

NEW ON PATERNALISM AND PUBLIC POLICY

THOMAS C. LEONARD

Princeton University

ROBERT S. GOLDFARB AND

STEVEN M. SURANOVIC

George Washington University

Bill New's (1999) thoughtful paper has performed the valuable service of clarifying the meaning and the policy implications of paternalism. His careful formulation delimits the domain of justified state paternalism. Having argued successfully, in our view, for a narrow ambit, New proceeds to identify situations that justify paternalism. This comment is written in the spirit of a friendly reformulation that refines and improves the specification of when paternalism is justified. Our argument is two-fold. First, we argue that New's formulation, properly understood, will not readily permit the paternalistic interventions he argues are justified. Second, we identify a class of potentially justified interventions that have paternalistic aspects, but which are neither strictly paternalistic nor market-failure remedies.

New defines paternalism as the non-consensual interference in the self-regarding decision making of an autonomous person, which is intended to further the person's welfare, where autonomous persons are adults not incompetent, incapacitated, nor under coercion (1999, p. 65). New's definition *excludes* as non-paternalistic the following kinds of policies:

- (1) consensual interferences, as when a taxpayer invites the government to over-withhold her income taxes as a means of 'forced savings';
- (2) non-consensual interferences with persons not deemed autonomous, such as the incompetent and the incapacitated, also known as 'soft' paternalism;

- (3) non-consensual interferences that do not intend to further the paternalized person's welfare, such as imprisonment of criminals or eminent domain takings;
- (4) non-consensual interferences where the paternalized person's decision is not self-regarding in that it imposes costs and benefits on other persons, such as Pigovian taxes on polluters;
- (5) policies that may influence an autonomous person's self-regarding decision making, but that do not non-consensually interfere, such as the provision of information concerning health or safety risks.

Two important ideas follow. First, as underscored by the last two examples, truly paternalistic policies are distinguishable from interventions intended to correct market failures, *including* market failures that arise from insufficient information. The non-paternalistic state, for example, can provide information on the health risks of ingesting saturated fats, but it cannot intervene to protect informed adults from gorging on steak and ice cream. An inadequate supply of information, as with all market failures, is the product of missing markets (New's term is 'deficient exchange'), not of inadequate reasoning. There is nothing inherently deficient in the ability of beef eaters to assess their own well being, so the 'intervention' of supplying missing information is non-paternalistic.¹

Second, a paternalist is logically required to believe that the intervener is *better* placed than the paternalized person to judge the latter's welfare. It is not enough to show that individual reasoning is inadequate, one must also show that the intervener, or its agents, can and will do better. Clearly, interveners may be unable or unwilling (or both) to do better. Following New, we focus primarily on whether interveners can do better, leaving aside the further matter of whether interveners will do better.

Having set out this formulation, New identifies three situations where an individual's reasoning can fail, a necessary though not sufficient condition for paternalistic intervention. They are: (1) inter-

¹ To defend New's distinction between information market failure and paternalism is beyond the scope of this comment, but, by way of illustration, consider the following example, where a paternalist and a liberal (non-paternalist) view information in a fundamentally different way. The case is *disinformation*. A paternalistic public health official who believes that even well informed adult smokers demonstrate a failure of reason might endorse a policy of greatly exaggerating the risks of smoking, as a means to the end of smoking reduction. Since the paternalist's goal is not to inform *per se*, but to promote paternalized persons' welfare as she sees it, disinformation can be consistent with paternalism. For the liberal official, in contrast, a policy of disinformation is incoherent, because the liberal's goal is to inform – to provide missing information – not to produce a particular outcome. For the liberal, fraud, even well intended fraud, cannot be justified, as it cannot improve self-regarding individual choices.

temporal choice settings, where the individual faces a weakness-of-will problem, that is, may succumb to temptation that is not in his long-term interest; (2) decision-making situations where the individual has little first-hand experience of potential consequences, especially low-probability events, and (3) complex decision-making situations, especially where technical expertise is an advantage.²

1. WEAKNESS OF WILL

Akrasia, or weakness of will, most often refers to a voluntary, purposeful action that conflicts with a better judgement contemporaneously held. For judgements with respect to welfare, the *akratic* person prefers x to y and yet chooses y . *Akrasia* arises most commonly in intertemporal settings, where a prior decision requires resolution in the face of temptation. At breakfast you resolve to diet. Your plan for dinner is thus no dessert rather than ice cream. At dinner your resolution fails and you eat ice cream.

