

Law and Decision Making:
Incorporating Internal Harm into Rational Choice Theory

Joshua Masters¹
New York University

One does not have to experience driving in peak-hour traffic in New Delhi for very long to realize that not all drivers obey traffic laws or rules of tort law. Indeed, in that city at least, those who do appear to be the exception rather than the rule. A road sign that reads “Wearing a helmet could save your life. Law or not”, suggests that motorcycle riders require some additional reason for action beyond the law. This raises important questions. Why is the fact that a law is a law not enough to dictate certain behavior for its subjects? How does a legal rule influence an agent’s behavior? In other words, when an agent decides whether or not to engage in behavior that may or would impose criminal or tortious liability, to what extent does that law influence the decision? Through the imposition of liability, law demands certain behavior; it creates an obligation. But the obligation of a law is not sufficient to accurately predict an agent’s response to it. Laws are often broken. As well as breaking traffic rules, some people murder, or negligently put others at risk, or do not pay taxes. What then explains responses to legal rules? This article endeavors to delineate an explanatory theory of behavior in response to legal rules, as distinct from a normative or predictive theory. The theory modifies the standard rational choice account by treating two distinct causes of internal harm as affecting an agent’s well-being.

An agent will have a limited set of possible choices that can be made. There is also a set of potential outcomes that may arise from each of the possible decisions. The agent will have certain beliefs as to the likelihood of each of those outcomes being realized as a result of a given decision. The agent will also have preferences over each of the potential outcomes.² Thus, in order to effect a decision, a law must either:

¹ B.Ec., University of New England, Australia; LL.B., University of Adelaide; LL.M., New York University School of Law. jdm452@nyu.edu.

² See Lewis A. Kornhauser, “The Normativity of Law” 1(1) *American Law and Economics Review* 3, 7 (1999).

- change the set of possible choices that can be made,
- change the potential outcomes arising from each possible decision,
- change the beliefs of the decision maker as to the likelihood of those outcomes being realized, or
- change the preferences of the decision maker.

The aim of this article is to thoroughly evaluate the basis of claims that law might affect each of these aspects of decision making. Part I considers the effect law has on an agent's choice set. The effect of law on both potential outcomes of behavior and an agent's beliefs about those outcomes is discussed in Part II. The potential for law to change preferences is considered in Part III. Part IV examines the mechanism by which law affects the above aspects of decision-making, namely by providing information to the agent. This article distinguishes between the relationship information has with external outcomes and the relationship it has with internal outcomes. Finally Part V reflects on the ramifications of the theory for the economic analysis of law. While criminal law offers the starkest application of the theory, it is also important to examine its more subtle operation in the context of tort law. Throughout the article, examples from each are used to explain various issues relating to the theory.

While much of this article's analysis of law's effect on behavior agrees with the standard rational choice theory, it concludes that norms have a more significant role to play than that theory allows. However, rather than creating exclusionary reasons, as it is said to do in obligation theory, they are better viewed as creating psychological harm if they are breached, and thereby amount to a mere incentive in the rational choice theory. Such an incentive can be created by law through changing the outcomes of a decision. The rational choice theory is thereby modified to include internal outcomes as incentives. An example of how the theory better predicts behavior is outlined in the final part.

Changing the Choice Set

This Part considers the possibility that law might change the set of choices available to an agent. A choice set is made up of all of the

possible actions that an agent could take at a given moment. That is to say, any action that is within the realm of possibility, rather than the actions to which the agent gives serious consideration. Outlining the choice set is of course an important first step in the explanation of an agent's behavior. The behavior of the agent must form part of the possible choices among behavior that are open to the agent. It is difficult to conceptualize the possibility that a law might introduce a behavior into the agent's choice set. Of course, a law might, for example, change the legal status or liability of the agent as a result of certain behavior, but making particular behavior possible that was not possible before is another matter.

It goes without saying that if certain behavior is not available to the agent, then the agent cannot behave in that manner. If the law then could somehow remove a target behavior from an agent's choice set, it would thereby render it impossible for the agent to behave in that way.

Raz's normative theory of behavior includes 'exclusionary reasons' for behavior, which, as the name suggests, exclude reasons for choices from the agent's deliberations.³ Could that theory be extended such that the obligation of law (at this point we need not distinguish between an obligation to follow the law, and an obligation to follow a particular law) acts to exclude choices from the choice set? For example, if it is illegal for me to cross the street at a set of traffic lights unless the 'walk' signal is lit, then the law would operate to remove the possibility of crossing the street when the light is not lit from my choice set. It is difficult to see that it does so. It is even more difficult to see that a rule imposing liability for negligence removes the possibility that an agent will act negligently. As noted earlier, many people break laws. So for them, the law has not excluded the illegal or tortious action from their choice set. It would be interesting though to consider whether any person might be affected in that way by law. However, even if we imagine an agent who never breaks the law, a 'good' citizen, it does not automatically follow that illegal or tortious actions have been removed from the good citizen's choice set. Consider a lawmaker who wishes to prevent fatalities occurring as a result of negligent operation of a motor-vehicle by attaching liability to such conduct. Even if the good citizen accepts the obligation to drive with due

³ Joseph Raz, *Practical Reason and Norms* (Oxford University Press, 1975).

caution, we cannot say that it is not possible for her to drive negligently, she has simply chosen not to do so. Consider now a city councilor who does not want citizens to walk on the grass in a park. If the councilor were to pave the grassed area, we would accept that walking on the grass has been removed from the choice set of the citizenry. It is simply no longer possible to walk on the grass, as there is no grass to be walked on. We might also accept that if the councilor were to erect an electrified fence with a 24 hour guard around the grass, then walking on the grass might have been removed from the citizen's (reasonable) choice set. Again, it would not be possible, or reasonably possible, for the citizen to walk on the grass. However, if the councilor instead left the grass intact and unguarded and passed a law making it illegal to walk on the grass, even for the good citizen, we find it difficult to say that the choice of walking on the grass is not available. We would not say that the good citizen has been forced to stay off the grass, but rather that she has chosen to obey the law and stay off the grass.

