

**POSITIVE CONSTRAINTS ON
NORMATIVE
POLITICAL THEORY**

Geoffrey Brennan
Research School of Social Sciences
Australian National University

and

Alan Hamlin
School of Social Sciences
University of Manchester

Abstract

We consider the relationship between positive and normative political theory, emphasising both the role of normative ideas in motivating political behaviour and particularly the constraints on normative theory imposed by positive theory. Our point of departure is Christiano's recent claim that the rational choice approach to positive and normative political theory is self-defeating. In arguing against this position, we outline a revisionist rational choice approach that we believe provides a bridge between positive and normative political theories, and clarifies the nature of the constraints on normative theory.

Introduction

What should be the relationship between positive political theory and normative political theory? This is the large question towards which this essay edges. The answer to this large question will depend, in part, on our understanding of the purposes of positive and normative political theory respectively. We will not dwell for long on discussing the range of potential purposes and will simply stipulate that a primary purpose of positive political theory is to explain observed political behaviour. We will make no attempt to specify what we mean by 'explain', or the differences between explanation, prediction and other related ideas, but we will stress that the central idea of explanation goes beyond mere description while leaving open a wide variety of approaches to positive political theory¹.

With respect to normative political theory, we identify three possible aims. First, analysing normative concepts, categories, ideals and intuitions aimed at revealing their true features and inter-relationships (if any). Call this the formal aim of normative theory. Second, deploying normative criteria or ideals to evaluate particular actions, policies, practices, reforms and institutions. Call this the evaluative aim of normative political theory. Third, justifying and advocating particular actions, policies, practices, reforms and institutions. Call this the practical aim of normative political theory². Clearly these three aims cumulate, in that the practical aim requires formal and evaluative consideration, and equally clearly normative political theory might be valuable even if the practical aim were not achieved – although we would suggest that its value would be sharply reduced. At first glance it may seem that the evaluative and the practical aims overlap, but it is what divides the practical from the evaluative that is crucial for our purposes: for this is precisely where 'positive political theory' connects with normative theory.

Almost everything we say in this essay relates to the practical aim of normative political theory and its relationship with positive political theory. This immediately reduces the scale of the big question that we started with, and makes our task more manageable, but we do not mean to imply that the relationships between positive political theory and the formal and evaluative aspects of normative political theory are either unimportant or secondary. Indeed we believe that something of what we say here carries over to these other areas, and that there is more to be said. But we shall not pursue those issues here.

Indeed, the specific trigger for this essay is more limited still in its range, and is provided by a recent essay by Christiano (Christiano, 2004) focusing attention on the relationship between the positive (explanatory) and the normative (practical) aspects specifically of rational choice political theory. A sharp statement of Christiano's thesis is that the explanatory theory advanced by mainstream rational choice theorists, if true, undermines the normative stance adopted by rational choice theorists, so rendering the overall theoretical position 'self-defeating'. Christiano goes on to argue that the attempt to escape this problem by 'revisionist' rational choice theorists – exemplified in Christiano's characterisation by Brennan and Hamlin (2000) – fails.

We wish to examine several aspects of Christiano's thesis in the process of developing a more general argument. In particular, we will advance three propositions. Our first proposition is that the problem identified by Christiano is overstated – the central issue is not, in our view, best thought of in terms of the categorical distinction between theoretical positions that are self-defeating and those that are not. Rather the nature of the relationship between positive and normative theory is one in which the positive constrains the normative to a greater or lesser degree. And while we agree that the degree to which the positive constrains the normative is indeed significant, this fact does not undermine either the positive or the normative aspect of rational choice theory. Second, we will argue that this issue is in no way restricted to the rational choice approach to political theory. The issue arises with only minor modification and with equal force in all plausible approaches to political theory. Third we will argue that the 'revisionist' moves laid out in earlier work and summarised below can be expected to reduce the degree of tension between explanatory and justificatory aspects of rational choice theory and therefore to open up an increased, but still limited, practical role for normative theorising in the rational choice tradition.

While we differ from Christiano on many details, we agree with him that the nature of the relationship between positive and normative theory – whether in the rational choice tradition or more generally – is a topic of great significance that is too often ignored by theorists, of both positive and normative stripe. Our overall aims, therefore, are to refine Christiano's arguments and to generalise them.

Before embarking on developing and supporting these three propositions, it is appropriate to sketch out the basic terrain - much of which is common ground between all parties to the debate. As a first step in doing so, it may be appropriate to locate our central concerns in

relation to another debate. Critics of ‘ideal theory’ often allege that ideal normative theory has become too remote from practical reality: that normative theory devotes itself to the analysis and specification of the ideal at the expense of the consideration of applications of normative thinking to practical problems or issues in the real (non-ideal) world. We would agree that this criticism has some force, but this point relates more obviously to the discussion of the formal aim of normative political theory and it is not the point that we wish to make here.

Our point is closer to a second line of criticism of ideal theory: that normative theory depends crucially on facts about the world and so cannot be divorced from at least some descriptive or explicatory considerations. We are concerned with a rather special subset of facts about the world: those that, taken together, provide the basis for positive political theory. Our concern is with just one part of the move from the ideal to the practical: the relationship between normative theory and positive theory. Of course, both bodies of theory operate at a level of abstraction from the real world, but our point is that the two bodies of theory should be seen as interconnected in at least some respects³.

This interconnection is, we believe, best seen as bi-directional. A part of our argument will be to sketch out the case for incorporating normative concerns into positive political theory, even where that theory is approached from the rational choice tradition; but our primary concern in this essay is the connection that flows from the positive to the normative. Stripped to the bare essentials, our argument is that, to the extent that the practical purpose of normative political theory is important, normative theory must take seriously the behavioural and motivational structure of political agents as summarised in positive political theory. Put otherwise, if normative theory is, at least in part, concerned with preaching, it should ensure that its message can have purchase with people like us.

