Paternalism in Policy: Prospects and Limitations of an Economic Analysis

Péter Cserne
PATERNALISM

METHODOLOGICAL PROBLEMS OF AN ECONOMIC ANALYSIS

Péter Cserne

As a matter of theory, paternalism draws attention to philosophical and methodological difficulties involved in the justification of the limitations on freedom of individual choice in market and non-market settings. After a brief overview of the conceptual and justificatory problems surrounding paternalism in moral and political philosophy, this paper analyses whether economic theory, in its current state possesses the analytical tools for assessing paternalism adequately. I discuss how economists refine their analytical tools and redefine the concept of paternalism in order to fit the latter in their own methodological framework. I will argue that if economic analysis should have an impact on rational debates on paternalistic legal policies, it should both reconsider its normative assumptions and rely more systematically on empirical research, especially on the psychology of human judgment and decision making. Philosophical and empirical insights both suggest that the issue of paternalism is an “uncertain case”. Outright antipaternalism should not be replaced by uncritical paternalism.
Paternalism - methodological problems of an economic analysis

PÉTER CSERNE (BUDAPEST)

Introduction

In recent years, paternalism as a justification for market regulation has raised interest within both European and American legal scholarship. There are several reasons for this, both academic and practical. The practical ones are related, among others to policy debates on consumer over-indebtedness in the US, and the revision of the *aquis* and the pros and cons of a harmonized and social private law within the European Union. This interest is nurtured by theoretical considerations as well. Terms like regulation and paternalism draw attention to the philosophical and methodological difficulties involved in the justification of the limitations on freedom of individual choice in market and non-market settings.

Related to these, it has been argued that empirical findings on human behaviour in various decision-making and choice situations may lead to conclusions, especially policy recommendations that are significantly different from the outcomes of traditional philosophical and economic arguments favouring freedom of choice. More specifically, in the law and economics literature, the question has been raised as to whether the traditional anti-paternalist view of mainstream economics based on “consumer sovereignty” remains valid if individual decision-makers (consumers, managers, etc.) are imperfectly rational or not fully informed. Furthermore, as some of these imperfections of judgment and choice behaviour characterise humans generally, legislators and regulators with the task of setting a legal framework for contracting, or judges and juries involved in individual legal disputes, are not immune from these biases either. Thus the question should be raised whether the traditional anti-paternalistic stance of law and economics has to be modified, or even replaced, by anti-antipaternalism: a limited and critical version of paternalism.

As some scholars argue, “bounded rationality pushes toward a sort of anti-antipaternalism – a skepticism about antipaternalism, but not an affirmative defense of paternalism.” (Jolls – Sunstein – Thaler 2000: 46)

In this paper I analyse whether law and economics scholarship, in its current state possesses the analytical tools for assessing legal paternalism adequately. First, I present the conceptual and justificatory problems surrounding paternalism. As the term originates neither from law nor economics but moral and political philosophy, it will be necessary to briefly enter the domain of philosophy. Second, I discuss how economists refine their analytical tools and redefine the concept of paternalism in order to fit the latter in their own methodological framework. I will argue that if law and economics should have an impact on scientific debates on paternalistic legal policies, it should both reconsider its normative assumptions and rely more systematically on empirical research.

Paternalism: a rough definition

At least since Kant’s and Mill’s ideas on paternalism have become widespread, philosophers have spent much paper and ink on defining the concept and evaluating the moral or legal permissibility (desirability) of paternalism (Kant 1793: A 236, Mill 2004 [1859]), Sartorius 1983, Kleinig 1984, Clarke 2002, Dworkin 2005, Garren 2006–07 and the references there). As to the concept, terminological controversies notwithstanding, there seems to be a reasonable degree of consensus in the literature about the use of the term. Thus, an act is considered paternalistic if the following three conditions are fulfilled: the paternalist (1) interferes with the subject’s freedom, (2) acts primarily out of benevolence toward the subject (i.e., his goal is to protect or promote the interests, good or welfare of the subject), and (3) acts without the consent of the subject. Although these three conditions could be explicated more thoroughly, for our present purposes it is enough to highlight a few implications of the definition.

