

# 1 Introduction

Should smoking be banned? Is it right for governments to prosecute those who help the terminally ill to kill themselves? Should individuals be compelled to save for their old age? Why do most countries require their citizens to wear seat belts in a car? Why do they also require motorcycle riders to wear helmets? Should sadomasochistic sexual practices between consenting adults be made illegal? Is it appropriate for government to regulate the content of popular foods so as to tackle the growth in obesity in the population?

All these are examples of what is becoming one of the major social questions of the twenty-first century: should the government save people from themselves? More specifically, are there circumstances when the state, or the government of the state, should intervene to protect individuals from the possibly damaging consequences of their own decisions, even if those decisions affect only themselves, and even if the individuals concerned made the decisions while in full possession of their faculties and of all the relevant information? In other words, can government paternalism be justified?

The debate does not stop there. Even if it could be demonstrated that there is a case for saving people from themselves, are there not serious risks involved in allowing the government to be the agent of paternalism? Does this not create a “nanny state,” invading the autonomy of the individuals concerned and potentially infantilizing them? Or, yet worse, by legitimizing a paternalistic government, are we actually creating a potentially tyrannical state, justifying its intervention in every aspect of our lives?

Finally, even if all these risks were accepted and it was agreed that a paternalistic intervention was called for, what form should that intervention take? There are a variety of ways the government can affect individual behavior: banning or otherwise legally restricting potentially damaging behavior; taxing it; or the currently fashionable idea of “nudging” or reframing the choices that individuals face. Do all these have an equal impact on individual freedom and autonomy—or are some less dangerous than others?

This book is an attempt to answer some of these questions. It begins with issues of definition. Before discussing possible justifications for paternalism,

## 2 Chapter 1

it is necessary to specify what is meant by the term. How should it be defined? What different kinds of paternalism can be identified? The book then considers the extent of paternalism, examining some current forms of government policy to see the extent to which their rationale or consequences may be described as wholly or partly paternalistic. Following these preliminaries, the book then addresses the central question as to whether government paternalism can be justified, and, if so, in what circumstances. In light of those justifications, we then examine some aspects of paternalism in practice, or what might be termed the policy and politics of paternalism.

More specifically, chapter 2 discusses the strengths and weaknesses of the various definitions of paternalism and paternalistic policies that political philosophers and others have put forward. These definitions usually have three components: there is interference in the individual's freedom; the intention of such interference is the promotion of the individual's own good; and there is an absence of individual consent. We argue that all these components present conceptual difficulties, but the major problem is with the first. Some forms of policies that seem undeniably paternalistic, such as opera subsidies or the newly fashionable nudge ideas derived from so-called libertarian paternalism, do not obviously interfere with individual freedom. In fact in such cases, as indeed in all cases of government paternalism, the essential characteristic is the government mistrusting the individual's judgment. It does not believe that, without the intervention, the individual will make the "right" decision—"right" in terms of promoting the individual's own good, at least as the government perceives it. Without this intervention, the individual's judgment, and the behavior resulting from that judgment, will fail to promote her own good, or at least not as much as that good would be promoted through the intervention. It therefore seems preferable to define paternalism, not in terms of the intervention itself or of its consequences, but in terms of (a failure of) individual judgment; and so we propose a definition of our own that does not refer to coercion but instead incorporates this view of the government's intention. In brief, we conclude that government intervention is paternalistic with respect to an individual if it is intended (a) to address a failure of judgment by that individual and (b) to further the individual's own good.

Chapter 3 discusses some of the confusions in the literature over the different terminologies used to describe various kinds of paternalism. We distinguish between a number of different types. Of these the most important is between *ends*- and *means*-related paternalism: that is, between paternalistic interventions whose intention is to replace the individual's judgment because the government does not approve of the individual's *ends*—the aims or outcomes that he seeks to achieve—and paternalistic interventions that arise because the government perceives problems with the judgment that the individual has made concerning the *means* that are appropriate for achieving

those ends and intervenes to assist the individual to overcome these problems and thus better to achieve his own ends. This distinction is important because later in the book (chapter 5) we argue that only means-related paternalism can be justified, and that ends-related paternalism has no place in a liberal democracy.

