Prices vs. Quantities with Principled Agents*

Anthony Heyes Sandeep Kapur

December 2006

Abstract

When making decisions principled agents, whilst not altruistic, may take account of the impact of their actions upon others. The implied self-restraint may serve partially to mitigate the externality problem. It may change the response of individuals to an externality tax or cause quantitative limits to be less than universally binding as a subset of agents choose to operate strictly within the limit. Under optimal quantity-based regulation welfare is everywhere improving in the propensity for principled behavior in the population; under optimal price-based regulation the relationship is non-monotonic. Price-based regulation is preferred when the regulated population is either highly principled or highly unprincipled, quantity-based regulation for intermediate cases.

Keywords: regulation, principled behavior, prices vs. quantities

JEL codes: H23, H3, L51

*Heyes is at Royal Holloway College, University of London; Kapur is at Birkbeck College, University of London. Contact details for corresponding author: Professor Anthony Heyes, Department of Economics, Royal Holloway, University of London, Surrey TW20 0EX (a.heyes@rhul.ac.uk). We are grateful to Ron Smith, Peth Tuppe and seminar participants at IEDI University of Toulouse and University of Heidelberg for helpful comments. The usual disclaimer applies.

1 Introduction

People and organizations often face decisions in which their private actions impact not just their own well-being but also that of others. The impacts on others are generically referred to as externalities.

It is well known that under standard assumptions individual decisions will maximize social welfare. Where the spillover is negative, agents – left to their own devices – will engage in too much of the activity, failing to take account of the costs that their actions impose on others. Where the spillover is positive, they will engage in too little. Varian (2002) provides a textbook treatment.

Such incentive problems abound in economic models and the real world, notably in environmental policy settings that are our primary interest in this paper. The resulting market failure provides a rationale for policy intervention and extensive effort has been spent investigating the merits of alternative policy instruments, in particular the efficiency of those based on prices (taxes, subsidies, *etc.*) versus those based on quantities (quotas, legal standards). Weitzman (1974) famously demonstrated that when the locations of the marginal social cost and marginal social benefit curves are known either kind of instrument, correctly applied, can achieve first best. The duality between the instruments was made more explicit in a clarifying comment by Laffont (1977).

Weitzman's analysis implies that any efficiency-based preference between price and quantity instruments must depend upon their performance under imperfect information, and he goes on to formalize the circumstances under which each is to be preferred. The basic model has been extended in a variety of directions. Williams (2002) extends the analysis to differentiate between tradable and non-tradable quantity instruments; Kaplow and Shavell (2002) allow for non-linear taxes; Stavins (1996) considers the effect of the uncertainty between costs and benefits being correlated; Montero (2002) allows for less-than-full compliance; Baldursson (2004) extends to account for irreversibility of environmental damage; Finkelshtain and Moledina (1997) incorporate political considerations by allowing for rent-seeking activity by interest groups. Each of these developments – and this is a representative rather than exhaustive list – adds breadth to our understanding of instrument choice and each can provide an alternative basis for preferring one class of instrument over the other.

While legal compulsion and monetary incentives can alter behavior, an individual's 'principles' might also lead him or her to take account in his decision-making of the impact of his actions upon others. In this paper we incorporate 'principled behavior' in a model of externalities with otherwise standard features. We do not replicate 'the Weitzman model', but explore how principled behavior on the part of some or all regulated parties might impact the optimal conduct of price- and quantity-based regulation, and the comparative merits of the two. In that sense our work is in the spirit of the literature that has followed Weitzman.

We develop a behavioral model in which the agent retains a conventional utility function, but in which behavior is influenced by his principles. An individual's principles are captured by an (exogenously determined) parameter that measures the extent to which he takes account of the impact that his activities have on others. At the point of making a decision the agent acts *as if* some part of the externality were privately borne, even though it never is. Agents take actions not because so doing maximizes their intrinsic utility, but rather that their principles cause them to alter their behavior - in particular to exercise restraint in the extent to which they engage in externality-generating activities.[1]

Before progressing to the model it is useful to underline that we do not seek to explain where 'principles' come from but rather treat them as exogenous characteristics of the individual, like capabilities, tastes and other attributes.[2] These may derive from religious conviction, social ethic, education, genetics or elsewhere. We do not claim that such principles will in general be privately efficient, nor that they will be evolutionarily stable, but simply acknowledge that they may exist and will influence behavior in decision environments characterized by externality.

We can note that various *ad hoc* respecifications of individual utility functions have been proposed to 'explain' behavior that departs from that of the strictly selfish agent as traditionally conceived (Cohen and Dickens (2002: 335)). People may be assumed to give weight to the utility of others in

[1]Williamson (2003) adopts a similar approach to principled behavior in his study of organizations, noting: "To be sure 'principled' behavior is an ad hoc move. It is not, however, something to which we are unable to relate, recognise and describe." (Williamson (2003: 17)). He goes on to relate such behavior to self-image.

