

PRUDENCE AND JUSTICE

DONALD W. BRUCKNER

Penn State University

Whereas principles of *justice* adjudicate *interpersonal* conflicts, principles of *prudence* adjudicate *intrapersonal* conflicts – i.e., conflicts between the preferences an individual has now and the preferences he will have later. On a contractarian approach, principles of justice can be theoretically grounded in a hypothetical agreement in an appropriately specified *pre-moral* situation in which those persons with conflicting claims have representatives pushing for their claims. Similarly, I claim, principles of prudence can be grounded in a hypothetical agreement in an appropriately specified *pre-prudential* situation in which those temporal parts of a person with conflicting claims have representatives as advocates of their claims. During the course of developing the prudential contractarian methodology, I consider a dispute between those who would see principles of justice as the outcome of a *choice* (e.g., Rawls) and others (e.g., Gauthier) who argue for viewing principles of justice as the outcome of a *bargain*. I contend that the reasons I adduce in favor of viewing principles of prudence as the outcome of a choice weigh equally in favor of viewing principles of justice as the outcome of a bargain.

INTRODUCTION

Agents occasionally disagree with each other: one person thinks one course of action ought to be pursued, while another person supports a very different course of action. A common cause of such disagreement is simply that what I judge good or valuable is not what you judge good or valuable. As Hobbes says,

whatsoever is the object of any man's appetite or desire that is it which he for his part calleth *good*; and the object of his hate and aversion *evil* . . . For these words of good [and] evil . . . are ever used with relation to the person that useth them, there being nothing simply and absolutely so, nor any common

This paper reviews and extends some of the material from my (2001). Farid Abdel-Nour, Ben Eggleston, Umut Ergun, and David Gauthier provided helpful comments on an earlier draft of this paper, for which I am very grateful. I am also indebted to two anonymous referees and Luc Bovens for some very helpful suggestions.

rule of good and evil to be taken from the nature of the objects themselves.
(Hobbes 1651: ch. 6)

So the rich person, judging it good, supports one social policy because of its effects, while the poor person, judging *it* good, supports a *different* policy because of *its* effects. I take the question of how to settle such interpersonal conflicts as the central question to be answered by a theory of justice.

Not only can agents disagree with *each other*, they can also disagree with *themselves* from one time to another.

Good and *evil* are names that signify our appetites and aversions, which in different tempers, customs, and doctrines of men are different; and divers men differ . . . Nay, the same man in divers times differs from himself, and one time praiseth (that is, calleth good) what another time he dispraiseth (and calleth evil). (1651: ch. 15)

So, at one time, you might think some course of action is best, only to come to judge at a later time that a different course of action is, or would have been, best. A common cause of this “intertemporal tussle” in which we find ourselves¹ is time-discounting, according to which the utility or disutility of an outcome removed from the present is discounted in proportion to how distant it is. Accordingly, tonight I might disagree with my tomorrow’s self whether I should stay up and watch one more movie. Although I know I have to get up early in the morning, tonight I discount the disutility I will cause myself tomorrow if I do stay up, and I reach the judgment that it is better to stay up than not. When tomorrow comes, I discount the previous day’s utility, judging that it would be better not to have stayed up. What I judged good last night is not what I judge good today. I take the question of how to act in the face of such intertemporal, intrapersonal disagreements as the central question to be answered by a theory of prudence.²

So a theory of justice aims to solve *interpersonal* conflict, while a theory of prudence aims to solve *intrapersonal* conflict. In this paper, I want to investigate how far we can get from this little, banal recognition. I have two principal claims to make here. First, I shall argue that just as principles of justice can usefully be thought of as arising from agreement in an appropriately specified *pre-moral* situation in which those persons with conflicting claims have representatives pushing for their claims, principles of prudence can usefully be thought of as arising

¹ This incredibly apt phrase is due to Strotz (1955: 171).

² Surely agents can face *simultaneous* intrapersonal conflicts, as when deciding what action to take here and now when none of the possible actions will have significant effects in the future – for instance, when deciding which flavor of ice cream to order on this occasion. In such a case there is no question of intertemporal conflict. This simultaneous intrapersonal conflict is not the type of conflict that I aim to address here. I am stipulating that prudence, or at least the part of it that I intend to address, is concerned with intrapersonal conflict over time. Thanks to Umut Ergun for pointing out the need to make this clarification.

from agreement in an appropriately specified *pre-prudential* situation in which those temporal parts of a person with conflicting claims have representatives as advocates of their claims. This first piece constitutes an extension of the contractarian methodology not to moral theory but to rational choice theory itself, and is the focus of the first two sections of this paper.

Since I aim to keep the focus here on methodology in rational choice and moral theories, I pause only briefly to indicate what principle of prudence I think would be chosen in the pre-prudential situation. In the third section, I turn instead to an important objection against *using* the contractarian approach to prudence in the first place. This objection claims that the contractarian prudential methodology takes the distinctions among the conflicting temporal parts of a person far too seriously, which commits me to a queer ontology of the person. My response to this objection puts the final pieces in place so that I can make my second major claim, that once we see the proper way to apply the contractarian method to the selection of principles of prudence, we can settle – or at least gain some insight into – a controversy that divides contractarian moral theorists. This is the controversy between those contractarians such as Gauthier who think that moral principles should be thought of as the product of a bargain and those such as Rawls who think that moral principles should be thought of as the product of choice in ignorance of one's identity, conception of the good, and place in society. In the final section, I shall argue that the considerations I adduce in the second and third sections in favor of viewing prudential principles as the outcome of a choice rather than a bargain weigh equally in favor of viewing moral principles as the outcome of a bargain rather than a choice.

If I am successful in making out my two major claims, then the banal similarity I noted above between justice and prudence solving interpersonal and intrapersonal conflict, respectively, takes us a very long way not only toward a contractarian theory of prudence but in settling an important question within contractarian moral theory as well.

1. THE CIRCUMSTANCES OF PRUDENCE

Rawls follows Hume in his account of the conditions under which questions of justice arise and calls them the *circumstances of justice*.³ The circumstances of justice are composed of two conditions, one subjective, the other objective (Rawls 1971: 126–7). The subjective condition is our

³ The relevant passages are in Hume's *Treatise*, Bk. III, Pt. II, Sec. II, and in his *An Enquiry Concerning the Principles of Morals*, Sec. III, Pt. I. Rawls reintroduces the circumstances of justice in (1958: §3) and (1971: §22).

“natural temper” of selfishness or limited generosity (Hume, *Treatise*: 486). The objective condition is moderate scarcity, in which there is not enough of what people want for each person to have as much as she would like. As a result of these two conditions, competing demands for scarce societal and natural resources are made, and the adjudication of the competing demands is made by justice.

Both Rawls (127–9) and Hume (*Treatise*: 494–5; *Second Enquiry*: 183–8) stress that each of these conditions is necessary for questions of justice to arise. If resources were so abundant that no one’s desires could go unfulfilled, then there would be no need for justice to settle conflicting claims. If, on the other hand, resources were *severely* scarce then justice would “give place to the stronger motives of necessity and self-preservation” (*Second Enquiry*: 186). Concerning the degree of generosity present in oneself and one’s fellows, if one were to “fall into the society of ruffians” in which “a desperate rapaciousness prevail[s]” (187), again, one would be forced to abandon justice and “consult the dictates of self-preservation alone” (*ibid.*). On the other hand, if there were “in the human breast perfect moderation and humanity” (188), then the constraints of justice would again have no place, for they would serve no purpose. So it is only when scarcity is *moderate* (neither non-existent nor severe) and generosity is *limited* (neither perfect nor like ruffians) that the circumstances of justice obtain.