One strategy for dealing with correctly anticipated weakness of will is pre-commitment. Ulysses lashes himself to the mast because he anticipates that he will otherwise succumb to the Sirens' song. This outcome is inferior to unimpeded listening to the Sirens while sailing past, but the sophisticated Ulysses knows that, once his preferences change, the first-best outcome becomes unattainable. A myopic Ulysses fails to anticipate weakness of will, and thereby wrecks his ship upon the rocks. In principle, then, the *akratic* person can be made better off by a limiting of her choices, which opens the door to paternalism.³

The paternalist, however, confronts two related information pro-

² New identifies a fourth category, emotional decision making, wherein strong emotion overwhelms rational decision making. Unlike the other three categories, New does not pursue this case further, nor do we beyond the following remarks. First, some instances of 'emotional choice', such as immediate gratification (i.e., where appetite is overwhelming) can be seen as a subset of the weakness-of-will case. Indeed a common reading of *akrasia* is that feeling overpowers reason: Ulysses's expected inability to steer his ship past the rocks can be seen, on this reading, as due to the Sirens' song, which emotionally overwhelms him. Second, for other instances, it is unclear what kind of interventions excessively emotional decision making would merit, given that emotional choice, by its nature, is unpredictable.

³ On one still influential reading, which dates to the Platonic dialogues, pre-commitment is irrational. If you truly prefer no dessert, then forgo the ice cream. If you truly prefer ice cream, then do not forgo. Indeed why foreclose some options? There are two possible answers. One, discussed in note 2, is that the choice is really no choice at all. The desire to eat ice cream is so overwhelming as to be effectively non-volitional. The other possibility, more persuasive, is that preferences change between decision and execution. The dinner-time self, unlike the breakfast-time self, prefers ice cream. Many rational choice theories insist that preferences be stable – in economists' jargon, they must be dynamically consistent – which rules out weakness of will, and obviates the pre-commitment strategies designed to help overcome it.

blems: how to distinguish the weak willed from everybody else, and, among the weak willed, how to distinguish sophisticated from myopic *akratics*. Consider the second problem first. The sophisticated *akratic*, such as Ulysses, knows that she faces weakness of will and has good reason to adopt pre-commitment strategies. Thus, billions are spent on fat farms, smoke-ending clinics, drug rehabilitation, fitness trainers, over-withholding of taxes and other less formal pre-commitment devices, such as buying cigarettes by the pack, avoiding streets with taverns, keeping sweets out of the house, and locating the alarm clock across the room. (Schelling, 1984).

It is certainly true that the state, by virtue of its monopoly on legitimate coercion, enjoys a kind of credibility that less powerful, contractual interveners do not. But there is no a priori reason to suppose that draconian sanctions are required to overcome weakness of will. The aim is to sufficiently raise the cost of temptation and to lower the cost of resoluteness. So while it is possible to imagine instances where only state-based coercion is sufficient – say, for the weakest of the weak willed – this by itself does not justify state paternalism.

Our own view is that the sophisticated *akratic* is best placed to select among and to ‘fine tune’ the many possible methods of self-restraint. A competitive ‘pre-commitment industry’ can, for the usual mundane reasons, be more responsive to their customers’ varying needs than can the state. The smoke-ending industry, for example, offers a variety of different strategies – scare tactics, meditation, hypnosis, discrete nicotine replacement therapy (gum), automatic replacement therapy (patch), psychoanalysis, gradual reduction, cold-turkey reduction, etc. Government policy is, by its nature, less flexible. A one-size-fits-all excise tax on cigarettes, for example, may have the virtue of deterring some youth smoking, but it is a blunt instrument that imposes large and regressive burdens on the vast majority of (U.S.) smokers who are not young and are less likely to be deterred.