[2]It is increasingly recognized by economists that a wide variety of character traits, exogenous to economic models, matter for policy analysis. Naevdal and Shogren (2004), for example, survey evidence that shyness is largely hereditary, has a physiological and measurable manifestation, and can impact a variety of economic decisions.

their own utility functions (pure altruism). Alternatively they may derive a 'warm glow' from the action of giving (impure altruism), as suggested by Andreoni (1991). Bernheim (1994) proposes a theory of conformity in which an agent derives a benefit from 'fitting in'.[3] Others have posited an idealized action (sometimes assumed, arbitrarily, to be Kantian - the one that would maximize social welfare if everyone in the population were to choose it) and assumed that agents suffer disutility from behaving in a way that departs from that idealized action. We will note the correspondence between our analysis and this after we have introduced some notation.[4]

Bernheim and Rangel (forthcoming) point to the potential frailty of approaches that involve manipulation of utility functions and argue that a better alternative may be not to change utility functions but to allow for other restraints in utility maximization: "In our view analysis should begin with the premise that choices can diverge from preferences". The approach we adopt is consistent with this – the principles that agents hold will constrain their behavior and cause them to make choices that do do maximize their own utility. But we do not intend the focus of this paper to be about the relative merits of the two approaches. It would be straight-forward to manipulate the utility function to justify such an approach – for instance, by endowing each agent with some 'god given' or hereditary belief as to the morally correct line of action and having him derive a sufficiently high psychic sanction from behaving in a way inconsistent with that belief – but we choose not to. Our focus here, though, is not to 'explain principled behavior'.[5] Rather we point to it as a feature of the world in which we live and ask how that feature should impact optimal policy for a given instrument, and instrument choice. The existence of principles can change the response of individuals to an externality tax or cause quantitative limits to be less than universally binding as a subset of agents choose to operate strictly within the limit. The scenario in which all agents are completely unprincipled remains nested here

[3]His pay-off is a weighted sum of intrinsic utility, determined by actions, and 'esteem', determined by what 'type' others believe him to be. Conformity arises if the weight on esteem is sufficiently high, such that agents are willing to sacrifice intrinsic utility in order to fit in with the societal norm.

[4]See Bilodeau and Gravel (2004) for another interesting model involving Kanitan ethics and behavior. In fact it would be plausible to think that views on idealized behavior will vary substantially across individuals.

[5]How we might try to influence or educate these traits raises interesting – but separate – questions. We also ignore the possibility that the choice of policy instrument, or the level at which it is, may itself influence principles, as suggested by, for example, Frey (1997).

as a special case.

2 The Model

An agent chooses a level $x \in \Re_+$ of some anti-social activity that imposes external costs cx on others. By way of examples we can think of a firm that discharges noxious substances into a river, anti-social parking in a neighborhood, the carriage of superfluous luggage onto airplanes, excessive watering of lawns during a period of water shortage, or noise pollution.

Agents differ in two important ways: (a) the benefit they derive from the activity, and (b) the extent to which they take account of the cost that their activities imposes on others.

In respect to (a), agent i derives benefit $\theta_i b(x_i)$ from the activity where $\theta_i \in \Re_{++}$, $b^{'}(x) \geq 0$, $b^{''}(x) < 0$. Thus, private benefit increases concavely in x. Concavity corresponds to an assumption of diminishing marginal utility to engagement in the activity (or diminishing marginal productivity of abatement effort).

In respect to (b), we introduce the notion of 'principled behavior'. Whilst the damage associated with the activity remains fully incident upon the third party, the principled individual takes account – to a greater or lesser extent – of that damage in her decision-making. More concretely assume that i chooses x to maximize

$$\theta_i b(x) - \alpha_i cx.$$

Thus in choosing his actions the principled agent acts *as if* some part α_i of the damage is incident upon herself, where $\alpha \in [0, 1]$.[6]

We don't assume that all agents are principled to the same degree (though they may be), nor that all agents behave in a principled way. An agent, then, is characterized by the pair (θ, α), which is private information. For simplicity, we assume that θ and α are distributed independently in the population, with (cumulative) distribution functions $F(\theta)$ and $G(\alpha)$. Importantly, we

[6] A Kantian decision rule (which has received some attention in the economics literature) is one that, if universalized, would maximize social welfare. It would be associated here with $\alpha = 1$, and the extent to which an individual operates according to an α less than 1 can be thought of as the extent to which they fall short of the Kantian ideal. Introspection might lead most readers to conclude that their own α is somewhere strictly between zero and one.

leave the form of G entirely general. The standard 'textbook' case is nested here as the special case in which $\alpha_i = 0$ for all i.

2.1 Principled behavior and welfare

In the absence of regulation an agent chooses x to maximize $(\theta b(x) - \alpha c x)$. Under the usual Inada conditions the solution is interior and implicitly defined by the associated first-order condition

$$\theta b'(x) - \alpha c = 0.$$

We denote this choice as $x^*_\emptyset(\theta, \alpha)$. The notation $\emptyset$ will be used throughout to refer to the setting in which there is no policy intervention. Later we will consider the use of quantity-based regulation and price-based regulation to influence behavior.

The net welfare contribution of an agent's chosen activity level x can be written as

$$w(\theta, x) = \theta b(x) - cx.$$

The level of activity that maximizes an agent's welfare contribution – call this level of activity $\hat{x}$ – must satisfy

$$\theta b'(\hat{x}) - c = 0.$$

An unregulated agent chooses $x^*_\emptyset(\theta, \alpha)$, so that his welfare contribution is

$$w_\emptyset(\theta, \alpha) = \theta b(x^*_\emptyset(\theta, \alpha)) - c x^*_\emptyset(\theta, \alpha).$$

Note that $x^*_\emptyset(\theta, 1) = \hat{x}(\theta)$, so that a 'fully-principled' agent chooses the first-best level of activity. In the absence of regulation, an agent with $\alpha < 1$ chooses activity level that exceeds the first-best level, so that $x^*_\emptyset(\theta, \alpha) \geq \hat{x}(\theta)$. Recall that the parameter α captures the extent to which principles induce behavior *as if* the externality were internalized. An agent's welfare contribution $w_\emptyset(\theta, \alpha)$ is increasing in α, with $\alpha = 1$ achieving the first best. Also, due to the concavity of $b(x)$, the optimal level $\hat{x}(\theta)$ is increasing in θ.