According to the definition above, paternalism implies doing something against the freedom of the subject without his consent. Under any theory which attaches value to freedom of choice, this lack of concern for consent needs justification. In fact, the idea of paternalism is precisely that there is such a reason: it refers to the protection or promotion of the interests (welfare, good) of the subject. If we follow this definition, paternalism is not essentially wrong: it has both justifiable and unjustifiable cases or forms. As to the moral justification of paternalism, there are at least three conflicting philosophical positions: deontological, consequentialist, and perfectionist.
Deontological theories take freedom as an intrinsic and primarily important value. It is based on the value of personal autonomy. Theorists who defend autonomy on this ground argue that people should have the right to choose self-harming actions because there is an intrinsic value to freedom of choice. Nonetheless, this anti-paternalist standpoint is often weakened or linked to various empirical conditions; thus even autonomy theorists accept paternalistic intervention in certain cases. In this way, autonomy theories require some kind of balancing. As Dan Brock has convincingly argued: “Paternalism is not, contrary to the common view, an issue that forces a choice between rights-based and consequentialist theories. […] [C]ontrary to appearances, [even autonomy theories like] Feinberg’s and VanDeVeer’s […] in fact require a balancing of respecting an individual’s autonomy against protecting his good in the way commonsense morality supposes.” (Brock 1988, S. 565)

Although, as John Rawls has argued, “all ethical theories worth of attention take consequences into account in judging rightness” (Rawls 1999: 26), from a consequentialist perspective, what matters for such a moral judgment are the consequences of an action or rule only. In the welfarist version of consequentialism, characteristic for standard law and economics and certain varieties of utilitarianism, the relevant consequences are the effects in terms of welfare. Welfarism, in this rudimentary form, has no fundamental objection against paternalism (see Burrows 1998; Zamir 1998).

To bring the theory closer to common moral intuitions and render it more workable as a policy guide, commentators usually supplement it with empirical claims about the suitability of different institutional mechanisms for promoting welfare or satisfying preferences. One such claim is that freedom of choice in a market setting promotes well-being. More weight can be given to the value of freedom by acknowledging that the freedom to make one’s own choices is a component of well-being. Consequentialism is even more flexible: theoretically, autonomy can even hold a central place amongst the goods constituting welfare. Nevertheless, as long as there are, or can be, other components of welfare: trade-offs between freedom and these other values might be necessary.

Individual welfare is usually defined by actual wants (desire theory of preferences). Sometimes, the analyst or policy maker takes objective elements of human well-being also into account. Indeed, when welfare is understood so generally as to include ideal preferences, i.e. preferences which one should have in light of some moral theory, this view is hard to distinguish from perfectionism. When based on an ideal theory of preferences, consequentialism comes close to perfectionism.

In its most abstract form, perfectionism is a moral theory which views the human good (flourishing, excellence) as resting on human nature. Perfectionism has an ideal for each human, namely that she develops her nature. It accepts self-regarding moral duties, thus it is concerned with what one should choose for herself. In contrast to other moralities which hold that the good is subjective and thus exclude any claims about what humans ought to desire, perfectionism has an objective theory of the good.

For our purposes perfectionism is more relevant as a political philosophy, i.e. a set of normative ideas about the aims of a political community, eventually formulated in legal rules. In this respect perfectionism holds that “the best government [is the one which] most promotes the perfection of all its citizens.” (Hurka 1993, S. 5.) This formulation reveals the close links of perfectionism to ancient political philosophy, especially its Aristotelian variant.

It should be noted, however, that in current usage the term perfectionism refers to a more diverse set of approaches. A few theorists self-consciously subscribe to an Aristotelian tradition (Gordley 2001, S. 180–185, 2007). Others hold personal integrity as a primary value (Kleinig 1984, S. 70–71). Even autonomy can be defended within a perfectionist framework, as in the case of Joseph Raz’ liberal perfectionism (Raz 1986, ch. 15). When Amartya Sen stresses the distinction between what one desires (the fulfillment of actual preferences) and what one has reason to value (in terms of a philosophical or moral view of the good), his views can be characterized as a perfectionist version of consequentialism (Sen 2002).