In that sense, if certain behavior remains theoretically possible, it remains within the agent's choice set. A legal rule cannot, without more, cause certain behavior that was once possible to become impossible, or vice-versa.

Consider now, a Federal government, which passes the Defense of Grass Act, which obliges all metropolitan councils to erect electrified fences with 24 hour guard around the grass. Could we say that the Defense of Grass Act has removed walking on the grass from the choice set of the citizen? We are considering how a legal rule affects the decision making process of the citizen. But in this example, the law has not removed the option of walking on the grass from the choice set. It was not until the obligation of the law, in the form of the fence and guard, was met, that walking on the grass became impossible. It was not the law that changed the choice set, it was the erection of the fence and hiring of the guard. It may be the first link in the causal chain, but does not change the choice set simply by virtue of its existence.

If law cannot change the choice set of an agent, it must affect decisions in a different way. The next Part considers whether a law

might affect the potential outcomes of certain behavior, or the beliefs of an agent about the likelihood of those outcomes.

Potential Outcomes and Beliefs

Categories of Outcomes

Before considering whether law can change potential outcomes or the beliefs of an agent, it is useful to categorize the types of outcomes that might be included in the set of potential outcomes. The economic theory of rational behavior holds that only well-being matters to the agent.⁴ Outcomes that affect the well-being of the agent are usually thought of as external outcomes. We can distinguish between internal and external outcomes. We might also call these outcomes psychological and non-psychological. An external, or non-psychological outcome is observable by the decision maker, members of a society and an independent observer of a society. Common examples related to the law are taxes, fines, incarceration, civil liability, etc. For example, each of those entities might after a period of observation, expect that a tortfeasor will be sued in a civil court and, if found liable, be required to pay damages, or that a murderer will be arrested by the police, tried in a criminal court, and if convicted, be imprisoned or put to death. Psychological outcomes will not be observable to the independent observer. For example, an independent observer will not, in the absence of the consequences outlined above know that a tortious act or murder is 'wrong'. The wrongdoer may however consider that a tortious act or murder is wrong if he accepts that norm as a guide to his behavior. Breaching that norm will cause some internal harm to the wrongdoer aside from the non-psychological consequences outlined above. The internal harm is not observable to any entity other than the wrongdoer.

⁴ See Lewis A. Kornhauser, 'The New Economic Analysis of Law: Legal Rules as Incentives', in *Law and Economics* 27, 42 (Nicholas Mercurio ed.,1989); Lewis A. Kornhauser, 'Interest, Commitment and Obligation: How Law Influences Behavior', in *Justice and Power in Sociological Studies* 208, 215 (Bryant G. Garth and Austin Sarat eds.,1998).

External outcomes might be anything that affects an agent's well-being, and will have different weights assigned to them by different decision makers, depending on their risk aversion and their preferences for the potential outcomes (their utility function). As a simplified example, assume that two people have the same preference for money and also share the same preference for arriving at a destination quickly, the person with lower risk aversion may choose to drive a car at a negligent speed if there is 99% chance of reaching her destination quickly and a 1% chance of causing damage for which she is liable, whereas the person with higher risk aversion might choose to not drive at a negligent speed. Furthermore, if we assume that the two people have equal risk aversion and equal preference for arriving at a destination quickly, the person with a higher preference for money may choose to not drive at a negligent speed, whereas the person with lower preference for money may choose to drive at a negligent speed. The external outcome that might be affected by a law relates to the external incentives created by the law. In the example above, if the law states that a person driving at a negligent speed is liable for any damages caused thereby, the external incentive to accept the obligation created by the law to not drive at a negligent speed is the threat of liability for any damages caused thereby. Of course, the decision maker will also take into account the likelihood of that liability being realized. That likelihood might depend on observing the verdicts in other like cases, the availability of witnesses, the skill of the defense attorney, etc.

Internal outcomes related to a law may be divided into two categories. The first, relates to the norm that overlaps the subject matter of the law. The second relates to the norm that the law should be obeyed. These internal outcomes depend upon the agent's acceptance or commitment to those rules.⁵ If the agent does accept them, then an outcome of breaching a legal rule will be internal harm. A harm resulting from not following the subject matter norm, or, a harm resulting from not following the law because it is a law. The harm is related to an agent's conscience. For example, if an agent thinks that people should always tell the truth and the agent

⁵ See Kornhauser's discussion of HLA Hart, 'The Normativity of Law', supra note 1.

lies, she feels bad because she has breached a norm which she subscribes to; she has suffered an internal harm.

It is worth considering whether the normative obligation to obey the law exists at all. In the context of a just government, the legal obligation to obey the law is said by Raz to appear to depend on the existence of an independent legal obligation (moral duties as he calls them), to obey each of the laws of the system.⁶ However, a normative obligation to obey the law creates an apparent paradox. Raz uses the example of murder to illustrate the paradox in terms of morality. The same might be said in terms of normativity. There is a normative obligation not to murder. But to refrain from murder simply because of a normative obligation to obey a law that prohibits murder is not a normatively correct reason to do so. The paradox states that it cannot be normatively incorrect to follow a normative obligation. However, as Raz notes, that type of paradox is overstated. There may be an obligation to follow laws in certain circumstances. It is not the case that if there are certain laws where there is no obligation to obey, then there is never an obligation to obey a law beyond an obligation related to the subject matter of the law. Raz provides examples of circumstances where there is an obligation to obey the law simply because it is law as situations where the judgment of government experts is better than the agent; where the government is better able to end a widespread and damaging practice; and where non-compliance will undermine the government's ability to do good.

