FOLLOW-THE-LEADER? MEASURING THE INTERNALISATION OF LAW

Shaun LARCOM, Luca A. PANZONE, Timothy SWANSON

FOLLOW-THE- LEADER?

MEASURING THE INTERNALISATION OF LAW[1]

Shaun Larcom[a]; Luca A. Panzone[b]; Timothy Swanson[c]

[a] University of Cambridge;

[b] Newcastle University;

[c] Graduate Institute Geneva (IHEID). Corresponding Author.

Abstract

Expressive law is said to induce compliance with stated principles without a price on non-compliance. We empirically assess this proposition, by attempting to dis-entangle the impacts of a legal change (a 5p charge on use of plastic bags), on individual choices. We do so by measuring both behaviours and attitudes across the first two months of the legal change, and by comparing the impacts across neighboring jurisdictions both with and without the change. Using mediation analysis, we find that the self-reported change in internal motivation explains only about 10% of the change in behaviour. Interestingly, we find that the scale of the sanction (charge) is both irrelevant (because jurisdictions without sanctions still exhibit changed behaviour) and important (because the size of the sanction signals the reasonableness of the law).

Keywords: expressive law; internalisation of law; behavioural; signalling.

JEL codes: Q5; D1; K1; C4.

---

1. Introduction

How much of real-world behaviour is driven by stated principles of law, and how much of it is motivated by the economic incentives implicit in law's sanctions? For example, do drivers stay stationary at red lights in the middle of the night because it is a social obligation to do so, or because of the fear of the sanction that might result if they do not?[2] That is, is this a principle being observed, or a price being avoided? And what is the nature of the price/sanction that we fear? Is it a fine, the potential loss of a license, or the opprobrium of other motorists? Clearly there is some of each occurring at every red light in every society. Is it possible to separate out between these various functions of law, and to understand better their interaction?

A survey of experimental literature on the issue provides two real-world examples that illustrate the potentially divergent impacts of sanctions as economic incentives. (Bowles and Polania-Reyes 2012) They cite the classic example of the Haifa school that introduced a fine against late pick-ups, inducing a rash of lateness.[3] And they contrast this with the example of the Irish plastic bag charge, in which the introduction of a nominal 15 cent charge on each bag resulted in an almost immediate and absolute reduction of 94% in the regulated activity. (Convery et. al. 2007) They argue that these examples, and many experimental results, demonstrate that law does not always operate in the manner of a system of prices applied to non-conforming behaviour. Law is sometimes seen instead to express the societal, principled view on appropriate individual conduct in particular circumstances – and penalties play a role ancillary to this expressed principle, sometimes in support and sometimes not.

Legal scholars such as Cooter (1998), McAdams (1997), Bar-Gill and Fershtman (2004), and Sunstein (1996) have long defined the direct impact of statements of legal principle as the '*expressive*

---

[2] And why do I avoid parking in disabled spaces in parking lots that are completely empty?

[3] There is a substantial literature demonstrating that this expressive effect may sometimes be observed by consequence of its occasional negation. The classic example is the natural experiment in Haifa of school-imposed late fees imposed on parents for the late collection of children (Gneezy and Rustichini 2010). In this example the parents responded to the late fee by increasing the rate at which they were tardy, using the sanction as the price of additional childcare rather than as a sanction or guide to anti-social behaviour. Similarly, new policy-making by the Boston Fire Department back-fired when they introduced a fixed schedule of "sick leave days", and fines for excessive leave. The response to the fines was an increase of more than 100% in the number of days claimed, especially around holiday times. (Belkin, 2002; Greenberger 2003) Fines for staying-on in hospital beds in Norway had the same effect (Holmas et. al. 2010). This tradition of potentially perverse responses to economic incentives harkens back to the example of payment for blood donations first cited by Titmuss (1971), and demonstrates that norms and principles (without express sanctions) can indeed have an own-effect.

function' of law: behaviour modification through a process of internalisation.[4] Here, when we speak of the *expressive role* of the written law, we will mean that part of written law which functions primarily by guiding the public towards selecting a stated "preferred action" as designated by the lawmaker for the implementation of the stated legal principle. In other words, expressive law is believed to have effect through mere recording of legal principle and preferred action, and consequent self-enforcement, and with little reference to the price placed on non-performance.[5] The question we want to address is the measurable extent to which such a phenomenon exists, and how it operates.

This paper does so by analysing the effects of the adoption of a law similar to the Irish tax in England a few years later. In order to ascertain the impact of the legal change, we use the existence of changes in neighbouring jurisdictions to perform difference-in-differences estimation. We also wish to estimate the extent to which the mere expression of the law had effect.[6] This we pursue through a staged survey of UK residents occurring immediately before, and then in approximately month-long intervals after the introduction of the policy. In this way we assess the relationship between stated differences in changed behaviour and changed values, and compare this to the overall response. We then undertake a mediation analysis, attributing that part of the changed behaviour that does not appear to be related to self-reported changes in "social or moral responsibilities" to the impact of the economic incentive.

---

[4] That is, the State makes a 'statement' about appropriate behaviour through legislative change that can then lead to a change in prevailing norms, preferences and beliefs. Etzioni (2000, page 167) observes that 'internalisation is a remarkable process through which imposed obligations (compliance with which must be forced or paid for) become desires'. Once a law is internalised, some scholars such as Cooter (1998) and McAdams (1997) suggest that a psychological penalty (e.g. guilt, shame) can apply to the act, which discourages the negative behaviour beyond the actual price. Hart (1997, page 89) considers that internalisation involves a shift in viewpoint, whereby one moves from 'observer' to a 'member'; then the member uses the law as a guide for his or her conduct and in determining what is good (or right) behaviour. Rode et al. (2015, page 271) highlight the coordinating role of law and norms: 'it's easier for environmentally minded car drivers not to speed on the highway when a general "speed" limit and fine prevents other drivers from overtaking them'.

[5] In fact, it is possible for the relationship between principle and sanction/price to be a perverse one, so that when a law or policy is introduced, and is perceived solely as a price (and not an obligation), then this expressive function may be lost and the policy may backfire. This is then the case (as in the examples cited above) where the parents, the patients and the patrolmen viewed the intended sanction as the simple unit cost for a single violation of the policy-based norm, and discontinued their observance of the expressed common duty.

[6] The English law included a price-based sanction of 5p for continuing use of plastic bags, and so it is necessary to acquire additional information from consumers to assess the extent to which they responded to the principle or the price.

Our findings are that consumers responded to the English law, with changes in behaviour that outweighed the reported changes in internalised norms; the law has some expressive effect but more of its impact is due to the more usual sorts of channels (information, prices).[7] In addition, we find interesting forms of interaction, both across jurisdictions and also across the different types of agents' responses. Across jurisdictions there is a clear impact of the expressive function of law, even when no economic incentive has changed, since behaviour alters even when sanctions do not. Nevertheless, across agents, the price used in the policy intervention appears to act as a guide to the level of internalised behaviour modification. That is, information on the quantum of the social obligation, as much as its existence, is important to both the price-based and the motivation-based responses. These conclusions support much of the recent theorising concerning expressive law, regarding its internal motivational and its informational aspects. (Benabou and Tirole 2011a)

We proceed as follows. In section 2 we review the literature on the two potential impacts of legal intervention, and set out our strategy for analysing this policy experiment. In section 3 we set out the basics of the model we use (from Benabou and Tirole 2011a) and describe its potential implications. In section 4 we describe the policy intervention and the survey conducted around it. In section 5 we set out the basic results from the survey and our analysis of what it implies for the various approaches to expressive law. In section 6 we provide a brief discussion of our results, and how they provide evidence on the measurement of the internalisation of legal expression. In section 7 we conclude.

2. Literature Review : Expressive Law and its Functioning

The classic literature on agent response to legal change has long suggested that the introduction of a new (legal) principle or policy can generate responses consisting of both price-based and

---

[7] Note that our definition of the expressive effect of written law separates out between the informational role of law and the written effect of law. Others have noted that the role of law may be to inform agents on the appropriate action to take in specified circumstances, with the informational role of law providing a basis for coordinating actions and thereby affecting behaviours. (Raz 1999) We are distinguishing between this function of law and the purely expressive function of law, which operates through altering internal values of agents rather than by altering the information base. See also the discussion of the difference set forth in fn 11, below.

internalised or inculcated effects.[8] However, the way that internalised motivations and economic incentives interact has always remained uncertain and unpredictable.[9] Over the years, numerous mechanisms have been identified that can generate this wide range of potential interaction effects (sociological, psychological, philosophical).[10]

We would like to understand better the extent to which expressive law has effect through economic mechanisms (information, inculcation), and to that end we are interested in three fundamental questions at the base of this literature. First, how does expressive law have effect, via the transmission of information or of values or both? Second, is it possible to measure the extent of the inculcation of values, i.e. the internalisation of law? And, finally how do the sanction and the legal change interact in producing the overall societal outcome? That is, our enquiry is an attempt to separate out between the expressive and the incentive-based impacts of legal change, as well as their interaction.

We will investigate the operation of expressive law through two distinct approaches that have developed in the economics literature: that relating to the inculcation of preferences and that concerning the dissemination of information.[11]

---

[8] We will be considering legal change as equivalent to many other forms of policy intervention, such as organisational change or change in enunciated standards. These have been examined in many different contexts over a long period. Titmuss (1971), Frey (1992), Gneezy and Rustichini (2000) and many others have highlighted how internalised motivations serve an important function in the implementation of policy, and that these motives can be altered through the introduction of monetary incentives. A related literature considers how social norms impact adherence to specific modes of behaviour, e.g. energy conservation, but looks less at how this relates to specific changes in enunciated standards. (Arimura et. al. 2016)

[9] The interaction can be positive ("crowding in") or negative ("crowding out"). While the economics literature has focused more on crowding *out* than *in* (Rode et al., 2015, Frey and Stutzer, 2008), mechanisms have been identified that generate increased internal motivation from the implementation of economics incentives.

[10] Mechanisms that have been identified include: control aversion and frustration generated by a third party seeking to alter their behaviour (Gneezy et al., 2011, Frey and Stutzer, 2008, Bowles, 2008, Goeschl and Perino, 2012); reduced internal satisfaction or 'warm glow' as individuals no longer feel good about themselves as they did when acting on a voluntary basis (Bowles and Polonia-Reyes 2012, Frey and Stutzer 2008); reduced image motivation from being unable to signal to others that they are a 'good person' (Gneezy et al., 2011); release from moral responsibility (Frey, 1992); and frame shifting (e.g. moving the decision from a moral to an economic one (Frey, 1992, Cardenas et al., 2000).