Generally, perfectionist moral theories either hold that one should strive for a good which is supra-individual (the overall utility, happiness, etc. of a community) or compare one’s actual actions and decisions with a morally or otherwise superior preference or value system. This might sound overly rationalistic. However, in the Aristotelian view, to find the right choice in a given situation is not a matter of demonstration, deductive logic or scientific knowledge.
Choosing rightly is a matter of prudence, or practical rationality. Still, the view that there are wrong, e.g. self-harming or immoral choices that should not be supported but rather discouraged by law is part of most perfectionist theories.

Paternalism is justified under this approach if there is an objective good, i.e., there is a difference between a right and a wrong action, and an individual's good can be promoted by the actions of another person. Contrary to general views, this does not imply or justify paternalism across the board. It is possible to argue that in many cases the (mature) individual is the best judge of how to attain the good or end-state. Also, in many cases a third party, especially the state is unable to promote this good better than the individual. These and other considerations set limits to paternalism even from a perfectionist perspective (Gordley 2001: 280–285).

On the other hand, from the point of view of virtue ethics, freedom of choice is valuable as a means to facilitate, or promote the goal of humans to lead a happy and worthwhile life. Within certain conceptual and empirical limits, perfectionism provides an argument for restricting individual choice in order to vindicate a theory of the good by impeaching immoral or ignoble choices.

Looking at the extensive philosophical literature on paternalism, it seems difficult to decide between the standpoints without discussing such far-reaching questions as the nature of the good or the meaning of free will. On the other hand, when discussing the pros and cons of a prima facie paternalistic intervention as a matter of policy, philosophical arguments are often indeterminate. Be that as it may, in the last decades, economics has gained an increasing impact in almost every kind of policy debate. In the following sections I discuss the methodological tools which economic theories use for analyzing paternalism.

**Paternalism in economic theory**

The traditional economic approach to freedom of choice and paternalism is a non-reflexive mixture of liberalism and utilitarianism. As such, it is ill-equipped to handle the problems which arise when these two principles collide. The potential conflict between welfare-maximization and autonomy draws attention to the non-welfarist dimension of the problem of paternalism. When this dimension is not taken into account, the whole problem of paternalism is reducible to a more or less sophisticated exercise in welfare-maximization. Within a strictly welfarist perspective, if the regulator knows better, he should decide in every case in the agent's place. In other words, the constraints to paternalism can only be pragmatic.

In contrast to the philosophical arguments discussed above, pragmatic anti-paternalist arguments draw attention to the side-effects and non-intended or counter-intentional consequences of paternalistic interventions. The intervention may be more costly or harmful than beneficial, either for the paternalised subject, or for third parties and the general public. An example of the first can be the double bind effect which refers to the problem that in many contexts prohibiting exchanges may actually worsen the plight of the individual whose welfare is central to the issue. For instance, banning prostitution may eliminate an income-earning option of poor women (Radin 1996). Third-party external effects include the over-inclusiveness of the legal rule (Schauer 1991), or the implementation costs. In regards to the unintended consequences, overly protective regulation always runs the danger that certain transactions become unprofitable and lead to the collapse of the market, or the segment which was supposed to be protected. Paternalism may be problematic in these cases because it unjustifiable burdens for others, but also because it might backfire, i.e. make worse off the very group of persons it intended to protect.

In another dimension, standard economic theory is often criticized on the basis that it accepts existing preferences as given. As the critique goes, economic theory does not offer "ethical criteria for disqualifying morally offensive, self-destructive, or irrational preferences as unworthy of recognition." If, to the contrary, it acknowledges some exceptions, as it usually does (e.g. in case of minors or mentally incompetent persons) then "some theory of paternalism is required, the contours of which are not readily suggested by the private ordering paradigm itself." (Trebilcock 1993, S. 21) Being reluctant to criticize or "launder" preferences or rethink their model of individual choice as a combination of information (beliefs) and preferences, economists use various strategies for the "eliminative redefinition" of paternalism in order to fit the problem in the standard models.
Strategies of “eliminative redefinition”

The economic literature on paternalism (see, e.g. Burrows 1998; Zamir 1998) strives to justify instances of reasonable, seemingly paternalistic regulations in several ways. For the sake of simplicity, they can be put into four categories: market failures, merit goods, non-standard preferences and non-welfarist objectives.