Corresponding to these subjective and objective conditions that comprise the circumstances of justice are subjective and objective conditions under which questions of prudence arise, which I shall call the *circumstances of prudence*. The subjective condition is that people put greater weight on the satisfaction of their present preferences than they do on the satisfaction of their past or future preferences. In short, people display a limited generosity toward their selves at times other than the present.⁴ The objective condition is again a condition of moderate scarcity, in which the individual does not have resources and opportunities adequate to satisfy fully all of his preferences at all times. When these two conditions obtain, the different temporal aspects of an individual make competing demands on the resources and opportunities for preference satisfaction at the individual’s disposal, and the adjudication of the competing demands is made by prudence.

As in the case of the circumstances of justice, each of these conditions is necessary. If there is no scarcity then prudence is unnecessary, and if there

⁴ At least this is true of the sort of person for whom questions of prudence arise. I refrain from wading through the vast psychological evidence for the claim that this condition obtains widely. This task is admirably performed by Ainslie (1992: esp. chs. 3 and 4), who explains a host of psychological phenomena, from addictions to pains, by appeal to temporal discounting.

is severe scarcity then prudence gives way to more immediate motives. If a person is equally concerned with his preference satisfaction at all times and so finds within himself no intertemporal conflicts, then prudence is again unnecessary, while if he is simply a rapacious ruffian with respect to himself in the past and future then prudence again gives way to other more immediate motives.⁵

When the circumstances of prudence do obtain, however, the individual who must decide what action to take here and now is presented with a practical problem. Money is a scarce resource that most people need at all stages of life. Yet a person who puts greater weight on present satisfactions might prefer to spend all of her income each month instead of saving anything for retirement. As the time for retirement approaches, however, she might find that this preference over courses of action changes. Such an agent is faced with the practical problem whether and how much to save for retirement. Opportunities are also scarce, in a certain way. For example, a graduating high school student might prefer cashing in on the opportunity now to take a job with an attractive salary rather than passing up that opportunity and spending four years in college instead. If he acts on that preference, however, he may come to prefer later that he did not take the job, for his opportunities for advancement and change of career would be greater had he gotten his degree and passed on the first opportunity. He, too, clearly faces a practical problem. Let these cases and the original example of deciding whether to stay up for another movie serve to indicate the sort of practical problem that is present in the circumstances of prudence.

Socrates recognized the practical problem in his diagnosis of the apparent phenomenon of weakness of the will as a failure to give future pleasures their full weight (*cf. Protagoras*: 356a–e). Rousseau recognized the problem in the savage whose soul “yields itself wholly to the sentiment of its present existence” and who “sells his Cotton bed in the morning and comes weeping to buy it back in the evening” (*Second Discourse*: 151). Hume also recognized the practical problem, tracing it to the greater “force and vivacity” with which present objects strike us in comparison to those that are remote in time (*Treatise*: 428). Among more contemporary thinkers, Strotz (1955), Nagel (1970), Elster (1979), Parfit (1984), Ainslie (1992), and Hampton (1998) are among those who have recognized the practical problem presented in the circumstances of prudence. I now want to cut to the chase and lay out the essentials of the application of the contractarian methodology to prudence.

⁵ Thanks to Michael Thompson for suggesting that the parallel between the circumstances of prudence and the circumstances of justice might be deeper than I had appreciated in an earlier characterization of the circumstances of prudence.

2. THE PRUDENTIAL ORIGINAL POSITION

One of the centerpieces of Rawls' (1971) contractarian method is the *original position*. This is a hypothetical meeting place of representatives of individuals who convene for the purpose of agreeing upon principles of justice. The original position is surrounded by a *veil of ignorance*, a device that blocks from the parties to the original position information about those they represent that is irrelevant from the standpoint of arguments for principles of justice. Information that is blocked includes information about "natural fortune or social circumstances" as well as information about "particular inclinations and aspirations, and persons' conceptions of their good" (18). "The aim," Rawls says, "is to rule out those principles that it would be rational to propose for acceptance . . . only if one knew certain things that are irrelevant from the standpoint of justice" (*ibid.*). Very much as Rawls employs the original position in his argument for principles of justice, I believe that a similar device, the *prudential original position*, can be used in an argument for principles of prudence. The aim of this section is to present an interpretation, or description, of the prudential original position that includes just information that is relevant from the standpoint of prudence. Other information must be blocked by a veil of ignorance.

As in the case of Rawls' original position, we settle upon the most plausible interpretation of the prudential original position through a process aiming at *reflective equilibrium*. This is a process by which we move back and forth between "generally shared and preferably weak conditions" (20) that we would like to impose on arguments for principles of prudence, and the principles that would be agreed upon in the prudential original position characterized by those conditions. If those principles fail to match our considered judgments about prudence, we must adjust either the conditions that characterize the prudential original position, or revise our considered judgments. We iterate this exercise until our description of the prudential original position "both expresses reasonable conditions and yields principles which match our considered judgments duly pruned and adjusted" (*ibid.*), in which case we have reached a state of reflective equilibrium. "It is an equilibrium because at last our principles and judgments coincide; and it is reflective since we know to what principles our judgments conform and the premises of their derivation [i.e., the conditions imposed in the original position]" (*ibid.*). In the interpretation of the prudential original position to follow, part of the discussion will be explicit reasoning toward reflective equilibrium, part will be implicitly so.⁶

⁶ The reader familiar with the literature on reflective equilibrium will notice by my applications of the reflective equilibrium method to follow that I advocate *wide*, as opposed

This section is divided into three main subsections. In §2.1, I offer a short gloss on the parties to the prudential original position and those they represent. In §2.2, I give reasons for adopting an interpretation of the prudential original position as a choice situation rather than a bargaining place. In §2.3, I give an account of what knowledge the parties have and do not have in the prudential original position, in other words, an account of the thickness of the veil of ignorance.

2.1 Representation

In the original, moral original position it is clear enough whom the parties represent. Each party represents a particular person living during a particular period of history, in particular social circumstances, with a particular set of talents, and with a particular determinate conception of the good. Of all of these particulars⁷ the parties are ignorant, but they are *not* ignorant that each represents such a concrete, particular person and has the task of securing her interests as well as possible under these conditions of ignorance.

According to the contractarian account of prudence I propose, prudence stands to intrapersonal conflict in roughly the same relation as justice stands to interpersonal conflict. So whereas the parties to the moral original position represent those whose conflicting demands must be adjudicated by justice, the parties to the prudential original position represent those whose conflicting demands must be adjudicated by prudence. As outlined in the previous section, a succession of temporally indexed and conflicting preferences constitutes the circumstances of prudence, the circumstances in which different temporal aspects of a single individual make conflicting demands as to what ought to be done by that individual. So in the case of prudence, those who are represented are not distinct individuals; they are *temporal aspects* of a single individual existing through time who is subject to temporal preference changes.

How, though, should we conceive of these temporal aspects? I shall have some more to say on this topic in §3.2, but for now we can think of each aspect as a partial image of a person at a particular time that presents a view of her that concentrates on the temporally indexed preferences she holds at that time. For tractability, as well as to remain realistic, we should suppose that each such temporal preference endures for some significant interval of time. Such a preference can last from seconds – as the urge some drivers have to kill other drivers – to days – as the preference to cancel a dentist appointment as the time draws near – to decades – as the preference

to *narrow* equilibrium. For the distinction and the justification of this approach, see the first several essays in Daniels (1996).

⁷ And others; see §§4, 24 of Rawls (1971) for a fuller account.

a person might have to spend all of her income each month that does not change until the time for retirement comes. Thus, with the assumption that each preference profile endures through a non-degenerate interval of time, we can think of each temporal aspect represented in the prudential original position as defined by a preference profile, along with the time through which that preference profile is held. In brief, those represented by the parties to the prudential original position are temporal aspects of a person situated in the circumstances of prudence.

Notice that while there is room in a Rawlsian original position for the parties to be *either* the individuals in the circumstances of justice themselves *or* representatives of those individuals, in the prudential original position the parties must be thought of as representatives of the individual as she and her preferences progress through time.⁸ I will often lapse into language that might suggest that an individual can go forward in time to meet with herself in the future. That is, I shall conflate the “time slice,” “time-defined,” or “temporal,” aspects of the individual, who surely cannot meet with one another with their representatives, the parties to the prudential original position. This conflation is only a matter of expository convenience. Also, where no confusion threatens, I will use “prudential original position” and “original position” interchangeably.