Aggregate welfare depends on the distribution of θ and α in the population. With $F(\theta)$ and $G(\alpha)$ denoting the distribution functions, in the absence of regulation we have

$$W_\emptyset = \int_\theta \int_\alpha w_\emptyset(\theta, \alpha) dG(\alpha) dF(\theta).$$

3 Quantity Regulation

Quantity regulation entails the regulator imposing upon each agent a maximum admissible level of the activity. If the externality is due to, say, a firm's air pollution, then the regulatory cap would set the maximum admissible level of emissions. If the externality follows from profligate use of water during periods of shortage then such an approach would involve quantitative limits on water use. The cap is assumed to be fully and costlessly enforced.

Recall that the welfare-optimizing level of activity $\hat{x}(\theta_i)$ varies with θ_i. If the regulator were able to observe agent-specific characteristics then he could implement the first-best outcome by imposing a vector of agent-specific regulatory caps equal to $\hat{x}(\theta_i)$. If θ_i is private information, the regulator can do no better than set a common regulatory cap $\overline{x}$ for all agents.

Importantly, the existence of agents with strictly positive α's implies the possibility that a uniform cap will not necessarily bind on all agents. Not everyone will live at the 'limits of the law'. Some, guided to restraint by their principles, will choose a level of activity strictly less than the cap.[7] More concretely the cap is binding on a particular agent if and only if $x^*_\emptyset(\theta, \alpha) > \overline{x}$. With such a regime of quantity regulation in place, then, the agent's optimal choice is

$$x^*_Q(\theta, \alpha, \overline{x}) = \min[x^*_\emptyset(\theta, \alpha), \overline{x}],$$

and the welfare contribution of the typical agent is

$$w_Q(\theta, \alpha, \overline{x}) = \theta b(x^*_Q(\theta, \alpha, \overline{x})) - c x^*_Q(\theta, \alpha, \overline{x}).$$

Aggregate welfare associated with any regulatory cap is

$$W_Q(\overline{x}) = \int_\theta \int_\alpha w_Q(\theta, \alpha, \overline{x}) dG(\alpha) dF(\theta).$$

[7] We are familiar with two other papers in which (for quite different reasons) caps may not bind on all regulated parties. Brozovic, Sunding and Zilberman (2004) assume an interior solution to a firm's unregulated pollution problem. This contrasts with the (more conventional) assumption that costs are everywhere decreasing in emissions, such that all firms produce at a point where any quantitative emissions cap is binding. Bandyopadhyay and Horowitz (2006) point to a 'safety margin' effect whereby a firm would stay inside a quantitative limit if its realised pollution discharge had a stochastic element (it could only control its pollution level noisily) and penalties for violation were sufficiently large. They use this as an alternative to the altruistic 'explanation' for why firms appear to overcomply.

The regulatory cap binds if the agent derives high benefit from the activity (large θ) or if the agent is sufficiently unprincipled (low α). For particular values of θ and $\overline{x}$, define the critical value $\underline{\alpha}(\theta, \overline{x}) \in [0, 1]$ such that the regulatory cap $\overline{x}$ binds for agents with lower α. Aggregate welfare can be expressed as

$$W_Q(\overline{x}) = \int_\theta \left[\int_0^{\underline{\alpha}(\theta,\overline{x})} w(\theta, \overline{x}) dG(\alpha) + \int_{\underline{\alpha}(\theta,\overline{x})}^1 w(x^*_\emptyset(\theta, \alpha)) dG(\alpha) \right] dF(\theta).$$

The first integral captures the welfare contribution of agents for which the regulatory cap binds, the second of those whose behavior is restrained by principles. Let the regulator choose $\overline{x}$ to maximize aggregate welfare. Assuming differentiability, and using the fact that $w(x^*_\emptyset(\theta, \underline{\alpha}(\theta, \overline{x}))) = w(\theta, \overline{x})$,

$$W'_Q(\overline{x}) = \int_\theta \left[\int_0^{\underline{\alpha}(\theta,\overline{x})} w'(\theta, \overline{x}) dG(\alpha) \right] dF(\theta),$$

where $w'(\theta, \overline{x}) = \theta b'(\overline{x}) - c$. The optimal regulatory cap is given by[8]

$$W'_Q(\overline{x}^*) = \int_\theta \int_0^{\underline{\alpha}(\theta,\overline{x}^*)} w'(\theta, \overline{x}^*) dG(\alpha) dF(\theta) = 0. \tag{1}$$

How does the presence of principled agents affect regulation? We compare the case analyzed here with the 'traditional' economic analysis that ignores any self-restraint due to principled behavior – the special case in which $\alpha_i = 0$ for all i – and find that the presence of principled agents allows for less stringent regulation.

Proposition 1. *The presence of principled agents raises the optimal regulatory cap.*

Proof: Let $\overline{x}^*(G)$ be the optimal regulatory cap for any non-generate distribution $G(\alpha)$. We know that it must satisfy the first-order condition

$$\int_\theta \int_0^{\underline{\alpha}(\theta,\overline{x}^*(G))} w'(\theta, \overline{x}^*(G)) dG(\alpha) dF(\theta) = 0. \tag{2}$$

[8]Note that the function $W(\overline{x})$ is concave in $\overline{x}$. To see why, note that $w_Q(\theta, \overline{x})$ is either independent of $\overline{x}$ or concave in $\overline{x}$. Integration preserves concavity in $\overline{x}$.