Chapter 4 discusses the prevalence of paternalistic elements in existing government policies. It argues that many of the justifications conventionally put forward for such policies—whether derived from the economic theory of market failure that identify interventions to achieve social efficiency or from various theories of equity or social justice—seem insufficient to justify both the scale of the government intervention and the form that the intervention takes. Hence it is not unreasonable to suppose that in these cases a strong element of paternalistic motivation is involved.

The book then moves into normative territory. It examines whether paternalism can ever be justified, and, if so, in what form and in what circumstances. Chapter 5 discusses arguments over paternalism derived from considerations of individual well-being. It points to an increasing volume of evidence from behavioral economics and psychology of what we term “reasoning failure”: the fact that individuals, in trying to achieve the end of improving their well-being, often make mistakes and do so in a systematic way. It considers four possible sources for such failure: limited technical ability, limited experience or imagination, limited willpower, and limited objectivity. The existence of these forms of reasoning failure means that there is an opening for *means-related* paternalism; that is, for paternalistic interventions that improve or replace the means by which individuals obtain their ends.

However, there is no similar accumulation of evidence that individuals make mistakes over their ends; that is, over the factors that contribute to their well-being. Indeed, since such ends are essentially value-driven, it is hard to see what form such evidence might take. Hence chapter 5 concludes by rejecting *ends-related* paternalism—that is, paternalistic interventions designed to change or to replace individuals’ ends or aims—but accepting the well-being case for *means-related* paternalism—that is, paternalistic interventions designed to help the individual to achieve her own ends when she does not have the means to do so as effectively herself.

Chapter 6 considers what is perhaps the major objection to all forms of paternalism, including means-related: that it harms or inappropriately restricts individual autonomy. This is often characterized in terms of the “nanny state”: the state is seen to treat its citizens as a nanny treats her charges, instead of as autonomous adults. If a paternalistic policy has a deleterious impact on an individual’s autonomy, then this adversely affects her citizenship rights. In addition, psychological theory suggests that it may also damage her well-being and her intrinsic motivation to change her behavior in

#### 4 Chapter 1

the areas affected. At the extreme, the argument asserts that treating people like children turns them into children.

The chapter assesses the arguments of the “soft paternalists” who endeavor to overcome these challenges by arguing that the individuals affected in fact have little autonomy to be violated. We demonstrate that, in almost all cases of paternalistic interventions, there is indeed an impact on autonomy, actual or perceived, and the nanny state challenge cannot be avoided in this fashion. However, unless an individual’s autonomy is regarded as an absolute right never to be violated, that is not the end of the argument. Rather, we need to trade off how much we are willing to allow the government to intervene in people’s autonomy against the amount of good (in terms of well-being) that can be promoted or harm prevented as a result of an intervention. This judgment in turn rests on the extent to which we believe that people actually fail to make adequate judgments in their own interests. This will inevitably result in the need to trade off the value of well-being against the value of autonomy in different situations. If a specific means-related paternalistic intervention delivers a large gain in an individual’s well-being with only a minor infringement of the individual’s autonomy, then the intervention is probably justified; but one involving a small gain in well-being but a severe diminution of autonomy is likely to be unacceptable.