In my view, the paradox is not merely overstated, but is, rather, illusory. That is particularly so if there is a normative obligation to follow all laws rather than a separate obligation to follow some individual laws. An agent can have a normative obligation to not murder and also have a normative obligation to follow the law, which includes the prohibition of murder. Each normative obligation offers a reason for action to the agent. Choosing to not meet one or the other of those obligation would have the outcome of causing internal harm to any agent who subscribed to those norms. It is too simplistic to say that choosing to not commit murder solely because it is against the law is normatively incorrect. In that example, we see at least two decisions. The first is to not use a normative obligation to refrain from murder as a reason to not

⁶ Joseph Raz, *Ethics in the Public Domain* 326-327 (Oxford: Clarendon Press, 1994).

murder. The second is to meet the normative obligation to follow the law. The first decision is normatively incorrect. The second decision is normatively correct. We might say that if the agent has a higher preference for meeting the first obligation than the second, then, on weighing the two aspects of the decision, on balance the decision was normatively incorrect. However, if the agent had a higher preference for meeting the normative obligation to follow the law than for meeting the normative obligation to refrain from murder, then the decision was, on balance, normatively correct. Of course, it is inherently appealing to accept that the normative obligation to not murder has greater weight than any normative obligation to follow the law. That causes the illusory paradox to be, on its face, appealing.

Thus, there are at least some circumstances (if the paradox is overstated), if not all (if the paradox is illusory), where there is an internal obligation to obey the law because it is a law, and a corresponding harm that is suffered if that obligation is not met by an agent recognizing the obligation.

The existence of a normative obligation to obey the law is best illustrated by the comparison of laws that permit behavior and laws that prohibit it. If a parking space is allocated through the use of a parking meter that charges a fixed amount, an agent will decide whether to park there based solely on external matters, including the costs and benefits of parking and paying the fee to her well-being. If, on the other hand, parking was prohibited in that space, and a sanction is in place equal to the fixed charge, we would expect less people to park in the space. This does not suggest that the law has acted as an exclusionary reason for acting (because parking remains possible), but rather that, external costs remaining equal, the internal harm caused by not following the law is an additional cost to be taken into account by the agent.

An interesting challenge to the idea of internal harm is posed by the consideration of the actions of corporations in similar permission/prohibition examples. If corporations treat fines and taxes differently, then the theory that individuals only do so because of a psychological harm is weakened. If corporations have no personality outside of their legal personality, how can a corporation

suffer psychological harm? If corporations do indeed treat permissive taxes and prohibitive fines differently, then it must be some other consideration that affects their behavior.

The possibility that a corporation would treat these types of incentives differently, might be explained in two ways. First, if the corporation breaks the law, its standing in the community will be damaged and so profits may be affected. Thus the behavior is a reaction to external outcomes. However, internal outcomes might also operate to affect their behavior. It is an individual or individuals who ultimately make the decision on behalf of the company as to whether or not it should break the law. In that circumstance, the behavior of the company is affected by the internal harm that the individual(s) would suffer. The harm may be diluted by the individual's removal from responsibility for the behavior, but may be there nonetheless.

In any event, it is not clear that corporations do treat permissive taxes and prohibitive penalties differently. The legal community has long understood that corporations, particularly large corporations, may treat fines as simply another cost of doing business.⁷ If that is the case, then no damage is caused to the theory of internal harm suffered as a result of breaking a law because it is a law. In any event, to the extent that any corporation does treat them differently, that treatment has been explained in a consistent manner above.

Internal outcomes are not usually considered to affect an agent's well-being. That is where the legal and economic theories of behavior in response to law differ.⁸ However, the distinction between, for example, physical or pecuniary harm on the one hand, and internal harm on the other, is unnecessary for the purposes of identifying outcomes that affect an agent's well-being. If that distinction is removed, the economic theory is capable of processing internal harm.

⁷ See, e.g., New South Wales Law Reform Commission, *Sentencing: Corporate Offenders*, Issues Paper No 20 (2001) [3.9]; John C. Coffee, Jr., "No Soul to Damn: No Body to Kick": An Unscandalized Inquiry Into the Problem of Corporate Punishment, 79 MICH. L. REV. 386 (1980-1981); Linda S. Bosniak, 'Exclusion and Membership: The Dual Identity of the Undocumented Worker under United States Law', 1988 Wisconsin Law Review 955,1015 (1988). Cf. Kornhauser, 'Interest, Commitment, and Obligation', *supra* note 3.

⁸ See Kornhauser, 'The Normativity of Law', *supra* note 1.

Having explored the different types of outcomes that face an agent, namely, external and internal outcomes, this article now considers whether law can change the outcomes that an agent might face, or the agent's beliefs as to the likelihood that such outcomes will occur.