None of this argues that state intervention *must* be inferior to alternative interventions, least of all for the weakest of will. But knowingly weak-willed individuals are a heterogeneous group and they therefore benefit from having choice among methods of self-restraint. More importantly, sophisticated *akratics* are better placed than the state to choose appropriate methods of self-restraint.

Now consider the myopic *akratic*, the individual who does not know that weakness of will is at work. This situation looks more promising for a paternalist, since the individual will not make self-restraining provisions. But if the myopic *akratic* cannot anticipate weakness of will, how is it that an intervening state can?

Preference-revelation schemes are of no use here, since the individual is not lying, she is ignorant. The state can presume that all

individuals who make choices with short-term benefits and long-term costs – for example, smoking, drinking, exercising or saving too little – *demonstrate* weakness of will. But such a presumption is clearly wrong, and therefore, would visit costs on all persons who are not weak of will or who already have self-control mechanisms in place. Broad state interventions risk *target inefficiency*, the danger that a policy will harm those it does not intend to affect, and miss those it seeks to affect, about which more in the next section.⁴

The sophisticated *akratic* knows that pre-commitment will save her from herself, so she will already have contracted for it. Additional state restraint will make her worse off. If the sophisticated *akratic* opts to employ the state – for example, over-withholds income taxes as a means of ‘forced savings’ – this is consensual and not paternalistic. Myopic *akratiks* are candidates for paternalism, but since they do not know this, it is difficult for the state to identify them, and any policies designed to reach them are therefore going to be broad-based, which will risk high target inefficiencies.⁵

2. UNFAMILIAR OR COMPLEX DECISION-MAKING SETTINGS

We grant New’s point that even individuals with complete information will sometimes make poor choices. This is one reason why individuals hire experts to help with unfamiliar or complex decisions, such as medical, legal or investment decisions. They consult experts not merely for information as an input to a decision (is the bone broken?), but also to make decisions on their behalf (what should be done to treat it?).

As before, it is useful to distinguish those who know they require expert help from those who do not. In the former group, it is reasonable to suppose that such individuals willingly abide by an expert’s decision. After all, we cede decision-making power only on the premise that the expert is indeed better placed. This is consensual, and therefore not paternalism as defined.

Okay, but what about those who *do not know* that they do not know enough, who lack the savvy to consult expert opinion? Surely such people, those who do not use seat belts, for example, are candidates for

⁴ The term ‘target efficiency’ derives from the public economics and policy literature, and refers to the extent to which actual policy results coincide with desired policy results. See, for example, Weisbrod [1970] 1977. Clearly, all policies, not only paternalistic policies, risk target inefficiency. But the point applies with special force here, because there is no incentive scheme that can induce myopic persons to reveal what they do not know.

⁵ Bill New has suggested to us that some persons may be uncertain about whether they will behave *akratically*. If pre-commitment expenditures are sunk, then the potentially *akratic* person may prefer not to pre-commit, gambling that her future self will prove resolute. This scenario turns on the extent to which pre-commitment expenditures are indeed sunk, and upon the relative likelihood of resolution and weakness of will, respectively.

compulsory intervention? New's point here is that, even with complete information on the risks of driving unbelted, some drivers still make the wrong choice, and the state can make them better off by compelling the right choice. We see a difficulty with this analysis, similar to that in the weakness-of-will case.

The first problem – target inefficiency – follows from the state's information problem. How can the state distinguish individuals who know they need expert help from those who do not? If it cannot successfully discriminate, which seems likely, universal interventions to benefit the ignorant – ignorant not of the relevant information, but of the fact that they are better off having others choose for them – risk injuring the non-ignorant.

Consider alcohol. Prohibition in the United States presumably had the virtue of protecting latent alcoholics – including those well-informed about the risks – from themselves. But this gain came at the cost of injuring others who were denied access to alcohol, in addition to violence, corruption and other costs that attend black markets. So was prohibition *qua* paternalism justified?