2.2 Choice Situation or Bargaining Place?

Some contractarian moral theorists (e.g., Rawls 1971) think that selection of moral principles in the original position ought to be modeled on individual choice under uncertainty. Others (e.g., Gauthier 1986) think that the original position ought to be modeled on a bargaining situation. In this section, I shall argue for a *choice* conception rather than a *bargaining* conception of the prudential original position.

Some preliminary remarks concerning the choice conception of the moral original position are necessary. Although it is often said, it is a bit misleading to say that the choice conception models the selection of moral principles on individual choice under uncertainty. When an individual faces a choice under uncertainty, she knows her preferences over actions given various states of the world, but she is uncertain as to what state of the world obtains. She might, for instance, prefer action A_1 to action A_2 if the world is in state S_1 and hold the reverse preference if the world is in state S_2 . In a choice under uncertainty she is wholly ignorant as to which state the world is actually in, lacking sufficient information to make any

⁸ We are not to imagine, of course, that these parties to the prudential original position will actually meet or that they have met at some time in the past. The original position is a heuristic device to help us arrive at an understanding of the requirements of prudence, not a historical event. Rawls (21, *et passim*) issues a similar qualification.

probability assignments to the possible states or even to judge that one state is more probable than the other.

Choice behind the veil of ignorance is not choice under uncertainty in the sense just sketched. Behind the veil of ignorance, the chooser is ignorant not only of the state of the world. She is ignorant as well of her *preferences*. She does know that people with preferences of certain sorts will have them better satisfied under some principles governing the basic structure of society than under others. Her task behind the veil of ignorance, according to the choice conception, is to choose moral principles that will best satisfy her preferences whatever they turn out to be. That is the sense in which choice behind the veil of ignorance is choice under uncertainty. It is a double uncertainty, both of the state of the world and of her preferences. These preliminary remarks about the choice conception will be important for the following discussion.

2.2.1 Against the Choice Conception. David Gauthier offers the following argument against the choice conception of the moral original position. A rational decision under uncertainty, he says, requires “a single, unified standpoint from which to establish a preference ordering” (1982: 159). The basis for rational choice, after all, is such a preference ordering (1985: 203), since it is with reference to the suitability of a choice for the satisfaction of one’s preferences that the rationality of a choice is judged. Hence, there is no basis for rational choice behind the veil of ignorance, since the veil blocks from the parties to the original position knowledge of their preferences.

We can understand Gauthier to have recognized part of what I pointed out in my preliminary remarks, that in a true choice under uncertainty one has knowledge of one’s preferences. But one does not have that knowledge behind the veil of ignorance. From this recognition, Gauthier draws the conclusion that rational choice behind the veil of ignorance is not possible.

Gauthier is probably right that choices of *actions* cannot rationally be made without knowledge of one’s preferences. Conduciveness to preference satisfaction is the correct standard to apply when determining the rationality of actions, so without knowledge of one’s preferences one is left without a standard to apply. This does not entail, however, that rational choice behind the veil of ignorance is not possible. As I also pointed out in my preliminary remarks, choosers behind the veil of ignorance are not choosing *actions*; they are choosing *principles*. And choice of *principles* of choice certainly can be made rationally without knowledge of one’s preferences. Moral principles are principles to guide our choices in cases in which our preferences conflict with the preferences of others. The rational standard to apply to a *principle* to resolve the conflict is its suitability for providing a resolution, *whatever* the preferences happen to be that are conflicting.

I need to stress that the rationality of a choice of a principle of choice is not completely removed from preference satisfaction. There is the following link between the suitability of a principle as a conflict-resolver and preference satisfaction. A party to the original position has the task of representing a particular set of preferences (though she does not know what they are) as well as she can. Her goal as a representative, therefore, is to see to it that these preferences be satisfied. It follows that she has a second-order preference that she be as effective as possible in seeing to it that the preferences she represents be satisfied. Given that the thickness of the veil has left her with knowledge of just this one second-order preference, it is with reference to the satisfaction of *this* preference that the rationality of her choice of a principle to resolve conflict is determined. Ignorance of first-order preferences over actions therefore does not entail that one has no knowledge of a preference with reference to the satisfaction of which one can make a rational choice of a principle of choice.

2.2.2 Non-maximizing Rational Choice. Gauthier might very well respond to this argument with the objection that there *is* no principle of choice that can guarantee the maximal satisfaction of this solitary preference, and this is precisely because of the ignorance of first-order preferences: “[N]o principle or set of principles . . . is best from every standpoint, or best advances one’s own good whatever that good might be” (1985: 204). Rational choice is maximizing (203). A situation, such as the original position, in which a maximizing choice cannot be made, is therefore a situation in which a rational choice cannot be made – not even a rational choice of a principle as opposed to a choice of a concrete, particular action. In such a situation, there is “no basis for anything but bargaining” (204).

The problem with this objection is that it fails to recognize the possibility of rational choices that are not expected utility maximizing. But certainly rational choice that is not maximizing is possible. Under certain conditions – such as the original position, in which protection against devastating and intolerable consequences is of the utmost importance – non-maximizing choice is quite rational.⁹ So we can see the objection just rehearsed as presenting a faulty dilemma: Either the original position is a choice situation in which the chooser can make a rational (= maximizing)

⁹ This is admittedly a bit dogmatic, but the argument for this claim is somewhat involved. One piece of the argument is given in my (1999), in which I argue against Harsanyi’s (1975) claim that the particular sort of non-maximizing choice that Rawls recommends in the original position is irrational. Another part of the argument is contained in my (2001) (the core of which is also given in my (2003)) where I argue that the particular non-maximizing choice that I recommend for a chooser in the prudential original position is rational. So at this point, I must simply beg the reader’s cooperation – or suspension of disbelief, as the case might be.

choice, or it is one in which there is a basis only for bargaining. The ignored third option is non-maximizing rational choice.

Let us take stock of the argument so far. Gauthier argues against Rawls' choice conception of the moral original position by claiming that rational choices cannot be made there. I have attempted to refute Gauthier's argument. The important point to mark here is that if I have been successful, then an objection parallel to Gauthier's objection against the choice conception of the moral original position does not go through against the choice conception of the prudential original position.

2.2.3 The Choice Model is More Appropriate for an Individual. The line of argument just presented was aimed at defending the choice conception of the moral original position – and, by analogy, of the prudential original position – against the charge that rational choice from behind a veil of ignorance is not possible. I now want to sketch a positive argument in favor of the choice model of the prudential original position.

An original position (whether moral or prudential) is a heuristic device that helps to inform our reasoning toward important (moral or prudential) principles. It is, however, a highly artificial device. So we will be well advised to settle on a conception of the original position that is best fitted for use in the selection of the *sort* of principle the device is supposed to inform. Now under both conceptions it is true that what one party will accept will be accepted by all the parties. This is because we can suppose that the parties are equally rational and would be convinced by the same arguments. Notice, though, that the bargaining conception of an original position relies quite heavily on there being other parties, with reference to whose claims one can make one's own claims. In order for there to be any *discussion* or *deliberation* on proposed principles of justice, there must be a multiplicity of parties who interact with each other and exchange ideas and claims through their deliberations. And in order for the parties to bargain in this sense, they must not all be identically situated; there must be "some differences in the perceptions or interests or knowledge or concerns of the discussants" (Sandel 1998: 129). These differences are present in the original position when conceived as a bargaining place.