The Terrain

Rationality and Revisionism

A central point of relevance in relation to rational choice theory concerns the specification of the content of rationality. In Christiano’s terminology, mainstream rational choice theory

“...adheres strictly to the thesis of *homo economicus*. In other words, it explains the operation of institutions and justifies the reform of those institutions under the

assumption that individuals normally maximise their own utility in every action they undertake.” (Christiano, 2004, p.123)

What Christiano (along with many other commentators) has in mind by ‘utility-maximisation’ is the commitment to narrow self-interest as the basis for rational choice, and by its act-by-act application of this motivation⁴. Of course, one might characterize rationality by the formal, structural properties of ‘utility functions’ rather than by their content, but we broadly accept that Christiano identifies a clear theme in the rational choice tradition. We will take the ‘mainstream’ position as one that emphasises a relatively narrow conception of self-interest; and understand the ‘revisionist’ rational choice theory alternative as one that seeks to move away from the strict conception of *homo economicus*. Christiano initially states that,

“...revisionists think of individuals, at least in a large set of cases, as not maximizing utility in every action but as adopting dispositions to act that maximise utility for the person as a whole.” (Christiano, 2004, p.123)

And we agree that this is at least part of what is at stake⁵. We conceive of a disposition as a piece of motivational apparatus that may be influenced by first or higher order desires in the long term, and which operates to condition or govern first order desires and behaviour in the short term – so that dispositions are, at least to some extent, subject to agent influence in the long term, but serve to commit action in the short term⁶. But this move from act-rationality to dispositional-rationality is, we think, only part of the ‘revisionist program’⁷. We also want to argue for a rather broader concept of ‘utility’ than is typically assumed in the mainstream approach to rational actor politics, where utility is taken to be restricted to self-interest. Much of what we have in mind here is outlined in the opening chapters of Brennan and Hamlin (2000) which makes the move from preferences to dispositions and also introduces the possibility of what are termed ‘somewhat moral motivations’. These moral motivations may be analysed in terms of desires and beliefs - the desire to act as morality requires and a set of beliefs about what morality requires in particular circumstances - and, clearly, since the desires and beliefs belong to the individual in exactly the same way as other desires and beliefs (for example, those that relate to mainstream self-interested preferences) there is no logical or formal difficulty in incorporating moral motivations within the framework of utility maximisation. However, while there is no formal difficulty, we should also recognize that the incorporation of such moral motivations may affect the substantive nature of utility maximization and, in particular, break the link to narrow conceptions of self-interest⁸. We make no strong claims about the precise substantive content of these moral motivations;

indeed, we suggest that such motivations will be unevenly distributed in the population both in terms of their precise content and in terms of their strength (relative, say, to standard preferences). Motivational (and therefore behavioural) heterogeneity is a hallmark of our understanding of the revisionist position. On this account, individuals are both rational and somewhat moral, with different individuals taking rather different views on what morality requires. Moral motivations sit alongside other motivations within any individual; they do not pre-empt or otherwise dominate other motivations but, we argue, their presence can make a difference to action both directly, and indirectly through dispositions. Accordingly, once somewhat moral desires are admitted, dispositions may also be at least somewhat moral.

There is a third aspect to the revisionist position: namely, the move from purely instrumental action so as to allow for elements of expressive behaviour. This expressive move does not depend upon either the dispositional or the moral moves, but we argue that it may interact with them. Essentially, the expressive argument recognizes that there may be benefits (utility) from speech acts (or equivalent acts) that merely express an opinion or view or preference, even if that act has no further consequences; and, in particular, even if that act plays no causal or instrumental role in bringing about the act or state for which a preference is expressed, or in realizing the underlying view, or persuading others of the underlying opinion, etc. Normally we might expect such expressive benefits to be small relative to instrumental benefits, and we might also expect the (speech) acts that realize the expressive benefits to be perfectly consistent with the acts that satisfy the underlying preference. But in at least some cases these connections come apart. The now standard example relates to voting in a mass election. The scale of the election makes the probability that your vote will have any instrumental impact vanishingly small, thus removing any instrumental reason for voting in any particular way, or indeed for voting at all. At the same time, it is easy to imagine that at least some individuals will derive expressive benefits from voicing political opinions or positions that are different from those that they would choose to bring about if they were instrumentally decisive. In these circumstances, we might expect expressive voting to carry implications for substantive political outcomes⁹.

In what follows we will take the revisionist position on rational actor political theory to include these three departures from the mainstream account - the dispositional, the somewhat moral and the expressive - and we will try to be clear as to which of these moves is doing the work at each stage of the argument.

Normative rational choice

The normative aspiration of rational choice political theory may be understood in terms of its recommendations for institutional design and reform. There is a contrast here between mainstream economics and rational choice political theory. Within mainstream economics the normative focus is often on the design of policies, understood as particular interventions by government or other agencies. A major aspect of the public choice critique of mainstream economics is that policies are properly understood as emerging from a political process – rather than as being directly chosen by an agent labelled ‘government’ that is assumed both to be motivated to serve (a specific notion of) the public interest *and* to have the power to implement whatever policies turn out to be optimal. In moving away from this ‘benign despot’ model of government, public choice theory and rational actor political theory more generally argues that the political process should be seen as the interaction of essentially rational individuals operating within a structure of particular institutional rules. Thus, attention is focussed on the properties of alternative sets of such rules, and of the policies that might be expected to emerge from their operation. Notice that the point here is not that rational choice theorists believe that policies are completely beyond the reach of normative advice; but rather that, until and unless one has an understanding of the policy-making process, its institutional context and the role of agency, it is impossible to understand the role that any such advice might play. Furthermore, there is the general point that normative recommendations about the institutional structure may be both more powerful and less idiosyncratic than normative theorising about specific policies.

Basic structure determinism

We now come to what Christiano terms ‘basic structure determinism’:

“What I mean by ‘basic structure determinism’ is the thesis that the development, maintenance, and decline of the basic structural institutions in society is determined by forces that are beyond the capacity of human beings to guide and design.” (Christiano, 2004, p.124)

So, basic structure determinism takes seriously the idea that institutions themselves (at least the basic institutions which include the institutions of politics) emerge as the unintended consequences of many actions by many individuals over time, so that the design or reform of such institutions is beyond the reach of any individual. In its pure or extreme form, basic

structure determinism denies any genuine agency at all in the process of institutional evolution. Note that the case of basic structure determinism is, in at least some respects, similar both to the argument just sketched in relation to normative policy advice and to the case of expressive voting sketched above: just as the argument on policy advice points to the need to understand a complex policy-making process rather than assume a ‘benign dictator’, and the argument for expressive voting involves the claim that the outcome of a mass election is beyond the influence of any individual voter; so basic structure determinism involves the claim that the design or reform of basic institutions is beyond the influence of individual agents¹⁰.