The traditional case, with unprincipled agents, amounts to one in which $G(\alpha)$ is degenerate at $\alpha = 0$, or that $G(0) = 1$. For this case any regulatory cap binds on all agents, and the optimal cap – call it $\widetilde{x}$ – must be such that

$$W'(\widetilde{x}) = \int_\theta w'(\theta, \widetilde{x})dF(\theta) = 0. \tag{3}$$

As $w'(\theta, \widetilde{x})$ does not vary with α, we can rewrite the above as

$$\int_\theta \int_0^{\underline{\alpha}(\theta,\overline{x}^*(G))} w'(\theta, \widetilde{x})dG(\alpha)dF(\theta) + \int_\theta \int_{\underline{\alpha}(\theta,\overline{x}^*(G))}^1 w'(\theta, \widetilde{x})dG(\alpha)dF(\theta) = 0. \tag{4}$$

To demonstrate that $\overline{x}^*(G) \geq \widetilde{x}$, assume, to the contrary, that $\widetilde{x} > \overline{x}^*(G)$. Noting that $w'(\theta, x) = \theta b'(x) - c$, due to the concavity of $b(x)$ we must have $w'(\theta, \widetilde{x}) < w'(\theta, \overline{x}^*(G))$. Then, comparing (2) and (4), it follows that the first term in (4) must be negative; if so the second term in (4) must necessarily be positive. Thus our assumption requires

$$\int_\theta \int_{\underline{\alpha}(\overline{x}^*(G),\theta)}^1 w'(\theta, \widetilde{x})dG(\alpha)dF(\theta) > 0. \tag{5}$$

Note that the marginal welfare contribution is negative for $\alpha > \underline{\alpha}(\widetilde{x}, \theta)$,

$$w'(\widetilde{x}, \theta) = \theta b'(\widetilde{x}) - c < \theta b'(\widetilde{x}) - \alpha c < 0, \tag{6}$$

and so also for $\alpha > \underline{\alpha}(\overline{x}^*(G), \theta) > \underline{\alpha}(\widetilde{x}, \theta)$. If so, (5) cannot hold. We have a contradiction. ■

So in this setting principled behavior can be regarded as a *substitute* for regulation – more of the former implies there should be less of the latter. The introduction of principled agents means that there is now an increased likelihood that any agent choosing a high level of x is doing so because it faces high costs of reducing its engagement in the activity (a high θ) and so makes any tightening of the cap more 'expensive' at each margin than was the case in the absence of principled actors.

4 Price Regulation

An alternative to setting quantitative limits on the activity is to impose a tax on it. If the externality is derived from air pollution then the tax would be on emissions of that pollutant. We assume throughout that the tax is fully and costlessly collected.

If the activity is taxed at rate t, the agent described by (θ, α) will choose x to maximize

$$\theta b(x) - tx - \alpha cx.$$

The first-order condition associated with an interior maximum is

$$\theta b'(x) - t - \alpha c = 0.$$

From concavity, we know that an agent's optimal choice, $x_P^*(\theta, \alpha, t)$ is increasing in θ and decreasing in t or α. Other things equal an agent will do more of the activity when he finds cutting back expensive, and will do less when he is highly principled or when the tax rate is high.

In this regime, the agent's welfare contribution is

$$w_P(\theta, \alpha, t) = \theta b(x_P^*(\theta, \alpha, t)) - cx_P^*(\theta, \alpha, t),$$

so that aggregate welfare

$$W_P(t) = \int_\theta \int_\alpha w_P(\theta, \alpha, t) dG(\alpha) dF(\theta) \tag{7}$$

varies with chosen tax rate t. The first-order condition[9] for the optimal tax rate t^* is

$$W_P'(t^*) = 0,$$

where

$$W_P'(t) = \int_\theta \int_\alpha \frac{\partial}{\partial t}[w_P(\theta, \alpha, t)] dG(\alpha) dF(\theta). \tag{8}$$

We have

$$\begin{aligned} \frac{\partial}{\partial t}[w_P(\theta, \alpha, t)] &= [\theta b'(x_P^*) - c] \frac{dx_P^*}{dt} \\ &= [t - (1 - \alpha)c] \frac{1}{\theta b''(x_P^*)}. \end{aligned}$$

[9] To see that the second-order condition holds note that $w_P(\theta, \alpha, t)$ is concave in t as the derivative

$$\frac{\partial w_p}{\partial t} = \left[\theta b'(x_p^*) - c\right] \frac{\partial x_p^*}{\partial t} = [t - c(1 - \alpha)] \frac{\partial x_p^*}{\partial t}$$

is positive for small t (i.e., $t - c(1 - \alpha) < 0$) and negative for large t. The integral, a convex combination of these concave functions, inherits the concavity.

The last equality obtains from manipulation of the first-order condition for the agent's maximum, and makes transparent the role that the tax play. The tax can be thought of as 'correcting' the insufficiency of principles in restraining the anti-social activity. Given heterogeneity in principles, the optimal corrective tax rate varies too: the ideal tax to face agent i would be $t_i = (1 - \alpha_i)c$. With completely principled behavior, when $\alpha_i = 1$, no taxation is necessary. In the absence of any principles, that is if $\alpha_i = 0$, the optimal tax rate is c, the standard Pigovian prescription. In general, for α that are positive but less than 1, the optimal tax rate is positive but less than c.

We note that if all agents were equally principled – that is if $\alpha_i = k$ for all i – then the regulator could implement first best by setting a tax equal to $(1 - k)c$. By appropriate manipulation of the tax rate the same first best level of welfare could be achieved regardless of the value of k.