On the choice conception, however, the veil of ignorance blocks out all these differences. As Rawls says, in his original position, "there is no basis for bargaining in [this] usual sense" (139). For "[t]here can be no bargaining among people who, even though they actually have conflicting ends, do not know what those ends are. The whole idea of bargaining thus becomes inapplicable and the choice of principles reduces to a choice by anyone in the original position picked out at random" (Barry 1995: 58). So on the choice conception the multiplicity of parties is inessential. All that is needed to power the device of the original position on the choice conception is one party who will be put in someone's place (a natural

person or a time slice), he knows not whose. So the sense in which the parties to a Rawlsian original position *discuss* or *deliberate* is not the sense in which individuals discuss issues with each other or the sense in which juries deliberate. Rather, the sense in which they discuss is the metaphorical sense in which one has a discussion with oneself, tossing ideas around in one's own head, so to speak; the sense in which they deliberate is the sense of deliberation that occurs prior to an individual decision. This is the only sort of discussion and deliberation possible behind a veil of ignorance.

This difference is quite important because it is more fitting that the selection of a principle by which one individual is to govern his choices through time be modeled as the choice of one individual, as in the choice conception of the prudential original position. The time slices of an agent who are represented in the prudential original position are not distinct individuals, unlike the persons who are represented in the moral original position. Since those represented in the moral original position, however, are distinct, it is perhaps more fitting that the selection of moral principles be modeled in the original position as a bargaining process.

Indeed, the more seriously the contractarian methodology I am developing here takes the distinctions among time slices, the more plausible becomes the charge that that methodology applied to prudence is implausible just for the reason that it treats the different temporal aspects of one individual as if they were distinct natural persons. I will have more to say in §3.2.3 about how seriously the contractarian prudential methodology takes the distinctions among time slices. I will argue there that the methodology I employ based on a choice conception of the prudential original position takes the distinctions just seriously enough, while a methodology based on a bargaining conception would take the distinctions too seriously. Since the discussion there is intimately connected to my solution to the objection to be presented in §3.1, I must end here with an incomplete argument for the choice conception of the prudential original position, issuing a promissory note to complete the argument in the course of my response to that objection. For now, the extent of the argument of this subsection is that the choice conception of the original position seems more fitting than the bargaining conception for the selection of prudential principles, since the former relies less heavily on there being other parties to the original position than the latter does, and it therefore presents a more appropriate framework for the selection of principles by which one individual is to govern his choices over time.

2.3 The Veil of Ignorance

Assuming that the arguments of the previous section are successful, we now need to go on to give an account of the information available to

the parties in the prudential original position, conceived as a choice situation rather than as a bargaining place. We will proceed by discussing the information that is given to the parties in the next subsection, and then the information withheld from the parties in the following subsections.

2.3.1 The Parties' Task. Deliberation in the original position cannot get off the ground if the parties do not know *what* they are to deliberate *about*. If they are only told that they are representatives and that they are to reach an agreement, they will not know if they represent players of bocce ball who are to reach agreement on the rules of the game or if they represent warring nations who are to reach agreement on the use of chemical weapons. Now just as the parties to the Rawlsian original position know "whatever general facts [about human society that] affect the choice of principles of justice" (1971: 137), so too parties to the prudential original position must know whatever general facts about human beings affect the choice of principles of prudence. And just as the life of a person takes place within the circumstances of justice, the life of a continuing self takes place within the circumstances of prudence, that is, under conditions of temporal preference change.

So the representatives must be told explicitly that they represent temporal aspects of a continuing natural person who is in the circumstances of prudence. Thus, they know that the preferences of some of them conflict with the preferences of others of them. In short, they know they are representatives in the sense spelled out in §2.1. The task they are given is to reach agreement on principles for regulating conduct in the circumstances of prudence. Not just any agreement will do, of course, since each is a *representative* whose role is to secure the interests of the time slice he represents as well as possible.

2.3.2 Temporal Preferences. In this section, I shall offer a *reductio* argument for the claim that a party to the original position is to know only the *indexical* temporal preference profile of the time slice he represents, as opposed to the *proper* temporal preference profile of that time slice. To see the distinction, recall that each party knows that his task is to secure his interests as well as possible in the circumstances of prudence.¹⁰ By his general knowledge of the circumstances of prudence, each party also knows that in those circumstances earlier time slices can negatively affect the opportunities for preference satisfaction of later time slices. It follows that each party knows that he has the preference that the time slices prior to him do not negatively affect his opportunities for preference satisfaction.

¹⁰ Notice that I am conflating the party to the original position and the time slice he represents here. The two are not the same. The conflation is only for expository convenience, as noted in the caveat at the end of §2.1.

This is an *indexical* preference, so long as he does not know to which time slices the description “the time slices prior to him” refers. There is no way of avoiding the parties having knowledge of this preference since, as I said, it follows from their mission and their general knowledge of the circumstances of prudence.

If, however, the time slices know the corresponding *proper* preferences, absurd things follow. If, for instance, the twenty-seventh time slice, S_{27} , knows not only that he has the (indexical) preference that none of the time slices prior to him foreclose his opportunities, but he knows as well that he prefers for none of the time slices S_1 through S_{26} to foreclose his opportunities, then the latter preference is a *proper* temporal preference. Knowing this preference, he would easily be able to deduce his temporal position, as would every other time slice be able to deduce his own. It is from this knowledge of his temporal position that absurd consequences follow. To appreciate the problem, consider a somewhat artificial model, in which there is some quantity Q of exactly one non-perishable but fully divisible good G available to the individual for consumption, beginning with time slice S_1 . Under the hypothesis that each has knowledge of his proper temporal preferences, each time slice knows his temporal position in the prudential original position, and the original position is thereby transformed into a very perverse bargaining place, in which earlier selves are *de facto* dictators over later selves in the following way.

S_1 knows his identity. Although he may not know his particular preferences (whether, for instance, he prefers lobster and wine or beans and cola for dinner), he does know that some quantity $q_1 \leq Q$ of G will be necessary to satisfy his preferences as fully as possible. Since he knows his temporal position, he knows that no earlier self can prevent him from taking exactly what he wants from G . S_1 therefore, in the absence of agreement, will take q_1 from G . S_2 , knowing his own temporal position, knowing that in the absence of agreement some quantity q_1 of G will be gone when his turn to choose comes, and knowing that there will be no one to prevent him from taking exactly what remains from G when his turn comes, will, in the absence of agreement, take that quantity $q_2 \leq Q - q_1$ from G that will maximally fulfill his preferences. We can, of course, continue inductively, supposing that if S_1 through S_{i-1} were to consume $q_1 + \dots + q_{i-1}$, then S_i would consume q_i subject to the constraint that $q_i \leq Q - (q_1 + \dots + q_{i-1})$.

Therefore, we see that in the absence of agreement the distribution of G that would result is $d = (q_1, q_2, \dots, q_n)$. In discussions of bargaining, such a distribution is called the *disagreement point*. What has happened is that by giving the time slices knowledge of their temporal positions by giving them knowledge of their proper temporal preferences, we have inadvertently turned the original position into a bargaining place.

In a bargaining situation such as this, it is usually assumed that if the distribution rationally agreed upon by the bargainers is different than the disagreement point, then that distribution must be at least as good for each bargainer as the disagreement point, and better for at least one of the bargainers. Such an n -tuple at least as good in each coordinate as d and better in at least one coordinate is said to be *Pareto superior* to d . If unanimous agreement on the bargaining solution is required, and I am assuming it is, then this Pareto requirement makes all the sense in the world. For if there were a bargainer who did worse under a proposed bargain than under d , she would rationally veto the proposed bargain, and this proposed bargain would thereby not be the rational solution to the bargaining problem.

