

# The welfare state and the rise of paternalism.

Gilles Saint-Paul  
Toulouse School of Economics<sup>1</sup>

## Introduction

We live in increasingly paternalistic societies; almost every day, somewhere in the developed world, a new law regulates what people can eat, drink, smoke, view, or read<sup>2</sup>; and transactions in goods, financial and labor markets are restricted, in many cases to protect one of the parties against its own mistakes. In my book *The post-utilitarian society* (forthcoming, Princeton University Press), I argue that those policies are partly influenced by parallel developments in the social sciences.

The traditional intellectual underpinning for government policies has been utilitarianism. I define utilitarianism as the combination of two ideas. First, policies should maximise a social welfare function that is an aggregate of individual preferences. Second, individuals act on their own so as to rationally maximize their objective function—thus, this objective function defines their preferences and at the same time drives their choices, and social welfare is an aggregate of those individual objective functions.

This approach allows for a large degree of redistribution, and other types of interventions as well if there are market failures such as externalities. But it puts restrictions on what government can do, in the sense that its actions cannot contradict individual preferences; government cannot impose on an individual something that he or she does not want, unless this allows the social planner to give some other agent what he wants, in such a

---

<sup>1</sup> This paper was prepared for the Mont Pèlerin Society congress, Stockholm, August 2009

<sup>2</sup> Many examples are discussed in David Harsanyi's *Nanny State* (Harsanyi, David (2007), *Nanny State*, New York: Broadway books)

way that the latter ‘gains more’ than the former, using the metric defined by the social welfare function.

Utilitarianism rests on revealed preferences, which makes it compatible with individual responsibility. Revealed preferences mean that if an individual chooses an action, he prefers it to any other available one. Consequently, the government would reduce welfare if (everything else equal) it were to paternalistically dictate what the individual can do, say with his own money. And, importantly, a seemingly undesirable outcome cannot be deemed unacceptable if it is the result of the individual’s own choices and if another available choice would have ruled it out. In other words, individuals are responsible for the consequences of their own decisions.

The view that individuals act rationally so as to maximize a unique, well-defined, objective function has been increasingly challenged by the social sciences. Even economics, which traditionally stuck to rational utility maximisation, is now increasingly viewing the individual actions as the outcome of a game played by multiple selves whose objectives are inconsistent with one another.

A prominent example of the multiple selves perspective is the so-called “hyperbolic discounting” approach to intertemporal preferences. Under the traditional view, intertemporal utility is restricted in such a way that, if the individual reoptimizes at some future date, he will decide to carry through his original plan, i.e. to enforce the choices he had made at a previous date. Under the hyperbolic discounting approach, the individual instead wants to renege on his past choices; this means that he has different incarnations at different dates, and that the utility function of the future incarnations is inconsistent with that of the current one; the current one wants to make different choices for future actions than the future incarnations who will be around when those actions are undertaken. Typically, proponents of hyperbolic discounting assume that the present is ‘salient’, which means that my current incarnation puts a higher

weight on today's consumption relative to tomorrow's consumption, that on tomorrow's consumption relative to my consumption the day after. That assumption then generates results such as 'undersaving'; for example, my incarnation tomorrow will want to save less for the day after, given its income, than I would like it to do. The approach then justifies policies such as compulsory contributions to public pension schemes, and so forth.

The view that individuals behave inconsistently, coupled with the utilitarian ideal of a state which maximises the sum of individual welfares, leads to what I call the post-utilitarian society. Because individual rationality no longer holds, in a post-utilitarian society, welfare can be increased by imposing actions on individuals that they disagree with (other than and beyond the sheer redistribution of income); furthermore, the whole foundation for individual responsibility collapses. The purpose of this paper is to discuss the consequences of such a collapse for the welfare state.

### **The welfare state and utilitarianism**

Traditional utilitarianism delivers intellectual foundations for the welfare state. This is because the utilitarian social welfare function aggregates preferences in such a way that it is typically desirable to redistribute resources between individuals. Utilitarianism means that the happiness of different individuals may be measured using a common currency – the marginal utility of the social planner. If, by that measure, individual A's consumption makes him or her happier, at the margin, than individual B, then total welfare is increased by redistributing consumption from B to A. Typically, this will be the case if A is poorer than B, as long as A and B are treated symmetrically, and their utility is concave, meaning that an additional unit of consumption increases utility by less if one is richer. Thus, the marginal utility of consumption makes the rich less happy than if that unit were consumed by a poor; this is the

central mechanism by which a utilitarian social welfare function provides some foundations for a welfare state. Another justification lies in providing insurance against adverse income shocks. In principle, though, private markets can play such a role; and even though insurance markets may function poorly because of informational problems, those can be solved by regulation rather than direct government provision. But some shocks are uninsurable by the market, such as innate medical conditions; in such cases, the government has to provide insurance – but in fact that form of insurance is very similar to redistribution.