The argument in support of basic structure determinism presented by Christiano is intended to reflect the mainstream rational choice position, and draws extensively on Hardin (1999). Essentially, the argument revolves around the conception of an institution as a co-ordination equilibrium in a game involving many individuals. Such co-ordination points emerge in the course of repeated play, and become institutionalized just because they are self-enforcing. Notice that this account, which is broadly similar to Lewis’s account of conventions (Lewis 1969), takes the rational action of individuals within political games as a starting point. The co-ordination equilibria that emerge as institutions may or may not be justifiable in terms of any particular normative criterion, but the point is that, since they are equilibria, no individual can rationally depart from the institutionalized behaviour, nor expect to shift society to a different equilibrium (assuming that other equilibria exist). And this is true regardless of the normative status of the equilibrium.

We will return to more detailed discussion of some aspects of this line of argument later, but for the moment we merely note the difference between this argument and the more contractarian line of argument often associated with public choice accounts of politics following James Buchanan. On that account, there is a distinction between everyday or in-period politics which might be thought of as the playing of the game, and constitutional politics where the rules of the game are considered and decided. On the more contractarian account, institutions do not simply emerge from repeated play, but may be directly influenced by individuals acting collectively. The difficulties faced by the contractarian account include explaining how individuals escape from the obvious regress of seeing the constitutional level of politics as just another in-period game.

Strategy

Christiano's strategy is first to argue that the explanatory aspect of rational choice theory supports the idea of basic structure determinism, and then to argue that if both rational choice theory (in its descriptive form) and basic structure determinism are true, then the practical normative aspect of rational choice theory is completely undermined - in the sense that its normative aspirations cannot be realized. Of course, it is still possible for rational choice theorists to use the normative framework to evaluate alternative states of the world, but there is no practical role for the normative theory to influence the world. The institutional and constitutional recommendations of normative rational actor political theory could not be effective – they could play no role in actually bringing about institutional change.

We agree with Christiano that the potential effectiveness of practical normative theory is an important focus of attention. If normative political theory, of whatever substantive content, were to be insulated from the world in the sense that it could not be action-guiding and so could have no impact on political behaviour or outcomes, this would seriously limit the value of normative political theory¹¹.

We also agree with Christiano's argument in formal terms: If the strict, mainstream interpretation of rational choice theory and basic structure determinism were both true, there would be no room for an effective normative branch of rational actor political theory. However, we do not believe that Christiano's argument is successful in substantive terms, and much of the remainder of this essay is devoted to exploring some of its shortcomings.

Proposition One – Categories or Continuity?

Our first line of argument concerns Christiano's ambition to show that rational choice theory is 'self-defeating' – that is, to cast the question in terms of a categorical distinction between those theories that are self-defeating and those that are not. We argue, to the contrary, that the issue at stake is better conceived in terms of a continuum that indicates the extent to which the positive features of political reality (whatever they may be) constrain the practical effectiveness of normative theory (of whatever variety). In other words, the 'constraints' posited by positive political theory are to be seen as matters of degree. With this view in place, we suggest that while arguments such as those deployed by Christiano clearly indicate the constraining nature of the rational choice approach to positive political theory, they fall far short of demonstrating that the constraint is absolute, so that the practical effectiveness of

normative rational choice theory is not undermined. For simplicity, in discussing this line of argument we will operate entirely within the mainstream tradition of rational choice theory, returning to our preferred revisionist account in later sections.

A starting point is that while Christiano phrases his basic claim and conclusion in absolute or categorical language, much of the argument is rather less cut-and-dried than that language implies. For example, on the one hand we have:

“...unlike some forms of determinism, basic structure determinism is a hard kind of determinism. It is not merely the case that agency is determined by external forces; in basic structure determinism, there is no agency at all, in the sense that the development of these institutions is not guided by human design. ...The development of political institutions is not up to human beings.” (Christiano, 2004, p.124)

which indicates an absolute or categorical stance. While on the other hand we have:

“The thesis of basic structural determinism is also compatible with people making marginal changes to the basic structural institutions. One might attempt to change aspects of the committee system in the United States Congress, for example, and in some cases succeed. And to this extent, there is still some room for rational choice theorists to make practically effective recommendations for change. But it is a highly limited space and certainly much more limited than rational choice theorists normally have in mind” (Christiano, 2004, p.125)

which is explicitly concerned with the extent of the determinism and the ‘space’ remaining for effective normative debate. In this way, Christiano seems to concede that it is appropriate to think in terms of the extent to which normative theory is constrained by features of the positive theory of political behaviour. And we are happy to agree that many rational choice theorists have assumed rather more ‘space’ than might actually exist. For the moment, we are content to accept the principal of continuity - that the constraining nature of positive political theory is a matter of degree and that these constraints do not entirely preclude the possibility of a space for effective normative debate. Later we will argue that this space may be larger than conceded by Christiano. But if there is any space at all, normative theory will have potential practical value.

But potential value is not actual value. Even if we all agree that there is some normative space remaining, normative rational choice theory might not be able to operate in that space. And there is at least one reason for supposing that this might be the case. Under the

mainstream interpretation, the agents operating in the world are all motivated exclusively by a relatively narrow conception of self-interest, regardless of their positions in the political structure (whether voter, politician, bureaucrat, policy advisor or whatever). It might be that the recommendations made by normative rational choice theory for institutional reform are such that they carry no weight with such agents, so that they can not be effective – it would be as if the recommendations were made in a language that no actual individuals can understand.

While the normative aspirations of rational choice theory are best understood in terms of its institutional recommendations, the fundamental normative criterion underlying rational choice theory might be crudely categorised as a form of normative individualism. By this we mean only that whether or not some particular institutional reform is considered, on balance, to be worthwhile will depend crucially upon that reform's impact on individuals and how well their lives go. We do not mean to place any specific limits on what is included in the set of things that make an individual's life go well (or better), and in particular we do not necessarily restrict this set to any particular notion of 'welfare', but we do suggest that a hallmark of normative rational choice theory is that it adopts an essentially individualistic approach to value, so that social or collective value supervenes on individual value¹². Given this normative individualism it is of course likely that any normative recommendation made by a rational choice theorist is likely to be heard, understood and accepted by at least some individuals – those who benefit from the reform. But there is no assurance that these individuals will be positioned so that they can be effective in acting on the recommendation, and the fact that it is a recommendation from rational choice political theory carries no weight either with them or with others. After all, almost all reforms that might seriously be contemplated might be expected to benefit at least some individuals, and the mere fact that some individuals support a reform on these grounds neither makes that reform more likely to be effected nor justifies it to the rational choice normative theorist. In this way, one might suggest that the recommendations of the rational choice normative theorist, if not literally unintelligible to all real individuals, will have no special or distinctive voice in the babble of self-interested debate.