In the case at hand, there is no feasible distribution that is Pareto superior to d . For suppose there were one, say $d^* = (q_1^*, q_2^*, \dots, q_n^*)$. Then there is some time slice S_j who does better under d^* than under d , so that $q_j^* > q_j$. But given this peculiar bargaining situation in which the bargainers choose in a known order and are free to choose what they want, as rehearsed above, each time slice is rational to choose the quantity of G from the remaining stock that will maximally fulfill his preferences. Thus, since S_j does better under d^* than under d , it must be that q_j does not maximally fulfill his preferences, which means that the stock of G that was remaining when S_j 's turn to choose came, $Q - (q_1 + \dots + q_{j-1})$, was not sufficient to fulfill his preferences. Thus, if more resources are available for S_j under d^* than under d , it must be that $q_1^* + \dots + q_{j-1}^* < q_1 + \dots + q_{j-1}$, which implies that some time slice prior to S_j does *less well* under d^* than under d (i.e., that there is an $i < j$ such that $q_i^* < q_i$), contrary to the supposition that d^* is Pareto superior to d . Thus, if the parties know their proper temporal preferences, then the rational solution to the bargaining problem will be the distribution d .

We can put the point of the preceding formalism as follows. If the parties to the original position know their proper temporal preferences, then they will know their temporal positions. If they know their temporal positions, then earlier time slices will have a complete bargaining advantage over later time slices, and it will be rational for earlier time slices to insist on the lion's share of the agent's resources and opportunities. So this solution allows living for the moment in the sense of a total lack of constraint on earlier time slices.

I said at the beginning of this section that I was offering a *reductio* argument for the claim that the parties are not to know their proper temporal preferences. The broadest structure of my argument so far is that if they do, then the rational resolution of conflict in the circumstances of prudence is living for the moment, which is absurd. In effect, we are now engaged in reasoning toward reflective equilibrium. We are testing one of the characterizations of the prudential original position to see if it

will result in principles that issue judgments that come close to matching our considered judgments.

I claim that the prudence of living for the moment does not match our considered judgments, that those judgments ought to be retained, and that the characterization of the original position ought therefore not to include knowledge by the parties of their proper temporal preferences. There are several serious problems with the principle of living for the moment. First, the principle does not take the problem of prudence seriously in that it does not resolve the conflicting demands that time slices make. Instead of offering a solution to the problem, this principle caves into its pressure. Second, the principle allows as prudent exactly the actions that we are most prone to judge imprudent. These are actions that are taken with no *foresight*, no regard for the morrow.¹¹ Third, the reason we judge imprudent the actions this principle allows seems to be that we regard temporal position as largely irrelevant to prudence. Prudence as living for the moment is the analogue of justice as the advantage of the stronger. The way in which early temporal position is irrelevant to prudence is similar to the way in which strength is irrelevant to justice.

This concludes my discussion of knowledge of proper temporal preferences. The parties are not to have that knowledge, for if they did, they would rationally agree on the prudence of living for the moment, which is absurd.

2.3.3 Temporal Identities and Particular Preferences. One lesson we can draw from the discussion of knowledge of proper temporal preferences is that the parties also cannot have knowledge of their temporal positions. This is because knowledge of proper temporal preferences is what gave the parties knowledge of temporal positions and the latter led to the transformation of the original position into a bargaining place where the rational solution was absurd. So the veil of ignorance must block knowledge of their temporal positions from the parties as well.

Any preferences over *particulars* that would give away the temporal positions of the parties must also be blocked. So, for instance, knowledge of a preference to go out carousing *tonight* and pay the price with *tomorrow's* exhaustion or knowledge of a preference not to have gone out carousing *last night* so as not to be paying the price *today*, would be prohibited. In general, any knowledge of the preferences of the time slices represented in the original position that would allow one representative to tailor the principles selected there to the advantage of the time slice he represents must be blocked by the veil of ignorance.

¹¹ There is a strong etymological connection between prudence and foresight.

2.4 Contractarian Prudence

That completes the development of the central ideas behind a theory of prudence approached from a contractarian perspective. We have modeled choice of principles of prudence in the intrapersonal analogue to the Rawlsian original position of morality, arguing for a choice conception of the prudential original position and illustrating some of the restrictions that it makes sense to impose on deliberation toward prudential principles. It takes another paper to show that this framework can actually support the work intended for it, namely, the selection of a reasonable principle of prudence. I should pause, however, to indicate briefly the principle that I think would be selected in the prudential original position and to sketch very succinctly some of the reasons in favor of choosing this principle.

Recall the Difference Principle, one of the principles of justice that Rawls claims it would be rational to choose in the original position. According to this principle, “[s]ocial and economic inequalities are to be arranged so that they are . . . to the greatest benefit of the least advantaged” (302). In brief, the Difference Principle makes things go best for the least advantaged members of society. I propose a principle of prudence that makes things go best for the least advantaged time slices of an individual. I quickly issue the caveat that this principle is *not* simply the parallel, in the intrapersonal case, to the Difference Principle. To see what the difference is, recall that the Difference Principle measures how well a person in society is doing based on her level of (expectation of) primary social goods – things like income that anyone would want to have more of rather than less of, no matter what her plan of life is (90ff.). So the Difference Principle maximizes the minimum level of primary social goods held by anyone in society.

By contrast to the approach based on primary social goods, I propose that we measure how well a time slice is doing by looking at her level of *regret*. The more regret a time slice has, the worse she is doing, while the less regret she has, the better she is doing. The least-well-off time slice, then, will be the one experiencing the greatest regret. Thus, in order to make things go as well as possible for the least-well-off time slice, we must *minimize* the *maximum* regret felt by any of the individual’s time slices. This is the Principle of Minimax Regret: an individual ought to organize the pursuits of his life in such a way that the maximum regret experienced by any of his time slices is minimized.

With this brief description of the Principle of Minimax Regret, at least two important questions immediately arise. First, how does a time slice measure her regret? To this question, we can say that a time slice ought to be allowed to measure regret however she sees fit. She might be concerned only with the shortfall of her current level of utility from what her utility level would have been had an earlier time slice made a different choice and call any such shortfall *regret*. Or she might care very little about her

current level of utility and instead measure her regret largely based on the shortfalls of past and future time slices.

A second important question asks what the nature of the argument is in the prudential original position in favor of this principle. Thinking of the parties to the original position as representatives of time slices, a time slice will measure the performance of her representative on the basis of how well the representative has succeeded in securing for the time slice what she cares about. Regret varies inversely with the time slice having what she cares about, so it is appropriate for the representative to select the principle of prudence that will guarantee that the largest level of regret is as small as possible. No other principle of prudence can guarantee a better evaluation of the representative's performance than the Principle of Minimax Regret.

As I noted at the outset, I aim to keep the focus here mainly on methodology, so I refer the reader to my companion paper (2003) for a much fuller account of the content of Minimax Regret and the arguments for it. I now want to consider an important objection, the response to which will lay a bit more groundwork for the promised lesson of this methodological investigation for moral contractarianism.

3. FRAGMENTATION

3.1 The Problem

There is a problem connected with the time slices that might be thought to infect the entire contractarian approach to prudence that I have developed. This is the question whether it is even sensible to conceive of a person as composed of a series of temporal parts. In particular, by identifying these parts, or time slices, with a temporal preference profile we seem to have created many distinct persons, none of whom is identical to the person we believe persists through (and in spite of) the temporal stages that contractarian prudence concentrates on. We seem to have taken the time slices too seriously, thus fragmenting the person in our thought to such an extent that only the fragments are real and the person they compose has somehow blown – or been analyzed – away. But certainly that is not plausible, at least if we hold the common sense belief that natural persons persist through the duration of human lives.

The root of this difficulty would seem to be the separate temporal preferences that serve to identify the time slices. Rawls claims that “[i]n the case of the individual, pure time preference is irrational: It means that he is not viewing all moments as equally parts of one life” (295). In another passage, Rawls writes:

Acting with deliberative rationality can only insure that our conduct is above reproach, and that we are responsible to ourselves *as one person over time*. We

should indeed be surprised if someone said that he did not care about how he will view his present actions later any more than he cares about the affairs of other people (which is not much, let us suppose). One who rejects equally the claims of his future self and the interests of others is not only irresponsible with respect to them but in regard to his own person as well. *He does not see himself as one enduring individual.* (422–3, emphasis added)

So a person who, at some time, weights utility removed from the present in some way that leads to a temporal conflict in preferences thereby fragments herself by failing to regard as identical her current self and her self that is removed from the present. Future or past moments count differently from the current moment, and the inference that Rawls makes is that so to count the importance of moments is to fail to view them as equally moments of one life, to fail to see oneself as one enduring individual.