Changing Outcomes and Beliefs

As a result of a decision by an agent, a certain outcome will occur. Different decisions may lead to different outcomes. However each decision has only one inevitable outcome. That outcome is made up of an almost infinite number of individual consequences. For example, if I choose to cross a street, individual consequences might be that I step in a puddle, am beeped at by a car, and make it to the other side quicker than I would have if I had chosen to cross the street at a later time. In a certain world, an agent will know the exact consequences of a decision, and so, will choose the decision that results in the most preferred outcome. However, the world is not certain. An agent is only able to identify consequences which she considers possible results of her decision, which may or may not include the consequences that make up the inevitable outcome. Thus, for each possible choice in an agent's choice set, there is a set of subjectively potential outcomes. Corresponding with the subjectively potential outcomes of each of the choices is a belief of the agent as to the likelihood that each potential outcome will occur. That in turn depends on the agent's belief as to the likelihood of certain consequences occurring. For example, consider a jurisdiction where the voluntary assumption of risk is a complete bar to recovery of damages in tort. If I decide to cross the street while there is a 'do not walk' sign lit, I may have beliefs about the likelihood that I will make it to the other side of the street safely. I may also have beliefs about the likelihood that I would be found by a court to have voluntarily assumed the risk of injury if I do not make it to the other side safely. If I think that there is a 50% chance of each occurring, then I will have the following beliefs about the likelihood of outcomes: 50% chance that I will make it to the other side safely; 25% chance that I will not make it safely to the other side and will not be found to have voluntarily assumed the risk of injury; 25% chance that I will not make it to the other side safely and will be

found to have voluntarily assumed the risk of injury. Of course, this simplifies matters in many ways. Not least because the occurrence of one consequence might affect the chance of another consequence occurring.⁹ Regardless, the example helps to illustrate how beliefs are compiled.

It was stated earlier that outcomes may be external or internal. On an objective level, we might conclude that law cannot possibly change the potential external outcomes of a decision. Legal rules can be thought of as coercive orders.¹⁰ The coercion may be a result of authority or of threat of punishment. Whatever form an order takes, it does not change what is physically possible and what is not. It may not be likely that a police officer will arrest me for selling apples if selling apples is legal, but it is possible. There is nothing to physically restrain the police officer from doing so. If the law made selling apples a criminal offence, it has not done anything to make my arrest physically possible. It has, however, merely made it more likely, assuming that the police officer obeys the command of the law and thus is inclined to enforce the law by arresting those who commit criminal offences. Thus, we might conclude that, on an objective level, the law does not affect the set of potential external outcomes.

Internal outcomes are another matter. In particular, a normative obligation to follow a law because it is law will indeed create a potential outcome, even at an objective level. Choosing to undertake behavior that will now be contrary to law, where it was not before, will have a new outcome: the harm caused internally to the agent by not meeting the normative obligation to follow the law. So we see that law can change the objective set of potential outcomes by the creation of new normative obligations. The use of the word potential is perhaps not necessary in this context. An agent either subscribes to a norm or not. If the agent does subscribe to a norm, then breaching it will definitely cause an internal harm. If the agent does not breach the norm, it will definitely not cause an internal harm. These are not situations where the agent will have varying

⁹ It is also unlikely that a pedestrian would consider rules of tort before deciding how careful to be crossing the street: See Gary T. Schwartz, 'Reality in the Economic Analysis of Tort Law: Does Tort Law Really Deter?' 42 *UCLA Law Review* 377, 429 (1994).

¹⁰ H.L.A. Hart, *The Concept of Law* 20-25 (Oxford: Clarendon Press, 1961) .

beliefs about the likelihood of those outcomes occurring. It is more interesting when considering situations where the agent is unsure as to whether a norm will be followed or not as a result of certain behavior. The uncertainty itself acts to mitigate any internal harm that is suffered should the agent reach the conclusion later than a norm was breached. Nonetheless, the agent is bound to suffer a greater internal harm if later it is concluded that a norm was breached, than if a norm was not breached. The uncertainty only mitigates, rather than removes, the harm. Thus, we might conclude that in circumstances where there is uncertainty about whether or not a norm will be breached, then the outcome remains uncertain to the agent.

The agent is concerned only with those outcomes of each possible choice that are foreseeable to the agent. Limitations of cognitive ability mean those outcomes are unlikely to be coincident with the set of objectively possible outcomes. The subjective nature of the definition of the set of potential outcomes might allow us to depart from the objective conclusion that law cannot affect that set. A decision is based only on foreseeable outcomes. Equally, unforeseeable individual outcomes play no role in the agent's deliberation. I say individual outcomes because an agent may give consideration to the potential for unforeseeable outcomes generally. (Rumsfeld's unknown unknowns). What is it then that makes an outcome foreseeable? In the common law world, a foreseeable event is an event that is not 'far-fetched or fanciful'¹¹. Even an event which has an infinitesimal chance of occurring is foreseeable.¹² In the High Court of Australia, Mason J (as he then was) has stated: 'Although it is true to say that in many cases the greater the degree of probability of the occurrence of the risk the more readily it will be perceived to be a risk, it certainly does not follow that a risk which is unlikely to occur is not foreseeable'. Of course, a court necessarily must describe an objective account of foreseeability, usually, based on some standard of the reasonableness. That necessity exists because only the agent will actually know what was and was not foreseen by her. A court has the task of describing a normative account of what should have been foreseen. Nevertheless, the likelihood of occurrence forms the basis of objective and subjective accounts of

¹¹ *Wyong Shire Council v. Shirt* (1980) 146 C.L.R. 140 (Australia).

¹² *Bolton v. Stone* [1951] UKHL 2 (U.K.).

what is foreseeable. We might conclude then that whether an outcome is foreseeable or not depends on the likelihood of its occurrence. There must be some real chance of it occurring in order for the outcome to be foreseeable. Pursuant to that test, my arrest for selling apples if selling apples is legal is unforeseeable because it is in my mind fanciful. I do not think there is a real chance of it occurring.

Thus, the set of foreseeable outcomes are defined by the agent's beliefs about the likelihood of a decision leading to a particular outcome. An outcome will be foreseeable and thus within the agent's subjective set of potential outcomes if the agent believes that its occurrence reaches a certain (very low) threshold of likelihood. A potential outcome that is too unlikely will be unforeseeable and thus not included in the weighing deliberations of the agent. It is not that such an outcome is considered and discarded as too unlikely, but rather, it is so unlikely that it does not enter the mind of the agent at all. Thus, for law to affect an agent's decision by changing the set of potential outcomes, it must either remove or introduce an outcome into the realm of foreseeability of the agent.