[11] The initial discussion concerning the role of expressive law-making focused on its potential for efficiency in aggregate rule-making (and as a substitute for individual decision making) by reason of minimising the *costs of contracting*. Perhaps the first modern discussion of the idea emphasised the view that the regulator is able in many instances to solve a general problem faced by many individuals in society, replacing the need for many multiple individual solutions of the same. (Raz 1999) This approach assumes that individual preferences are

2.1 Expressive Law as the Inculcation of Values

The idea of *inculcation* is that governments may shift preferences in order to have a group of agents self-motivated to implement and enforce the expressed law. Perhaps the first exponents of the modern idea of individual preference-based adoption of legal constraints were Kaplow and Shavell (2007). They argued for the recognition of the internal values of "guilt" and "virtue", which are intrinsic values associated by government inculcation with individual action or in-action in a categorical situation. The regulator is somehow shifting individual preferences in the direction of the preferred action in a categorical (i.e. well-defined) situation. Here, expressive law functions as a pricing system, but one that is internal to the individual decision maker (self-imposed rather than through monitoring and external penalties).[12]

The incorporation of these direct forms of utility (that flow from action or inaction) clearly make law a system for altering individual preferences. The authors assume that both guilt (negative increment of utility for inaction) and virtue (positive increment of utility for action) may be inculcated into individual objective functions by expenditures on "inculcation". In this view of expressive law, individual objectives become an amalgam of both the direct utility received from specific acts in given situations and the indirect utility received from conformity with stated legal principles. A

---

relatively stable in the face of expressive law, but that the law-making process involves the use of asymmetric information (on costs, benefits of an act) on the part of the regulator/Principal to solve an "average problem" on behalf of many agents, thereby avoiding the costs of individual information acquisition and calculation. It points to the problem of Type I and Type II errors – on account of individual costs deviating from those of the aggregate – meaning that the general result may not be efficient for the individuals in the tails of the distribution. (Shavell 2011).

Another early proponent emphasised that the expressive role of law provides a principled basis for the coordination of behaviour, such as red-green stop lights that solve asymmetric choice problems. These would then be recognised as the basis by which efficient coordination may be achieved. (MacAdam 1997) Similarly, in experimental contexts, there may be a potential "focal point" aspect to providing statements of "principles", in that players in games may view common knowledge of morals or principles as potential solution concepts. (Dal Bo and Dal Bo 2015) In this case legal principles are probably being viewed as aiding cooperation by filling in missing terms in contracts, when transactions costs are high. (Gneezy and Rustichini 2000)

[12] Perhaps a closely related notion is the Bowles and Polania-Reyes (2011) concept of social preferences, in which individuals receive utility both through individual preferences and also through performing acts with social consequences. Although these authors are not principally interested in the idea of expressive law and how it operates, they do express the opinion that in some important contexts law and policy operates by express shifting of these social preferences.

related idea is that non-compliance can have a momentum of its own, and that law must build in additional penalties against this externality.[13] (Shavell 2011)

Other authors imply that individual preferences are pliable, but that they will tend to follow average conduct rather than government admonition. (Lin and Yang 2006) In this case, the government is not guiding society by shifting preferences so much as it is recognising the capacity for conformity to aid in compliance. Then the function of expressive law is to serve as the conduit for official statements of common standards of behaviour.[14] The analysis remains the same, i.e. the role of expressive law is to shift individual choices by shifting individual preferences, but the vehicle is through statements of "common standards" rather than legal obligations. In both cases, there is some internal mechanism that then causes the individual to shift its own choices toward that standard or obligation.

2.2 Expressive Law as Dissemination of Information

The alternative to the idea that expressive law operates through inculcation of individual preferences is that such law operates through *information supplementation*. One obvious example is via information dissemination that directly impacts reputation, for example through "naming and shaming".[15] (Besley and Ghatak 2016)

There are other more indirect ways in which governmental dissemination of information may impact individual choice, derived from the extension of motivational theories previously applied

---

[13] Shavell (2011) has raised the idea of the compliance externality, in which the behaviour of potentially noncompliant individuals is witnessed, and reinforces or encourages the choices of others (toward noncompliance). Penalties may then have to be increased to internalise the costs of un-monitored non-compliance. Non-compliance becomes an additional socially harmful act, in addition to the one that is specified in the law, and so penalties must account for both potential forms of deviation.

[14] A related notion is that of culture or conformity, in which individual preferences have a tendency to move towards the norm or average conduct in a given situation. The idea of a tendency to conformity (a dynamic externality) has also been raised as a possible reason for statements of principles. If individuals automatically wish to move toward the "average" or norm, then this tendency may be exploited by simply stating the desired outcome or norm (rather than the price for deviation). (Lin and Yang 2006)

[15] If reputation affords an important conduit through which utility is received, it has been argued that the consequence of publically honouring or shaming an individual for non-observance of a social obligation can be a very effective form of implicit sanction in some contexts, e.g. in public hospitals, rendering the need for explicit penalties redundant. (Besley, Bevan and Burchardt 2009; Kreps, 1997, Besley and Ghatak, 2016)

within organisations. (Benabou and Tirole 2011a; Benabou and Tirole 2006) It is also based on the idea of individual morality and reputation developed by the same authors. (Benabou and Tirole 2011b) In this framework the agent receives direct utility from conforming with the expressed legal principle in an amount varying with its individual type, which is drawn from a society-wide distribution that is known to the regulator/principal but not necessarily to the agent. The agent receives intrinsic value relative to its own valuation of the action, and also from the aggregate actions of all others consistent with the legislative principle (i.e. the sum of social behaviour that is in accord with the expressed legal principle). Finally, and most distinctively, the agent also receives a benefit from the enhancement of its reputation by taking the indicated action – to the extent that the action changes the expectation of its perceived type in the perception of the remainder of society.

Here the government operates not by shifting preferences (which remain stable) but by disseminating information (over which it may have an advantage). Given that individual choices generate own value, both in interaction with other individuals' choices and also by others' observation of own-choice, the government is able to impact individual returns by providing information on the societal distribution of preferences. Specifically, the principal may be able to use the fact that the sanction or penalty imposed for non-conformance implies information on the distribution of intrinsic values. For example, a high penalty implies that the desired behaviour is individually costly to elect, while a low penalty implies the opposite.[16] For this reason, the penalty chosen by the government is an important instrument in expressive law – as much so as the statement of the principle of law itself.

2.3 Related Empirical Literature

There is a large empirical literature that relates to these issues, with virtually innumerable experiments on related questions. (See, for a recent survey of experimental results, Bowles and Polania-Reyes 2012) Perhaps the closest example of an experimental study pursuing the same ideas as those here is Dal Bo and Dal Bo (2015). There they introduce an interesting series of experiments that analyse the extent to which moral suasion or sanctions are important to cooperation. Again, they

---

[16] Low sanctions indicate that the requested action is "reasonable" in nature, and so low-cost to individuals.

are pursuing this idea in an experimental setting, and observing the extent to which players cooperate in game situations when they are reminded of important principles of cooperation as moral suasion (e.g. "Golden Rule"). They find that moral suasion is indeed effective in itself as a motivating mechanism, but that it interacts positively with sanctions (the ability to punish non-cooperators) in producing a non-transitory impact. In a final application, they find that moral suasion works in a non-informational sense, i.e. that players perform more cooperatively when they are informed of the guiding principles for cooperation even though they know that the other players are not. They also find that the availability of (private) sanctions interacts positively with the effect of moral suasion.

This raises the empirical question of how penalties are seen to interact with principles in various field studies: Does the effectiveness of law vary with the economic incentive being introduced? The studies on this question regarding tax compliance and evasion are voluminous. (Kleven et. al. 2011; Fellner et. al. 2013; Dwenger et. al. 2015) These studies sometimes find that evasion of the legal duty can increase with the sanction. (Berger et. al 2015) A recent field experiment with library late fees found that notices that incorporated information on penalties increased compliance. (Apesteguia et. al. 2013) A similar result was found with regard to TV license payments. (Fellner et. al. 2013). [17]

In general, these are studies implemented in regard to already existing laws, examining the question of how various forms of sanctions impact upon levels of non-compliance.[18] We are pursuing a slightly distinct question: How does the quantum of the penalty impact upon the *expressive effect* of the legal principle at the time of its adoption?[19] That is, we wish to see if it is possible to measure the pure impact of the expressive function of law. And we wish to understand better how the two distinct

---

[17] There is also a related literature on how public and private enforcement inter-relate. A recent experiment found that public sanctions were important to the effectiveness of private enforcement measures. (Romaniuc et. al. 2015) And, the closely related study by Dal Bo and Dal Bo (2015) came to a similar conclusion regarding the interaction between moral suasion and private enforcement.

[18] A study that falls somewhere in-between is Boyer et. al (2016) in which constituents subject to an unenforced existing law are advised (in various treatments) of the law's existence or of its voluntary nature. This is a somewhat more ambiguous environment within which to test the expressive nature of law, simply because the authorities do not seem to be very well persuaded themselves of the legal principle being expressed (i.e. that payment of church levies is required conduct).

[19] Of course the expressive function of law may have effect at any point that the government makes announcements or admonitions, but it seems that the purest measurement of its impact would be possible at the time of initial adoption.

motivations (price and principle) interact to produce a joint societal outcome. We turn now to those tasks.

3. Modelling the Impacts of Expressive Law: Alternative Hypotheses

In this section we briefly set out the information-based model used by Benabou and Tirole (2011a) in their description of the functioning of expressive law, and then we use it to describe the alternative ways in which law might work.

This is a standard Principal/Agent sort of framework in which the principal designates an act (a) together with an incentive (y) for the agent's choice of the act. (Benabou and Tirole 2006) The agent does so at an individual cost (c). Here the "agent's preferences (should).. encompass the three key ingredients of intrinsic motivation, extrinsic incentives and (social or self) esteem":

$$U(v, e, \mu) = (v + y - c)a + \epsilon \, \Sigma a + \mu(E(v \,|\, a)) \qquad (1)$$

where v represents the agent's intrinsic motivation, and represents the value that the individual places on the performance of the indicated action (a).

The agent also derives a benefit (ϵ) from the aggregate supply of the expressed action a. This is the society-wide external benefit from the individual choice of the specified act. This term represents the potential for society-wide momentum towards the act. This is because knowledge of the existence of this joint benefit could imply the joint optimality of the choice of action (a) (with c low enough). The mechanisms that could make such collective action feasible might include mutually-enforced monitoring or avowed custom.[20]

And finally, the agent receives a reputational benefit μ from any improvement in its imputed individual character or "type" by reason of the choice of the action. This reputational benefit is acquired through public performance of the act, thereby revealing the individual's intrinsic value for the act.

---

[20] Of course this depends on whether the mechanism in the society arise to make the jointly optimal outcome attainable, as the individual optimum would be of course to free-ride on others' choice of a.