With heterogeneity in α, however, optimal policy requires choosing a tax that optimizes the welfare impact of distortions across agents. At the optimum an interval of the most principled agents in the population will face a tax that is 'too high', an interval of the least principled face a tax that is too low. It is clear that at the optimum, the tax rate will not exceed c, which leads us to:

Proposition 2. *The presence of principled agents lowers the optimal tax rate.*

Proof. The first-order condition for the optimal tax rate is

$$W_P'(t^*) = \int_\theta \int_\alpha \frac{1}{\theta b''(x_P^*(\theta, \alpha, t^*))}[t^* - (1 - \alpha)c]dG(\alpha)dF(\theta) = 0.$$

The traditional case, with unprincipled agents, amounts to the degenerate case with $G(0) = 0$. If $\alpha = 0$ for all agents, the optimal tax rate is simply c. For a non-degenerate distribution $G(\alpha)$, we must have $[t^* - (1 - \alpha_i)c]$ strictly positive for some α_i and negative for other α_i in the support of distribution G. [To see why, recall that $b'' < 0$ due to the concavity of b; if $[t^* - (1 - \alpha_i)c]$ was strictly positive (negative) for all α_i, the integral in the above expression would be strictly negative (positive)]. Let the support of α_i be $[\alpha_{\min}, \alpha_{\max}]$. If

$$t^* - c + \alpha_{\min} c < 0.$$

we must have $t^* - c < 0$, as long as $\alpha_{\min} \geq 0$ (that is, as long as principles are never 'perverse'). ■

5 Prices vs. Quantities in the Control of Principled Agents

How do the two kinds of regulation compare? How does the insertion of principled agents into a model of externality with otherwise standard features lead us to reconsider the comparative merits of quantity- versus price-based policy?

Before considering policy with principled agents it is useful to note the 'default' result that would obtain in the absence of principled behavior,

Remark 1. *Absent principled behavior (that is when $\alpha_i = 0$ for all i) the maximum achievable level of welfare is weakly* ***greater*** *under price-based regulation than under quantity-based regulation. It is strictly greater if there is heterogeneity in θ_is.*

That is the optimal tax delivers higher welfare than does the best available quantity cap. In this world a Pigovian tax set equal to marginal external damage c implements first best. Any heterogeneity in compliance costs θ_i ensures that the best available cap cannot match that. In such and environment, prices are better than quantities as regulatory instruments.

Our interest in this paper, however, involves allowing for dispersion in α's – variation in the extent of principles amongst agents – and in that setting ranking instruments is less straightforward.

Let W_Q^* denote aggregate welfare when the regulatory cap is chosen optimally. That is, $W_Q^* = W_Q(\overline{x}^*)$. Similarly, let $W_P^* = W_P(t^*)$. It is clear from the analysis of previous sections that these optimized values vary with the distributions $G(\alpha)$ and $F(\theta)$. The following is straight-forward to establish:

Proposition 3. *The maximum achievable level of welfare under quantity-based regulation may be higher or lower than the maximum achievable under price-based regulation. That is, $W_Q^* \lesseqgtr W_P^*$.*

Proof. By example. To show the possibility that $W_Q(\overline{x}^*) > W_P(t^*)$, let the distribution $F(\theta)$ collapse around some value $\overline{\theta}$ (that is, let $\theta_i = \overline{\theta}$ for all i). Then the optimal cap is implicitly defined by

$$\overline{\theta} b'(\overline{x}^*) - c = 0.$$

Imposing this cap directly achieves first best outcome, while the tax regime does not as long as there is some heterogeneity in α.

To establish the possibility that $W_Q(\overline{x}^*) < W_P(t^*)$, let the distribution $G(\alpha)$ collapse around some value $\overline{\alpha}$ (that is, let $\alpha_i = \overline{\alpha}$ for all i). Then the optimal Pigovian tax is implicitly defined by

$$t^* = (1 - \overline{\alpha})c$$

Imposing this tax achieves first best outcome, while the quantity-based regime does not as long as there is some heterogeneity in θ. ■

In contrast to Remark 1, once we allow for the possibility of heterogeneity of α, variation in principles across various actors, we cannot in general establish the superiority of one type of instrument over the other. The proof makes transparent why this should be the case by focussing on extreme cases. Removing variability in compliance costs but leaving some variability in principles implies that quantity-based regulation can implement first best, while price-based regulation cannot. Conversely removing variability in principles but leaving some variability in compliance costs implies that price-based regulation can implement first best, while quantity-based regulation cannot.

Other things equal price-based regulation is better at handling technological heterogeneity (variation in θ) whilst quantity-based regulation is better at handling what we might call 'motivational' heterogeneity (variation in α). The implicit preference for price-based regulation in the setting described in Remark 1 followed not from the *absence* of principles but rather from their homogeneity (that all the α's were of common value, rather than the fact that the common value was zero).

To understand the variation in the performance of the two regimes, it is instructive to consider a further special case. In particular fix some non-degenerate $F(\theta)$ and consider the case in which each agent is characterized by one of two values, $\alpha \in \{\alpha_L, \alpha_H\}$, with $0 \leq \alpha_L < \alpha_H < 1$. To make things easier, and without further loss of generality, we can set $\alpha_L = 0$. We refer to those characterized by $\alpha_L = 0$ as 'unprincipled', and those characterized by $\alpha = \alpha_H$ as 'principled', and denote by π the proportion of principled agents in the population.

With slight abuse of notation we will write W_Q^* and W_P^* as functions of π, as $W_Q^*(\pi)$ and $W_P^*(\pi)$. The functions have a number of characteristics. First note that $W_P^*(0) = W_P^*(1) = W^{FIRSTBEST}$. The level of welfare delivered by the best available tax is the same when $\pi = 0$ as when $\pi = 1$. In the first case the optimal tax is c, whereas in the second it is $(1 - \alpha_H)c$, but in either case setting the tax correctly delivers the first-best level of welfare. In case of departures from these extreme cases, price-regulation cannot deliver the first-best outcome: $W_P^*(\pi) < W^{FIRSTBEST}$.