Once a welfare state is established, it faces a crucial problem: How should one deal with people who put themselves in need out of their own choices? Not only some people would do so absent a welfare state due to their own preferences, but redistribution clearly increases the incentives for such a behaviour since it alleviates its consequences – this is what we call the *moral hazard* problem.

The moral hazard problem raises ethical issues and economic issues. The ethical issue is: are people who are in need as the outcome of their own decisions worthy of ‘our’ concern? The economic issue is: how should the welfare state be designed to deal with moral hazard?

Because of the assumption of individual rationality and its corollary, revealed preferences, traditional utilitarianism provides a strong answer to those issues – and that answer is that individual responsibility should be enforced.

Consider first the ethical issue. We want to redistribute in favour of those who are worse-off. By revealed preference, somebody who is in need due to his own past choices cannot be worse-off than if alternative choices had been made. Therefore, he does not deserve public support more than anybody who was initially in the same situation but has made different choices. By that metric, it is even immoral to tax a minimum wage earner to support, say, a person who was originally wealthier but became indigent because of some addiction such as drugs. In other words, welfare has to be compared between individuals not at a point

in time, but in an intertemporal fashion. Those who are poorer now because they saved less in the past, enjoyed a greater consumption then; they are not worse-off in terms of their total intertemporal flow of utility, compared to those who made thriftier choices.

Consider now the economic issue, i.e. that of designing the redistributive system. Standard utilitarian economics<sup>3</sup> has extensively studied the moral hazard issue associated with the welfare state in a branch of the literature which we may call “dynamic social contracting theory”. This research says that to mitigate the moral hazard problem one should condition current redistribution to the whole history of the individual’s past outcomes, so as to give that individual the strongest possible incentives to undertake the right actions. This implies imposing high future penalties for ‘inappropriate’ behaviour today; thus, for example, having saved too little today would typically reduce one’s future claims to welfare, everything else equal. The reason is that we can give people a high *total* level of *expected* utility while at the same time imposing very low utility in *some* future situations, provided they are compensated by sufficiently high utility in other situations. This allows us to impose very large taxes on the outcomes that are likely to result from the actions we want to deter, and thus to approximate the first best allocation of resources quite well.

This point is studied in a paper by Atkeson and Lucas (1992). They consider a world where people are sometimes “needy” and should consume more when they are in such a situation. If the government could directly observe whether people are needy it would just redistribute money to them. The problem is that everybody then has an incentive to claim that they are “needy”, and would do so if that condition were not directly observable. The authors show that the optimal social contract involves redistributing in favour of the needy today in exchange for a permanent reduction in their subsequent utility, implemented through a permanent reduction in their claims to the welfare state. While the setup is somewhat abstract,

---

<sup>3</sup> See for example Atkeson and Lucas (1992), Fahri and Werning (2005), Grochulski and Kocherlakota (2007).

this can be interpreted as a welfare state that would redistribute less to those that have saved less in the past, since saving less is indeed correlated with being in need. In their model, the unluckiest individuals – those that have experienced a sequence of bad shocks, become poorer and poorer and inequality widens over time; and this occurs despite the presence of a ‘welfare state’ in the form of a utilitarian social planner who wants to insure the representative individual against adverse shocks. Tolerating these bad outcomes is optimal because they are discounted by their low probability in the representative individual’s evaluation of his welfare and their counterpart is that the moral hazard problem is solved, which means that effort and productivity are high and so is the amount of money that can be redistributed.

By conditioning on past actions, the utilitarian social planner discriminates between different kinds of “needy” agents; those who have overconsumed in the past are entitled to a lower level of support, in order to enforce the incentive not to end up on welfare on purpose. This can be interpreted as the optimal design of social insurance enforcing a certain level of individual responsibility.