A critical difference in the Benabou-Tirole model is the manner in which sanctions impact behaviour. In the standard Beckerian model of law, an increase in government penalties/sanctions will always increase compliance with the law. In the Benabou-Tirole model, however, sanctions serve an additional signalling function. Any given individual reveals information regarding its own individual type (v_i) by its choice regarding action (a). Given a specific sanction (y^*), v^* represents the corresponding cut-off value of v for the marginal individual performing the act given this level of sanction. That is, all individuals following the injunction of an expressed law will have type $v > v^*$, where:

$$v^* : \ v - \ c + y - \mu \Delta (v^*) = 0 \qquad\qquad (2)$$

In this formulation, the $\Delta(v^*)$ represents the change in reputation accruing to each individual electing action a, and depends primarily on the societal distribution of intrinsic preferences $g(v)$.

The final two results we borrow from these authors concern the manner in which conduct alters with incentives (i.e. dv^*/dy) – in the case where the government recognises the importance of the particular form of the distribution of preferences. First, where the distribution of preferences is known, there is a *social multiplier effect* associated with the imposition of a sanction (y) that derives from the fact that a higher sanction implies a more uncooperatively skewed distribution of intrinsic motivation. That is, the higher sanction conveys by its relative value alone that the proportion of the population that will elect to implement the action in the absence of an incentive is relatively small. This implies a supply of conduct curve that is declining with an increasing subsidy (or increasing in sanction cost).

$$a(y) = g\big(v^*(y)\big) \frac{1}{1+\mu(\Delta'(v^*))} \qquad\qquad (3)$$

Implicit within the above model of individual preferences is the fact that the choice of penalty should take into account that individuals will receive both an individual benefit from compliance (the intrinsic benefit) and a reputational benefit from compliance (the change in perception). For this reason, a minor sanction/penalty will have a larger impact than would be

expected given just the price being charged. So - the first point of an informational theory of sanctions is that relatively low sanctions will have a relatively greater impact on choice.[21]

In fact, it is possible under the Benabou-Tirole model for sanctions to be zero, i.e for self-enforcement to be the only form of enforcement required. That is, if the distribution of preferences is such that the expressed legal principle calls out a relatively low level of intrinsic preference for giving effect to that act (i.e. $v^*(a)$ is relatively low), then that level of supply of the desired conduct would occur without sanction.

A corollary to this finding is that, given that the expressed conduct is relatively self-enforcing, there are information-based reasons to avoid placing a significant sanction on non-compliance. First, a sanction or penalty will have a much-reduced impact in the context of conduct that would be perceived as "reasonably-occurring" (i.e. there is a small potential benefit from the penalty). Secondly the scale of the penalty will be interpreted as a *signal,* conferring information on the distribution of social preferences, and so should be kept low in order to re-affirm the beliefs that the distribution of societal preferences is skewed towards compliance.[22]

So, in summary, the Benabou-Tirole model assumes that the individual responds to legal expressions in a certain way on account of the intrinsic and reputational values the individual possesses. In addition, information on the distribution of individual intrinsic preferences is important, because it indicates the change in reputational value available from choosing conforming conduct. A government recognising these responses is able to manipulate individual choice through

---

[21] And if there is some spillover between contexts, then there is an additional multiplier related to this informational advantage – the "informational multiplier" which combines with the above social multiplier to produce the overall effect of expressive law. This informational multiplier provides a similar impact via information's impact on the distribution of preferences, and its capacity to spillover to other contexts.

[22] Finally, if the scale of a penalty applied in one context may imply something about societal preferences in other contexts, and so there may be an informational *spillover* between legal contexts, then there may be additional reason to adjust the penalty. In this case, it is important that the penalty applied be calibrated to convey the information on societal preferences that is generally desirable, not just in the individual instance. This means that a relatively small penalty is able to convey some information about the distribution of societal preferences in general.

both its expression of legal principles and also its election of the incentive/sanction elected for the law's implementation.

Testable Hypothesis 3.1   Existence of Expressive Law

In the absence of any possible governmental threat of legal sanction, the mere announcement of a principle of law may have some effect by reason of either a) changes in intrinsic values adequate to engage self-enforcement (i.e. $v*$); and/or b) reputational values that accrue when an individual chooses conduct consistent with the expressed principle.   Preference shifting would have affect primarily via the first terms in (1), i.e. the creation of some new intrinsic values for a specified act.[23] If law functions primarily as information, then the effect of such a legal change is to indicate a new type or category of behaviour in which reputational value may accrue, i.e. the latter terms in (1). Individual values remain stable but the law provides an indication of an area in which individual reputations matter.

Actions that are chosen consistent with an expressed legal principle but in the absence of any additional or new threat of official sanction may be viewed as the consequence of behaviour motivated by the impact of expressive law (i.e. evidence of existence of expressive law).

Testable Hypothesis 3.2 Informational or Inculcative Channel for Expressive Law Effect

The second basic hypothesis arises out of the way in which sanctions interact with the law. As described above, the role of penalties in the informational model of expressive law is as a signalling device. The sanction calls out a cut-off level of individual costliness, with a low penalty indicating the extent to which conformance is expected to be of little cost (or constraint) to individual choices.

On the other hand, in any other model of law (expressive or otherwise) the penalty should vary not with the intrinsic values of individuals, so much as with the potential harm to society. And,

---

[23] Or, equivalently, the creation of some internal sanctions or rewards for conformance with the specified conduct, in accord with Kaplow and Shavell (2007).

in any other model of law's impact (i.e. the Becker model or the inculcated value model), the government's choice of sanction will be viewed as a means of supplementing the incentives for compliance – to some extent the higher the sanction, the better the compliance.[24] In this approach, the sanction acts as a price on specific individual choices, and so would be viewed as a constraint.

The populace subject to a new penalty would view the sanction as constraining if they are acting in conformity with the standard model of law's effect. The populace subject to a new penalty within the informational model of expressive law would see it as relatively un-constraining, i.e. the populace would perceive the sanction to be "low".

The establishment of penalties/sanctions that are perceived to be low from the perspective of the regulated agents is consistent with the Benabou-Tirole informational model of expressive law, where the establishment of higher penalties is consistent with other models of law-making (e.g. Becker, Shavell).

4. Measuring Internalisation of Legal Change: Survey Method

We now proceed to our attempt to measure some of these various impacts of expressive law, and to distinguish and to separate out between these various effects. Similar to what had occurred in Ireland ten years before, the introduction of a mandatory charge of 5p on "plastic bags" occurred in England on 5 October 2015.[25] Prior to this policy intervention, the regulation of this commodity in England was undertaken on a voluntary basis.[26] The bag tax went into effect for the expressed purpose of "eliminating the unnecessary use of this item".

---

[24] There are of course reasons to discourage the use of high levels of sanctions, such as the existence of Type I and Type II errors regarding the question of guilt, but these have little pertinence in the examination of issues of self-enforcement.

[25] From this date supermarkets and other large retailers in England were required by law to charge at least 5 pence for single-use plastic carrier bags. The legislative change followed the introduction of mandatory charges being introduced in Wales in October 2011, Northern Ireland in April 2013, and Scotland in October 2014.

[26] Prior to the introduction of the mandatory charge, the government used voluntary schemes to try to reduce the use of plastic bags. In 2008 the British Retail Consortium signed a voluntary agreement which saw the distribution of plastic bags almost halve in 2009 compared to their peak in 2006 of 12.1 billion (House of Commons, 2014, WRAP, 2015). With the exception of one retailer (Marks & Spencer), who introduced a voluntary charge on plastic bags (discussed below), retailers agreed to reduce uptake of new plastic bags by making them less accessible to consumers. However, despite an initial reduction, plastic bag consumption resumed, increasing by approximately 4 per cent per annum to 8.5 billion in 2014 (WRAP, 2015). In explaining

The government also took a proactive role in expressing the reasoning behind a proposed change in conduct. The government stated that single use plastic carrier bags 'take longer than other bags to degrade in the environment, can damage wildlife, and are extremely visible when littered in our towns, parks and the countryside' (UK Government, 2015). The government report went on to cite benefits of £6 million per year in savings from litter clean-up costs and £1.3 million per year in carbon savings.[27] The policy intervention was publically announced as a social responsibility, and publicised widely as such prior to its introduction.[28]

Importantly for our purposes, the charge was first introduced in Scotland, Northern Ireland and Wales prior to its introduction in England, and it was previously adopted (as a voluntary measure) by one English grocer (Marks and Spencer). For this reason, there are obvious control and treatment jurisdictions across the UK, providing an excellent opportunity to observe the impact of the policy as it is sequentially introduced.

To measure the impact of the legal change on English consumers, data on the consumption of plastic bags and on consumer attitudes was collected through three successive cross-sectional surveys of the UK population. The required charge for plastic bags came into force in England on October 5th, 2015. To assess consumption and attitudes before the introduction of the charge, the first survey took place on the 29th-30th of September 2015[29]. A second survey collected data one week after the introduction of the policy, on the 13th-14th of October 2015. A final round was then set seven weeks after the introduction of the carrier bag charge (24th-25th November, 2015) to capture the mid-term impact of the policy on behaviour. This survey, identical across periods, collected information on the consumption of plastic bags, plus personal attitudes towards plastic bag use and its regulation. To obtain nationally representative samples, the survey used a quota-sampling strategy that set quotas for

---

its decision to introduce the charge, the UK Government (2015) cited the fact that the use of plastic carrier bags had increased five years in a row.

[27] In addition, retailers are required to donate the funds raised to 'good causes', resulting in a projected amount of £73 million per year to be given to charity.

[28] These are two of the conditions that Bowles and Polania-Reyes (2012) argue are necessary to create this expressive function of law, publicity and public observation of non-compliants.

[29] It should be noted that the introduction of charge was known in advance, so it is possible that some internalisation could have occurred prior to the charge actually coming into force. Therefore, the results capture the effect of the charge actually being implemented.

age, gender, and regional distribution of the population to match actual UK statistics. A median test shows that respondents across the three periods also do not differ significantly in terms of income (p=0.474), and education (p=0.716). Finally, all participants completed the survey within 5 minutes.

Before the introduction of the charge, plastic bags were free to take and they were not scanned at the till (they had no bar code) nor recorded by the cashier. To ensure the availability of the same consumption data before and after the policy was introduced, the survey used a standard recall question (Browning et al., 2003): consumers were asked to report the number of carrier bags they used or acquired on their last grocery shopping trip.[30]

The idea of our experiment is to use the legal change regarding plastic bags to assess both how bag usage and how consumer attitudes were changing over the first two months of the new regime; in doing so, we then wish to separate out between those effects of the law caused by inculcation of values and those caused by the law itself (including the penalty). In order to accomplish this objective, each survey collected information on preferences/attitudes (on scales from zero to 100).