So, for the purposes of the agent's decision making process, the set of potential outcomes is defined by subjective measures. If a legal rule can change the beliefs of the agent, then the agent's set of subjective potential outcomes may also be changed. Consider the sale of apples example again. I must decide whether or not to sell apples. If selling apples is legal, then, even though it is physically possible that I could be arrested by a police officer for selling apples, it is so unlikely, that it is not a foreseeable outcome of a decision to sell apples. The chance that it will happen is so insignificant that the possibility of it will not enter my mind at all. Thus, in my deliberations, I will not place any weight on the possibility that I could be arrested. If selling apples was made a criminal offense, and I was made aware of that information, that I might be arrested by a police officer is now likely enough that the possibility of that outcome enters my mind as a foreseeable result of selling apples. Consequently, my decision may be affected, depending on my preference for not being arrested and the weight given to other considerations.

I noted above that possible outcomes identified by the agent may not include the inevitable outcome, either because the inevitable outcome was far-fetched, or because the agent has mistakenly overlooked a potential outcome. In an uncertain world, the agent does not know the inevitable outcome. Consequently, whether or not a law affects the inevitable outcome is not necessarily of any consequence in the agent's decision-making process.

The agent's beliefs about the likelihood of particular outcomes are informed by information that is available to the agent. Uncertainty is a result of imperfect information. Therefore, on the assumption that law cannot change what is physically possible, in order for a legal rule to change the foreseeability of an event, and thus introduce it or remove it from the set of potential outcomes of an agent, it must do so by changing the information available to the agent.

Of course, an agent's behavior may also be affected by a law that changes an agent's beliefs about the likelihood of an outcome that is already within their set of potential outcomes. For example, consider the difference between a comparative negligence regime and a contributory negligence regime. In a comparative negligence regime, a plaintiff is barred from recovery of damages if her own negligence is deemed to have contributed a certain amount to the cause of those damages, say 50% or more. In a contributory negligence regime, a plaintiff is barred from recovery of damages if her negligence is found to have contributed to the cause of those damages at all. If I were a particularly contemplative person and drove a car in low visibility with slightly worn brakes, I might consider the likelihood of my being barred from recovery of damages if I were involved in a crash with a car driving on the wrong side of the road without lights. In a comparative negligence regime, I might believe that there was a small likelihood that I would be barred from recovery of damages. However, if the regime were changed to contributory negligence, then I would believe that there was a greater likelihood of being barred from recovery of damages. That change in belief may change my choice to drive even though it has not introduced or removed the potential outcome of being barred from recovery from my set of potential outcomes.

So whether it changes the set of potential outcomes or not, the law can affect an agent's behavior by changing the agent's beliefs about the likelihood of potential outcomes. A law can do this by creating new information, or by bringing existing information to the agent's attention. How a legal rule might do this is explored below in Part IV. Before that, I consider whether the law might affect an agent's behavior by changing the agent's preferences.

Changing Preferences

The preferences of an agent rank outcomes. In a certain world, an agent will prefer the inevitable outcome of one decision to the inevitable outcome of another decision. In an uncertain world, preference cannot be simply reduced to a preference for the outcome of one decision over the outcome of another. In an uncertain world, the set of subjectively potential outcomes for each decision are ranked. Then the agent will combine preferences over the potential outcomes, beliefs as to likelihood of occurrence, and her risk aversion, to obtain an all things considered conclusion. As discussed above, normative obligation is incorporated in this calculation through the inclusion of internal harm as an outcome over which an agent has preference.

As it was with the choice set, it is difficult to conceptualize how a law might change the preference of an agent. If, in a given circumstance, I prefer x over y, how can a law result in my preferring y over x in the same circumstances? If a motorcyclist prefers the feeling of driving with no helmet, can a law result in the motorcyclist preferring the feeling of wearing a helmet?

There are three claims that might be made as to how law could change the preferences of an agent. First, if a law forms habits in the agent, those habits might lead to a change in preference. For example, if I dislike wearing a helmet, but wear it for a period of time because I may be barred from recovery of damages if I do not, I may grow to like the feeling of wearing a helmet. Then even if that threat is removed, say by the negligence regime changing from contributory to comparative or pure comparative negligence, I continue to wear the helmet because I am used to the feeling, and perhaps prefer the warmth or feeling of safety that it offers. The strength of this claim is undermined by its causal relationship. On

the first day of application of a rule that would bar recovery of damages for not wearing a helmet, the law has had no effect on my preference for wearing a helmet. It is at best an indirect effect on my preferences. The distinction is important particularly if the explanatory theory is to have any application in a design project. If the change in preference requires me to wear the helmet consistently over time, then the other effects of law must be the sole consideration in design of a law that will cause me to do that in the first place. For example, the primary, or direct, effects of law, such as changing beliefs through the provision of information, is required before any effect on my preference can be effected. Nonetheless, habituation makes the best claim as to how the law can change preferences. It is important to note that the habit does not simply change my preference for the behavior that the law encourages. The habit must change my preference for one or more of the outcomes of the behavior. In the helmet example, I might have changed my preference for the feeling that I get when wearing a helmet, rather than a preference for wearing the helmet. If it were otherwise, that would confuse my decision with my preferences.

The second claim relates to the normativity of law. If law is respected by the agent as a norm, it will alter her motivations in some significant way. The claim might be, for example, that if my religious beliefs require me to protect my head by wearing a helmet at all times, then my preference for wearing helmets will change. Similarly, if the law is accepted as a norm, it too will change my preference by altering my motivations. In my view, however, this simply confuses preference changing with changing the set of outcomes. If I accept the wearing of a helmet as a norm, I wear it, at least in part, to avoid the internal harm from not abiding by the norm. That says nothing about whether I nonetheless prefer the feeling of not wearing a helmet.