## **The post utilitarian welfare state**

Clearly, modern welfare states do not resemble the kind of dynamic social contracts we have just described. On the one hand, individual responsibility is not enforced and there is little conditioning on past actions. For example, eligibility to assistance income, in most countries where it exists, is unconditional. Similarly, people who engage in hazardous behaviour enjoy the same public health coverage as the others. On the other hand, instead of welfare entitlements being conditioned on past actions, we observe preventive measures that

restrict those actions, and such restrictions typically apply also to those individuals who are not eligible for the corresponding welfare entitlements, or who are willing to forfeit them in exchange for a relaxation of those constraints. Hence we observe paternalistic policies such as taxing and banning smoking or imposing compulsory participation in public pensions.

In my view, the prevalence of those policies illustrates the post-utilitarian nature of the principles guiding modern redistribution: Under such principles, the self who is being punished for putting himself voluntarily out of jobs by not receiving unemployment benefits, or into indigence by not saving enough in the past, is not the same person as the one who did these harmful things in the past. He is instead suffering from an “externality” imposed on him by the decisions of his own past incarnation<sup>4</sup>. It then makes sense that social insurance covers such an adverse shock. If my past self is different from my current self, there is no difference between that situation and having lost one’s job involuntarily due to the boss’s decision to close the plant—the current self is simply not responsible for the decisions of the past self.

Under the traditional, rational view of the individual, somebody who saves little, thus putting himself in a needy position in the future, is just revealing a strong preference for the present as opposed to the future. Instead, under the post-utilitarian view, such an individual is putting his future incarnations in a difficult situation, thus forcing them into welfare dependence and at the same time imposing a fiscal “externality” on the tax payer. One can show that a post-utilitarian social welfare function can reflect the assumption that future incarnations are different individuals from current ones through an extra weight given to future consumption by the social planner, in addition to the weight it has in the current incarnation’s objective function. This extra weight implies that any incentive scheme which penalizes current behaviour by reducing the utility of some future incarnation will in fact harm the social welfare function *more* than it harms the individual. Consequently, by pursuing

---

<sup>4</sup> Most of the argument being made here would carry through if instead of multiple selves, there were unitary individuals but the government would instead be subjected to its own commitment problems that would preclude conditioning transfers on past behavior.

dynamic incentive schemes such as the ones discussed above, the social planner is shooting in its own feet; the social cost of these schemes is larger than if the social planner were strictly utilitarian. Fahri and Werning (2005) in a similar context show that this mechanism puts severe bounds on the degree of inequality that the social planner is willing to tolerate to induce incentives, in contrast to the result of Atkeson and Lucas that the optimal scheme delivers an ever-increasing degree of inequality.

At the same time, under a post-utilitarian social welfare function, it is more valuable to solve the moral hazard problem by imposing restrictions on preceding incarnations of welfare-dependent individuals in order to reduce the costs that their decisions will impose on their own future incarnations, and consequently on future tax payers.

Consider the following example: an individual lives for two periods and ‘we’ want to insure him against adverse shocks. In the first period he is subjected to a “needs” shock, which means that he values consumption more, the greater this shock. If he consumes more he saves less and has less wealth in the second period. In the second period he is subjected to an insurable income shock which makes his income from wealth volatile. To simplify ideas let us assume that in the first period the individual only has two alternative choices: save vs. not save, the latter yielding more consumption in period 1, but less income in period 2, than the former. Let us also assume that the needs shock in the first period can only put the individual in one of two situations: needy vs. not needy.

Clearly, a naive full insurance scheme in period 2 would allow everybody to have the same consumption; but under such a scheme, the not needy would then have no incentive to save. We can improve on that by conditioning insurance in period 2 on whether or not the individual has saved in period 1. If we want the not needy to save, it must be that the consumption they will get in period 2 is larger than that of the needy by a sufficient amount. From a utilitarian perspective, this is costly to the social planner since it reduces the welfare

of the needy relative to the not needy and the latter have a larger marginal utility of consumption<sup>5</sup>. This is the price we have to pay to give everybody the proper incentives to save. This option is the ‘conditional insurance’ one and is basically what dynamic social contracting theory prescribes. An alternative ‘paternalistic’ approach would be to force people to save in period. This would again allow full insurance, like the naïve scheme, but the total level of savings and therefore of consumption in period 2 would be higher. The downside of this scheme is that people in the needy state are forced to save despite that they value current consumption a lot. Which option is preferred therefore depends on how this cost compares to the benefits.