Minor additional restrictions ensure convexity of $W_P^*(\pi)$. So the welfare delivered by the (optimal) tax regime in non-monotonic, in fact U-shaped, in π.

What of W_Q^*? First we can note that $W_Q^*(0) < W_P^*(0)$ and $W_Q^*(1) < W_P^*(1)$. At each extreme price-based regulation delivers the first best, a level of welfare that quantity-based regulation cannot match. For this example we can establish the following:

Proposition 4. $W_Q^*(\pi)$ *is (weakly) increasing in* π.

Proof. Fix $\overline{x}$ at any arbitrary value, not necessarily the optimal level. Unprincipled agents would like to choose (infinitely) high x for any $\theta > 0$, but are capped at $\overline{x}$.[10] Principled agents fall into one of two groups: those with high θ face a binding cap while those with low θ choose $x_\emptyset^*(\theta, \alpha_H)$ strictly below the cap. Let $\theta_c(\alpha_H, \overline{x})$ denote the critical value that demarcates the latter two types.

The welfare contribution of all capped agents – whether principled or unprincipled – depends on the regulatory cap

$$\overline{w}(\theta, \overline{x}) = \theta b(\overline{x}) - c\overline{x}. \tag{9}$$

The welfare contribution of agents whose choices are not capped is

$$w^*(\theta, \alpha) = \theta b(x_\emptyset^*(\theta, \alpha)) - c x_\emptyset^*(\theta, \alpha). \tag{10}$$

For given $\overline{x}$ and π, aggregate welfare is

$$\begin{aligned} W_Q &= (1-\pi)\int \overline{w}(\theta, \overline{x})dF(\theta) + \pi\left[\int^{\theta_c} w^*(\theta, \alpha_H)dF(\theta) + \int_{\theta_c} \overline{w}(\theta, \overline{x})dF(\theta)\right] \\ &= \int^{\theta_c} \left[(1-\pi)\overline{w}(\theta, \overline{x}) + \pi w^*(\theta, \alpha_H)\right] dF(\theta) + \int_{\theta_c} \overline{w}(\theta, \overline{x})dF(\theta). \end{aligned}$$

[10] In any non-trivial quantity regulation problem, the regulatory cap $\overline{x}$ must constrain the choices of a subset of agents. What happens if α_H is relatively low and the θs relatively large so that the optimal cap constrains all – not a strict subset of – agents? Modifying the steps in this proof implies that variations in π do not affect the outcome or associated welfare levels. Therefore we have a W_Q^* curve that is flat in π. This possibility is subsumed in the qualifier, 'weakly' increasing in the statement of the Proposition.

We claim that $w_Q^*(\theta, \alpha_H) > \overline{w}_Q(\overline{x})$ for $\theta < \theta_c$.[11] If so, an increase in π, keeping $\overline{x}$ fixed, will raise $W_Q(\overline{x}, \pi)$.

The final step in the argument is as follows. If $W_Q(\overline{x}, \pi)$ is increasing in π, it means that the graph of $W_Q(\overline{x}, \pi)$ (as a function of $\overline{x}$) must lie everywhere above the graph of $W_Q(\overline{x}, \pi')$ whenever $\pi > \pi'$. If so, the maximized value of the function – that is, its value when $\overline{x}$ is set optimally – is increasing in π. ■

So an increase in the fraction of principled agents in the population increases the level of welfare delivered by the optimally-calibrated quantity instrument. Whilst the Proposition is presented for the current example, the result is in fact much more general. Proposition 5 in the Appendix argues that any upward shift in the distribution of principles in the population improves welfare under optimal quantity regulation.

Figure 1 here

Figure 1 plots W_Q^* and W_P^* as π varies. It is plotted for the most interesting case, namely that in which W_Q^* and W_P^* cross. We cannot rule out (but ignore here) the possibility that for some distributions $F(\theta)$ and $G(\alpha)$ it may be that W_Q^* lies everywhere below W_P^*.

What policy conclusions can we draw?

Remark 2. *Price-based regulation is better than quantity-based regulation when the proportion of principled agents in the population is either sufficiently high or sufficiently low. Quantity-based regulation is better for a range of intermediate values.*

[11] To see why, note that, by construction, the constraint does not bind in this range, so that $x_\emptyset^* < \overline{x}$. We note that

$$\theta b'(x_\emptyset^*) - c < \theta b'(x_\emptyset^*) - \alpha_H c = 0.$$

The inequality in this relation follows from the fact that $\alpha_H < 1$, and the equality comes from the optimality of $x_\emptyset^*$ for α_H. If so, $w(x)$ is decreasing in the neighborhood of $x_\emptyset^*$, and for $\overline{x} > x^*$, we must have $w_Q^*(\theta, \alpha_H) > \overline{w}_Q(\overline{x})$. In words, optimizing with less than perfect principles results in excessive activity relative to the social optimum. If so, $x_\emptyset^*$ lies on the declining section of the welfare contribution curve: for any $x_\emptyset^* < \overline{x}$ welfare $w(x_\emptyset^*)$ must be higher than welfare $w(\overline{x})$.