The third claim is that the law can change preferences through the provision of certain types of information. For example, if the law were able to provide me with the information that helmets improve my safety, the feeling of wearing a helmet might change from being a feeling of discomfort, to a feeling of safety, thus changing my preference for the feeling of wearing a helmet compared to not. However, this claim too appears to suffer from confusion as to the

true underlying change. It may be that the information simply changes my beliefs about the likelihood that I will suffer significant damage if I ride a motorcycle while wearing a helmet. My preference for the feeling of wearing a helmet to not wearing a helmet has been confused with my decision to wear the helmet or not to wear the helmet.

So, we see that preferences might only be changed indirectly by law, by causing habits to be formed by those who decide to abide by the law for other reasons.¹³ Once the preference is changed, the other effects of the law are of less importance in affecting behavior.

Law and Information

It was found above that to affect an agent's behavior, a law must change the information available to the agent, which must change the beliefs of the agent about the likelihood of a potential outcome occurring. This may or may not introduce a potential outcome into the subjective set of potential outcomes of the agent. Over time, if those changes are able to form a habit of the agent to behave in the way the law requires, then the agent's preferences for certain outcomes related to that behavior may also change. To understand how the law might change the information available to the agent, consideration of the categories of outcomes for which the likelihood beliefs of the agent might change, that were outlined above, is important. That will enable us to consider the type of information that might affect those beliefs.

Information and External Outcomes

Sanctions are the most obvious way to change information available to an agent concerning external outcomes. A law describing certain behavior and outlining the liability that will be incurred for undertaking that behavior provides the agent with the information that the behavior may lead to liability of a certain type. The existence of the law is one type of information. The agent will know that the law's existence will make it more likely that legal mechanisms will be implemented to either stop her from continuing that behavior

¹³ But cf Kornhauser, 'The New Economic Analysis of Law', supra note 3, 44.

and/or to punish her for it in the way specified in the law. The size of the sanction is important information. The agent will usually prefer a smaller sanction over a larger sanction. Therefore, a larger sanction will be given greater weight in the agent's deliberation and is thus more likely to change the agent's behavior. This type of information is one that not only changes the beliefs and subjective set of outcomes of the agent, it also probably affects the inevitable outcome of the decision. To be effective, the law must adequately define the behavior subject to the law, the obligation, and the sanction.

The law can also provide information that will assist the agent to make a better informed decision. This relates to external outcomes in a different way to sanctions. For example, a law may require the road traffic authority to publish information about the safety benefits of helmets. This type of information will act to change the beliefs and possibly the subjective set of outcomes of the agent. It will not however, have any effect on the inevitable outcome of a decision. For example, if I now know that a helmet is capable of saving my life in a crash, that does nothing to make the wearing of helmets more or less safe. Of course, this type of effect on behavior is secondary in nature, which raises certain problems discussed later. If I choose to wear a helmet as a result of the information, that is a result primarily of the authority reacting to the law, rather than in direct response to the law itself.

Either of those types of information change the beliefs of an agent, and potentially change her set of subjective potential outcomes, and, if habit forming, her preferences over the set of subjective potential outcomes.

Information and Internal Outcomes

It was found above that two types of internal outcomes relate to behavior that may break a law. Psychological harm suffered by not following a norm that overlaps with the subject matter of the law, and psychological harm suffered by not following the norm that the law should be followed.

The information provided by law relating to the second type of internal outcome is that the behavior required by the norm now extends to the subject matter of the law in question. In other words, the norm that I should follow the law has subject matter that equates to subjects dealt with by the law. That subject matter will change as laws change. A new law on subject matter not previously dealt with by law will inform an agent that the norm now covers that subject. The agent already knows that not following that norm will cause internal harm, assuming she subscribes to the norm, and so will now believe with certainty that internal harm will be caused if she behaves in the manner now prohibited. For this result, the law merely needs to set out the subject matter and the obligation. A sanction is not required.

Information provided by the law might also relate to the first type of internal outcome, that is, the harm suffered by the agent when not following a norm that is coincident with the subject matter of the law. It can do this if a law can lead to the creation of a norm covering the subject matter of the law. That is, a norm that would continue to exist without the law itself. A thorough discussion of norm creation is beyond the scope of this paper, however, it might for example occur as a result of the enactment of a law because of the symbolic value of the law.¹⁴ The relation, even if it did exist, however, is secondary. The information must be information of the type(s) discussed above in order to change behavior through either external means or by the extension of the subject matter covered by the norm that the agent should follow the law. Like in the formation of habits, the process must be initiated by a direct change in behavior by the law.

Ramifications for Economic Analysis of Law

We have seen that law can affect behavior by changing beliefs, which may also create new subjective potential outcomes if those outcomes are not already within the agent's set of subjective potential outcomes. Law may also create an inevitable outcome where the agent subscribes to the norm that the law should be

¹⁴ See, e.g., Shubha Ghosh, 'Where's the Sense in *Hill v Gatenay 2000*? Reflections on the Visible Hand of Norm Creation', 26 *Touro Law Review* 1125, 1129 (1999-2000).

followed. This may change behavior as a result of internal harm that an agent would suffer if she breached that norm. Indirectly, the law might change behavior through the creation of new norms that coincide with the subject matter of the law, such that the norms would remain in the absence of the law, which in turn create new inevitable outcomes; or through the formation of habit, which can in turn change the preferences of the agent.

Drawbacks for the Predictive Project

Attempts to use this explanatory theory to predict behavior will encounter difficulties in at least three areas: quantifying internal harm, predicting secondary effects, and irrationality.