Of course, this issue has long been recognised by rational choice theorists in the tradition of James Buchanan¹³. There are (at least) two types of issue involved in the choice of institutional and constitutional rules. On the one hand, the very idea of the choice of rules is somewhat attenuated since an essential feature of a constitutional or institutional rule is its

relative fixity – if such rules were not seen as fixed (at least in the short to medium run) it is difficult to see how they could play the required role of constraining and shaping choice and behaviour. On the other hand, if individuals rationally choose the rules under which they operate, an obvious regress beckons: if individuals are modelled as self-interested in their choice of actions, they will presumably be just as self-interested in their choice of rules, unless constrained by some deeper rules, and so on. And no normative advice from rational choice theorists will bear on this issue.

The basic response to these two issues from mainstream public choice theory has been to suggest that the former issue might help to resolve the latter. If rules are to be quasi-permanent then in those rare moments when I face the choice of rules I have to evaluate alternative rules over the long term, when I cannot be sure of the impact any particular rule may have on me or my interests. This uncertainty yields a veil of ignorance which is argued to distance constitutional choice from self-interest by ensuring that each individual adopts a more impartial standpoint. This distinction between constitutional choice and in-period political choice is basic to the normative aspect of rational choice political theory in that it holds out the prospect of the relatively impartial choice of institutional and constitutional rules by which in-period politics and the choice of policy might be influenced.

Of course, we accept that the setting of constitutional choice falls a long way short of the full veil of ignorance that might be required to establish an entirely impartial standpoint.

Christiano rightly reminds us¹⁴ that even constitutions have a relatively short expected life, and are in any case open to interpretive change, and so may not always be regarded as once-and-for-all choices by constitution makers. Equally, the typical constitution maker will not necessarily imagine herself as occupying a potentially wide range of roles under the constitution to be chosen. Nevertheless, we believe that the shift to the constitutional perspective is a shift in the direction of impartiality – albeit a modest one.

And this shift to impartiality - however modest in scale - opens up the space available to the normative theorist. At its most basic, the theorist may consider which institutional reforms are desirable when seen from this more impartial perspective. To be sure this is mostly an analytic role based on an ability to analyse the operating characteristics of alternative institutional arrangements, but it will also require more obviously normative skills in articulating the relevant idea of impartiality. A clear example of this type of normative, analytic logic in action is provided by the Rae-Taylor theorem on majority voting (Rae, 1969; Taylor, 1969). This theorem sets up the problem as the choice of a voting rule by an

individual who expects the chosen rule to be used on a range of issues where they are not certain of whether they support or oppose the individual propositions to be voted on.

So far then, we have one role for the normative theorist – that of informing institutional choice or institutional reform from behind a veil of ignorance that shrouds the actual outcomes of those reforms as they will play out into the future. We now turn to consider a rather different role – the role of political entrepreneur or professional politician.

In many areas of life we rely on the professional incentives faced by individuals rather than their personal preferences. This is the basic point of the famous Adam Smith quotation, “It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest” (Smith, 1981, book 1 chapter 2). A butcher faces a professional incentive to build and maintain a reputation for quality and we can rely on this to at least some extent. And we might think that the same applies to professional politicians, at least in a tolerably democratic society: in striving to build a career as a politician, they will face incentives not to pursue policies or reforms that they particularly like as an individual, but rather to construct packages of policies and reforms that will advance their careers and one leading possibility here is to offer advantage to significant groups of voters.

Now, the introduction of such professional politicians makes a difference to the way in which we conceive of the operation of positive political theory. On the account offered by Christiano, following and extending Hardin, social and political situations are seen as games in which all individuals participate in a broadly symmetric way – each pursuing her own self interest, so that the resultant equilibrium could not be said to be substantially intended or brought about by any of them. But if we replace this picture with one in which some agents act as political entrepreneurs or professional politicians in shaping the agenda and structuring the game, we might conclude that these aspects of the game’s formation were just as significant as the mere fact that the final outcome could be seen as a co-ordination equilibrium. In economic markets most individuals, most of the time, have no direct or intentional effect on the market outcome or equilibrium, but some individuals do have direct and intended effects on some markets – they are the entrepreneurs who offer something distinctive that creates, shapes or otherwise influences a market (for better or worse). In the same way, in a more structured political world it might be true that while most individuals, most of the time, have no direct or intentional impact on either policy choice or institutional reform, it might still be the case that some individuals do have such effects. This point further weakens the claim of basic structural determinism. The essential point here is that, while the

idea of co-ordination equilibria as a source of institutions is a key part of a rational choice theoretic account, it is not the only part.

This point leads to a further line of concern with Christiano's argument and, in particular, with that part of Christiano's argument that revolves around what it might mean for a positive political theory to be 'true'. Recall that the form of Christiano's argument is that if both mainstream rational choice theory and basic structure determinism are true as descriptive theories, then normative rational choice theory is undermined. But this formulation places a considerable burden on the idea of a true theory - a burden that is, in our view, unwarranted. Most obviously, the form of Christiano's argument requires that the truth of the positive theory implies that the theory is a complete and accurate account of reality. Complete in the sense that there is no systematic residual political behaviour left unexplained by the theory: and accurate in the sense that the explanation offered cannot be improved upon by recourse to additional factors. This does not mean that the theory has to be capable of perfect prediction – the theory could (and presumably should) allow for any non-systematic or random elements that may be relevant, and randomness may be an important feature of at least some aspects of behaviour¹⁵.

But we do not recognise this idea of truth as the key ambition of rational choice political theory in general or of that aspect of the theory used to support the idea of basic structure determinism in particular; and we would certainly deny that rational choice theorists must be committed to the descriptive truth of either theory in this sense. To the contrary, rational choice theory is grounded in the idea of theory as model; where a model – especially an idealised, theoretical model – is, by necessity, an abstraction from the truth. Rational choice theory – any theory – is seen not as the embodiment of truth, but as a useful perspective on reality that focuses attention on an important feature of that reality without denying that other features may also play significant roles. So, while a rational choice theorist should be committed to the idea that her theory captures a significant aspect of the truth, she need not (and should not) take the view that it *exhausts* the truth.