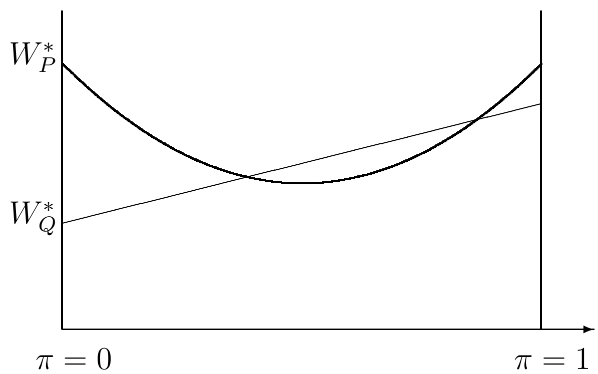

Figure 1: Welfare comparison: price vs. quantity regulation

In terms of Figure 1 we note that there are intervals of π close to zero and one where $W_P^* > W_Q^*$. This reflects that the tax instrument performs best when principles are homogeneous in the population. With unobservable heterogeneity in the level of principles the regulator is obliged to set a tax based on average responses – one that is too high than would ideally be set for a principled agent and too low than would ideally be set for an unprincipled. When there is greater dispersion in the level of principles – values of π close to the middle of the unit interval amount to greater dispersion – the losses due to those departures are their greatest. More generally, Proposition 6 in the Appendix shows that a mean preserving spread of $G(\alpha)$ decreases the value W_P^*.

If we conceive of principles as developing through time (say, due to increased environmental awareness in the population) then as π rises from a low base we can note the possibility of a reversal in instrument choice – switching from price instruments to quantity instruments at one threshold, and then back at some higher threshold.

The impact of greater technological heterogeneity – an increase in the dispersion of θ – is more complicated. Unlike α, the parameter θ enters the welfare function directly, as θ scales the benefit function $b(x)$. A mean preserving spread of $F(\theta)$ causes both W_P^* and W_Q^* to *rise* for a given distribution of principles.[12] Regardless, we know from the proof of Proposition 3 that quantity regulation achieves first best when there is no dispersion in θ. For a range of plausible specifications that we have explored, price regulation comes to out-perform quantity regulation when and only when the distribution of θ is sufficiently dispersed, though we cannot claim this as a general result.[13]

[12]This follows from the fact that the optimized welfare contribution function w^* is convex in θ. The mechanics are best understood from an example. Let $b(x) = \ln x$, and consider the benchmark unregulated case. The optimal choice is $x^* = \frac{\theta}{\alpha c}$ and the optimized welfare contribution

$$w(x^*) = \theta \ln\left(\frac{\theta}{\alpha c}\right) - \frac{\theta}{\alpha},$$

is convex in θ, as $w_{\theta\theta} = \frac{1}{\theta} > 0$. Aggregate welfare W^* is an integral of this convex function, and an increase in dispersion of θ causes the expected value to rise. This sort of argument can be shown to apply more generally.

[13]This claim is easily demonstrated say for the case where $b(x)$ is logarithmic or of the constant relative risk aversion form and the distributions $F(\theta)$ and $G(\alpha)$ are uniform, but probably holds for a wider class of specifications.

Remark 3. *For a range of plausible specifications price regulation dominates quantity regulation when and only when the distribution of θ is sufficiently dispersed.*

In Figure 1, an increase in the dispersion of θ causes graphs of both W_P^* and W_Q^* to shift upwards. Our qualitative result is preserved: that quantity regulation may be preferable for an intermediate range of parameter π. Whether greater technological uncertainty expands this intermediate range may be sensitive to the choice of specification.

6 Conclusions

There is an everyday notion of principled behavior, an important element of which would surely be the taking into account the impact that ones actions have on others. In contrast to the notion of altruism, in which the utility of one agent enters as an argument in the utility function of another, we conceive of principles as a restraint on maximization of intrinsic utility.

We do not seek to explain where principles come from (though we have noted the existence of other scholars who have tried), but rather take their existence and distribution in the population as given and ask how recognizing the prevalence of principled behavior leads us to change our views on how externality-generating behavior should be regulated. Our analysis also ignores the possibility that principles may be endogenously shaped by the choice of policy instrument, or the level at which corrective policies are imposed. At one level the answer is simple: if agents differ in the extent to which they recognize corrective policies, the broad qualitative conclusions of our model survive. Exploring this possibility would be a natural path for further development.

Loosely we can think of the model presented here as featuring two types of heterogeneity – technological (variation in θ's) and motivational (variation in α's). In our setup price instruments tend to be good at handling the former but less so at handling the latter. When there is no motivational heterogeneity the price instrument is unambiguously preferred, but once it is introduced the ranking of instruments becomes more complex.

If tax is the policy instrument in use then the existence of principled agents in the population leads to a lower tax being optimal. If quantitative caps are in use then the optimal cap is higher. Under optimal quantity-based regulation welfare is everywhere improving in the propensity for principled behavior in the population; under optimal price-based regulation the relationship is non-monotonic. Price-based regulation is preferred when the regulated population is either highly principled or highly unprincipled, quantity-based regulation for intermediate cases.

References

1. Andreoni, J. (1991). "Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving," *Economic Journal 100: 466-77.*

2. Baldursson, F. M. (2004). "Prices versus Quantities: The Irrelevance of Irreversibility," *Scandinavian Journal of Economics* 106(4): 805-25.

3. Bernheim, B. D. (1994). "A Theory of Conformity," *Journal of Political Economy* 102: 841-77.

4. Bernheim, B. D. and A. Rangel (forthcoming). "Behavioral Public Economics: Welfare Economics with Fallible Decision-makers," forthcoming in *Economic Institutions & Behavioral Economics Eds.* Peter Diamon and Hannu Vartiainen: Princeton University Press.

5. Bilodeau, M. and N. Gravel (2004). "Voluntary Provision of a Public Good and Individual Morality," *Journal of Public Economics* 88(3): 645-66.

6. Brozovic, N., D.L. Sunding and D. Zilberman (2004). "Taxes versus Quatos *or* Taxes versus Upper Bounds" Working Paper in Agricultural & Resource Economics, University of California, Berkeley.