Quantifying Internal Harm

While the quality of response to internal effects of law are easy to establish, their quantity is more difficult to predict or measure. Measurement of the effect of sanctions is a task that is familiar to an economist. However, the internal harm suffered as a result of not complying with a norm is more difficult to quantify. Cooter suggests that the harm can be equated with the intrinsic value of obeying a law.¹⁵ In his view, the decision as to whether or not to obey the law is dependent on three measures: the net cost of obeying the law, the net cost of disobeying the law and the agent's willingness to pay to obey the law. In other words, he measures the internal harm suffered by an agent by the amount of money the agent is willing to pay to avoid it. The last measure thus converts the internal harm caused by breaking the law into a monetary value. However, Cooter thus ignores the operation of two distinct types of internal harm, that which results from the existence of two separate norms discussed above. By condensing internal harm to a single measure, the effect of the law is mixed with the internal harm that would have been suffered in any event from an agent not following the norm which coincides with the subject matter of the law. It would be difficult, if not impossible for an agent to separate her willingness to

¹⁵ Robert Cooter, 'The Intrinsic Value of Obeying a Law', 75 *Fordham Law Review* 1275 (2006-2007).

pay to follow the norm that she should not negligently maintain lighting in a stairwell because it is a law from her willingness to pay to follow the norm that she should not negligently maintain lighting in a stairwell because it is wrong. How then, can we measure the effect that law has on an agent through its creation of the outcome of an internal harm?

It is possible, if we continue the assumption that the norm that individuals should follow laws because they are laws does not attach itself to each law individually, but rather to a body of law as a whole. If that holds, then the internal harm will be identical for each law. All that is required then to measure that harm is a law, the subject of which does not coincide with any other norm subscribed to by the agent. An agent's willingness to pay to follow that law will be the equivalent of that agent's internal harm that would be suffered as a result of breaking a law because it is a law. Of course, the internal harm related to a subject matter norm could be calculated by asking the same question of the agent concerning a law, the subject of which does coincide with the other norm. The second type of internal harm will be the difference between the two willingness to pay figures.

For example, imagine a new law that says that citizens should wear blue wigs on Fridays, and there is then no pre-existing norm of that content. Agent A is willing to pay x dollars to follow that law. There is also an old law that says that citizens should wear red hats on Saturdays, which coincides with a norm in society that red hats should be worn on Saturdays (perhaps because that law has existed for long enough to create the norm). A is willing to pay y dollars to follow that law, where following that law will also have the effect that A meets the obligation of the subject matter norm. In that example, the internal harm suffered by A should she not follow any law is x . The total internal harm suffered by A if she does not wear a red hat on Saturday is y . The internal harm suffered by A because she did not follow the norm of wearing a red hat on Saturdays (rather than because it is law) would be $y-x$.

Another problem raised by Cooter is that some outcomes are incommensurate with money, which is usually the common measure used by economists.¹⁶ This incommensurability might be the result

¹⁶ Ibid.

of practical or theoretical problems. The practical problem is that the measure requires agents to identify the amount of money they are willing to pay for a given outcome. That measure is obviously limited by the wealth of the agent. For example it makes little sense for an agent with total wealth of \$1,000 to say that they are prepared to pay \$10,000 for a given outcome. Cooter states that willingness to pay only measures value so far as an agent is able to pay. However, that limitation might be overcome by asking not how much an agent is willing to pay, but rather, how much they would have to be paid in order for them to be willing to not follow the law.

The theoretical problem raised by Cooter is that there are intrinsic reasons why some goods cannot be given a monetary value. Death, disfigurement, loss of reputation and permanent injury are offered as examples of such goods. However, in the context of the explanatory theory, this problem does not apply. The only effect of law to which this problem could conceivably apply is the internal harm suffered when not following the norm that the law should be obeyed. However, the fact that many people break the law suggests that the internal harm is small enough to be commensurable with money. Even if some people do treat that norm as an exclusory reason for decision, and as such, no amount of money could compensate them for breaching that norm, the problem does not appear to affect the explanatory or predictive project. Indeed, it seems to make prediction far easier. If an infinite weight is placed on the avoidance of that harm, then any set of outcomes where the norm is not breached is preferred to any set of outcomes where the norm is breached. In such situations, the behavior of the individual is predicted with perfect accuracy. Problems might arise however, where two or more outcomes are incommensurable. If that were to occur and the outcomes always acted in the opposite direction, then predictions using the willingness to break the law test are unhelpful.

Predicting Secondary Effects

Further difficulties in the use of this explanatory theory for prediction occur when considering measurement of the indirect effects of law on behavior: providing information for better informed decisions, preference changing through the formation of habits and norm creation. Each requires the passage of significant

time, which makes prediction less accurate. Of course, the necessity for the passage of time creates complexities, but so does the secondary nature of law as the cause of the behavior.

One difficulty caused by the secondary effects is the dilutive effect of removal from the law itself. The efficiency of law's influence will dissipate as the agent is removed from the law. If a law in its primary application is unsuccessful in changing behavior in the manner intended, then its indirect, or secondary influence is less predictable.

Of course, the change in behavior accorded to the primary influence of a law will also be unpredictable to the extent that the predictive project is imperfect, but those inaccuracies will be magnified as the distance between law and agent increases. It seems that those magnified inaccuracies can only be removed if an assumption is made that individuals will do as the law says. However, such an assumption is unrealistic. Thus, any prediction based on the formation of habit or creation of norms is bound to be inaccurate on some level.