If positive rational choice theory is understood in this way, as a useful but necessarily incomplete model of relevant aspects of the world, with the idea of institutions as co-ordinating equilibria as a particular example of this approach to theory, further space for the normative aspirations of rational choice theory is opened up. Of course, this might be a slightly uncomfortable argument for some rational choice theorists, since it suggests that the effectiveness of their normative approach depends, at least in part, on factors not explicitly

accounted for in their positive models. But that is a different point, and one which we defer until our discussion of proposition three below.

To return to the theme outlined at the start of this section, concerning the language of constraints and feasibility and its interpretation in various settings: we suggest that, in the context of practical normative theory, constraints and issues of feasibility are most often taken to operate as logical counters or requirements which are digital in nature, but we suggest that it is appropriate to focus less on the digital, binary or on/off nature of such requirements and rather see issues of feasibility more as analogue, continuous or ‘plausibility’ counters. Once this shift of interpretation is accepted, so that the relevant debate becomes one of determining the degree to which any particular constraint binds, the importance of even mild relaxation in such constraints becomes apparent.

Proposition Two – Sauce for the goose

In this section we turn our attention to examine the relationship between the linkages connecting the positive and the normative, and the particular case of rational choice theory. After all, Christiano’s claim is that rational choice political theory is self-defeating, not that all political theory is self-defeating. The impression that Christiano’s account may give is that the self-defeatingness problem is unique to rational choice theory¹⁶. But we think all positive theories are in the same boat here: and that the relation between the normative and positive strands of political theory is essentially independent of the content of the positive account. We will proceed in several steps, first holding fixed the positive account of politics to consider variations at the normative level, and then allowing alternative positive accounts of politics.

So, initially (and despite our arguments in the previous section) we want to take as our starting point Christiano’s extreme case in which both mainstream rational choice political theory, and the hard version of basic structural determinism, are true in the strong sense. In these circumstances, it is difficult to see how any variety of normative political theory could be effective in terms of its practical aim. Of course, as was the case with rational choice normative theory, it would be perfectly possible for any variety of normative political theory to serve as an evaluative criterion. There is no bar to theorists of the relevant variety debating which actions, policies or institutional rules would be “best” according to the theory in question, but neither debates of this form, nor any other aspect of the normative theory could

be effective in *bringing about* the identified actions, policies or institutional rules. The existence of the normative theory could have no political impact.

The point here is simple enough: the full specification of the nature of the political world in descriptive terms identifies all of the factors that can be effective in influencing outcomes¹⁷ in that world. If we stipulate that the political world is thus and so, and this specification involves the claim that no normative argument is relevant in determining political outcomes, we effectively undermine the practical relevance of normative political theorising of all varieties.

It might seem that this claim is too strong, and that even a fully deterministic specification of political life would allow of normative theorising of at least some type – a type that was in some relevant sense integrated with the nature of the positive specification. This is not so, but the point usefully underlines the distinction between the evaluative and practical roles of normative political theory. In the case in which descriptive rational choice theory and basic structure determinism together fully characterise the political world, the normative theory most obviously linked to, and integrated with, this underlying positive account would be a form of Paretianism. Within this normative scheme, alternative political outcomes A and B could be compared in terms of their impact on individual lives and how well they went, with A being better than B if and only if everyone's life went as least as well in A as in B, and some people's lives went better in A than in B. And it may well be the case that, at least under some circumstances, theorists could demonstrate that the political world as specified tended to produce outcomes that were, indeed, Pareto optimal in this sense. But even if all of this were true, it would still be the case that the Paretian normative theory would be ineffective in influencing action, simply because the fact that some action (or policy or institutional rule) would be recommended by the Paretian normative theory could not count as increasing the probability of that action (policy, rule) actually occurring in the world. No individual (or group of individuals, or office holder) would be influenced by the normative theorising – even if they were fully aware of it.

The situation here is directly comparable to the standard discussion of the one-shot prisoner's dilemma. Each prisoner has a dominant strategy of defection, and this is so even on the assumption that each is fully informed about the game – including the fact that both prisoners would be better off if both co-operated. The information concerning the theoretical availability of a Pareto improvement has no weight in the actual play of the game, since the play of the game is already fully determined by other factors. We can use the Pareto criterion

to label the equilibrium of the game as inefficient, but such labels do not in themselves offer practical or behavioural routes to efficiency.

Now, notice that this general problem is in no way dependent on the positive specification of the political world being based on rational choice theory. Any positive political theory that serves to identify the causes of political outcomes, to the extent that it succeeds, limits the scope for practically effective normative political theory. To put the same point in another way, unless a positive political theory explicitly includes some means by which appropriate normative political theory may operate, so that there is a clear operational linkage between the normative theory and the relevant descriptive political theory, the positive theory will constrain and, in the limit, undermine normative theory.

At one level, this should be obvious enough. The normative theories that build on the idea of equality, or of rights, for example, could hardly be thought to be effective if everyday political debate and decision making did not also refer to equality or rights. But the issue is rather more subtle than that. It is not sufficient for the appropriate normative language to be used in the real world: that language must be effective in normative terms. To revert, for a moment, to the case of mainstream rational choice theory at the descriptive level, it may still be the case that rational actors – including rational politicians – invoke ideas of equality or of rights in political debate. But, if the narrowly self-interested idea of rationality is taken to be ‘true’, such invocations must, by definition, simply provide a rhetorical screen for more narrowly self-interested or professional motivations by those concerned. It might be that rights are generally upheld, or that equalising policies are supported, but the explanation for these facts would not lie in the existence of recognised normative theoretical claims surrounding equality or rights, but rather in the specifics of the (true, by assumption) positive model of political behaviour. Now, of course, rhetoric may actually play an effective political role – we do not mean to imply that rhetoric is necessarily empty – but we do want to argue that if rhetoric is effective, then this fact should be built into the relevant positive political theory, so that the rhetorical appeal to equality or rights might itself be explained.

Now, as we argued in the previous section, we do not believe that the rational choice model of politics at the descriptive level fully constrains normative theorising, and more generally we do not believe that the best available model of politics at the descriptive level (regardless of the extent to which this model might be characterised as a rational choice model), fully constrains normative political theorising. But we do believe that descriptive political theory substantially constrains normative political theorising, and that this constraint imposes a duty

on normative theory to demonstrate how it might be effective. This duty implies in turn that normative political theory needs to engage with positive political theory in an explicit manner.