Therefore, the objection goes, merely entering the contractarian framework by considering a succession of time slices with the accompanying succession of temporal preferences really is to be committed to an absurd ontology. On this ontology, the time slices are reified, given a life of their own, in the queerest way. They are real, distinct parts of an individual – or worse, they simply are distinct individuals – who compete for resources and opportunities for preference satisfaction. Instead of a view of the self as a harmonious whole, or at least as cooperative parts of a whole with continuing plans and mutual interests, the view of the self presupposed by the contractarian framework for prudence is at best schizophrenic.

3.2 Time Slices

In order to respond to the fragmentation problem, we need to give greater philosophical attention to the time slices and the role they play in contractarian prudential thought. My purpose in this section is to provide an account of the time slices of an individual that regards each time slice as a non-exhaustive *aspect* of the person who persists over time.

3.2.1 Aspects. Let us begin with an ordinary person. Call her *S*. Let us regard her simply as a *natural person*, a *continuing self*. We can pick out this notion of her with language that is a bit less philosophical by calling her also an *actor* or an *agent*, though we must take care to divorce these terms from their slightly more specialized connotations of someone who impersonates, acts on behalf of, or represents another. So I mean for us to think of *S* as something *very* ordinary: A natural person who persists over time and who is capable of action.

Now when I say that we are beginning with this ordinary notion of a natural person, I mean that we ought to begin this discussion of the proper conception of a time slice by taking the notion of a natural person

as conceptually prior to the notion of the time slice. Certainly much of what has gone before, because of the mode of investigation that the contractarian methodology employs, might have made it seem as though it were the other way around: That the time slices were primary, and the conception of the natural person had to accommodate the conception of the time slices. But this investigation has been about how a *person* ought to behave. So the person ought still to be primary – though if the work here is on track, it is fruitful to conduct the investigation through the eyes of the time slices.

Given the person S as primary, the question we ought to ask is: What is the relationship between S and her time slices, the S_i s? The answer is that the time slice S_i is an *aspect* of S at time t_i . She is an aspect of S in a way similar to the way in which my vocal properties comprise an aspect of me. There are, of course, other aspects of me: my hair color, my personality in social situations, the activities I most like to engage in to stay fit. That aspect of me comprised by my vocal properties, however, might be the most relevant aspect of me to consider for some particular purpose, say, for assessing my qualifications for a job as a radio announcer.

Similarly, the temporally indexed preference profile of S_i at t_i is what is relevant to examine from the standpoint of prudence. After all, as I claimed in §1, it is the conflict of such time-indexed preference profiles that constitutes the circumstances of prudence. Carefully note, however, that this is not to say that S_i is just the *same* as S at t_i . There is more to S at t_i than the properties of S_i that are taken into account when thinking about prudence. She might indeed have prudential reasons that are not reflected in her temporal preference profile. There is also more to me than my vocal properties, even though that aspect of me is what is most important in assessing my qualifications for the job.

There does not seem to be anything objectionable in talking about aspects of a thing in the way I am proposing. As I have been saying, considering a certain aspect or set of properties of a thing seems wholly appropriate given a particular aim. Just as a time slice presents an aspect of a natural person, a photograph of a building presents an aspect of the building, though not a complete characterization of it. Different photographs of the building might present different aspects, depending on our purposes. An aerial photograph might be most appropriate for showing its location in relation to others, while a ground-level photograph including the surrounding buildings is most appropriate for a comparative analysis of their external architectural features. Photographs presenting other aspects might be relevant for still other purposes. I hope to have isolated in the time slices with their temporal preference profiles the aspects of an agent that are relevant to the problem of prudence.

So we begin conceptually with an agent S , a natural person. We then isolate certain aspects of her, her temporal aspects, that are relevant to the

problem of prudence, how to act in the face of temporal preference change. These aspects are incomplete, in that they do not present a complete view of S at any time. They present just the aspects of her that are relevant to the problem of prudence.

3.2.2 *But Why Not More?*. One might agree that there is nothing wrong with talking about aspects of things as I am proposing, but ask nevertheless: Why not include more in the aspects of a person at a time? Surely persons can have reasons other than those that derive from their temporal preferences. So should we not instead conceive of a time slice S_i as presenting a more complete picture of S at t_i , with more of her reasons, temporal and prudential?

The reason we do not conceive of S_i as presenting a more complete picture of S at t_i is that we do not yet know what ought to be true about S at t_i – in particular, what her prudential reasons are – until we decide what prudence requires. We do not yet know how S ought to pull her motivational states together, as it were, until our theory of prudence tells her how to do that – until our theory of prudence tells her, that is, what reasons she has at t_i in addition to those that derive directly from her temporal motivational states.

So our methodology begins with the primary entity S . She has a practical problem, because she suffers from a succession of motivational states over time. We slice her up, as it were, and consider these successive motivational states in isolation. This slicing is done before we have determined what prudence requires. Thus, whatever prudential reasons S might have are not yet reflected in the reasons we have isolated in the time slices via the temporal preference profiles of those slices. Our methodology *then* goes on to ask how the conflict in motivational states ought to be settled, and we bring in the contractarian machinery. With that machinery, we identify what prudence requires. So it is only *after* we show what principle would be selected in the original position that we know what prudence requires of S . So it would beg the question we seek to answer about the requirements of prudence to conceive of S_i as presenting a picture of S at t_i that includes her prudential reasons.

Thus, the very short answer to the question why we do not present S_i as a more complete picture of S at t_i , rather than only an aspect of her, is that if we did, we would be supposing that the problem of prudence – how to act in the face of temporal preference changes – was already settled. Only once the problem is solved can we consider an aspect of S at t_i – call her S_i^* – who presents a complete picture of S at t_i , a picture that includes not only her temporally indexed motivational state, but also her *settled* motivational state that she comes to by way of considering the requirements of prudence.

This way of thinking of time slices of individuals resembles the way a certain kind of contractarian moral theorist thinks of rational agents.¹² Such a theorist begins with the preferences of a rational agent and then argues that it is rational for her to constrain her pursuit of those preferences, whatever they are, according to some moral principles that would be selected in an appropriately specified pre-moral situation. Thus this sort of contractarian moral theorist begins by considering only one aspect of the rational agent by beginning with just some of the reasons she truly has – those that come directly from her preferences – and then shows that she has more reasons than just those. Were it objected that this sterilized image of the rational agent is an incomplete picture of her – for it leaves out the moral reasons she has – the proper response would be that to begin with a complete picture would be to assume that the question of what reasons she has to behave morally had already been answered. But that issue has not yet been settled when the contractarian opens the investigation. Whatever moral reasons a rational agent might have are not yet reflected in the reasons the contractarian moral theorist has isolated – pre-moral-theoretically – in the rational agent via her preference structure. That is how the methodology works.¹³

3.2.3 How Seriously to Take Distinctions. A further point we need to note before the fragmentation problem can be solved is that by the very conception of the prudential original position sketched in §2, the contractarian methodology at once takes the perspective of every time slice *and* of the agent or continuing self as a whole. It is clear how the methodology takes the perspectives of the time slices. It does this by giving them representation in the original position, where the selection of a principle of prudence takes place. The aim of the present section is to elaborate and explain how it is that the contractarian prudential methodology also takes the perspective of the continuing self.

To begin, we must briefly review a dispute in contractarian moral theory regarding the proper conception of the moral original position. We have already discussed one aspect of this dispute in §2.2.3, where

¹² The kind of theorist I have in mind is more like Gauthier than Rawls.