Irrationality

Furthermore, an agent will not necessarily act rationally. Conduct may be pursued that does not, or does not appear, to be in the best interests of the agent. While it is tempting to disregard such conduct as simply irrational, such conduct may instead point to a gap in our theory of rational behavior. By way of example, Sobel has considered the destructive behavior of fired employees, which does not appear to be in their material interest.¹⁷ He suggests that such behavior might be explained in many rational ways if we do not require the employee to be in pursuit of maximum monetary income. One such explanation is that such an employee might attempt to maximize a utility function that depends not only on his own income, but also that of his employer. By damaging the employer's income, the employee increases his own utility.

Of course, to some extent, complete rationality is not possible due to limited access to information or limited computational abilities of

¹⁷ Joel Sobel, 'Interdependent Preferences and Reciprocity', 43 *Journal of Economic Literature* 392.

agents.¹⁸ To compensate, an agent may apply approximations, or 'rules of thumb'. Such rules will not always arrive at the best decision possible, but may produce better decisions on average over time than random decisions, or lead to a more desirable outcome than would be obtained if the agent spent sufficient time to gather the necessary information and make calculations.

The explanative theory will be unhelpful in predicting individual circumstances where behavior appears irrational. However, unless agents are consistently or often irrational, irrationality causes less significant problems for the theory than the other difficulties described above. In particular, irrationality will be less significant in predicting average effects of laws on behavior than the potential difficulties associated with quantification of internal harm and prediction of secondary or indirect effects. Irrational agents may react to laws in a way that causes some behavioral outliers, the significance of which are negated upon aggregation with the behavior of all affected agents.¹⁹

Improvements made by the theory

The theory takes ideas from the obligation theory (the idea that norms create exclusionary reasons for action) and incorporates them into the rational choice theory, by viewing obligations as creating the potential for internal harm. That harm is ranked alongside other outcomes as a preference of the agent. Indeed, agents who do behave as the obligation theory predicts might simply be said to suffer such great psychological harm when breaching a norm, that the avoidance of such harm is always preferred and thus outweighs other reasons for action. The theory is also better at explaining why an apparent obligation resulting from one norm might be overcome by another; a breach of each norm causes different levels of psychological harm, which might be weighed against one another and the other outcomes to find a preference.

¹⁸ See Herbert A. Simon, 'A Behavioral Model of Rational Choice', 69 *Quarterly Journal of Economics* 99.

¹⁹ See Gary S. Becker, 'Irrational Behavior and Economic Theory', 70 *Journal of Political Economy* 1, 13.

To illustrate, I will use the example of a driver approaching a red traffic light. First, imagine the driver is taking her father, who has suffered a heart attack, to hospital. She knows that the quicker she arrives at the hospital, the more likely her father is to survive. She is under a normative obligation to arrive at the hospital as quickly as possible. At the time, other cars are approaching the intersection perpendicular to the driver and face a green light. Nonetheless, the driver speeds through the red light, narrowly avoiding the other cars, in order to arrive at the hospital earlier than she would if she had waited for a green light or for the other cars to pass. We see here an example of competing normative obligations. On one hand, the driver is obligated to get her father to hospital as quickly as possible, on the other, the driver has a normative obligation to other drivers to drive safely. Furthermore, the driver has an obligation to follow the law, which includes driving with due care and stopping at a red light (in the circumstances, the subject matter of the law overlaps with the normative obligation to drive safely). The driver's external considerations include the possibility of a fine for driving through a red light, the possibility of crashing, being held liable for damages suffered by other drivers and causing herself and her father injury, and the effect that her father's eventual fate will have on her well-being.

In this example, the standard rational choice theory does not take into account the potential for internal harm that is caused by driving negligently and not stopping at the red light. However, in light of the driver's decision to drive negligently and go through the red light, the consequences for predicting her behavior were slight. However, the consequences become more obvious as the degree of emergency decreases. Before considering those situations however, it is worth noting that the explanatory theory outlined in this paper is better able to explain the decision in the face of competing normative obligations than the obligation theory. The driver would simply suffer a greater internal harm by not meeting the obligation to get her father to hospital as quickly as possible than she suffered by driving negligently and proceeding through the red light, thereby breaking the law and not driving in a safe manner. Thus she has a greater preference for the internal outcomes associated with driving negligently and proceeding through the red light than not doing so. That preference is so high that it has also outweighed the potential

external outcomes and the competing normative obligations, which would cause different internal harm.

Now consider the situation as the level of emergency decreases, for example, if the driver's father was suffering from a non-life threatening injury, or if she was simply running late for a flight, or for work, or merely wanted to shorten her daily commuting time. Over the range of these circumstances, as we attempt to predict the effect of the law prohibiting a driver from driving negligently or from proceeding through a red light, we expect the standard rational choice theory to underestimate the effect of the law, as it will ignore the potential for internal harm, instead focusing on external considerations such as a potential fine or tortious liability. As such there will be a range of emergency where the standard rational choice theory will predict that the driver will drive negligently and proceed through the red light, when in fact she does not because of her reluctance to suffer the internal harm associated with not following a law. The explanatory theory outlined in this paper will better predict her behavior by allowing for the avoidance of that internal harm as an additional incentive.

Conclusion

This article sets out a theory of the manner in which law can affect human behavior. However, the explanatory theory described holds significant difficulties for those wishing to use it as the basis for a predictive theory. The complexity of the issues involved and uncertainties created by such issues as internal harm, secondary effects and irrationality suggest that a predictive project would need to further improve the explanatory theory. Nonetheless, it should offer better results than the standard rational choice account by incorporating the effects of internal incentives, otherwise ignored by the standard rational choice theory as not relating to well-being.

Acknowledgments: The author wishes to thank Professor Lewis A. Kornhauser (New York University School of Law) for his helpful comments on an earlier draft of this article. The author retains responsibility for any errors or inadequacies remaining.