Market failures

The first approach sticks with consumer sovereignty and revealed preferences. Paternalism is redefined or “explained away” by showing that the policy in question serves to prevent externalities or other market failures. Not questioning consumer sovereignty, this eliminative redefinition of paternalism is arguably the natural way to treat the substantive problem in economic theory. Indeed, at first glance it is relatively easy to incorporate paternalism in mainstream economic analysis: we just have to identify specific transaction costs and/or informational imperfections and asymmetries which lead to a market failure. Limits of freedom of contract, like the judicial control of standard form contracts, labor law or consumer protection can be analyzed in terms of these (now) standard economic concepts. These limits are economically justified to the extent that they remedy market failures. In this way not only the case for freedom of contract, but many of its limits can be explained in relatively narrow economic terms, by neither relaxing the rationality assumptions nor recurring to fairness arguments (Mitchell 2002). In fact, the biggest “advantage” of this approach is that the conflict between welfare and autonomy does not come to surface.

Merit goods

The second approach clearly and almost openly faces this conflict and solves it in favor of welfare. In the 1950s Richard Musgrave introduced two new concepts in the theory of public finance: merit wants and merit goods (Musgrave 1987). These concepts serve to formalize the welfarist idea of paternalism with regard to certain publicly provided or publicly subsidized goods. In the case of merit wants or goods, the welfare function that the policymaker has to maximize on behalf of the individuals is modeled formally differently from the way the individuals themselves (are modeled to) perceive and reveal their preferences. In this way, “consumer sovereignty” is openly questioned and abandoned in favor of a supra-individual assessment as to how much one should receive and consume of certain “merit goods”. Although merit goods often have public good characteristics and their provision can also be driven by redistributive concerns or justified by asymmetric information, Musgrave has insisted that there can be a separate reason, independent of all these standard economic ones for overriding individual preferences. In case of merit goods, individual choice is overridden in the name of what a particular (political) community considers worthwhile for individuals to have or be able to do. This reveals the perfectionist flavour of the concept.

“Irrational” preferences

The third category of redefinitions comprises those more or less sophisticated models of preference-formation and decision-making which introduce specific ad hoc assumptions about the preference structure of individuals. Examples include choice models based on path-dependent preferences, dynamic inconsistency or the multiple self. The methodological goal here is to analyze certain conflicts between autonomy and welfare with a minimal deviation from mainstream economic theory (rational choice theory).

When economists analyze market behavior, they not only implicitly rely on a standard of voluntariness but also explicitly on a standard of rationality. According to rational choice theory, weakness of will, “sour grapes” mechanisms etc. are irrational behavioral patterns. This view implies that the preferences of real-world individuals should be measured on a normative scale. The actual or revealed preference structure of individuals is compared to an ideal or rational preference structure of an abstract model construct, the rational decision-maker. Irrationality in this sense may justify intervention.

To some extent, rationality is treated in these models like autonomy in freedom-maximization ‘models’. One’s preferences are to be respected if and only if they truly and consistently reflect one’s desires. But this is not always the case. The rational self can be in conflict with other features of the very same person or in other terms with different selves. These models re-conceptualize paternalism as a multiple-self problem or even as an ‘intra-personal externality’ or ‘internality’ problem.
In sum, in certain cases, paternalism is justified by the following argument: limiting freedom of choice is instrumental to the defence of the true self of the subject against one’s weakness of will or judgment errors. Eventually, these limitations may even increase the subject’s rationality and/or autonomy. As we will see below, the same behavioural patterns are assessed within psychological theories and “behavioural law and economics” in a rather different way.

