Honesty in a Regulatory Context
- Good Thing or Bad?

Anthony Heyes
December 15, 1997

Abstract

It is often taken for granted that if more firms were innately honest or ethical in the way they behaved, this would be a good thing. In this paper we use the example of environmental regulation to show that such an claim cannot, in general, be sustained. If regulation is by pollution tax we show that - once optimal agency response is taken into account - social welfare is non-monotonic in the proportion of firms that report emissions honestly. The choice of policy instrument may itself be characterised by “reversals” with command-and-control methods being preferred for intermediate values of population honesty, a tax system being preferred at the extremes. This means that if - because of the spread of “ethical shareholding” or for whatever reason - the honesty of the corporate population increases through time, we should not be surprised to see at first a switch away from market-based instruments, and then a switch back.

The model is argued to be consistent with a number of the stylised features of American regulatory history. It may also provide a defence of the USEPA’s willingness to tolerate a “dysfunctional culture of dishonesty” amongst those it is supposed to police (Yergin (1991)). Though environmental regulation is used as an example for the purposes of exposition, the results are of more general interest.

*Department of Economics, Royal Holloway College, University of London, Egham, Surrey, TW20 0EX. Electronic mail: aheyes@econ.royal.ac.uk.
1 Introduction

It is often taken for granted that if more firms were inherently honest or ethical in the way they behaved this would be a good thing. In this paper we use the example of environmental regulation to show that such an claim cannot, in general, be sustained. If regulation is by pollution tax we show that once optimal agency response is taken into account - social welfare is non-monotonic in the proportion of firms that report emissions honestly. The choice of policy instrument may itself be characterised by "reversal" with command-and-control methods being preferred for intermediate values of population honesty, a tax system being preferred as the extreme. This means that if - because of the spread of "ethical shareholding" or for whatever reason - the honesty of the corporate population increases through time, we should not be surprised to see at first a switch away from market-based instruments, and then a switch back.

A straight-forward heterogeneity argument underpins the key results. Consider a simple model of a tax regime based on (imperfectly verified) self-reports in which tax play an incentive role, as in the case of a tax designed to internalise an externality. It is not too surprising that such a regime may perform best - once optimal agency response is factored in - either when everyone is inherently honest or when everyone is inherently dishonest. When everyone is inherently honest the agency can achieve first-best by setting the tax-rate equal to marginal damage (i.e. as the Pigovian level). Suppose, instead, that everyone is inherently dishonest and is able to get away with under-reporting some fraction of their true emissions, the agency can achieve the same outcome simply by raising the tax-rate to (m/1 - d) such as to maintain the effective rate. The regime performs less well when the population is a mixture of some people who are routinely honest (i.e. do not exploit opportunities for evasion presented to them) and others who are not. Whilst the framework developed in this paper is somewhat more complex than this - incorporating differences in compliance costs, evasion opportunities etc - the basic story is unchanged.

The model is argued to be consistent with a number of the stylized features of American regulatory history. It may also provide a defence if the EPA's willingness to tolerate a "culture of dishonesty" amongst those it is supposed to police (Yaeger (1991)). Though environmental regulation is used as an example for the purposes of exposition, it is apparent that the results are likely to be of more general interest.

It is conventional in economics to model agents as rational choose - understanding with laws or under-reporting tax liability when enforcement parameters are such as to make such behaviour financially profitable. There is compelling empirical and anecdotal evidence, however, to suggest that some individuals and firms are inherently honest or ethical. A key to the scale of 'honesty' in a compliance context can be had by setting all of the enforcement parameters (fine of penalty, probability of apprehension etc) in econometrically estimated compliance functions equal to zero. So doing invariably implies a positive level of 'unassisted' compliance which can be substantial (see, for example, Magat and Viscusi (1980) and Jones (1990)). Limited attempts have been made to model the implications of routine honesty in enforcement contexts. Grunig, Riehmanen and Wilde (1990) included a sub-group of routinely honest agents ("ethical compliers") in a standard income tax enforcement framework. Varying the prevalence of routine honesty in their model has no effect on optimal policy nor expected net tax revenue. Eehrend and Fedzials (1990) use numerical simulations to show (e.g. Yaeger (1991)).
\[ SL(x_{CC}) = x(x_{CC}) + \delta x_{CC} \]

where \( x_{CC} \) is implicitly defined by
\[ c_{\delta}(x_{CC}) = -d. \]

Under a CC regime regulation is input-based and \( a \) and \( \beta \) are not relevant. The EPA sets (optimally) the pollution target to be satisfied and the technology to be used in achieving it, and polices it directly.