Specifically, the survey focused on ascertaining attitudes (and their shifts) reflective of various facets of the individual characteristic of *intrinsic motivation*: a) **Personal Motivation**, the intrinsic pleasure of an act (Kreps, 1997, Besley and Ghatak, 2016), measured here by assessing the level of agreement with the statement "Minimising the number of plastic carrier bags when I shop for groceries is important to me, regardless of any benefit or inconvenience that may result to me"; b) **Moral Obligation** (Bowles and Polania-Reyes, 2012), measured by the statement "Shoppers have a moral obligation to minimise the use of plastic carrier bags"; c) **Desired Behavioural Change** (Geisinger and Stein, 2016), captured here through the statement "Plastic carrier bags are currently overused"; d) **Control Aversion** (Bowles and Polania-Reyes, 2012), captured here by the statement "The government should not interfere by requiring retailers to charge for plastic carrier bags"; and e)

---

[30] Self-reported measures of behaviour are known to carry a downward bias (Schwarz, 1999, Browning et al., 2003), particularly when consumption is considered socially undesirable (e.g. alcohol in Feunekes et al., 1999). Consumption from recall can be accurate if the questionnaire is designed correctly: recalling a specific event, which has occurred recently, and having no time restriction to answer the question has been shown to give more precise measures of behaviour (Schwarz and Oyserman, 2001). As a result, respondents were asked to base their recall on their last grocery shopping trip and were given a scale to facilitate the task. The same recall approach was used to generate behavioural data.

Free Riding (e.g., Andreoni, 1988) measured by the statement "Other customers will continue using plastic carrier bags even if I stop using them".

We posit that these attributes capture various facets of the same inherent individual tendency to accept and/or self-enforce the indicated act (halting use of plastic bags). The hypothesis is that the first three attitudinal variables will be positively related to increased self-compliance with the law, while the latter two variables will be negatively related to increased self-compliance.

The impact of the plastic bag charge initially will be identified using a difference-in-difference estimator across treatment and control jurisdictions. In fact, this type of estimator requires variability in the sample along two dimensions: firstly, the sample should contain information of behaviour before and after the policy is introduced; secondly, in both periods the sample should include a number of individuals who are not affected by the policy when this is implemented, which represent a control group against which changes are compared (Bertrand et al., 2004). The present dataset contains information on plastic bag consumption before as well as after the charge was introduced. The treatment-control dimension of the estimator is captured by variation within UK supermarkets. First, the market presented institutional differences: specifically, Wales introduced a plastic bag charge on the 1st of October 2011, followed by Northern Ireland on the 1st of April 2013, and from Scotland on the 20th of October 2014. At the same time, one retailer chain introduced the charge voluntarily: Marks & Spencer started charging 5p for plastic bags in the UK using an explicit environmental rationale from February 2008 and donating the resulting profits to charity[31].

UK supermarkets represent a setting that lends itself to a quasi-experimental design: the plastic bags charge impacted only those English consumers who did not shop in Marks & Spencer (the 'treatment group'). The remainder of the sample (Scottish and Welsh consumers, and English customers of Marks and Spencer) constitute the control group, already subject to the charge. Our

---

[31] See http://corporate.marksandspencer.com/blog/stories/carrierbagcharge, and

http://www.thisismoney.co.uk/money/bills/article-3254567/Clothes-shops-charge-5p-carrier-bag-Monday.html

17

results show that around one third of consumers had shopped in Marks & Spencer in the four weeks prior to the survey[32], and the sample shows no major shift in the choice of retailer brand, with consumers on average visiting about 4 different retailer brands per month. This distribution and the sampling strategy set out before indicates that the sample can be expected to be reasonably representative of UK consumers, and that it remains relatively stable across the survey period.[33,34]

5.   Results: Estimating Impact, Existence, and Internalisation[35]

The results reported here concern the manner in which the legal change impacted aggregate societal behaviour (5.1), the existence of expressive law (5.2), and the extent to which this impact is attributable to value inculcation (5.3).

5.1 General Impact of the Legal Change

A first question related to the introduction of the charge concerns the measurable impact of this legal intervention, measured in terms of plastic bag consumption (new and re-used). Table 1 shows that on average *total* plastic bag use per shop in the UK decreased mildly in the weeks following the introduction of the charge, going from 4.49 to 4.18 units per person. On the other hand, average consumption of *new* plastic bags fell about 40%, from around 2.55 units per person per trip to just over 1.44 bags.

Table 1 also reports changing attitudes toward the behaviour being regulated in the lower part of the table. In general, respondents' attitudes all move in the direction supporting the intervention between surveys 1 and 2, but then show less movement afterwards.   The only behaviour registering

---

[32] Notably, only 2 individuals in each round of surveys shopped exclusively in Marks & Spencer, indicating that only 0.2% of the population had only access to stores that charged for plastic bags use.

[33] The choice of control group is crucial to obtain unbiased estimates using a difference-in-difference estimator, due to unobservable preferences for the environment before the plastic bag charge was introduced. Specifically, a control group that erroneously includes individuals who hold strong preference for the environment before the legislative change is introduced would cause the failure to reject a false null hypothesis (a type II error) by showing little change after the policy is introduced. Similarly, the incorrect inclusion of individuals who are initially unexposed to the charge in the control group would result in a significant change in consumption in response to the legislative change, favouring the incorrect rejection of a true null hypothesis (a type I error).

[34] Note that we have tested for robustness, in regard to our selection of treatment and controls, by repeating our exercise, with little difference in results. See Appendix B for a description of our robustness checks and results.

[35] For a complete presentation and discussion of the results of our survey and analysis, we refer the interested reader to our working paper on this case study. Larcom, Panzone, and Swanson (2016).

consistent change over the course of the survey is the respondents' concerns about free-riding, which declines consistently. So, in general, the announced legal change appears to have an immediate impact on attitudes, which remains relatively constant thereafter. We interpret the on-going change in perceptions regarding free-riding as an indicator that information continues to arrive to inform respondents in regard to this attitude.

Table 1: Mean behavioural and attitudinal variables in the sample

|	Variable	Survey 1	Survey 2	Survey 3
Behaviours	**Bag used – All**	4.4878	4.3689	4.1836
	Bag used – New	2.5467	1.9067	1.4447
	Reusable bag – carried	0.6233	0.7178	0.7478
	Reusable bag – purchased	0.2344	0.3122	0.2777
	Bin liners purchased	0.2056	0.2400	0.2290
Support for the charge	**Government is right**	0.664	0.722	0.698
	Fair charge (£)	5.049	7.083	7.149
Attitudes	**Personal motivation**	64.742	66.888	65.349
(Intrinsic Value)	**Control aversion**	46.107	40.823	43.016
	Behavioural Change	71.328	73.309	69.166
	Free-riding	70.872	67.503	66.982
	Moral	68.293	70.072	68.478
	Observations	900	900	904

5.2 Existence of Expressive Law

We turn now to our first hypothesis (3.1): the existence of a discernible impact of expressive law. Interestingly, we can see this immediately, looking at the impact of the legal change across jurisdictions, both those with and those without the new sanctions.

Figure 1 shows the distribution of the consumption of both new and all plastic bags in the three survey periods in the differing jurisdictions.[36]

---

[36] A Kolmogorov-Smirnov test indicates that the distribution of total plastic bags used in the sample did not change significantly over time, while the distribution of new bags shifted toward zero. Figure 1 shows that individuals previously unexposed to the mandatory charge showed both a decline in total as well as new plastic bags (p-values for trend equal to 0.022 and 0.000, respectively).

Figure 1: Impact of Policy over Time – Control and Treatment Jurisdictions
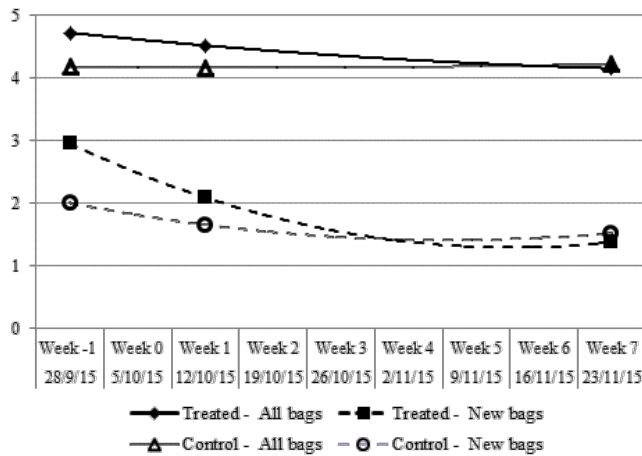
The two dashed lines at the bottom correspond to the average effect of the policy in both those jurisdictions that previously had a charge in place (and so were unaffected by the new charge) and those jurisdictions that previously had no charge. With the information surrounding the policy initiative, the intervention had the effect of significantly reducing the uptake of new plastic bags (p-value for trend = 0.000) in the *control group* (i.e. consumers previously paying for their use). Consumption of new plastic bags declined much more rapidly in the treatment group, dropping from 2.95 to 1.38 while the control group's consumption fell from an average of 2.00 to 1.52 new bags, but both movements are significant.

Interestingly, over the period of the survey, the effect of the intervention was to bring the two groups (control and treatments) into alignment in overall impact – both arriving at approximately an average of 1.5 new bags per weeks, from very different starting points. This result – of controls and treatments moving together in response to legal announcements – might seem to be spurious, but for the fact that it has been noted to exist in previous studies of the same tax along the same borders. (Portinga et. al. 2013) It would appear to be a likely artefact of the relatively free flow of information across these borders, even though sanctions are changing at differing times, so that the impact of the expression (and announcement effect) of the laws may be discerned separately from the impact of the introduction of penalties. The overall effect – of expression and sanction – appears to be the same along the border, when both are brought into alignment.

We would like to see precisely how things are occurring in the treatments and controls. To do this, we undertake difference-in-difference analysis to compare outcomes within control and treatment groups, in regard to impacts on behaviour (engagement in the regulated activity) in those jurisdictions where new sanctions were introduced compared to those jurisdictions where no new sanctions existed.