Beyond welfarism

Finally, there are such heterodox economic approaches that criticize mainstream economic theory for the reduction of every normative instance to preferences over outcomes and suggest this simplistic view be remedied. For instance, some economists highlight the need for including freedom of choice in economic models (see, e.g. Sugden 1998, Van Hees 2002. This line of research can be extremely helpful for an economic analysis of the non-welfarist dimension of paternalism. Here I only briefly characterize this “freedom of choice” literature. Originating from Amartya Sen’s seminal article on the impossibility of a Paretian liberal (Sen 1970), there is now an emerging branch of literature in social choice theory which searches methods and modeling techniques for the incorporation of the dimension of freedom of choice into formal economic (social choice) models. The intuition behind this line of research is that the extent of opportunities, i.e. the number and diversity of alternatives open to an individual might be valuable in and of itself. The freedom to choose among alternatives might have some value independent of the intrinsic value of these alternatives. This idea is hardly new in political and moral philosophy. Still, the philosophical literature on autonomy, liberty or paternalism usually lacks the conceptual rigor, or the degree of formalization, that would make such arguments directly amenable to economic analysis. Conversely, economists find it difficult to incorporate relevant and important philosophical ideas into their analysis while ever they cannot translate them to their own formalized language. Authors active in the freedom of choice literature, besides searching for formal methods to measure the extent of freedom, usually argue for the importance and normative superiority of a non-welfarist metric of well-being. In regards to paternalism however, this line of research has no clear and direct policy implications.

The need for empirical foundations

From the discussion of economic analysis of paternalism so far, it has become clear that the mainstream economic approach has to face two problems. First, it remains controversial from a normative point of view whether the function of law can be reduced to the maximization of individual preferences. Second, there remains a methodological difficulty: the growing amount of evidence on biases and other irrational behavior calling the scientific fruitfulness of ad hoc modeling of bounded rationality into question. To put it differently: in cases where people systematically make suboptimal choices, paternalistic intervention may be justified. To determine the appropriate scope and technique of intervention, systematic empirical research on bounded rationality is necessary. However, the justification of interventions cannot be based simply on revealed individual preferences as the ultimate normative benchmark, because the concept of preference has also become questionable.

As Robert Sugden has argued with regard to political philosophy on the one hand and economics on the other, albeit in a slightly different context:

“Neither approach is grounded in empirical hypotheses about human psychology or human nature. Political philosophy is concerned with good reasons: the question of how, as a matter of psychological fact, good reasons motivate people to act is left unanswered. But …economic theory … does not concern itself with the psychology of motivation either. It simply asserts the a priori postulate that, for each individual, there is a well-defined set of self-interested preferences, on which that individual invariably acts. When pressed, economists usually defend this assumption on the grounds that it corresponds with the requirements of rational behavior. In other words, they appeal to a notion of good (prudential) reasons and do not concern themselves with the question of how such reasons motivate. […] However, an empirical social science has to rest on empirical foundations.” (Sugden 2004, S. 210).

I consider Sugden’s critique regarding philosophy and economics mainly justified. Empirical research is crucial in order to answer questions about the best possible way to design legal rules, be they paternalistic or otherwise, and in the discussion of the reasons for paternalism. This implies the relevance of psychological research for economic analysis. In fact, this line of research has become increasingly respected, even popular both within economics in general and law and economics in specific, under the somewhat misleading label behavioural (law and) economics (Korobkin – Ulen 2000, Sunstein 2000, Englerth 2004, Camerer et al. 2004, Engel et. al. 2007).
In recent decades, there has been a large volume of ongoing empirical research in this direction, both by psychologists and economists. It has been shown in thorough and extensive empirical studies that human behaviour systematically deviates from the precepts of expected utility theory and rational choice theory in general. Human decision-making, choice and judgment reaching behaviour are often characterized by loss aversion, the endowment effect and framing effects. Furthermore, people commit systematic judgment errors when assessing probabilities. By using mental shortcuts called heuristics (availability, representativeness etc.) their judgments might become biased and vulnerable to manipulation – self-serving bias, hindsight bias, over-optimism, unstable risk-assessment being the most well-known examples.