The shortcomings of CC regulation in this context are the usual ones. For a given distribution of abatement effort choice of technique is not cost-effective - some fraction of firms \( F(I) \) (i.e. those with \( \delta i \leq 1 \)) use the back-stop technology despite knowing a cheaper way of doing the same thing. Furthermore, the distribution of effort across firms is inefficient - no mechanism exists for firms with lower marginal cost to reduce emissions more than do those with higher marginal costs. The calibration of the CC regime and the level of welfare delivered do not depend upon \( \beta \).

2.2 A pollution tax regime

A tax regime does not share these problems, but has other ones of its own.

We will differentiate between "cheaters" and "non-cheaters" (C and NC respectively). A cheater is a firm which under-reports its emissions under a tax regime. To do this it must be (a) able and (b) willing to do so. Cheaters and non-cheaters constitute proportions \( a(1-\beta) \) and \( (1-a(1-\beta)) \) of firms respectively. Note that analytically there is no difference between firms which report truthfully because they are inherently honest, and those that do so because they have no option. The distinction will matter when we come to comparative statics, however, and will also motivate some of our later discussion of alternative policy instruments.

Assume firm \( i \) is a cheater. For a given tax rate \( t \), its problem is to choose a level of emissions to minimize the sum of abatement expenditures plus tax payments, i.e. to minimize
\[ F(x_i, t|C) = \min\{\theta_i, 1\} c(x_i) + t(1-m) x_i. \]

The interior solution \( x^*(t|C) \) is implicitly defined by
\[ \min\{\theta_i, 1\} c(x^*(t|C)) = -(1-m)t \]

The analogous problem faced by the non-cheater is to minimize
\[ F(x_i, t|NC) = \min\{\theta_i, 1\} c(x_i) + t x_i \]

such that \( x^*(t|NC) \) is implicitly characterized by
\[ \min\{\theta_i, 1\} c(x^*(t|NC)) = t \]

Equation (7) is conventional in the case of a fully-enforced tax. The firm chooses a level of emission such as to equate the marginal cost of abatement (the left-hand side) with the marginal cost of tax, the marginal cost term being conditional on an optimal choice of technique. Equation (6) is the same condition adapted to take account of the fact that the effective marginal tax rate for a C is \((1-m)t\).

It is apparent that, other things being equal, a C emits more than does a NC since the effective tax rate it faces is lower. Choice of technique in equilibrium is cost-effective - each firm chooses the least cost way of doing the abatement it does. The distribution of abatement effort across firms remains inefficient. Combining (6) and (7) yields
\[ \frac{c(x^*(t|NC))}{c(x^*(t|C))} = \frac{1}{(1-m)} \]

- marginal abatement costs are not equalized across firms. An NC abates more than does an identical C.

The incentive effects of tax changes are qualitatively conventional. Implicit differentiation of (6) and (7) yields
\[ \frac{\partial x^*(t|C)}{\partial t} = \frac{(1-m)}{\min\{\theta_i, 1\} c(x^*(t|C))} \]

and
\[ \frac{\partial x^*(t|C)}{\partial t} = \frac{1}{\min\{\theta_i, 1\} c(x^*(t|C))} \]

respectively. Both derivatives are negative: A marginal increase in \( t \) induces a decrease in emissions from both types of firms, the reduction is greater, externals parties, for an NC.
2.2.1 The EPA's tax-setting problem

The EPA chooses $t$ to minimise an expected social loss function:

$$SL(t|a, \beta) = (1 - \alpha(1 - \beta)) SL(t|NC) + \alpha(1 - \beta) SL(t|C)$$  \hfill (10)

where

$$SL(t|NC) = \int_0^\infty \min[\delta, 1] \alpha^*(t|NC) d\delta$$

$$SL(t|C) = \int_0^\infty \int_0^\infty \min[\delta, 1] \alpha^*(t|C) d\delta d\theta$$

and $SL(t|C)$ is defined analogously. The optimal tax is implicitly defined as a function of $\beta$, the proportion of honest firms in the population, by the first-order condition associated with the EPA's problem:

$$(1 - \alpha(1 - \beta)) \frac{dSL(t^*(\beta)|NC)}{dt} + \alpha(1 - \beta) \frac{dSL(t^*(\beta)|C)}{dt} = 0$$  \hfill (12)

which, after substitution and simplification becomes

$$\alpha(1 - \beta) \int_0^\infty \frac{\partial \alpha^*(t^*(\beta)|C)}{\partial t} f(t) d\delta t^*(\beta)(1 - \alpha) = d$$

$$= \alpha(1 - \beta) \int_0^\infty \frac{\partial \alpha^*(t^*(\beta)|NC)}{\partial t} f(t) d\delta t^*(\beta) - d$$  \hfill (13)