I now show that the paternalistic option is more likely to be preferred under a post-utilitarian social planner who considers future incarnations as distinct, than under a strict utilitarian who simply aggregates the intertemporal utility of individuals. To see this, assume that the social planner is a strict utilitarian. Its weight on period 2 consumption, relative to period 1 consumption, is the same as in the private agents’ utility function. If this weight goes up, which can only be the case if it also goes up in the individuals themselves put more weight on the future, the social planner cares more about period 2 consumption and therefore is willing to pay more to achieve full insurance in period 2; and the cost of the paternalistic option, i.e. a lower consumption in period 1, is lower. Does this mean that the paternalistic option is more likely to prevail? No, because the individual also puts a greater weight on the future, and that makes it also easier to induce incentives. The individual is more sensitive to his lower future consumption if he does not save, and this allows the government to induce him to save while having a lower consumption differential in period 2 between the needy and the not needy. In other words, conditional insurance insures the individual better. So this option also improves and there is no presumption that paternalism is more likely to prevail.

---

<sup>5</sup> Equivalently, we can say that the expected utility of a representative individual who does not know whether or not he will be needy, i.e. acts under a veil of ignorance, is reduced by this gap in consumption. So the argument can be cast in a pure insurance fashion rather than a redistributive one.

Assume now that because of multiple selves the social planner puts a greater weight on future consumption than the individual. The difference between the two is the weight given on the period 2 incarnation. If we increase that difference while holding the preferences of the period 1 incarnation unchanged, this means that the two incarnations are more distinct – if the difference is zero then the two incarnations are a single individual. Since the total weight on period 2 consumption is then larger in the social welfare function, we care more about the level of insurance in period 2, which again increases the value of the paternalistic option. On the other hand, period 1 incarnation, which makes the relevant arbitrage between savings and consumption, still values future consumption at the same level as before, so incentives do not improve and nor does the conditional insurance option. Clearly, then, the paternalistic option is more likely to be chosen, the greater the weight of the period 2 incarnation as an independent individual in the social planner's utility, that is, the more the selves are 'multiple'. Overall, this suggests that paternalistic tools are more likely to be chosen to implement redistribution and social insurance, the less individuals are considered as unitary and rational in the social welfare function.

This explains why we observe that modern welfare states are increasingly reluctant to deny benefits to individuals who put themselves in welfare-dependency out of their own choice and rely instead on coercive preventive measures. Thus we see unconditional social assistance coupled with compulsory contributions to pension systems; taxes on tobacco and the proposal of a tax on "junk food", but free medical treatment of the associated subsequent health problems; and free delivery of heroin (or methadone) and needles to addicts, while its consumption is illegal for the non-addicts.

### **The welfare case for paternalism**

While a post-utilitarian approach makes the design of the welfare state more paternalistic, relative to the utilitarian one, the interaction between paternalism and redistribution also goes in the other direction: the perceived value of paternalistic policies is itself greater, the greater the redistributive concerns. The underlying argument is that the poor are more likely to be behaviourally challenged, because such failures are a cause of poverty (if one can't commit to wake up in the morning to go to work, he or she is more likely to end up on the dole). These views are well represented in a 2004 *Guardian* article by pro-paternalist journalist Jackie Ashley<sup>6</sup>

“This is a time of year and a moment in politics when the issue of ministerial lectures is particularly timely. (...) Hardly a day passes without new government-approved warnings on cholesterol, stress, sleep, sugar and exercise. New laws on using mobile phones while driving have come in. Ministers look on approvingly as the British Medical Association raises the question of whether there should be a general ban on smoking. There's been a "salt summit" on the use of coronary-threatening quantities of salt in processed foods. This is also a time of year when rocketing credit card debt has become a big public issue, and yet offers to borrow more money drop through the letterbox daily. As Britain's overweight, indebted families look bleakly ahead into 2004, the government is on hand to teach them family values too. A raft of new measures aimed at the parents of anti-social children is now available to the police and courts.

Only yesterday another female minister, Margaret Hodge, was being hammered in the papers for approving funds for Parentline Plus, an email advice service for parents. The children's minister argued that good parenting could boost children's performance

---

<sup>6</sup> <http://www.guardian.co.uk/society/2004/jan/01/futureforpublicservices.comment>

at school by more than 10% and that "the festive season often highlights family problems and stresses". Hodge's crime, according to her critics? "Nanny-state meddling ... increasing interference by the state in people's everyday affairs."

It is time to stand up for the "nanny state"(...) And also, in general, for the state's right and duty to involve itself in questions of diet, health, family budgets and good parenting.