For instance, several hundreds of studies raised fundamental doubts about the assumptions regarding preferences in economic theory. The framing effect shows that in many situations the concept of preference itself is indeterminate. In a descriptive psychological sense, the assumption that preferences are autonomous and stable is false. Several other features and mechanisms of human judgment and choice have been also described which contradict the assumptions in neoclassical economics (rational choice theory). Cognitive limits and emotional biases are ubiquitous. These should be, and have been, theoretically explained and modelled within neuro-sciences and cognitive psychology. To note, in terms of these disciplines, the issue is not one of biases or anomalies but the understanding of the very way human minds work. Nevertheless, these empirical results have an obvious relevance from a normative point of view.

Interested in how law works and how it should work, law and economics should rely on these insights. The question is, how. Is it possible to build a competing (and possibly superior) version of legal theory based on the analysis of these psychological phenomena? There are many obvious difficulties impeding such an endeavour. The main criticisms against behavioral law and economics (and the counter-arguments) can be summarized in the following extremely simplified way (Rachlinski 2006).

(1) The phenomena described in this line of research are not real. The empirical results do not have internal validity. True, these mechanisms are context-dependent. Controversy in the literature remains about the magnitude and subsequent significance of these effects. Still, some of the effects have been studied in several hundreds of experiments and their significant presence has been confirmed in a wide range of contexts. The internal validity of the empirical research seems warranted.

(2) The phenomena analyzed and the results presented in the empirical studies are not relevant to the real world. This line of research has no external validity. With regard to the external validity of the empirical facts uncovered by behavioral decision theory, it has been empirically confirmed that people (consumers, managers, judges, etc.) “fall prey” to certain choice and judgment anomalies systematically and repeatedly. This happens not only in one-shot laboratory experiments but in real-world situations and despite the possibility of learning effects.

(3) The phenomena disappear when sufficient monetary incentives are provided. When stakes are high, people do calculate and choose as rational choice theory would predict. To be sure, the responsiveness of biases to monetary incentives varies. In some cases, this incentive effect is observable. However, it is unclear through which psychological mechanism it is driven and how predictably it happens. Furthermore, “small-stake” situations may be also economically or socially relevant.

(4) There are private institutions that cope with biases. This is, in fact the case. Biases are so varied and complex that it is often the case that individuals themselves are in the best position to cope with their own irrational tendencies and deficiencies. In fact, this is what happens when people hire experts, use self-binding techniques etc. However, this is not an argument against a theory explaining how this debiasing works and when it is expected to be used. Rather it provides a reason why legal paternalism should be designed so as to take into account this institutional context. Law should encourage, or at least avoid crowding-out private debiasing mechanisms.

(5) The errors compensate, even cancel out each other, thus on the macro (market) level their effect cannot be observed. This might be the case for random errors in many cases. But the distribution of biases is typically not random – they do indeed follow some well-defined patterns. It is enough to refer to the literature in “behavioral finance” in this respect (Thaler 2005).
Finally, the last argument claims that there is no coherent theory to encompass all the empirical observations on human judgment and choice behavior.

In fact, behavioral law and economics does not rest on a single theory or definition of bounded rationality that would allow the observations on human behavior to be axiomatized. Despite prospect theory having been suggested by Kahneman and Tversky as the alternative to expected utility theory (and still other candidates arising in the literature, e.g. Gigerenzer 2002), in this matter the research has remained inductive to a considerable extent. There are specific behavioral regularities which are highly relevant and can be modeled at a lower or middle level of abstraction. Nevertheless, the fact that behavioral law and economics cannot fully explain every main feature of human decision making and judgment in terms of a single general theory, should not be seen as a failure or "a problem to be solved”. It is rather the general condition of scientific research, its reason being in the nature of empirical knowledge.

While the objections against behavioural law and economics are either not convincing or not decisive in dismissing it as a line of research, it is clear that empirical findings alone cannot justify legal intervention. The normative implications (and policy conclusions) are not straightforward. Psychological insights suggest that the justifiability of paternalism is an “uncertain case” – outright anti-paternalism should not be replaced by uncrirical paternalism (Rachlinski 2003, Trout 2005, Blumenthal 2007). In short, there are strong arguments both for and against paternalism. The behavioural findings may lead, in certain circumstances, to normative conclusions (policy recommendations) that are significantly different from those arising from the traditional anti-paternalism of mainstream economics. The consequences for normative analysis, however, are not simply the uncritical endorsement of paternalism.