We define plastic bag consumption of individual i in group s (control, treatment) at time t as Y_{ist}. The total impact of the policy on the target consumers can then be written as the difference in consumption of individuals previously unexposed to the charge compared to the same difference in individuals who were previously exposed to the charge (Imbens and Wooldridge, 2009), as:

$$\beta = \left[Y_i(s=1,t=1) - Y_i(s=1,t=0)\right] - \left[Y_i(s=0,t=1) - Y_i(s=0,t=0)\right] \qquad (2)$$

Equation (2) represents a difference-in-difference (DID) estimator of the impact of the legislative change. This effect can be estimated by the regression (Bertrand et al., 2004):

$$Y_{ist} = A_s + B_t + \beta I_{st} + \gamma D_{ist} + \varepsilon_{ist} \qquad (3)$$

where A refers to group-specific fixed effects, capturing systematic differences between control group and the initially unexposed group (e.g. different plastic bags needs); B refers to time-specific fixed effects that capture the presence of the plastic bags charge if $t > 0$; and ε is the error term. In equation (3), I refers to the interaction term between treatment dummy (equal to one for those initially unexposed to the charge) and the legislative change dummy. Then, β captures the impact of the legislative change, measuring the change in consumption of people exposed to the plastic bag charge for the first time relative to the control group who were previously exposed to the charge.[37]

Table 2 presents the results of the difference-in-difference estimation, which indicates the change in law had a large and significant impact on the use of new plastic bags, which declined by around 1.7 plastic bags per person per trip after seven weeks, a significant drop from the 0.5 reduction in consumption one week after the policy was introduced. For purposes of the existence of expressive

---

[37] Importantly, the presence of different time periods causes problem of serial autocorrelation of the residuals ε_{ist} (Bertrand et al., 2004); this problem can be significantly mitigated by clustering residuals by period (Wooldridge, 2003).

law, the interesting result is that the impact of the time of announcement (Period 2, Period 3) had an equally significant impact in control jurisdictions as in the treatments.[38] There is an additional impact in the treatment jurisdictions (due perhaps to their initially higher level of use of plastic bags), but the announcement of the new law is seen to have a clear impact in the absence of any added sanction.[39]

Table 2: Impact of the charge in treatment and controls

	Bags used – All		Bags used – New	
	Coefficient	S.E.	Coefficient	S.E.
Intercept	6.7367***	0.2833	7.5591***	0.6990
Treated group	0.1868***	0.0485	0.4591***	0.0967
Period 2	-0.0340***	0.0098	**-0.9721***	0.0532**
Policy, time 2	**-0.1031***	0.0209**	**-0.4889***	0.0205**
Period 3	-0.0906***	0.0099	**-1.7025***	0.0649**
Policy, time 3	**-0.5162***	0.0184**	**-1.6806***	0.0672**
Age	-0.0341***	0.0029	-0.1071***	0.0126
Male	-0.4219	0.2603	0.6578	0.4930
England	Baseline		Baseline	
Northern Ireland	-0.7842***	0.1915	-1.6539	1.5472
Scotland	-0.5414***	0.1322	-2.1254***	0.7217
Wales	-0.7634***	0.1727	-2.0140***	0.1998
Sigma	3.2360***	0.1035	4.7323***	0.1655
Income dummies	Yes		Yes	
Education dummies	Yes		Yes	
Retailer dummies	Yes		Yes	
Observations	2704		2704	
Pseudo R^2	0.0187		0.0523	
Log-likelihood	-6387.20		-4786.52	

Note: residuals are clustered by period. Significance is as follows: * = 0.10; ** = 0.05; *** = 0.01. Treated Group refers to respondents within jurisdictions of legal change. Period i refers to respondents to ith survey. Policy, time i refers to respondents within treated group in period i.

---

[38] Of course the period fixed effects report the effects of all changes occurring in the 2 weeks between survey 1 and survey 2, and the 5 weeks between survey 2 and survey 3. Given that the most significant event impacting plastic bag users in these periods was likely to be the new law and the announcements regarding the same, it is reasonable to ascribe these period effects to these announcements.

[39] As mentioned above, and given a common national parliament and a common broadcasting network, it is not surprising that there would be some spillover of the information on the new legislation, into those neighbouring jurisdictions where the new sanctions would not apply. Other researchers have noted the same effect regarding these legal changes along this border. (Portinga 2013)

5.3 Mediation Analysis : Measuring the Internalisation of Law

We have seen that the legal change impacted aggregate behaviour regarding the regulated activity (reduction in new plastic bags), and that it did so in both control and treatment jurisdictions. This evidence supports the belief in the existence of expressive law (at least in the neighbouring jurisdictions). Now we would like to explore how much of the observed change in behaviour might be attributable to the inculcation of values, received at the time of the expressed principle of legal change.

We do this by considering the impact of the introduction of the carrier bag charge on reported attitudes and internal motivations, as reported by our survey respondents. Table 3 presents the estimates of the impact of the plastic bag charge on social preferences using the difference-in-difference estimator of equation (3). There are numerous interesting observations to be made about how attitudes are shifting. First, and unlike the case of behaviours, the positive attitudes (intrinsic motivation, desired behavioural change, moral engagement on issue) are shifting only in the treated jurisdictions. Most of the reported attitudinal shifts in controls are insignificant.[40]

Interestingly, the controls show positive shifts in their responses regarding negative attitudinal changes (control aversion, free riding concerns); these are usually significant and often larger than those of the treatments. One of the most important changes occurs in the way consumers view the role of government in the newly regulated area. In particular, control aversion immediately declines once the policy is introduced, with a successive large drop seven weeks after the legislative change. And the controls show further declines in this reported attitude. As a result, it appears that consumers increasingly believe that the government is correct to intervene in this market. The lapse of time after initial introduction of legal change appears to continue to erode attitudes of resistance, even if it does not encourage on-going inculcation.

In the treated jurisdictions, the legal change is associated with reported attitudes favouring intrinsic values that are significant and reasonably stable. There is a slight tailing off of the reported

---

[40] The one exception being the response to the second survey regarding moral engagement.

attitude favouring intrinsic motivation and moral engagement (of about 20-25%), while the desire for

behavioural change remains almost constant.

Table 3: Impact of the plastic bag charge on reported attitudes and preferences

	Control aversion		Behavioural Change		Intrinsic motivation		Free-Riding		Moral engagement	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Intercept	51.8345***	3.6953	60.8068***	2.5423	50.7474***	3.5548	70.0639***	4.5790	57.5085***	2.4310
Treated group	4.1625	1.5764	-4.2205***	0.2811	-3.2999**	0.4514	2.4737	1.6793	-3.2559**	0.3543
Period 2	-4.5058***	0.3624	-0.8863*	0.2771	0.9003	0.3837	-4.2076***	0.3642	-0.6791	0.2394
Policy, time 2	**-1.1249****	0.2396	**4.8902*****	0.3573	**1.9385***	0.4715	**1.4318***	0.4137	**3.8656*****	0.3000
Period 3	0.6467	0.4385	-4.6855***	0.1156	-0.1183	0.1826	-1.6267**	0.3290	-1.5049**	0.3145
Policy, time 3	**-6.6851*****	0.2493	**4.6112*****	0.1838	**1.4126****	0.2956	**-4.0491****	0.4226	**2.7825****	0.3344
Age	-0.0811*	0.0209	0.1460**	0.0222	0.2496**	0.0373	-0.0012	0.0357	0.2029**	0.0252
Male	0.5353	0.3979	-2.8051	1.0889	-4.1977*	1.0080	-2.1356*	0.6870	-3.6170**	0.5849
England	Baseline		Baseline		Baseline		Baseline		Baseline	
Northern Ireland	-11.2902	5.1570	4.7572*	1.4603	6.6748	3.6127	-1.8025	4.1799	7.6470**	1.4892
Scotland	-8.7500	4.3285	-1.1668	2.1487	2.7410	1.2506	-4.1099	2.8891	5.3522	2.7872
Wales	-9.6300	3.3188	-0.5367	3.2079	3.6311**	0.4613	-3.0269	3.2062	3.0098	1.1710
Income dummies	Yes		Yes		Yes		Yes		Yes	
Education dummies	Yes		Yes		Yes		Yes		Yes	
Retailer dummies	Yes		Yes		Yes		Yes		Yes	
Observations	2,704		2,704		2,704		2,704		2,704	
Log-likelihood	-13,265.01		-12,582.43		-12,793.79		-12,156.82		-12,604.62	
R²	0.0530		0.0381		0.0484		0.0370		0.0418	

Note: residuals are clustered by period. Significance is as follows: * = 0.10; ** = 0.05; *** = 0.01. Treated Group refers to respondents within jurisdictions of legal change. Period i refers to respondents to ith survey. Policy, time i refers to respondents within treated group in period i.

We have seen that this change in law resulted in movements among consumers both in regard to changes in behaviour (the amount of the regulated activity) and changes in attitudes (indicated facets of intrinsic valuation of the legal change). To ascertain the contribution of changed values to changed behaviours, we use *mediation analysis* (e.g. Baron and Kenny, 1986, MacKinnon et al., 2007).

To measure the extent of the internalisation of the legislative change, behaviour is modelled as a system of two equations. Firstly, consumption is modelled by extending equation (3) to adjust for the role of motivation *X* on demand as

$$Y_{ist} = A'_s + B'_t + \beta' I_{st} + \delta X_{ist} + \gamma' D_{ist} + \epsilon_{ist} \tag{4}$$

Similarly, internal motivation *X* can be modelled as

$$X_{ist} = C_s + F_t + \alpha I_{st} + \varphi D_{ist} + u_{ist} \tag{5}$$

where *C* and *F* reflect group-specific and time-specific motivational fixed effects.

There are two effects: a *direct effect* and an *indirect (inculcated value) effect*. The introduction of the plastic bag law and charge has a direct effect that reduces the use of plastic bags by *β'* units. The charge also has an indirect effect: the legislative change increases (decreases) the average internalised motivation in the sample by *α* units - e.g. increasing (decreasing) the belief that plastic bags are overconsumed in aggregate - because each unit of motivation influences behaviour by *δ* units, the charge will further reduce (increase) consumption by *α·δ* units.

Merging equations (4) and (5), we then obtain the aggregate equation

$$Y_{ist} = (A'_s + \delta C_s) + (B'_t + \delta F_t) + (\beta' + \delta \alpha)I_{st} + (\delta \varphi + \gamma')D_{ist} + (\delta u_{ist} + \epsilon_{ist}) \tag{6}$$

which is equivalent to equation (3). In equation (6), the total effect of the policy is

$$\hat{\beta} = (\beta' + \delta \alpha) \tag{7}$$

which is the sum of a direct effect (β') and an indirect effect ($\delta\alpha$). The standard error of the indirect

effect equals $\sigma_{\alpha\delta} = \sqrt{\sigma_\alpha^2\delta^2 + \sigma_\delta^2\alpha^2}$ (Krull and McKinnon, 2001).[41]

The results of the mediation analysis are reported below in Table 4. These indicate that the change in plastic bags use is much larger than that attributable to changes in reported intrinsic values (i.e. motivational effect). Specifically, one week after the charge was introduced, the effect of the policy attributable to reported changes in attitudes accounted for a reduction of around 0.15 new plastic bags per person per trip, against a non-attributed reduction of 0.34 new bags, so that non-attributed change accounted for 69% of the total reduction. Seven weeks after the policy was introduced, the attributed effect of the policy increased slightly in absolute value, causing a reduction of 0.19 new plastic bags per person per trip, against a non-attributed reduction of around 1.5 new bags, so that the non-attributed change represents 89% of the total reduction in behaviour.

This suggests that other effects grow much more than the inculcated values over time, both in absolute and relative terms. Importantly, after seven weeks the largest motivational contribution to compliance comes from a reduction in control aversion as well as a reduction in the concerns of free-riders. Respondents are indicating that their values *for* the policy are moving less than the reduction of the values they had reported previously that had been biased *against* the new policy.