(...) Affluent families have more freedom to deal with the consequences. They buy the fresh food which isn't loaded with sugar and salt. They buy mounds of expensive fruit for their children. They pay subscriptions for gyms and health clubs, and go on pricey diets. They are better-educated about health issues and they buy the parenting books. They can respond to the latest "must-have" (always a lie) advertising for electronic goodies without falling into crippling debt. Oh yes, and they can buy in domestic help – nannies and their successors.

(...) Just as in urban America, it is the poorer British who are most visibly damaged by the downside of consumerism. This is pretty simple. If you can't afford the expensive consolations of winter travel or dinners out, then the homelier treats of chocolate, burgers and alcopops are that much harder to resist. Fat now means poor. Pop Idol's 15-stone Michelle is a working-class girl who stands, and wobbles, for millions. Pass a group of kids smoking and the odds are they will be female and poor. And if poorer families fall down on the parenting front, they can't buy in foreign girls to help plug the gaps.

Who really thinks that a centre-left government should stand aside and do nothing? The whole point of progressive politics is to stand with the most vulnerable people.

Today that means helping counter the great commercial, short-termist forces that rain down. (...) Today, in trying to help all families struggle with consequences of consumerism and social change, these ministers are doing a vital and noble job.”

There are two ideas in the preceding quotes. First, behavioural issues like obesity and smoking are a problem of the poor, therefore one should implement more paternalism, the more one cares about the poor. Paternalism is no longer only motivated by intra-personal redistribution, but also by inter-personal redistribution. Second, the causality between poverty and behavioural issues goes both ways, because behavioural issues can be solved by activities like acquiring education, going to the gym, and so forth, and the rich are in a better position to obtain these goods<sup>7</sup>. Thus commitment is no longer a matter of self control, it is now supposed to involve resources and therefore is now regarded as a privilege.

Let us discuss the merit of those arguments. To a traditional utilitarian, the question should be re-cast as follows: why should we redistribute to the poor in any other way than a pure monetary transfer? The paternalists in fact argue in favor of a sort of in-kind redistribution, which consists in helping the poor by controlling their behavior. As such, the view that commitment involves resources does not invalidate the standard economic argument that all redistribution should be in monetary form. After all, the poor can indeed use their transfers to the welfare state to purchase whatever commitment technologies they need. The other argument that it is precisely those individuals who have commitment problems who are likely to be poor is more serious. It means that they will spend the money given to them on the ‘wrong’ goods. However, if they do so, it means that this benefits at least one of their incarnations. To make a case for paternalism we actually have to go back to intra-personal

---

<sup>7</sup> The view that healthy food is more expensive, suggested by the author, is largely a myth. As far as proteins are concerned, many species of fish such as mackerel or sardines are un-expensive; the price of salmon and trout has fallen dramatically as a result of the ability to breed them. And if the obesity crusade is to be believed, people eat too much and could therefore substitute quantity for quality at constant total costs.

redistribution and argue that another incarnation is harmed and that the social planner gives sufficient weight to that self.

But that is not the end of the story, because the paternalistic also has to explain why such restrictions should be imposed on everybody rather than the poor only. The above quoted article is silent on this issue presumably because the author does not value the welfare of the non-poor. But if we believe, as she does, that behavioural issues are essentially a matter with the poor, then it is unclear why all the paternalistic interventions that we observe also affect the rich. Why isn't paternalism just targeted to the poor, as it originally was with private charities and early welfare states? Even if the social planner is a pure Rawlsian who only cares for the poor, the rich would in principle be happy to trade higher taxes against less restrictive regulations; if this trade-off leaves them indifferent, the State will have more money to redistribute to the poor, who will then be strictly better-off. So why should the rich who can buy diets, read books about cancer and obesity, and subscribe to a gym, have to face the same distorted set of prices as the poor? One obvious answer is that it is difficult to discriminate between the two because there would be some arbitrage going on between the rich and the poor that would prevent the targeted paternalism from working. But this issue is not mentioned in the article and I am afraid that the correct explanation is that individual freedom is simply no longer part of our contemporary values; which is not surprising if its original foundations were purely instrumental.

## **Conclusion**

This article has discussed how the welfare state favours increasingly paternalistic policies, as society's conception of the individual moves away from that of unitary, rational people, and

considers them instead as made of multiple selves plagued by behavioural issues. One mechanism is that individual responsibility loses legitimacy and efficiency in a post-utilitarian society, leading to a substitution of coercive prevention measure for incentive schemes. Another mechanism is that behavioural issues are more likely to be concentrated at the bottom of the distribution of income, implying that policy makers will value paternalism more, the greater their redistributive concerns.