The prima facie arguments for paternalism seem obvious: when law can promote the interests of humans by reducing their biases, increasing their autonomy, it should. But why should the law not always intervene despite systematic irrationality? Besides traditional skepticism toward governmental regulation and the pragmatic antipaternalist arguments discussed above, there are some specific psychological reasons for not intervening in people's choices (cf. Buckley 2005: 63-101).

First, research has revealed that the behavioural regularities (biases) are highly context-dependent. At present, this context-specificity argues against drawing general normative conclusions for policy.

Second, as mentioned above, what are called “biases” in the light (or shadow) of rational choice theory are in fact (in terms of psychology) components embedded in a complex decision-making mechanism. In this complex cognitive and emotional system, there are several interactions between these mechanisms. It is possible that one bias may temper another. When such interactions are neglected, debiasing might make the overall result worse than the initial situation.

Third, learning effects can be at work. In a dynamic perspective, regulation may lead to the inhibition of learning and negatively affect rational and autonomous choice in the future. This provides a dynamic or developmental argument against paternalism. This dynamic effect, already stressed in the philosophical literature (Mill 2004 [1859], Feinberg 1986: 241), has been now demonstrated in empirical terms as well (Klick–Mitchell 2006).

To be noted, this third consideration provides a reason why the law should deliberately deviate from a "psychologically adequate” view of man. Arguably, if rational and autonomous choice is accepted as a normative ideal, then the law should counterfactually slightly “overshoot” with its assumptions concerning both. This should be done in order not to simply map and thus stabilize biases but to leave space for learning and development (Eidenmüller 2005).

Three regulative ideas

In arguing for protective rules, legal commentators often refer to empirical data about the vulnerability of consumers to biases and manipulation. The psychological, pragmatic, economic and philosophical arguments, and counter-arguments, discussed in the previous sections should make supporters of legal paternalism cautious. An uncritical, across-the-board support of paternalism is unwarranted. Recently, law and economics scholars have suggested a

---

1 Feinberg 1986: 24: “If adults are treated as children, they will in time come to be like children. Deprived of the right to choose for themselves, they will soon lose the power of rational judgment and decision. Even children, after a certain point, find better not be ‘treated as children,’ else they will never acquire the outlook and capability of responsive adults.”
number of more or less sophisticated approaches how to take into account both the behavioral insights and some of the counter-arguments when designing legal rules or policy in general. I briefly recall here three of these regulatory ideas: asymmetric paternalism, libertarian paternalism, and debiasing through law.

“Asymmetric paternalism”
In an influential article, American psychologists, economists and law professors argue for the partial rehabilitation of the idea of paternalism. “A policy is asymmetrically paternalistic if it creates large benefits for those people who are boundedly rational while imposing little or no harm on those who are fully rational.” (Camerer et al. 2003: 1219). More concretely, they claim that legal interference with private choices is justified (in a firm/consumers setting) if:

\[(p \cdot B) - [(1-p) \cdot C] - I + d\Pi > 0,\]

where \(B\) denotes the net benefits to boundedly rational agents, \(C\) is the net costs to rational agents, \(I\) stands for the implementation costs, \(d\Pi\) denotes the change in firms’ profits, and \(p\) is the fraction of consumers who are boundedly rational (all other consumers are supposed to be fully rational). (Ibid.)

The commentators who argue for ‘asymmetric’ paternalism on a welfarist basis also compare real-world agents with the fully rational individual as assumed in orthodox economic models. They go on to say that bounded rationality is something which should be regulated in a similar way to externalities. Here one has, first, to suppose the existence of a true ‘inner self’, characterized by such desires and beliefs which are normatively undisputed. This is problematic because it is not clear which features of the empirical self should be respected and protected. Also, asymmetric paternalism is a purely consequentialist argument. As such, it is open to criticisms for not taking autonomy seriously. Read in a different way, this formula only illustrates the structure of the problem of legal paternalism. It does not serve to measure and quantify these variables, but to highlight who are the beneficiaries, and who are the cost bearers of a paternalistic intervention. These costs and benefits should be assessed more precisely in specific contexts.