Table 4: Mediation Analysis - direct and indirect effects of law on behaviour

			ALL BAGS		NEW BAGS	
		Period	Effect	S.E.	Effect	S.E.
Direct effect of policy		1	-0.0511**	0.0238	**-0.3382***	**0.0325**
		2	-0.4524***	0.0504	**-1.5111***	**0.1139**
Indirect effect via	Personal motivation	1	-0.0108	0.0124	-0.0425**	0.0177
		2	-0.0079	0.0090	-0.0310**	0.0123

---

[41] Notably, mediation analysis identifies causal effects only if two conditions hold (Keele et al., 2015). First, the charge has to be randomly assigned, i.e. the treatment variable D in equation (4) has to be unrelated to unobservable preferences for plastic bags after adjusting for covariates. Secondly, attitudes must be exogenous to the behaviour given the remaining covariates, i.e. the main relation is structural. These points are discussed in more detail in section 7, when some robustness checks test whether these two conditions hold.

Control aversion	1	-0.0093**	0.0037	-0.0240***	0.0068
	2	-0.0554***	0.0184	-0.1428***	0.0272
Behavioural Change	1	-0.0285	0.0213	-0.0552*	0.0284
	2	-0.0268	0.0200	-0.0520*	0.0266
Free Riding	1	-0.0083***	0.0028	-0.0169***	0.0052
	2	0.0236***	0.0045	0.0479***	0.0068
Moral obligation	1	0.0062	0.0064	-0.0120**	0.0060
	2	0.0045	0.0046	-0.0087**	0.0044
Total Indirect effects	1	-0.0507*	0.0259	**-0.1507*****	**0.0350**
(Changed Intrinsics)	2	-0.0620**	0.0294	**-0.1867*****	**0.0408**
Total effects		-0.1018***	0.0352	-0.4889***	0.0478
		-0.5145***	0.0583	-1.6977***	0.1210

Significance is as follows: * = 0.10; ** = 0.05; *** = 0.01.

Our mediation analysis demonstrates that there is a clear inculcative effect of law-making in this instance, as measured by the self-reported changes in values of those individuals subjected to the new law. Although only able to explain about 10% of the overall change in behaviour, the associated change is clear and significant. Interestingly, the inculcated values that would positively relate to the self-implementation of the law (morals, motives, desired change) are smaller in size and duration than those more negatively related to the expression of new laws (control aversion, aversion to free riding). The ongoing and increasing reductions in these attitudes go a long way towards explaining the extent of the self-enforcement of this law.[42] [43] Law has a measurable own-effect due to inculcation, and it appears to be increasingly accepted over time due to withdrawal of resistance to the intervention.

5.4 Ascertaining the Role of Sanctions in Expressive Law

We now turn to the second testable hypothesis concerning the role of sanctions in the functioning of law. We have put forward two models regarding sanctions: the standard model views sanctions as prices or constraints on individual choices, while the second (informational) model views sanctions as information on the reasonableness of the behaviour being admonished by the government. The first model would indicate that regulated individuals would believe that the new price constrains choice, and would feel constrained by it. The second model would indicate that regulated individuals would perceive the new sanction as information that the admonished act or

---

[42] For robustness checks, see Appendix B and discussion therein.
[43] See Appendix C for the final table of results of our analysis.

behaviour is reasonable and relatively un-constraining; that is, the sanction would be perceived to be low.

Evidence that the charge was perceived in this way is demonstrated in Figure 2, which shows the reported "fair charge" from the perspective of the survey respondents in the treatment jurisdiction. This charge commences at about the level of the actual charge (4.51 relative to 5.00), but then increases to about 60% more than the actual charge over the course of the survey (8.45 after 7 weeks). This is evidence that the populace subjected to the new sanction find the requested behaviour to be reasonable, and that they perceive the new charge to be relatively "low". This is consistent with the informational model of expressive law.

Similarly, the change in the attitude of "control aversion" is also evidence that the requested behaviour is perceived to be un-constraining. In Tables 3 and 4, we see that the single attitude moving most over the course of the survey relates to declining concern over government control over individual choice in this sphere. The decline in the concern over government intervention over plastic bags indicates that the populace finds the requested behaviour to be reasonable, and so relatively un-constraining. Alternatively, the populace could become habituated to the intervention, and so reduce any resistance over time. In general, we believe that this is evidence that the population took the combination of principle (use fewer plastic bags) together with the price (5p per bag) as an indication that this was reasonable behaviour, and a reasonable request for the government to make.[44]

In sum, we find that the response to the charge in this case study is indicative of the existence of an effect of expressive law, most consist with the informational model put forward by Benabour and Tirole (2011a). The announcement of the combined law-with-charge had a clear impact in jurisdictions with and without the charge, indicating that the charge represented information more than potential penalty. And the charge was viewed as below the "fair charge", i.e. low from the

---

[44] In another study, in the context of grocery shopping, we found the opposite result in regard to consumer response to government intervention, i.e. the populace did not like anyone "messing with their groceries". (Perino et. al. 2014) For this reason, it appears that the resistance to intervention does vary with context as well as with time, and so the more likely reason for reduced resistance (from these two examples) is likely to be the perceived reasonableness of the intervention.

perspective of consumers. This makes sense from a self-enforcement perspective, only if the charge is supposed to convey reasonableness of behaviour rather than threatened penalty.

Figure 2: Evolution of the reported "fair charge" over time
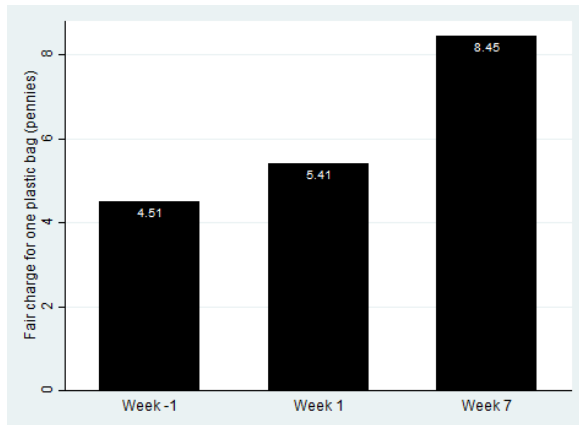
Note: figures refer to averages. Using a one-tailed t-test, differences between period 1 and the two post-policy periods are significant at 4% (period 2) and 6% (period 3) for treated, and insignificant (period 2) and significant at 9% (period 3) for controls. Kruskal-Wallis rank test for treated: chi2 = 12.330 (p = 0.0021). Kruskal-Wallis rank test for control: chi2 = 7.117 (p = 0.0285).

6. Discussion

This study sheds light on whether expressive law "exists", and the extent to which it is linked to the inculcation of values expressed within the law. Many commentators cited the legal change related to the introduction of a plastic bag charge in Ireland as an example of the type that might register a significant impact via changed preferences. (Bowles and P-R 2011; Benabou and Tirole 2011a) We have investigated the impact of a similar change in law in England in 2015.

We find that indeed the change in law has fundamentally changed the way a significant proportion of consumers report their own values in the area of the new law. Our results provide some evidence of an internalisation process, clearly indicating that over time consumers adapted to the legislative change once it was introduced, and supported the legislator in increasing the regulatory space within which it operates. They both supported the lawmakers' expansion of its field of regulatory activity, and also reduced their concerns over the ability for the law to be effective (due to free-riding concerns). Clearly this form of support for the lawmaker's intervention represents one important shift in values.[45]

Another important enquiry concerns how much the legal change resulted in the inculcation of changed values, or whether values remain the same but other processes are operating. Our main result in this regard derives from our mediation analysis which finds that the change in behaviour attributable to changed values commences initially at about 31% of the total behaviour change, and then ends at 11% after two months of the law. The remainder of the impact of the law is not attributed to shifts in their underlying preferences/attitudes.

So we then assume that residual change in behaviour is information-sourced, either a response to the change in price/sanction or to some other information associated with that sanction. What

---

[45] This finding is important: as pointed out by Bilz and Nadler (2014), legislative change does not always lead to the internalisation of the values the legislation aims to promote. Indeed, it is well known that negative reactions to legal change can occur, leading to reduced compliance (Carbonara et al., 2012, Acemoglu and Jackson, 2014) and of course massive evasion, as in the tax context.

determines the response of agents to a law-and-sanction combination: its potential application to them (via governmental monitoring and penalties) or the public opprobrium attached to a perceived violation?

The most interesting result in this regard concerns how the legal change resulted in responses from both *the control* and the treatment groups, and that they reached the same overall outcome. We demonstrated that the legal change here had a significant impact on both jurisdictions, those with and without the threat of new sanction. Thus, the control jurisdictions' response to changed information (rather than changed penalties) explains about one-quarter of the observed behavioural change.[46]

In fact, there is something potentially interesting in the fact that all jurisdictions attained the same equilibrium after the introduction of the law, despite the fact that the new sanction was introduced in only England. In Figure 1 it is seen that the impact of the legal change guided both groups to the same outcome, with the controls starting from a lower level but both groups converging over the course of the survey. Somehow the expression of the law is determining not only the direction of the response, but some measure of the scale of the response.

This would all seem to be consistent with an informational model of expressive law. Perhaps the expression of the law is indicating direction of change, and the information on the sanction (rather than the threat of the sanction itself) is determining in part the scale of the response. Additional evidence for the idea that sanctions act as information rather than prices in this context is to be found in the relatively significant response to a relatively low sanction price (5p). This is consistent with the idea that the expressed legal principle together with the low sanction would signal to the citizenry that this is one type of behaviour that should be viewed as "respectable" (Benabou and Tirole, 2011). Together the movement of behaviour in these circumstances is consistent with the idea that consumers responded to legal change "leadership" in guiding the direction in which behaviour should change, while small sanctions served as an indicator that this was something that "just ought to be done".

---

[46] See Figure 1, which shows that average consumption fell from 3 to 1.5 in treatment jurisdictions, but from 2 to 1.5 in control jurisdictions.

7. Conclusion

Why do laws take effect?   Do we sit at red stoplights in the middle of the night on account of the existence of intrinsic values, the fear of societal opprobrium, or the threat of governmental penalties?

It has been hypothesised that a new law can have both effects, sometimes expressing the common duties of a populace and other times providing prices for regulated behaviours. We have reported on an experiment to test this hypothesis in the context of the introduction of a plastic bag charge in England in 2015. In short, we find that the announcement of a bag tax in England did indeed have an immediate impact on behaviour, changing the way in which agents used these items in a very short amount of time.

We have also argued for the existence of evidence of the impact of expressive law in this case. This is demonstrated in part by the significant impact of the new law in the control regions of the study, where the information on the law was known even though there were no new sanctions applied within those jurisdictions. Given that information readily crossed these borders even though the sanction did not, this would seem to be a clear example of the expressive impact of law.

In addition, we have attempted to measure the extent to which the expressive impact of law was attributable to the inculcation of new values (together with the expression of the law). Again, this attribution is possible, however the scale of the behavioural change attributed to attitudinal shifts by respondents was relatively minor, contributing only about ten percent of the total impact in changed behaviour over the two months of our survey.