¹³ David Gauthier has pointed out that this analogy between moral and prudential contractarianism is imperfect. In the moral case, it is the parties to the agreement who acquire new reasons by virtue of the agreement. In the prudential case, since the time slices are not properly speaking *agents* (as is clear from the discussion in §3.2.1), *they* do not acquire new reasons by virtue of the agreement. Rather, the agent of which they are aspects acquires prudential reasons. Despite this disanalogy, the analogy does help to make clear why S_i does not present a picture of S at t_i that includes her prudential reasons.

the question was whether the prudential original position ought to be conceived as a bargaining situation or as a choosing place. There the issue was which model is more appropriate for the selection of a principle by which one individual is to guide her choices over time. I gave an incomplete argument for the claim that the choice model is more appropriate, based on the recognition that the multiplicity of parties to the original position conceived as a choice situation was inessential. I can now fulfill the terms of the promissory note I issued then to complete that argument in the course of my response to the fragmentation problem. To see the rest of the argument, we now must take up another aspect of the dispute about how to conceive the original position. In the moral case, the issue arises as the question of how seriously to take the distinctions among persons. In the prudential case, the question is how seriously to take the distinctions among time slices.

Rawls complains against “the utilitarian view of justice... that it does not matter, except indirectly, how [a given] sum of satisfactions is distributed among individuals” (26). By throwing all of society’s resources into one utilitarian pot, we are left with “no reason in principle why the greater gains of some should not compensate for the lesser losses of others; or more importantly, why the violation of the liberty of a few might not be made right by the greater good shared by many” (26). This strikes Rawls as an objectionable view of justice, leading him to charge that “[u]tilitarianism does not take seriously the distinction between persons” (27). Rawls’ own view of justice as fairness attempts to remedy this defect by requiring that persons’ basic liberties be equal and then requiring that deviations from the equal distribution of social and economic goods be to the advantage of everyone.

Gauthier claims, however, that Rawls does not go far enough in recognizing the distinctions among persons. Indeed, given the thickness of the veil of ignorance, the distinctions among persons are all but erased in Rawls’ original position. Gauthier thinks in particular that there are two sources of distinction that are relevant to rational agreement on moral principles, knowledge of which is not possessed by the parties to the Rawlsian original position. First, Gauthier claims that a rational agreement on terms of cooperation must leave an individual at least as well off as she would be in the absence of an agreement. Society is, after all – in Rawls’ own words – “a cooperative venture for mutual advantage” (4). So if the venture is to be to each person’s advantage, it must consider how individuals would fare in the absence of an agreement. To do that, initial factor endowments must be considered, and so the distinctions among the initial endowments of different individuals are relevant distinctions. This proviso that Gauthier places on the agreement on terms of cooperation stands in direct opposition to Rawls’ claim that “a hypothetical initial

arrangement in which all the social primary goods are equally distributed" provides the "benchmark for judging improvements" (62).¹⁴

A second relevant distinction among individuals for Gauthier is the contribution that each individual makes to the fruits of social cooperation, via her talents. This opposes Rawls' claim that the natural talents of individuals are to be regarded as common assets (101). Gauthier concludes that Rawls' theory falls victim to the very charge that he leveled against utilitarianism, for Rawls' theory fails to take seriously the distinctions among persons by "not taking seriously the individuality of persons" which "lies in their capacities and preferences" (1986: 254).¹⁵ We can now turn to the question of how seriously the theory of contractarian prudence I have advocated here takes the distinctions among time slices. I shall argue that it takes them just seriously enough.

By failing to advocate the maximization of the aggregate utility of the temporal aspects of the agent, the contractarian methodology here takes the distinctions among temporal aspects seriously – in fact, just as seriously as Rawls takes the distinctions among persons. The time slices do, after all, have conflicting preferences. If they were part of an individual whose preferences remained unchanged throughout his entire lifetime, then it might make sense to maximize the aggregate utility of this agent's temporal aspects.¹⁶ But they are not part of such an individual, so some consideration would seem to be due to their differences in preferences. The contractarian methodology gives them this due consideration. Hence this methodology is not subject to the analogue of Rawls' criticism of the utilitarian view of justice, i.e., that it fails to take seriously the distinctions among time slices.

Nevertheless, the methodology I advocate does not take the distinctions among time slices *too* seriously. I have sketched an account of the temporal aspects of an agent according to which they are just that, aspects. They are not separate individuals or self-sufficient beings who have claims based on what they bring to the bargaining table or on what they can contribute to the fruits of cooperation. For a contractarian

¹⁴ My gloss of the role of initial factor endowments for Gauthier is very slim; for a full account see ch. 7 of Gauthier (1986).

¹⁵ I have mainly followed Gauthier's own account of his disagreement with Rawls from his 1986 (245ff.). For more on this disagreement about how seriously to take the distinctions among persons, see his 1985 as well as 1984a (181) and 1984b (273).

¹⁶ Rawls certainly thinks this. A main component of his criticism against classical utilitarianism (22–7) is that it attempts to extend "to society the principle of choice for one man" (27) for his choices over time. For this claim to have any plausibility at all, Rawls must be assuming that this one man has static preferences so that it makes sense for him to maximize the sum of his satisfactions; otherwise the intrapersonal analogue to the problem of interpersonal comparisons of utility stands in the way. Other uncritical advocates of this utilitarian view of prudence include Sidgwick (1907), Trebilcot (1974), and Bricker (1980).

prudential theory to take the distinctions among temporal aspects of individuals as seriously as Gauthier claims a contractarian moral theory ought to take the distinctions among persons would be to take the distinctions far too seriously. We do not think it irrational or imprudent for an agent to eat dinner, even though it was earlier aspects of her who had to make the dough. All else held constant, however, we do think there is something morally wrong for persons who make no contributions to production to take a free lunch. So there is a great difference between (on the one hand) distinctions between different aspects of one person and (on the other) distinctions between different persons. To take the distinctions among time slices too seriously would be to treat different time slices as different persons, which they are not.

Contractarian prudence steers a middle course, taking the distinctions among time slices just seriously enough. It takes them more seriously than a utilitarian view of justice takes the distinctions among persons. It recognizes that they are aspects of a whole person, but since they are not aspects of a person with unchanging preferences, it takes their distinct temporal preferences seriously. It avoids taking the distinctions too seriously, by avoiding a bargaining conception of the prudential original position. By throwing all of the resources of the temporal aspects of an individual into one pot in the limited, non-utilitarian way it does, the prudential original position takes the perspective of the agent as a whole. That was the aim I announced at the beginning of this section – to show how the agent as a whole is reflected in the prudential original position. More than this, we can finally conclude that the choice conception of the prudential original position is more appropriate than the bargaining model for the selection of prudential principles, so I have satisfied the terms of my earlier promissory note.

3.3 The Fragmentation Problem Solved

There is a happy concomitant result. Really, the result I now want to note is a consequence both of the way the self as a whole is reflected in the choice conception of the prudential original position and of the aspectual account of the time slices that I gave above. Given the way the self as a whole is reflected in the prudential original position, the fragmentation problem, which claimed that the conception of the time slices invoked by contractarian prudence involves a commitment to the view that they are distinct individuals, is seen not to be a serious worry.

Surely a bargaining conception of the prudential original position would lead to a real fragmentation of the self, but the choice conception merely underlines the different aspects of a self without reifying them. Indeed, we saw quite clearly in §2.3.2 that a certain bargaining conception of the original position leads to the principle of living for the moment. *That*

conception of the original position and its resulting principle of prudence sanction an agent allowing his temporal aspects to run rampant. That conception and the resulting principle take the distinctions among an agent's temporal aspects far too seriously, making them real entities with real claims in a way that they clearly are not.