Proposition Three – Revisionism rides again

So far, we have argued:

1. that Christiano's original case was overstated in claiming that normative rational choice theory is self-defeating, while agreeing that the content of rational choice theory at the descriptive level must constrain the effectiveness of normative theory in the rational choice tradition;
2. that the same general issue affects all approaches to political theory, and points to the generality of the positive theory/normative theory issue.

We now turn to our third and final proposition: that the move from mainstream to revisionist rational choice theory carries with it a reduction in the severity of the constraints imposed by descriptive theory, so expanding the space available for practically effective normative political theory. While the style of normative political theory that we have in mind is generally of the institutional and constitutional type already indicated as characteristic of the rational choice theory approach, our argument is more general: to the extent that something like a revisionist rational choice theory account of politics is a significant part of the truth in accounting for political behaviour, we argue that this opens up significant space for normative theory of any type, providing only that that normative theory engages in an appropriate way with the motivations of rational (in our revisionist sense) individuals in political settings.

The first and most obvious step in support of this proposition relates to the aspect of revisionist rational choice theory that incorporates somewhat moral motivations into the rational calculus of individual agents. This step directly connects the positive to the normative insofar as individuals' moral motivations line up with the structure of the relevant normative theory. Essentially, by incorporating a specifically normative perspective within individual agents we allow of the possibility that normative argument can be effective simply by persuading individual agents of what morality requires and so influencing their chosen actions.

But, once again, a possible effect is not necessarily an effect. We still have to overcome at least two further hurdles to transform the potential impact of somewhat moral motivations into an actual impact. The first hurdle is the possibility, noted earlier, that moral motivations may often be swamped by more self-interested motivations. Clearly, if moral motivations are present, but are so weak that they rarely, if ever, drive behaviour, they will not be sufficient to build an effective channel from the normative to the positive. The second hurdle is provided by the structure of the collective action problem that lies at the heart of both the argument concerning basic structure determinism (even in its softer and more acceptable form) and the more general rational choice interpretation of political life. The issue here is that, even if individuals are partly motivated by appropriate normative concerns, and this motivation impacts on their individual action, this may still not be sufficient to prove effective at the level of political outcomes. We will tackle these two hurdles in turn.

We have emphasised that our interpretation of the revisionist position places the desire to act as morality requires alongside other desires – indeed we might suggest that one could summarise desires into just a few similarly high level desires, such as the desire to act as prudence requires, supported by sets of beliefs about what morality and prudence require in particular circumstances, rather than the indefinitely long list of specific desires and preferences that are more normally considered. Such a move might seem to shift the focus of attention from desires to beliefs, but actually we think that it illustrates the difficulty in constructing a clear dividing line between desires and beliefs. To have a desire provides some support for the belief that you have that desire; and beliefs often include a degree of intentionality: for example, to believe that something is good is close to believing that you should desire that thing. Anything like a satisfactory discussion of the structure of desires and beliefs would take us too far from our major theme, but we want to register the view that both desires and beliefs may be the subject of rational scrutiny as well as being the inputs to rational choice. This is not to say that certain desires and beliefs are rationally required (although we do not rule this out) but rather that we think that desires and beliefs may be subject to evaluation against both epistemic and structural criteria.

With this in mind, and bearing in mind our earlier indication that heterogeneity of desire and belief in relation to morality should be expected, it is by no means clear that our revisionist position requires us to accept that moral motivations will normally or typically be overwhelmed by prudential or other desires. This is essentially an empirical matter and one where we will not offer any real evidence. But we will argue that the two other moves

associated with our revisionist position – the expressive and the dispositional – will systematically increase the effectiveness of moral motivations in the arena of politics, amplifying whatever basic moral motivation may be present¹⁸.

In the case of the dispositional move, the basic point is simply that the shift to the dispositional perspective removes political action from the domain of act-by-act evaluation against the full range of desires and admits the possibility that at least some actors will be committed to political dispositions that embody a moral perspective. At this point we need to say a little more about the nature of dispositions (drawing on Hamlin, 2006). We distinguish between three types of disposition: specific conditioning dispositions; general conditioning dispositions; and modal dispositions, where the key issues are the nature and scope of the commitments involved. In the case of specific conditioning dispositions, the commitment is substantive, but is essentially narrow and focussed on a particular domain – an example might be the commitment to vegetarianism where the commitment clearly indicates the course of action to be taken but the range of decision situations in which the disposition is relevant is both clear and relatively small. In the case of general conditioning dispositions, the commitment is again substantive, but the domain is broad and perhaps even universal – an example might be the disposition to tell the truth, or the disposition to be cautious. In the case of modal dispositions, the commitment is procedural in the sense that the actor commits to adopt a particular decision making perspective when confronted with choices in a particular domain, but does not commit to specific substantive actions. A moral disposition over some domain might be seen as a modal disposition in that it commits the individual to consider decisions in the relevant domain from a moral perspective, and act as morality requires.

Once the possibility of dispositions that privilege the moral calculus over the prudential calculus in a particular sphere of decision making is recognised, it remains to argue that the political domain has characteristics that lead us to believe that specifically *moral* dispositions are likely to be particularly relevant there. This argument is provided by an application of the general idea of relative prices. In the arena of democratic politics, the impact of any single individual who is not a professional politician on the overall political outcome is, at best, very limited. This fact reduces the *ex ante* prudential impact (favourable or unfavourable) of any political action. Thus, compared to a situation in which the individual is decisive, prudential considerations will be backgrounded, so that moral considerations are likely to take on greater relative prominence. And this is true both at the level of disposition choice and at the level of action. We would expect moral dispositions - of the modal type - to arise more often

in the domain of democratic politics than in other domains where individual action is more directly linked to outcomes. And we would expect moral motivations to play a greater role in action when the opportunity cost of behaving morally is reduced – as it characteristically is in large number electoral settings.

So, both in the case of uncommitted action (that is, where dispositions do not apply), and in the case where dispositions replace act-by-act evaluation, we have good reason to believe that the setting of democratic politics will tend to reduce the relevance of individually prudential considerations and increase the relevance of moral considerations as guides to political behaviour. This is not to say that politics will be dominated by moral action, or that what counts as moral action in any particular political context is unambiguous; but the argument is sufficient, we believe, to indicate that the revisionist version of rational choice theory with its combination of somewhat moral motivations, expressive behaviour and a dispositional approach to commitment can offer a coherent account of a connection between normative considerations and political behaviour.