7. Cohen, J. L. and W. T. Dickens (2002). "A Foundation for Behavioral Economics," *American Economic Review* (Papers & Proceedings), 92(2): 335-38.

8. Cropper M. L. and W. E. Oates (1992). "Environmental Economics: A Survey," *Journal of Economic Literature* 30(2): 657-740.

9. Fehr, E. and S. Gachter (2000). "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14(3): 159-82.

10. Finkelshtain, I. and Y. Kislev (1997). "Prices versus Quantities: The Political Perspective," *Journal of Political Economy* 105(1): 83-100.

11. Frey, B.S and F. Oberholzer-Gee (1997), "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding- Out," *The American Economic Review*, 87(4): 746-755.

12. Heyes, A. G. (2001). "Honesty in a Regulatory Context: Good Thing or Bad," *European Economic Review*.

13. Kaplow, L. and S. Shavell (2002). "On the Superiority of Corrective Taxes to Quantity Regulation," *American Law and Economics Review.*

14. Laffont, J.-J. (1977). "More on Prices vs. Quantities ," *Review of Economic Studies* 44(1): 177-182.

15. Montero, J. P. (2002). "Prices vs. Quantities with Incomplete Enforcement," *Journal of Public Economics* 85(5): 435-454.

16. Shogren, J. F. and E. Naevdal (2004). "Genetic Variability, Economic Behavior and the Formation of Social Norms," Discussion Paper in Economics Number 228, Woodrow Wilson School of International and Public Affairs, Princeton University.

17. Stavins, R. N. (1996). "Correlated Uncertainty and Policy Instrument Choice," *Journal of Environmental Economics and Management* 30(2): 218-32.

18. Varian, H. (2002). *Intermediate Microeconomics: A Modern Approach*, Sixth Edition, W. W. Norton.

19. Weitzman, M. (1974). "Prices versus Quantities," *Review of Economic Studies* 41: 471-91.

20. Williamson, O. E. (2003). "Human Actors and Economic Organization," Berkeley Working Paper in Economics: California.

21. Williams, R. (2002). "Prices vs. Quantities vs. Tradable Quantities," NBER Working Paper No. 9283.

A Appendix

Consider any two distributions, $G_1(\alpha)$ and $G_2(\alpha)$. We say that $G_1(\alpha)$ first-order stochastically dominates $G_2(\alpha)$ if $G_1(\alpha) \leq G_2(\alpha)$. Then G_1 describes a more principled population, as G_1 puts more mass on the higher values of α.

Proposition 5. *Let $W_Q^*(G)$ denote the level of welfare when the cap is chosen optimally given distribution G. If $G_1(\alpha) \leq G_2(\alpha)$ then $W_Q^*(G_1) \geq W_Q^*(G_2)$.*

Proof. To begin, fix $\overline{x}$. For fixed $\overline{x}$ and given any θ, the cap binds for some $\alpha \leq \alpha_c(\theta)$. The welfare contribution of agents with capped choices is

$$\overline{w}_Q(\theta, \overline{x}) = \theta b(\overline{x}) - c\overline{x}, \tag{11}$$

while that of agents whose choices are not constrained by the cap is

$$w_Q^*(\theta, \alpha) = \theta b(x_\emptyset^*(\theta, \alpha)) - c x_\emptyset^*(\theta, \alpha). \tag{12}$$

Aggregate welfare is

$$W_Q(G) = \int_\theta \left[\int_0^{\alpha_c(\theta)} \overline{w}_Q(\theta, \overline{x}) dG(\alpha) + \int_{\alpha_c(\theta)}^1 w_Q^*(\theta, \alpha) dG(\alpha) \right] dF(\theta). \tag{13}$$

Note that $w_Q^*(\theta, \alpha) > \overline{w}_Q$ for $\alpha > \alpha(\theta)$. To see why, note that in this range, the constraint does not bind, so that $x_\emptyset^* < \overline{x}$. Further, given the optimality of $x_\emptyset^*$, we have

$$\theta b'(x_\emptyset^*) - \alpha c = 0.$$

If so, we must have $\theta b'(x_\emptyset^*) - c < 0$ for $\alpha < 1$. If so $w(x) = \theta b(x) - cx$ must be decreasing in the neighborhood of $x_\emptyset^*$. Given that $\overline{x} > x^*$, we must have

$$\theta b(x_\emptyset^*) - c x_\emptyset^* > \theta b(\overline{x}) - c\overline{x}.$$

Consider how changing the distribution $G(\alpha)$ affects W_Q for a fixed $\overline{x}$. If $G_1(\alpha)$ first-order stochastically dominates $G_2(\alpha)$, it attaches more weight to relatively high values of α. Given that $w(x_\emptyset^*) \geq w(\overline{x})$, $W_Q^*(G_1) \geq W_Q^*(G_2)$, as required.■

Proposition 6. *A mean preserving spread of $G(\alpha)$ decreases the value W_P^*.*

Proof. Follows from the fact that $w_p(\theta, \alpha, t)$ is concave in α. To see why, note that the derivative

$$\frac{\partial w_p}{\partial \alpha} = \left[\theta b'(x_p^*) - c\right] \frac{\partial x_p^*}{\partial \alpha} = [t - c(1 - \alpha)] \frac{\partial x_p^*}{\partial \alpha}$$

is positive for small α (i.e., $t - c(1 - \alpha) < 0$) and negative for large α. A mean preserving spread of $G(\alpha)$ lowers $W_P(t)$ for a fixed tax rate t. As a mean preserving spread of the distribution $G(\alpha)$ causes the entire function $W_P(t)$ to move down, the maximized value of this function – that is, W_P^* – must necessarily fall. ■