Finally, we have attempted to assess the way in which penalties and principles interact in the overall impact of expressive law. Here we find that the consumers subject to the charge reported it to be less than the "fair charge" for violating the new law. This evidence is consistent with the view that individuals see the proposed behaviour as reasonable (and so most individual behaviour should be

consistent with the legal principle), and it is consistent with the view that the government wished to propose a low charge that caused consumers to see the range of behaviour as such.[47]

In conclusion, the short answers to the questions we raised at the outset of this paper are as follows: We estimate that approximately 10% of all individuals avoiding plastic bags are doing so on account of a change in intrinsic motivation. The other 90% of individuals are doing so on account of the effect of the charge. Is the charge effective on account of the price being charged, or because of the possibility of public opprobrium to continued use of bags? The evidence here indicates that most people found the charge to be reasonable and un-constraining, and so would probably find behaviour inconsistent with it to be unreasonable. Therefore, most people are probably acting on the belief that public opprobrium will result from behaviour inconsistent with the new law.

So, we conclude that the two systems (expressive and economic) work together within society's conception of law. Changes in law that incorporate economic incentives (such as fines, sanctions or charges) also may act in part as purely expressive law, altering the way that the populace views its own duties and obligations. More interestingly, it appears that the two systems are not entirely distinct. Prices within the economic system also seem to act as information within the system of social preferences and reputations. The systems seem to work together in guiding society – in terms of direction and scale of change. Legal principles can point in the direction for societal change, while the sanctions and penalties within the system guide it toward its final equilibrium.

---

[47] However, the extent to which the change in behaviour was driven by the price charged or the reputational factors concerned is not discernible from this experiment. But clearly the price charged was relatively low (5p) and the change in behaviour relatively dramatic (50% decline), and so it is likely that information played a significant role here.

References

Acemoglu, Daron, and Matthew O. Jackson. 2014 "History, expectations, and leadership in the evolution of social norms." *The Review of Economic Studies*: 82(2): 423-456.

Akerlof, George A. 1980. "A theory of social custom, of which unemployment may be one consequence." *The quarterly journal of economics* 94(4): 749-775.

Andreoni, James. 1988. "Why free ride?: "Strategies and learning in public goods experiments." *Journal of public Economics* 37(3): 291-304.

Apesteguia, Jose, Patricia Funk, and Nagore Iriberri. 2013. "Promoting rule compliance in daily-life: Evidence from a randomized field experiment in the public libraries of Barcelona." *European Economic Review* 64: 266-284.

Arimura, Toshi H., Hajime Katayama, and Mari Sakudo. 2014. "Do social norms matter to energy saving behavior? Endogenous social and correlated effects." Tokyo Center for Economic Research (TCER) Paper No. E-76 (2014).

Bar-Gill, Oren, and Chaim Fershtman. 2004. "Law and preferences." *Journal of Law, Economics, and Organization* 20(2): 331-352.

Baron, Reuben M., and David A. Kenny. 1986. "The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." *Journal of personality and social psychology* 51(6): 1173-1182

Belkin, Douglas. 2003 "Boston firefighters sick-or tired of working." *Boston Globe* 18.

Roland Bénabou & Jean Tirole, 2006. "Incentives and Prosocial Behavior," *American Economic Review*, vol. 96(5), pages 1652-1678, December

Benabou, Roland, and Jean Tirole. 2011a. "Laws and norms." *National Bureau of Economic Research Working Paper* No. w17579.

Roland Bénabou and Jean Tirole, 2011b. "Identity, Morals, and Taboos: Beliefs as Assets" *The Quarterly Journal of Economics (2011) 126 (2): 805-855*

Besley, Timothy J., Gwyn Bevan, and Konrad Burchardi. 2009. "Naming & Shaming: The impacts of different regimes on hospital waiting times in England and Wales." *Mimeo.*

Besley, Timothy J., and Maitreesh Ghatak. 2016. "Market Incentives and the Evolution of Intrinsic Motivation." *Mimeo*.

Berger, Melissa, Gerlinde Fellner-Röhling, Rupert Sausgruber, and Christian Traxler. 2016. "Higher taxes, more evasion? Evidence from border differentials in TV license fees." *Journal of Public Economics* 135: 74-86.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics*, 119, 249-275.

Bliz, Kenworthey, Janice Nadler . 2014. Law, Moral Attitudes, and Behavioral Change. In: Zamir, E. & Teichman, D. (eds.) The Oxford Handbook of Behavioral Economics and the Law. Oxford: Oxford University Press.

Boyer, Patrick, Denger, Nadia and Rincke, Johannes. 2016. "Do norms on contribution behavior affect intrinsic motivation? Field experimental evidence from Germany. " *Journal of Public Economics*, 144, 140-153.

Bowles, Samuel. 2008. "Policies Designed for Self-Interested Citizens May Undermine "The Moral Sentiments": Evidence from Economic Experiments". *Science*, 320, 1605-1609.

Bowles, Samuel, and Sandra Polania-Reyes. 2012. "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature*, 50, 368-425.

Browning, Martin, Thomas F. Crossley, and Guglielmo Weber. 2003. "Asking consumption questions in general purpose surveys". *The Economic Journal*, 113, F540-F567.

Carbonara, Emanuela, Francesco Parisi, and Georg Von Wangenheim. 2012. "Unjust laws and illegal norms." *International Review of Law and Economics*, 32, 285-299.

Cardenas, Juan Camilo, John Stranlund, and Cleve Willis. 2000. "Local environmental control and institutional crowding-out." *World Development* 28(10): 1719-1733.

Cooter, Robert D. 1998. "Expressive Law and Economics". *The Journal of Legal Studies*, 27, 585-607.

Convery, Frank, Simon McDonnell, and Susana Ferreira. 2007. "The most popular tax in Europe? Lessons from the Irish plastic bags levy." *Environmental and Resource Economics* 38(1): 1-11.

Dal Bó, Ernesto, and Pedro Dal Bó. 2014. "Do the right thing: The effects of moral suasion on cooperation." *Journal of Public Economics* 117: 28-38.

Disney, Kate, Julian Le Grand, and Giles Atkinson. 2013. From irresponsible knaves to responsible knights for just 5p: behavioural public policy and the environment. In: OLIVER, A. (ed.) Behavioural Public Policy. Cambridge: Cambridge University Press.

Dolan, Paul, Matteo M. Galizzi, and Daniel Navarro-Martinez. 2015. "Paying people to eat or not to eat? Carryover effects of monetary incentives on eating behaviour." *Social Science & Medicine*, 133, 153-158.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 2016. "Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany." *American Economic Journal: Economic Policy*.

Etzioni, Amitai. 2000. "Social norms: Internalization, persuasion, and history." *Law and Society Review*: 157-178.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler. 2013. "Testing enforcement strategies in the field: Threat, moral appeal and social information." *Journal of the European Economic Association* 11(3): 634-660.

Feunekes, Gerda IJ, Pieter van't Veer, Wija A. van Staveren, and Frans J. Kok. 1999. "Alcohol intake assessment: the sober facts." *American Journal of Epidemiology* 150(1): 105-112.

Frey, Bruno S. "Pricing and regulating affect environmental ethics." 1992. *Environmental and Resource Economics* 2(4): 399-414.

Frey, Bruno S., and Alois Stutzer. 2008. Environmental morale and motivation. In: LEWIS, A. (ed.) The Cambridge Handbook of Psychology and Economic Behaviour. Cambridge: Cambridge University Press.

Geisinger, Alex C., and Michael Ashley Stein. 2015. "Expressive Law and the Americans with Disabilities Act." *Michigan Law Review* 114: 1061-1079.

Gneezy, Uri, and Aldo Rustichini. 2000. "A Fine is a Price." *The Journal of Legal Studies*, 29, 1-17.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and why incentives (don't) work to modify behavior." *The Journal of Economic Perspectives* 25(4): 191-209.

Goeschl, Timo, and Grischa Perino. 2012. "Instrument choice and motivation: Evidence from a climate change experiment." *Environmental and Resource Economics* 52, no. 2 (2012): 195-212.

Greenberger, Scott S. 2003. "Sick day abuses focus of fire talks." *Boston Globe* 17.

Hart, H. L. A. 1997. The Concept of Law (Second Edition), Oxford, Oxford University Press.

Holmås, Tor Helge, Egil Kjerstad, Hilde Lurås, and Odd Rune Straume. 2010 "Does monetary punishment crowd out pro-social motivation? A natural experiment on hospital length of stay." *Journal of Economic Behavior & Organization* 75(2): 261-267.

House of Commons. 2014. Plastic Bags. In: committee, H. O. C. E. A. (ed.). London.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent developments in the econometrics of program evaluation." *Journal of economic literature* 47(1): 5-86.

Kaplow, Louis, and Steven Shavell. 2007. "Moral rules, the moral sentiments, and behavior: toward a theory of an optimal moral system." Journal of Political Economy 115(3), 494-514.

Keele, Luke, Dustin Tingley, and Teppei Yamamoto. 2015. "Identifying mechanisms behind policy interventions via causal mediation analysis." *Journal of Policy Analysis and Management* 34(4): 937-963.

Kleven, Henrik Jacobsen, Martin B. Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez. 2011. "Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark." *Econometrica* 79(3): 651-692.

Kreps, David M. "Intrinsic motivation and extrinsic incentives." 1997 *The American Economic Review* 87(2): 359-364.

Larcom, Shaun, Luca Panzone, and Timothy Swanson. 2016. A price is a guide: The English plastic bag charge and measuring the internalisation of law, C-EENRG Working Paper. University of Cambridge.

Lin, Chung-cheng, and C. C. Yang. 2006. "Fine enough or don't fine at all." *Journal of Economic Behavior & Organization* 59(2): 195-213.

Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan. 2011. "Mechanism experiments and policy evaluations." *The Journal of Economic Perspectives* 25(3): 17-38.

MacKinnon, David P., Amanda J. Fairchild, and Matthew S. Fritz. 2007. "Mediation analysis." *Annual review of psychology* 58: 593-614.

McAdams, Richard H. 1997. "The origin, development, and regulation of norms." *Michigan Law Review* 96(2): 338-433.

Perino, Grischa, Panzone, Luca, and Swanson, Timothy. 2014. "Motivation crowding in real consumption decisions: who is messing with my groceries?" *Economic Inquiry*, Vol. 52, Issue 2, pp. 592-607.

Poortinga, Wouter, Lorraine Whitmarsh, and Christine Suffolk. 2013. "The introduction of a single-use carrier bag charge in Wales: Attitude change and behavioural spillover effects." *Journal of Environmental Psychology* 36, 240-247.

Raz, Joseph. 1999. *Practical reason and norms*. Oxford University Press: Oxford.

Raz, Joseph. 1999. The Obligation to Obey: Revision and Tradition. In Edmundson, William A. (ed.), *The Duty to Obey the Law*. Rowman & Littlefield: Lanham, Md.

Rode, Julian, Erik Gómez-Baggethun, and Torsten Krause. 2015. "Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence." *Ecological Economics* 117: 270-282.