So, we argue that we can clear the first hurdle by combining the various feature of the revisionist account of rational choice theory. In effect the dispositional and expressive arguments serve to amplify the effectiveness of moral motivations in the particular context of democratic politics so that the ‘space’ for effective normative debate is enlarged.

The second hurdle we set ourselves was that even if individuals are motivated to at least some extent by appropriate normative concerns, with that motivation amplified so that it impacts on their individual action, this may still not be sufficient to prove effective at the level of political or social outcomes. We see this hurdle as the real testing ground for both positive and normative political theory. To some extent we can refer again to our earlier discussion of the idea of political entrepreneurs, and our earlier suggestion that the idea of co-ordination equilibria does not exhaust the set of political mechanisms – but these appeals do little more than suggest that the hurdle might be overcome at least sometimes, rather than offering any clear argument. Indeed, it would seem implausibly optimistic to think that there might be some general argument to indicate that such a hurdle can always be overcome – that normative ideas once agreed and appropriately internalised can always be translated into the relevant actions and social outcomes. So here lies the irreducibly political aspect of the interrelationship between positive and normative political theory: even when the normative argument is settled and the political agents are (reasonably) appropriately motivated, still there is still the question of actually generating the normatively indicated outcome. Clearly,

institutions (as well as individual motivations) matter here, and the focus on institutional design within the rational actor tradition reminds us of the way in which thinking normatively about our institutions can help us. But just as our actual individual motivations will not be perfect, so institutions will also be imperfect. The political problem is essentially a ‘second best’ problem, and ideal theory is a poor guide.

Finale

Politics is a matter of the feasible and the desirable, with the interaction between feasibility and desirability identifying the best that can be achieved. Discussion of the feasible is provided by positive political theory (together with other physical and social sciences operating in explanatory and descriptive modes). But it is important to note that feasibility issues arise in a variety of forms; at the beginning of this essay we distinguished between the debate on ideal versus non-ideal normative theory and our own concern with the relationship between normative and positive theory. These distinctions reflect two aspects of feasibility, one concerned with what might be termed the practical or pragmatic feasibility of normative theory, the other concerned with what we might term structural or theory-feasibility of normative theory. A particular ideal normative theory may be pragmatically feasible or applicable if it is capable of generating a policy recommendation that responds to a specific real world issue. For example, a particular version of egalitarianism might be judged to be pragmatically feasible in a particular case just insofar as it produces a clear policy recommendation in respect of, say, the issue of the distribution of particular health care resources. But such a clear policy recommendation might still be theory infeasible in our terms if that recommendation did not engage appropriately with the operation of the political system, so that the fact of the normative endorsement itself did no real work (and could do no real work) in bringing about the recommended policy. This is not a matter of the technical feasibility of the policy – it might be accepted that the recommended policy could be implemented – but rather it is a matter of the political will to act as the relevant normative theory requires – a ‘will’ that may be directly the will of individuals acting in political roles, or may be the product of the institutional structure.

Our general claim is simply that issues of theory feasibility – the extent of coherence between normative and positive theory – are important, and that positive political theory (of whatever variety) therefore implies considerable constraints on normative theory if normative theory is

to be effective in its practical, action-guiding sense. While these constraints are significant, they do not, even in the case of a positive political theory that is based in mainstream rational choice theory, completely undermine the role of normative political theory.

More specifically, we have argued that the revisionist account of rational choice political theory – which generalises the mainstream analysis to allow of somewhat moral motivations, expressive behaviour and a dispositional approach to rational choice, offers a clear route by which normative political theory and positive political theory can be connected, with each constraining the other to an appropriate degree.

We finish with what we hope may be a helpful story. Imagine an anthropologist (*A*) studying a remote civilisation – the *X*. Having studied the language, *A* notes that the *X* engage in normative political debate – debate, that is, about what should be done – although *A* also notes that not all of the *X* are engaged in this debate, that the debate includes a number of rather different positions, and that the substance of the *X*'s moral debate seemed wholly alien to anything that might count as morality in her own society. Let the dominant moral code of the civilization under study be *M1* and the moral code of the anthropologist's own society be *M2*. In her role as a field anthropologist, *A* studies whether the behaviour of individuals can be explained, at least in part, by the normative code *M1*. To the extent that *M1* contributes to explaining behaviour, *A* might reasonably conclude that the *X* internalise their normative debate, so that the requirements of morality take on at least some motivational force. She might also note that this force varied across individuals and across circumstances. If, on the other hand, *A* found that *M1* played no explanatory role in understanding the behaviour of the *X*, she might reasonably conclude that normative debate among the *X* was practically ineffective, and she might go on to suggest that the normative debate itself might be understood in other terms (including formal and evaluative terms). Note that *A* is engaged in an essentially positive or descriptive analysis throughout. It is true that the subject matter of her study includes the normative debate among the *X*, but her study is not itself normative. Specifically, *M2* – her own moral code – has played no role.

We see our discussion here as related to that of the anthropologist. Obviously, we are not concerned with a remote civilization, and we cannot pretend to be external observers. Nevertheless, our basic question has been the nature of the connection between normative political debate and actual political behaviour – at the general level of abstraction at which general theories (positive and normative) are constructed. The anthropologist was concerned primarily with the connection from the normative to the positive – with the animating

question being: do normative considerations contribute to the explanation of political behaviour? And one aspect of our revisionist rational choice position is to allow normative concerns to be considered for their explanatory power within a positive model of political behaviour. But we are also concerned with the extent to which the understanding of positive political theory feeds back into and constrains normative analysis. Here then we must reverse the direction of the anthropologists study and investigate the extent to which the **X**'s own understanding of positive politics informs and constrains their normative debate and their political and moral code **M₁**. To the extent that **M₁** attempts to take proper account of what is politically feasible in making its various recommendations, this feedback from the positive to the normative is, we think inevitable. And we think that the two dimensions of relationship – from the normative to the positive, and from the positive to the normative – must be viewed together, if we are to arrive at an overall political theory that is coherent.