It is straightforward to infer from [13] that $t^*(\beta) = d$ if all firms are honest, the optimal tax is Pigovian - and that this achieves first-best for the usual reasons.\(^{14}\)

If all firms are dishonest, the picture is more complicated. If all firms were C's the EPA could achieve first-best by setting a tax $t = [\delta/(1 - \alpha)]$ thus ensuring that every firm faced an effective tax rate $d$. Setting $\beta = 0$ does not, however, imply this since a probability $(1 - \alpha)$ that a firm will not have the scope to cheat even though it has the inclination. $t^*(\beta)$ is implicitly defined by substituting $d = 0$ into [13]. Implicit differentiation yields

$$\frac{d}{dt} t^*(\beta) = \text{sign} \left( \frac{dSL(t^*(\beta)|a, \beta)}{dt} \right) < 0$$  \hfill (14)

- the optimal tax is everywhere decreasing in $\beta$ - implying that $t^*(\beta) > t^*(1) = d$. This is intuitive.

2.3 Honesty - good thing or bad?

Of interest to us here are the welfare impacts of variations in $\beta$. Expected social loss conditional on the operation of an optimally calibrated tax regime can be written

$$SL(t^*(\beta)|a, \beta) = (1 - \alpha(1 - \beta)) SL(t^*(\beta)|NC) + \alpha(1 - \beta) SL(t^*(\beta)|C)$$  \hfill (15)

Complete differentiation yields (after application of the envelope theorem)

$$\frac{dSL(t^*(\beta)|a, \beta)}{dt} = \alpha [SL(t^*(\beta)|NC) - SL(t^*(\beta)|C)]$$  \hfill (16)

It is apparent from the structure of the model that

$$SL(t^*(1)|C) > SL(t^*(1)|NC) = \frac{dSL(t^*(1)|a, 1)}{dt} < 0, \forall a > 0.$$  \hfill (17)

When $\beta = 1$ the EPA is able to implement a first-best outcome by setting a Pigovian tax $t = d$. Any decrease in $\beta$ implies an increase in $t^*$ and an increase in expected social loss. Without additional restriction we are unable, however, to provide an unambiguous sign for the analogous derivative evaluated at $\beta = 0$. It is shown in an Appendix, however, that

\(^{14}\)This can be seen by setting $a = 1$ and $\beta = 0$.\n
8
\[ \frac{\partial}{\partial \beta} \left( \frac{\partial S_L(C(\beta))}{\partial C} \right) < 0, \forall \alpha, \beta > 0. \]  

(13)

These results together imply the following:

**Proposition 1** The relationship between the expected social loss associated with an optimally-calibrated tax regime and \( \beta \) can be described, qualitatively, by one of the two curves in Figure 1.

The more ‘interesting’ case is the non-monotonic one. Given the results in [17] and [18] it is apparent that a necessary and sufficient condition for this to be the applicable one is that \( \frac{\partial^2 S_L(C(\beta))}{\partial C \partial \beta} < 0 \). Moving to some \( \beta \) interior to the unit interval implies an injection of heterogeneity which makes regulation more difficult and generates a welfare loss. The EPA sets \( \beta \) to balance the marginal dead-weight loss due to honest firms doing too much abatement against that due to dishonest firms doing too little. At the resulting \( \beta(0) > \beta(2) > \beta(1) \) no firm faces the “correct” marginal incentives.