Romaniuc, Rustam, Katherine Farrow, Lisette Ibanez, and Alain Marciano. 2016. "The perils of government enforcement." *Public Choice* 166(1-2): 161-182.

Schwarz, Norbert. 1999. "Self-reports: how the questions shape the answers." *American psychologist* 54(2): 93-105.

Schwarz, Norbert, and Daphna Oyserman. 2001 "Asking questions about behavior: Cognition, communication, and questionnaire construction." *American Journal of Evaluation* 22(2): 127-160.

Shavell, Steven. 2012. "When Is Compliance with the Law Socially Desirable?." *The Journal of Legal Studies* 41(1): 1-36.

Sunstein, Cass R. 1996. "On the expressive function of law." *University of Pennsylvania Law Review* 144(5): 2021-2053.

Titmuss, Richard. M. 1971. The gift relationship: from human blood to social policy, New York, Pantheon Books.

UK Government. 2015. Carrier bags: why there's a charge. [Online]. London: Department for Environment Food and Rural Affairs. Available: https://www.gov.uk/government/publications/single-use-plastic-carrier-bags-why-were-introducing-the-charge/carrier-bags-why-theres-a-5p-charge [Accessed 15 January 2016].

Wooldridge, Jeffrey M. 2003. "Cluster-sample methods in applied econometrics." *The American Economic Review* 93(2): 133-138.

WRAP. 2015. WRAP announces new carrier bag figures. [Online]. Available: http://www.wrap.org.uk/2015_carrier_bag_figures. [Accessed 15 January 2016].

Zhao, Xinshu, John G. Lynch, and Qimei Chen. 2010. "Reconsidering Baron and Kenny: Myths and truths about mediation analysis." *Journal of consumer research* 37(2): 197-206.

APPENDIX A: QUESTIONNAIRE

Demographics

1) Where do you live in the UK? (England; Northern Ireland; Scotland; Wales)

2) What is your age?

3) What is your yearly income (before taxes)? (Below £ 15,000; £ 15,000-£ 20,000; £ 20,001-£ 30,000; £ 30,001-£ 40,000; £ 40,001-£ 50,000; £ 60,001-£ 60,000; £ 50,001-£ 70,000; Above £ 70,000; I prefer not to say)

4) What gender do you identify with? (Male; Female)

5) What is the highest level of education you have attained? (Basic education; Secondary education; University degree – undergraduate level (e.g. BSc, BA); University degree – Master level or equivalent (e.g. MSc, MA, MRes); University degree – Doctoral level (PhD); Other - please specify)

Behavioural questions

6) In which of the following retailers have you been grocery shopping in the last month (choose as many as you need)? (Asda; Aldi; Co-operative; Iceland; Marks & Spencer; Morrisons; Lidl; Sainsbury's; Tesco; Waitrose; Small local retailer; Other – please specify)

7) Think about the last time you used carrier bags in your grocery shopping trip (these can be your own bags, as well as new ones from the store). How many bags did you use? (if above 15, please indicate 15)

8) Of the bags you used, how many new carrier bags did you get from the store? (if above 15, please indicate 15)

9) Do you think it is right for the government to make retailers charge customers for plastic carrier bags? (Yes; No)

10) What do you think a fair charge for plastic carrier bag should be (in pence)?

Attitudes

On a 100-point scale (0 = "I strongly disagree"; 100 = "I strongly agree"), what is your level of agreement with the following statement? (randomised order)

11) Minimising the number of plastic carrier bags when I shop for groceries is important to me, regardless of any benefit or inconvenience that may result.

12) I feel good when I don't use new plastic carrier bags when I shop for groceries because it helps the environment.

13) Charging for carrier bags and giving the profits to a good cause (e.g. donating to a charity) will ensure that plastic carrier bag use is no longer a problem for our society.

14) The government should not interfere by requiring retailers to charge for plastic carrier bags.

15) Plastic carrier bags are currently overused.

16) Other customers will continue using plastic carrier bags even if I stop using them.

17) Shoppers have a moral obligation to minimise the use of plastic carrier bags.

Complementary/Substitute behaviours

18) Thinking about the past 7 days, how much money did you donate to a good cause, in pounds? (e.g. donating to a charity)

19) Thinking about today, did you have a reusable carrier bag with you? (Yes; No)

20) Thinking about your last grocery shopping trip, did you buy any long-life/reusable bags? (Yes; No)

21) Finally, thinking about your last grocery shopping trip again, did you buy plastic bin liners? (Yes; No)

Appendix B : Robustness Tests

The difference-in-difference results presented in the paper estimated the impact of the legislative change on a group of consumers not exposed to the charge relative to a reference group who were already paying the charge. The choice of reference group might influence the results observed in this analysis: in a difference-in-difference analysis, the impact of a policy on a target population is estimated relative to a control population of individuals unexposed to the policy. In particular, there may be a fundamental problem of self-selection whereby the treatment allocation is not truly random (given X and D). This potential self-selection entails that environmentally-friendly consumers belong to one the two groups before the legislative change occurs (Keele et al., 2015). This point is unlikely to characterise the choice between Scotland, Northern Ireland, or Wales, given the common environmental policies. However, it is possible that M&S's policy over plastic bags attracted a number of environmentally-conscious customers (see e.g. Disney et al., 2013). However, only two respondents shopped exclusively in M&S before the charge was introduced, and the number and type of attitudinal controls used in the analysis could be sufficient to address the problem. In this section we run a series of robustness checks to test the validity of the results that have been presented.

Specifically, the robustness checks entail repeating the analyses above for three alternative control groups. The first alternative is to observe change in England only (where the new policy was introduced), removing Scotland, Northern Ireland, and Wales; in this case, the control group refers only to shoppers in M&S, who were already being charged for plastic bags for environmental reasons. The second option is to remove individuals shopping in M&S from the analysis altogether, comparing inter-country differences as a purely natural experiment; using this control group would remove the potential self-selection of environmentally motivated shoppers who may prefer M&S because of their interest in the

reduction of plastic bags waste. A third option is identical to the control of the main analysis, but also includes shoppers in Lidl and Aldi, as their customers were also paying for plastic bags before the policy was introduced, although not on environmental grounds (and profits were not donated to a good cause). The last option removes M&S customers from the control group, whilst keeping them in the sample.

Table A2 presents the DID estimates using the different control groups. The impact of the plastic bag charge did not vary qualitatively, albeit some coefficients do differ noticeably. The consumption of new plastic bags decreased with the same trend regardless of the control group used. However, when considering England alone, the charge resulted in an increase in new plastic bags used after one week, a result suggesting that at this point plastic bags use by English M&S customers declined faster that other shoppers. Intrinsic motivations and moral engagement increased in all samples, although these changes remained significant only when M&S consumers were either omitted or not included in the control. Free riding aversion presented less stable results one week after the introduction of the charge, but this variable dropped after seven weeks in all versions. Finally, the impact of the charge on control aversion and on information on the intention of the government did not vary significantly, with control aversion declining (except when considering England alone in period 2) and information increasing in all scenarios. These results indicate that removing other UK countries from the control would have lost significant information relevant for the DID estimator; while removing M&S customers from the control or from the sample would have caused similar but larger estimated effects than otherwise. In summary, the control used in the result section seems to provide realistic estimates, and the key conclusions seem robust to various potential definitions of the treatment group.

Table A1: Robustness checks for the difference-in-difference estimator

	Policy	Initial results		Only England		No M&S		Incl. Lidl & Aldi	
	Period	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
New bags	2	-0.4889***	0.0205	0.1927***	0.0675	-1.4313***	0.1096	-0.4327***	0.0208
	3	-1.6806***	0.0672	-0.9592***	0.0571	-2.2725***	0.1332	-1.5901***	0.1053
Personal motivation	2	1.9385*	0.4715	0.7245	0.5807	0.5952	0.3671	0.6742	0.4266
	3	1.4126**	0.2956	0.1241	0.2897	3.7048***	0.5886	0.7595	0.6567
Control aversion	2	-1.1249**	0.2396	3.1111***	0.2913	-4.5759**	0.6886	-0.6506	0.3477
	3	-6.6851***	0.2493	-2.7340***	0.2213	-12.6123***	0.6713	-5.1179***	0.7343
Moral engagement	2	3.8656***	0.3000	2.6306**	0.4869	2.6093**	0.4888	1.8281**	0.2484
	3	2.7825**	0.3344	0.3142	0.4274	5.1677**	0.5961	1.3421	0.6139
Behavioural Change	2	4.8902***	0.3573	6.0493**	0.6678	0.8882	0.3925	2.3384***	0.2368
	3	4.6112***	0.1838	3.8163**	0.5280	3.8536*	0.9008	3.6199***	0.3505
Free Riding	2	1.4318*	0.4137	4.5929***	0.2073	-1.3000	0.5144	-1.3692*	0.3866
	3	-4.0491**	0.4226	-0.9458	0.4073	-6.3554***	0.7910	-4.9482***	0.3440
Observations		2,704		2,265		1,827		2,704	

Significance is as follows: * = 0.10; ** = 0.05; *** = 0.01. Note: for the "New bags" equation, results refer to the regression without attitudinal variables.

Appendix C:

Table A2: Impact of the plastic bag charge – mediation analysis

	Bags used – All		Bags used – New	
Intercept	7.2524***	0.3153	9.2980***	0.5899
Treated group	0.1312**	0.0633	0.2590***	0.0846
Period 2	-0.0206*	0.0123	**-0.9128***	**0.0855**
Policy, time 2	-0.0511**	0.0238	**-0.3382***	**0.0325**
Period 3	-0.1283***	0.0271	**-1.7465***	**0.0711**
Policy, time 3	-0.4524***	0.0504	**-1.5111***	**0.1139**
Age	-0.0315***	0.0034	-0.0986***	0.0123
Male	-0.4736*	0.2546	0.4741	0.4670
England	Baseline		Baseline	
Northern Ireland	-0.6376**	0.2555	-1.1964	1.4084
Scotland	-0.4941***	0.1732	-1.9598***	0.7145
Wales	-0.6829***	0.1418	-1.7560***	0.1625
Control aversion	0.0083***	0.0027	**0.0214***	**0.0040**
Personal motivation	-0.0056	0.0063	**-0.0219***	**0.0074**
Moral obligation	0.0016	0.0016	**-0.0031**	**0.0015**
Joint Action	-0.0058	0.0043	**-0.0113**	**0.0057**
Free Riding	-0.0058***	0.0009	**-0.0118***	**0.0012**
Sigma	3.2082***	0.0913	4.5524***	0.2077
Income dummies	Yes		Yes	
Education dummies	Yes		Yes	
Retailer dummies	Yes		Yes	
Observations	2,704		2,704	
Pseudo R^2	0.0223		0.0694	
Log-likelihood	-6363.72		-4700.20	

Note: residuals are clustered by period. Significance is as follows: * = 0.10; ** = 0.05; *** = 0.01.