The answer depends on the benefits and costs, and that is the point. Paternalistic policies, like all policies, have costs. Even if it is unambiguous that prohibition benefits a class of paternalized individuals, it does not follow that the policy is justified, most clearly where there are large target inefficiencies. In sum, then, a justified paternalistic policy must meet not only the 'better placed' test, it must also do more good than harm.⁶

This amendment to New's argument sheds light on paternalistic interventions that are popular even in liberal societies. Take mandatory seat belts. Let us stipulate to New's claim that most people view seat belt laws as 'a reasonable restriction on liberty' (1999, p. 63). The puzzle is why this paternalistic policy is tolerated, while others are not. Is it solely because, as New implies, the decision to use seat belts involves consideration of low-probability events, and that even well informed people will systematically fail to reason properly? We think not, and this points to a second difficulty.

There are many decision-making settings similar to the decision of whether to buckle up, because novelty and complexity are ubiquitous. But the number of paternalistic interventions deemed reasonable is far smaller, a difference we find suggestive. We want to argue that the case for mandatory seat belts does not rest on the fact of complex or

⁶ These considerations clearly do not exhaust the policy issues at stake with prohibition. Prohibition may well have been motivated more by moralistic than paternalistic concerns, and there are, as with drunk-driving, externality considerations as well. Consistent with New, our focus is on the paternalistic aspects of state intervention.

unfamiliar settings alone. Belts are deemed a reasonable restriction on liberty because the social costs of the intervention – costs borne by the target group, costs borne by the non-targeted, and the costs of enforcement – are seen as especially small relative to the benefits.⁷

Mandatory belts are deemed low cost, because, for the target group, the cost of compliance is seen as low. The failure to buckle up is taken as probative evidence of a failure to reason precisely because compliance is seen as low cost. And, importantly, target inefficiency is also likely to be small, given that the non-targeted group is already buckling up. We can assume enforcement costs are modest.

The low-cost presumption clearly will not be correct for all persons; some adults, perhaps those who experience significant discomfort, can rationally opt to ride unbuckled. We are not arguing that the low-cost presumption is correct, only that the rationale for paternalistic intervention involves a pragmatic belief that it is. The case for compulsory seat belts, we are suggesting, goes forward because the social costs of such a paternalistic policy are deemed low, a claim which is narrower than New's suggestion that the case for compulsory seat belts is justified by a complexity or unfamiliarity alone.

To see this, consider a similar decision – to drive or not – that is more complex than the decision to buckle up or not. Driving is quite risky; about 40,000 people die annually on U.S. roads. But no one considers, much less proposes banning driving, or mandating tanks. Given the greater decision-making complexity, why not? We suggest that, unlike seat-belts, it is because the costs are deemed too high, indeed, high enough to justify a yearly toll of 40,000. The target class (those who travel in cars) would clearly bear enormous costs with a ban on driving, and the non-targeted class (those who do not travel in cars) would face vastly higher prices for most goods. The enforcement costs would be similarly high. Notwithstanding the spectacular benefits – 40,000 lives saved annually – a ban on driving is simply not worth it.

The policy dangers here are of the slippery-slope variety. Between interventions deemed *prima-facie* reasonable (mandatory seat belts) and interventions deemed *prima-facie* unreasonable (a ban on driving) lie the more difficult intermediate cases. Consider prizefighting. Some states ban prizefighting on the model of mandatory seat belts – a restriction on liberty justified by a failure of reason, with high costs and low benefits – who would rationally risk brain damage for a modest paycheck? Yet other states permit prizefighting, implicitly acknowledging that while the costs are indeed high, so too, unlike seat belts, are the benefits of risking those costs. Prizefighting offers rewards to skillful

⁷ To these welfare costs one might also add, as New suggests, the cost entailed by the diminution of freedom itself.

pugilists; not all boxers become punch drunk; and thousands of spectators enjoy a prizefight. Whether intermediate cases such as prizefighting are closer to the mandatory-belt model or to the driving-ban model will depend upon the costs of paternalistic intervention.