For \( \alpha < 1 \) things are more ambiguous. It is apparent from the structure of the model (and straightforward to show formally) that \( SL(\beta(0), 0) > SL(\beta(1), 1), \forall \alpha < 1 \) - the best implementable tax regime works ‘better’ when all firms are honest than when all are dishonest. This is because when all are evaluated the EPA still has to contend with the fact that the regulated population contains both \( C \)’s and \( C_0 \)’s (the proportion \( 1 - \alpha \) of dishonest firms will be \( 1/1 \)’s because they are not able to evade, even though they would like to) whilst if all were honest all will be \( 1/1 \)’s. It is apparent that from a base of \( \beta = 1 \) a marginal decrease in \( \beta \) should reduce welfare. More interesting is that from a base of \( \beta = 1 \) a marginal increase in \( \beta \) does likewise. This is not necessarily a surprising result. When all firms are dishonest the EPA anticipates this by setting \( \beta \). With \( \beta \) set on this basis the replacement (in a large population) of one dishonest firm by one honest one reduces welfare become, ceteris paribus, that firm can be expected to be doing socially excessive levels of abatement.

9.2.1 Variations in \( \beta \) and the choice of regulatory instrument

It is also interesting to think about the possible implications of variations in \( \beta \) for the EPA’s choice of regulatory instrument and we do this by reintroducing the command-and-control option described in Section 2.1. Expected social welfare increases as \( \beta \) increases, and increases for a given \( \beta \), but with a slope that depends on the other parameters.

12This is a crude approximation to the more plausible variation of \( \theta \) with \( \beta \). In the latter model \( \theta \) is a non-monotone function of \( \beta \).

13It is, in this sense, the distinction between the two types of social gain - those without the externality (the ‘honestest’) and those with the externality but without the ability - affects the slope of the command-and-control option described in Section 2.1. Expected social welfare increases as \( \beta \) increases, and increases for a given \( \beta \), but with a slope that depends on the other parameters.
loss per firm under this regime is, recall,
\[ SL(x^A_{cc}) = c(x^A_{cc}) + dx_{cc} \] 
where \( x^A_{cc} \) is implicitly defined by
\[ a(x^A_{cc}) = \text{d}. \]

Neither the calibration of the instrument (i.e., choice of \( x^A_{cc} \)) nor the expected welfare from its operation are sensitive to \( \beta \).

\[ SL(x^A_{cc}) \text{ and } SL(t^B(\beta; x, \beta)) \text{ can be plotted on common axes (Figure [2]).} \]

The relative location of the horizontal line \( SL(x^A_{cc}) \) depends upon the productivity of the backstop technology. We know that \( SL(x^A_{cc}) > SL(t^B(\beta; x, 1)) \) since \( F(\beta) > 0 \) (recall our assumption that at least some portion of firms have access to an abatement technology cheaper than the backstop). Given continuity in \( \beta \) of both of the loss functions this can be stated thus:

**Proposition 3** When a sufficiently large fraction of firms are honest, a tax regime performs better than a command-and-control regime.

In Figure [2], for example, sufficient means \( \beta > \beta_3 \). It is also apparent that if \( SL(x^A_{cc}) > \text{Max}(SL(t^B(\beta; x, \beta) \text{ then a tax-based regime will always be preferred. More interesting is the scope for "reversals" in instrument choice.} \)

**Proposition 4** If \( \text{Max}(SL(t^B(\beta; x, \beta)) > SL(x^A_{cc}) > SL(t^B(0); x, 0)) \text{ then a tax regime performs better than a command-and-control regime for values of } \beta \text{ sufficiently large or sufficiently small. For intermediate values of } \beta \text{ the command-and-control regime will perform better.} \)

This is the case illustrated in Figure [5]. Tax is the preferred instrument for values of \( \beta \) less than \( \beta_3 \) or greater than \( \beta_4 \). Expected social loss conditions on optimal instrument selection and calibration is - for the interesting case (that involving reversals) - piecewise non-monotonic in \( \beta. \)

\[ \text{(This is why we have chosen to refer to } CC \text{ as the "benchmark" regime in this paper.} \]

\[ \text{Notice that this condition is necessarily satisfied if } \alpha = 1 - \text{ if the ability to evade is universal.} \]

**3 Implications**

The key advantage attributed to pollution taxation is the technological flexibility it offers to regulators - empowering them with choice of techniques and their implied ability to "handle" technological heterogeneity. This is captured in the model presented here: In the absence of the variation in firms' technological "outside option" command-and-control would invariably be the preferred method of regulation.

The novelty in our analysis is to incorporates another type of heterogeneity - what we might refer to as "motivational heterogeneity" - which a tax-based regime is comparatively bad at handling. The intuition is best understood by considering the \( \alpha = 1 \) case in which all firms have access to the means of compliance. In that case it is apparent that when all (or a sufficient fraction of) firms are either honest or dishonest then a tax regime can be expected to perform well, and certainly no worse than the optimally-calibrated command-and-control alternative. It is for intermediate values of \( \beta \) that command and control - implying, as it does, the stipulation of a technology that for some firms is not cost-effective - may work better. For values of \( \alpha < 1 \) the trade-offs are somewhat more complex, but the basic intuition for the non-monotonicity of welfare in \( \beta \) remains the same.