“Libertarian paternalism”
While “asymmetric paternalism” suggests policies that protect boundedly rational individuals while not (significantly) burdening others who do not need protection, “libertarian paternalism” draws attention to the different methods and techniques of this protection. More precisely, “libertarian paternalism” suggests policies that respect the autonomy of boundedly rational people to the extent possible, but nevertheless help them to avoid making bad choices (Thaler – Sunstein 2003, Sunstein – Thaler 2003). It advocates paternalistic interventions mainly in the form of default contract rules, opting out policies and “menus.”

“Debiasing through law”
The third regulatory ideal, suggested by Christine Jolls and Cass Sunstein is more ambitious. Instead of searching for legal rules which are adaptive to judgment and decision errors, it aims to reduce the occurrence of boundedly rational behavior at the first place (Jolls – Sunstein 2006). In addition, the novelty of this view is to suggest debiasing by exploiting or at least relying on bounded rationality itself. As already mentioned, psychologists have found several instances where biases interact, and more specifically, offset each other. In these cases, an intervention aimed at the reduction of only one of the biases can actually worsen the overall result. Now, the interaction between compensating biases may be deliberately designed and used by policymakers as well. Jolls and Sunstein illustrate with various legal examples, how legislators can make use of the presence of one psychological mechanism (e.g. the availability heuristic) in order to counteract the self-detrimental effects of another (e.g. over-optimism).

One of their examples is the regulation of the safety of consumer products (Jolls – Sunstein 2006: 215). It is a well-established fact that consumers are asymmetrically uninformed about the safety features of most of the products they purchase. Mandatory disclosure requirements are a wide-spread legislative response to this kind of informational asymmetry, both in the United States and in Europe (Grundmann – Kerber – Weatherhill 2001). In principle, this is a preferable way of intervention, as information provision respects individual choice. In practice, disclosure does not necessarily work well. One reason for this is that people do not process the information in a rational way. Not only that they deviate from Bayesian theory. Even if their understanding (estimations) of the general risk probabili-
ties is correct in a statistical sense, they are overly optimistic about the occurrence of safety risks in their own specific case. In the psychological literature this is called over-optimism bias.

The idea of debiasing through law is to remedy this bias by making use of another, the availability heuristic. By using vivid and personified examples, the tendency towards underestimation of personal risks can be compensated. If people are confronted by a story about a recent real-life case of harm caused by a defective product, they become more aware or even overly aware of the risks involved.

To be sure, this rudimentary idea of debiasing through law should be refined in many ways before making actual use of it in information regulation. Even if as a matter of psychology, the argument is sound, and we put normative concerns about autonomy and manipulation aside, a traditional legal scholar would definitely have a large number of doctrinal and systemic arguments as to why the idea cannot be easily put in legal form.

On the other hand, the basic notion behind this regulatory idea is plausible, some would even say trivial. The way information is communicated matters. Consequently, information should be provided to people in a way that allows them to process it properly. Or if they process it poorly, the way of provision should be calibrated accordingly.

**Summary**

In order to be taken seriously as an empirical social science relevant for public policy, economic theory should be open both to philosophical arguments and empirical results. Law and economics, as an application of economic methodology to legal issues, should reconsider both its normative foundations and be open to insights on human behaviour. Philosophical and empirical insights both suggest that the issue of paternalism is an “uncertain case”. Outright anti-paternalism should not be replaced by uncritical paternalism. Even if we are merely interested in policy implications, empirical research is inconclusive because it does not provide a normative standard. The normative standards are the subject of a different kind of discussion. With regard to the philosophical positions, we have seen the practical convergence of autonomy and consequentialist theories in balancing (Brock 1988). There are also good arguments that some conceptions of perfectionism and liberalism are compatible (Marneffe 1998) and perfectionism can also be combined with economic theory to some extent (Deneulin 2002, Buckley 2005). This suggests the possibility of finding an overlapping consensus regarding the proper role of paternalism in specific legal domains. This issue, however, is beyond the scope of the present paper.

**References**


Sartorius, Ralf (Hg.) (1983): *Paternalism*. Minneapolis: University of Minnesota Press