3. REMEDIAL PATERNALISM

There is a final point we think worth making. Here we offer a potentially justified state intervention that is not paternalism as New defines it but neither is it soft paternalism, nor does it involve market failure or redistribution. Consider a child neglected or abused by its parents. Is the state justified in intervening on behalf of the child?

Perhaps, yes. With children, paternalism is a given, since they are not autonomous. Ordinarily, it is supposed that parents are the *best* placed among alternative interveners. This latter assumption obviously is challenged when parents fail to act paternally, giving rise to the present question.

What makes this case interesting is that state intervention here has a paternalistic aspect to it, but it is neither hard nor soft paternalism. It is not 'soft' paternalism alone, because, while motivated by the child's welfare, such an intervention is designed to restrain the parents. The state intervenes not to protect the child from itself but from its parents. But it is not hard paternalism either, because the parents' decision making did not necessarily entail a failure of reason. On the contrary, the neglect might well have been rational from the parents' perspective. The state may be involved precisely because the parents have been excessively self-regarding.

And neither is there a market failure. Child abuse is generally considered a crime, not a tort. Abuse, unlike, say, pollution, is not merely the adverse byproduct of an otherwise beneficial process. We need another term for this category of justifiable intervention, perhaps 'remedial paternalism'. It may be helpful to consider some other examples that may call for remedial paternalism, the better to consider its potential scope.

(1) In some jurisdictions children cannot legally travel in cars without being placed in special safety seats. (2) Some religions (e.g., Christian Science) mandate prayer instead of conventional medical treatment for certain life-threatening illnesses. It is sometimes argued that, where a child has such an illness, the state should require treatment against the parents' wishes. (3) There are studies that find a correlation between parental smoking and childhood asthma and other respiratory diseases. Some advocates of smoking regulation argue that the state should therefore intervene to restrict smoking by adults in households with children. In listing these examples, we do not necessarily endorse

intervention. Our point is to show that 'remedial paternalism' comprises a set of non-trivial cases where a justified intervention has a paternalistic aspect and does not involve market failure or redistribution.

The last example of our list reminds us that interveners' estimates of cost must also be reliable. As it happens, the epidemiological literature on the dangers of environmental tobacco smoke is controversial, because its findings are largely inconclusive. Since an intervention is always costly in some measure, it is obviously important to have well founded estimates of its expected benefits.

4. SUMMARY

Our arguments essentially offer two amendments. New's formulation notes that paternalistic policies are distinct from market failure remedies, and are justified when the state is better placed than the paternalized individuals. Our first amendment requires that the state also be better placed than other interveners. This condition is less likely to be met when individuals are sophisticated – they know that they need help with self-restraint or with decision making.

When individuals do not know that they need help – they fail to anticipate weakness of will, or fail to reason successfully – state intervention requires distinguishing such persons from others, some of whom are behaving quite similarly. Such distinctions are difficult in practice because the members of the target group are ignorant of the fact that they need help. As such, any state interventions are likely to be broad-based, which gives rise to our second amendment: that benefits to targeted persons exceed intervention costs, which are costs imposed on non-targeted persons, costs imposed on targeted persons, and enforcement costs. We can be most confident that the state is better placed when the scientific consensus about costs is well established, and when the estimated costs of intervention are low. Finally, in addition to these amendments, we identify a kind of potentially justified state intervention, what we have called 'remedial paternalism'.

REFERENCES

- New, Bill. 1999. Paternalism and public policy. *Economics and Philosophy*, 15:63–83
- Schelling, Thomas S. 1984. Self-command in practice, in policy, and in a theory of rational choice. *American Economic Review: Papers and Proceedings*, 74(2):1–11
- Weisbrod, Burton A. [1970] 1977. Collective action and the distribution of income: a conceptual approach. In *Public Expenditure and Policy Analysis*, 2nd edn., pp. 105–31. Robert H. Haveman and Julius Margolis (eds.). Rand McNally