The significance of the analysis depends, to a large extent, upon the interpretation applied to \( \beta \).

Suppose that \( \beta \) evolves exogenously through time, influenced by broader social trends etc., then we should expect a responsive regulatory agency to adjust to take account of such evolution. This response might involve recalibrating a given instrument or - if one of the critical thresholds is traversed - switching between instruments. A long-term upward trend in the evolution of \( \beta \) - driven by, for example, the spreading influence of ethical investors - may generate "reversals" in instrument choice, with an initial switch away from market-based instruments followed by a later switch back towards them.

What of regulatory attitudes towards changes in \( \beta \)? This depends upon the starting point and the size and direction of the change - certainly the simple view that more honesty is always good cannot be sustained. Starting from \( \beta = 0 \) (the "defeat" assumption in conventional economic analysis) marginal increases in the parameter are an impediment to the agency in its work such that "pious" ethical firms might legitimately be regarded as a
Appendix

With obvious adaptation of notation we can write

$$S_L^*(\theta, \beta) = \{1 - \alpha(1 - \beta)\}, S_L^*(\theta, \beta|NC) + \alpha(1 - \beta), S_L^*(\theta, \beta|C)$$  \hspace{1cm} (A.1)

Noting that the partials $\partial S_L^*/\partial \theta = 0$ then, after application of the envelope theorem:

$$\frac{dS_L^*(\theta, \beta)}{d\theta} = S_L^*(\theta, \beta|NC) - S_L^*(\theta, \beta|C)$$  \hspace{1cm} (A.2)

It is apparent the structure of the model that:

$$\frac{dS_L^*(\theta, \beta|NC)}{d\theta} < 0.$$  \hspace{1cm} (A.3)

We can write

$$\frac{dS_L^*(\theta, \beta|C)}{d\theta} = S_L^*(\theta, \beta|NC) - S_L^*(\theta, \beta|C)$$

which is of ambiguous sign. (Need to characterize ambiguity as in Proposition 2 here).

Differentiating (A.2) gives

$$\frac{\partial}{\partial \theta} \left( \frac{dS_L^*(\theta, \beta)}{d\beta} \right) = \frac{\partial}{\partial \theta} \left( \frac{\partial S_L^*(\theta, \beta|NC)}{\partial \theta} - \frac{\partial S_L^*(\theta, \beta|C)}{\partial \theta} \right)$$  \hspace{1cm} (A.4)

Let the expression in brackets on the right-hand side of (A.4) be denoted $\Omega$. Then [A.4] will have the sign opposite to $\Omega$ (since $\partial S_L^*/\partial \theta < 0$). Now,

$$\frac{\partial S_L^*(\theta, \beta|C)}{\partial \theta} = - \frac{\partial S_L^*(\theta, \beta|NC)}{\partial \theta} \{1(1 - f) - d\},$$  \hspace{1cm} (A.5)

whilst Equation [13] implies

$$\frac{\partial S_L^*(\theta, \beta|NC)}{\partial \theta} = \frac{(1 - \beta) S_L^*(\theta, \beta|C)}{\partial \theta}$$  \hspace{1cm} (A.6)

such that

$$\Omega = - \frac{\partial S_L^*(\theta, \beta|NC)}{\partial \theta} \{1(1 - f) - d\}\frac{(1 - \beta)}{\beta} - 1 \hspace{1cm} (A.7)$$


---

18The initial motivation for this paper came from comments made during part of an interview I conducted with an EPA economist at the center of another piece of research in 1996: "...there is an expectation that people will play the game, get away with what they can, and then not necessarily such a big thing. As long as you realize you're playing a game then it's not necessarily all such a big problem".

19For good empirical and institutional evidence the reader is referred to Berger (1991: 296-298).

14
which is positive (note from Equation [13] that \((\ell(1 - f) - d) < 0\) for all \(\alpha > 0\).)

Given continuity of \(SL(t^*(\beta, \beta))\) in \(\beta\), (A.3) and (A.7) together are sufficient to prove Propositions 1 and 2.

Figure [1]
Bibliography


