistribution and Property-Owning Democracy* (New York: Oxford University Press, 2017).

Tomasi, J., *Free Market Fairness* (Princeton: Princeton University Press, 2012).

Van Parijs, P., *Real Freedom for All* (Oxford: Clarendon Press, 1995).

Voorhoeve, A., "Scanlon on Substantive Responsibility", *Journal of Political Philosophy* 16 (2008), 184-200.

Voss, T., "Game-Theoretical Perspectives on the Emergence of Social Norms", in *Social Norms*, ed. M. Hechter and K. D. Opp (New York: Russell Sage Foundation, 2001).

Waldron, J., "John Rawls and the Social Minimum", in *Liberal Rights: Collected Papers 1981-1991* (Cambridge: Cambridge University Press, 1993).

Waldron, J., "It's All for Your Own Good", *The New York Review of Books* (9 October 2014).

White, S., *The Civic Minimum: On the Rights and Obligations of Economic Citizenship* (Oxford: Oxford University Press, 2003).

Wiens, D., "Motivational Limitations on the Demands of Justice", *European Journal of Political Theory* 15 (2016), 333-52.

Wilkinson, T. M., "Nudging and Manipulation", *Political Studies* 61 (2013), 341-55.

Williams, A., "Liberty, Liability and Contractualism", in *Egalitarianism: New Essays on the Nature and Value of Equality*, eds. N. Holtug and K. Lippert-Rasmussen (Oxford: Oxford University Press, 2007).

Wolff, J., "Fear, Blame, Shame, and the Regulation of Public Safety", *Economics and Philosophy* 22 (2006), 409-27.

Wolff, J., "What is the Value of Preventing a Fatality?", in *Risk: Philosophical Perspectives*, ed. T. Lewens (London: Routledge, 2007).

Wolff, J., "Five Types of Risky Situation", *Law, Innovation and Technology* 2 (2010), 151-63.

Wolff, J., "The Precautionary Attitude: Asking Preliminary Questions", *Hastings Center Report* 44 (2014).

Wolff, J. and A. De-Shalit, *Disadvantage* (Oxford: Oxford University Press, 2007).