Chapter 7 focuses on a significant recent development that relates directly to the trade-off between well-being and paternalism: that of so-called libertarian or asymmetric paternalism and the associated nudge policies. These are government interventions that seek to change the context in which people make choices—the “choice architecture,” in the term of their principal proponents, Richard Thaler and Cass Sunstein (2008, 3)—so as to nudge them to make decisions in the direction that the government wants. Examples include the automatic enrollment of employees in pension plans, so that individuals who do not wish to participate in these plans have to make a conscious decision to opt out of them; an opt-out organ donation scheme where, instead of people having to carry a card to indicate their willingness to donate their organs in the event of a fatal accident, they have to carry a card to signal that they are *unwilling*; and the positioning of healthy foods in a cafeteria so that they are the first item that customers encounter, not the last. All these can lead to substantial changes in individual behavior, with far more employees saving appropriately for their old age, with many more organs becoming available for donation, and with people eating more healthfully. Yet they appear to achieve these changes without affecting autonomy or freedom, since they leave the actual choices that people face untouched.

Critics of nudge policies have disputed the contention that they have no impact on autonomy, claiming that they work best when they are unper-

ceived, and hence that they involve a degree of trickery or deception that inevitably reduces autonomy. We address these criticisms in the chapter, arguing that they can be partly resolved by introducing various transparency mechanisms. We conclude that at least some nudge policies can indeed significantly raise well-being and can do so with only minor infringements of autonomy.

Ultimately, the usefulness of all these arguments (as indeed of any philosophical argument) will depend on how they “cash out” in practice. Chapter 8 assesses actual paternalistic policies, some of which are already in place, some of which are proposed, against the criteria we have put forward to assess: the impact on well-being and on autonomy. We examine three areas where paternalistic interventions could be (and/or have been) considered: smoking, pensions, and assisted suicide. In the cases of pensions and smoking, we argue that there is evidence of significant reasoning failure and that therefore some form of intervention is justified, provided that the impact on autonomy can be minimized. Of the possible forms of intervention in those areas, we consider legally restrictive interventions, financial incentives, and libertarian paternalistic proposals. We conclude that restrictions tend to score badly overall, but that financial incentives and the opt-in/opt-out plans score well. With respect to assisted suicide, we argue that there is relatively little evidence of reasoning failure for the individuals concerned, and hence that paternalistic intervention to prevent assistance is not justified.

Chapter 9 considers what might be termed the politics of paternalism. In previous chapters we demonstrate that there is a case for paternalistic intervention by the government to address individual reasoning failure, and we describe ways in which this may be done to maximize the benefits of intervention in terms of improving individual well-being while minimizing the cost in terms of the impact on individual autonomy. However, these contributions on their own are not enough to provide an unanswerable case for paternalistic interventions in every situation of individual reasoning failure. For that would require demonstrating that, in the relevant circumstances, the government *can* make better decisions than the individual, and also that it *will* do so. Neither of these is obviously correct. The government, after all, is not some abstract benevolent entity but is itself a collection of individuals—politicians, civil servants, and advisers—who interact with one another in a variety of ways. These individuals are likely to be subject to the kinds of reasoning failure that we have previously ascribed to some of the people engaging in self-damaging behavior. Even if they are not subject to such failures, they may have their own agenda, being rather more concerned with maximizing their own well-being than with the well-being of the citizens whose interests they are supposed to be serving. In terms of a metaphor that one of us has used elsewhere, some may be self-interested knaves, not public-spirited knights (Le Grand 2006).

## **6 Chapter 1**

Chapter 9 argues that the government can indeed raise the well-being of individuals who suffer from reasoning failure, even when allowance is made for possible reasoning failure among those individuals who constitute the government. However, democratic mechanisms must be put in place to ensure that the latter do not pursue their own agenda and turn the paternalistic state into an instrument of authoritarianism. In particular, we argue for a retrospective endorsement of the policy concerned, with either a vote taken in the representative assembly or a referendum.

The final chapter summarizes the book's arguments and uses them to address what might be viewed as the central questions with which we have been grappling: Is a paternalistic government necessarily a nanny state that infantilizes its citizens and illegitimately erodes their autonomy? Or could it be a helpful friend that promotes their well-being at minimal, if any, cost to autonomy? For the answers to these questions, read on.