References

- Brennan, G. and Buchanan, J.M. (1985) *The Reason of Rules*, Cambridge: Cambridge University Press.
- Brennan, G. and Hamlin, A. (2000) *Democratic Devices and Desires*, Cambridge: Cambridge University Press.
- Brennan G. and Lomasky L. (1993) *Democracy and Decision*, Cambridge: Cambridge University Press.
- Brennan, G. and Pettit, P. (2004) *The Economy of Esteem*, Oxford: Oxford University Press.
- Brennan, G. and Southwood, N. (2007) Feasibility in Action and Attitude, in T. Rønnow-Rasmussen, B. Petersson, J. Josefsson and D. Egonsson (eds.) *Homage à Wlodeck: Philosophical Papers Dedicated to Wlodeck Rabinowicz*, www.fil.lu.se/homageawlodeck
- Broome, J (1991) *Weighing Goods*, Oxford: Blackwells.
- Buchanan, J.M. and Tullock, G. (1962) *The Calculus of Consent*, Ann Arbor: University of Michigan Press.
- Christiano, T. (1996) The Rule of the Many: Fundamental Issues in Democratic Theory, Boulder, Co.: Westview Press.
- Christiano, T. (2004) Is Normative Rational Choice Theory Self-Defeating?, *Ethics*, 115, 122-141.
- Cohen, G. A. (2003) Facts and Principles, *Philosophy and Public Affairs*, 31(3), 2003, 211-245.
- Copeland C. and Laband, D. (2002) Expressiveness and voting, *Public Choice*, 110, 351-363.
- Forbes, H.D. (2004) Positive Political Theory, in Gaus, G.F. and Kukathas, C. (eds.) *Handbook of Political Theory*, London: Sage.
- Hamlin, A. (2006) Political Dispositions and Dispositional Politics, in Eusepi, G. and Hamlin, A.P., *Beyond Conventional Economics: The Limits of Rational Behaviour in Political Decision-Making*, Aldershot: Edward Elgar.
- Hardin, R. (1999) *Liberalism, Constitutionalism and Democracy*, Oxford: Oxford University Press.

- Lewis, D. (1969) *Convention: A Philosophical Study*, Cambridge MA: Harvard University Press.
- Mason, A. (2004) Just Constraints, *British Journal of Political Science*, 34, 251-268.
- Rae, D.W. (1969) Decision Rules and Individual Values in Constitutional Choice, *American Political Science Review*, 63, 40-56.
- Schuessler, A.A. (2000) *A logic of expressive choice*. Princeton: Princeton University Press.
- Smith, A. (1776, 1981) *An Inquiry Into the Nature and Causes of the Wealth of Nations*, ed. R. H. Campbell and A. S. Skinner, Indianapolis: Liberty Fund.
- Taylor, M.J. (1969) Proof of a Theorem on Majority Rule, *Behavioural Science*, 14, 228-31.
- Tyran, J.R. (2004) Voting when money and morals conflict: An experimental test of expressive voting. *Journal of Public Economics* 88, 1645-1664.

Notes

This essay grew out of discussion at a colloquium held at Duke University in April 2005 and has also benefited from discussion at a conference in Manchester in October 2006, and at the Oxford Political Thought Conference, St Catherine's College, Oxford in January 2007. We thank participants in all of these discussions and particularly Tom Christiano, Russell Hardin, Loren Lomasky, Allen Buchanan, Doug Maclean, Jerry Postema, John O'Neill, Hillel Steiner, Jonathan Quong, Susan Mendus, Alex Voorhoeve, Chandran Kukathas, Keith Dowding, David Owen and David Miller for their comments. Normal caveats apply.

¹ For a discussion of positive political theory and its relationship to both political science and rational choice theory, see Forbes, '2004.

² Further distinctions are, of course, possible. For example, within the practical aim of normative theory we might distinguish between action-guiding and attitude-guiding roles, see Brennan and Southwood, 2007.

³ The argument that normative theorising is (or should be) independent of any issues of feasibility is put in the context of theories of justice by Cohen, 2003. For discussion and an argument that at least some 'facts' about human nature may constrain normative theorising see Mason, 2004.

⁴ Christiano exemplifies sophisticated mainstream rational choice theory by reference to Hardin, 1999.

⁵ We note that even in strictly self-interested modelling in the game-theoretic tradition, it is not individual actions that are considered as the options of choice but rather 'strategies' that may be complex combinations or actions contingent on the behaviour of others. Nevertheless, we believe that the move for 'acts' to 'dispositions' is distinct from the move from 'acts' to 'strategies'.

⁶ For a fuller account of dispositions, see Hamlin, 2006.

⁷ And Christiano agrees, later in his essay he discusses other aspects of the revisionist position; although it is not clear that he seems them as constitutive features.

⁸ We should stress that we do not see morality in sharp contrast to self-interest in all circumstances.

⁹ For more detailed discussion of the expressive argument see: Brennan and Lomasky, 1993; Schuessler, 2000; Brennan and Hamlin, 2000. For empirical investigations of expressive voting behaviour see Copeland and Laband, 2002, Tyran, 2004.

¹⁰ Note that the important aspect of basic structure determinism is its denial of agency; it is not part of basic structure determinism to deny that political structures may evolve in a manner that is stochastic rather than determinative in the statistical sense.

¹¹ We repeat that do not intend to imply that the action-guiding aspect of normative political theory is necessarily the most important aspect. We would also underline our view that most normative theorists act as if they expect or hope that their theories may have some impact in the world.

¹² For related discussion see Broome, 1991, whose ‘principal of personal good’ is individualistic in this sense since if something is good, it must be good for (at least) someone.

¹³ At least since Buchanan and Tullock, 1962, for an extended treatment see Brennan and Buchanan, 1985.

¹⁴ In private correspondence, see also Christiano, 1996, especially chapter 4.

¹⁵ So that the theory may be stochastic rather than deterministic in the statistical sense, see note 10 above.

¹⁶ This thought is encouraged by the fact that elsewhere Christiano pursues the idea of bringing positive political theory to bear on the elaboration of a normative political theory – see Christiano, 1996.

¹⁷ Here and in the following discussion, we use ‘outcomes’ in a broad sense to include procedural and other aspects of political life such as whether certain decisions were made democratically, or the extent to which rights are respected, etc.

¹⁸ A further addition to the set of revisionist ideas might be the idea of ‘esteem’ as motivator, see Brennan and Pettit, 2004.