Thus, the contractarian methodology, with a properly characterized original position and sensible principle of prudence, leaves the agent intact. The prudential original position is an analytical device that attempts to "do prudence to" our motivational contours, but it does not commit us to the view that the temporal aspects of an agent are distinct individuals. According to the theory of contractarian prudence presented here, the prudent agent merely takes notice of his temporal preferences in order to take them into account (and perhaps control them) in prudent decision-making. This contractarian methodology and this agent take his temporal aspects with their temporal preferences seriously, though not seriously enough for the charge that his self is thereby objectionably fragmented to hold water.

4. PRUDENCE AND JUSTICE

The pieces are now in place to draw some conclusions for contractarian moral theory. The bulk of the discussion so far has been aimed at the development of the contractarian prudential methodology. I want to claim now, though, that the reasons I have given for the *choice* conception of the *prudential* original position as superior to the *bargaining* conception weigh equally in favor of the *bargaining* conception of the *moral* original position as superior to the *choice* conception.

Recall, first, the arguments I gave in §2.2. There I claimed that Gauthier's argument against the possibility of rational choice behind a veil of ignorance *à la* the Rawlsian original position was not successful. I needed to make this argument to show that a similar argument against the Rawlsian-type prudential original position that I favor did not go through. Nevertheless, what surfaced was a recognition that an original position conceived as a choice situation (a *choice* original position, for short) was more appropriate to settle questions of intrapersonal conflict within an individual than an original position conceived as a bargaining place (a *bargaining* original position, for short). Now it does not follow immediately that a bargaining original position is more appropriate for settling questions of interpersonal conflict between individuals, for it might be that a choice original position is more appropriate for both tasks. But I now want to argue that the reasons I adduced in favor of the choice original position for arriving at principles for intrapersonal conflict also weigh in favor of a bargaining original position for arriving at principles for interpersonal conflict.

In §2.2.3 I pointed out that under either conception of the original position, what one party rationally accepts will be rationally accepted by all the parties. Nevertheless, the bargaining original position relies more heavily on there being other parties with reference to whose claims one can make one's own claims. The multiplicity of parties in the choice original position is inessential. All that is needed to power that device is one party who will be put in someone's shoes, though he does not know whose (or which time slice's) shoes they will be. Now the moral original position is a heuristic device that is intended to inform our reasoning toward moral principles. But since it is such an artificial device, we must be careful in arriving at a conception of the moral original position that is best suited for use in the selection of the *sort* of principle the device is supposed to inform. Since those represented in the moral original position are distinct individuals it is more fitting that the selection of moral principles be modeled in the original position as a bargaining process, which better reflects the multiplicity of the parties involved in morality than does a choice original position. So that is the first reason favoring a bargaining interpretation of the moral original position. Relying on the multiplicity of parties as it does, it presents a model better suited for the selection of moral principles.

A second new reason in favor of the bargaining moral original position comes from the discussion in §3.2.3. I reviewed there the controversy between Gauthier and Rawls as to what distinctions between persons are morally relevant in deciding upon moral principles in an original position. I showed that the choice conception of the prudential original position takes the distinctions between time slices just seriously enough, appropriately "doing prudence to" an individual's motivational contours. But if this is right, that the choice conception of the prudential original position builds in the right amount of distinction between time slices, then the choice conception of the moral original position treats distinct individuals as if they were just so many aspects of one person. That is obviously not so. So I conclude, with Gauthier, that the Rawlsian choice original position falls to an objection similar to the objection Rawls mounts against the utilitarian view of justice, that it fails to take seriously the distinctions among persons.

I want to stress, though, how this second argument of mine fits in with Gauthier's. I am inclined to believe that Gauthier has successfully shown the relevance of particular distinctions among individuals – initial factor endowments and contributions to the fruits of social cooperation – that Rawls ignores. However, even if one were to disagree with Gauthier regarding the relevance of these specific distinctions for the selection of moral principles, she could still agree with my claim that the Rawlsian choice original position is defective for the selection of moral principles, for the reason that *that* original position treats individuals as if they were just

so many aspects of one natural person. So my argument against the choice conception of the moral original position is independent of Gauthier's.

CONCLUSION

There is a deep similarity between prudence and justice. Prudence stands to intrapersonal conflict in roughly the same relation that justice stands to interpersonal conflict. If my development of the contractarian methodology based on a choice conception of the original position for the investigation of prudential principles has been on track and my reasons for adopting that methodology over its competitor are good reasons, then another result has been uncovered. This result is that there is a deep-*ish* similarity in the methodologies that are suitably employed in investigating principles of prudence and principles of justice. Both are contractarian methodologies, but one variety is more suitable for a theory of prudence and the other is more suitable for a theory of justice.

REFERENCES

- Ainslie, G. 1992. *Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person*. Cambridge University Press
- Barry, B. 1995. *Justice as Impartiality*. Oxford University Press
- Bricker, P. 1980. Prudence. *The Journal of Philosophy*, 77:381–401
- Bruckner, D. 1999. Harsanyi on the use of Maximin in the original position. Paper presented at the Eastern Division meeting of the American Philosophical Association, December 27–30, Boston.
- Bruckner, D. 2001. Contractarian prudence. Ph.D. diss., Department of Philosophy, University of Pittsburgh
- Bruckner, D. 2003. A contractarian account of (part of) prudence. *American Philosophical Quarterly*, 40:33–46
- Copp, D. and D. Zimmerman (eds.). 1984. *Morality, Reason and Truth*. Rowan and Allanheld
- Daniels, N. 1996. *Justice and Justification: Reflective Equilibrium in Theory and Practice*. Cambridge University Press
- Elster, J. 1979. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge University Press
- Gauthier, D. 1982. On the refutation of utilitarianism. In Miller and Williams (1982)
- Gauthier, D. 1984a. Justice as social choice. In Copp and Zimmerman (1984). Reprinted in part in Gauthier (1990). Page references are to the latter.
- Gauthier, D. 1984b. The incomplete egoist. In McMurrin (1984). Reprinted in Gauthier (1990). Page references are to the latter.
- Gauthier, D. 1985. Bargaining and justice. *Social Philosophy and Policy*, 2:29–47. Reprinted in Gauthier (1990). Page references are to the latter.
- Gauthier, D. 1986. *Morals by Agreement*. Clarendon Press
- Gauthier, D. 1990. *Moral Dealing: Contract, Ethics, and Reason*. Cornell University Press
- Hampton, J. 1998. *The Authority of Reason*. Cambridge University Press
- Harsanyi, J. 1975. Can the Maximin Principle serve as a basis for morality? A critique of John Rawls's theory. *American Political Science Review*, 59:594–606
- Hobbes, T. 1994/1651. *Leviathan*. Edwin Curley (ed.). Hackett Publishing Company

- Hume, D. 1978/1739. *A Treatise of Human Nature*. L. A. Selby-Bigge and P. H. Nidditch (eds.). Clarendon Press
- Hume, D. 1975/1777. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. P. H. Nidditch (ed.). Clarendon Press
- McMurrin, S. (ed.). 1984. *The Tanner Lectures on Human Values*, vol. V. Cambridge University Press
- Miller, H. and W. Williams (eds.). 1982. *The Limits of Utilitarianism*. University of Minnesota Press
- Nagel, T. 1970. *The Possibility of Altruism*. Princeton University Press
- Parfit, D. 1984. *Reasons and Persons*. Oxford University Press
- Plato. 1997. *Complete Works*. John M. Cooper (ed.). Hackett Publishing Company
- Rawls, J. 1958. Justice as fairness. *Philosophical Review*, 67:164–94
- Rawls, J. 1971. *A Theory of Justice*. Harvard University Press
- Rousseau, J. 1990. *The First and Second Discourses and Essay on the Origin of Languages*. Victor Gourevitch (trans.). Harper and Row
- Sandel, M. 1998. *Liberalism and the Limits of Justice*, (2nd Edition). Cambridge University Press
- Sidgwick, H. 1907. *The Methods of Ethics*, (7th Edition). Macmillan
- Strotz, R. H. 1955. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23:165–80
- Trebilcot, J. 1974. Aprudentialism. *American Philosophical Quarterly*, 11:203–10