Chapter 4 Imperfect Information

In our discussion to this point we have assumed that there is perfect and symmetric information among all agents involved in the design and execution of environmental policy. This has allowed us to refer to damage and cost functions that are fully observed by the regulator, firms, and households alike. Among other things, the full information assumption helped us establish the equivalency of auctioned permits and emission taxes, and more generally suggested that a particular pollution reduction goal can be obtained using one of several economic incentive instruments – all of which have similar efficiency properties. Though it is pedagogically useful, a perfect information assumption is a clear departure from reality. In this chapter we begin to consider how imperfect information may alter the effectiveness and efficiency properties of the policy instruments.

Imperfect information plays a large role in environmental economics, and as reviewed by Pindyck (2007), it comes in many guises. In this chapter we focus on a particular type of uncertainty: the regulator's inability to fully observe aggregate pollution damage and abatement cost functions. Importantly, we assume for this analysis that firms know their own abatement cost functions with certainty. Once we acknowledge that these functions must be estimated by the regulator (inevitably with error), it is natural to ask about the extent to which estimation error influences the performance of different regulatory approaches. Casual intuition suggests estimation errors will result in imprecise policy, but will not cause us to systematically favour one approach over another. Investigation of the extent to which this intuition holds began with Weitzman's (1974) classic paper, and has continued in various forms to this day. The unifying theme in this large literature is a search for policy instruments that minimize the ex post inefficiencies that occur due to the regulator's uncertainty. A related literature beginning with Kwerel (1977) examines ways that the regulator can design mechanism that resolve the uncertainty, thereby eliminating ex post inefficiencies. In this chapter we examine these two strands of literature.

In focusing on uncertainty related to abatement cost and damage functions, we abstract for now from other types of imperfect information. Importantly, in what follows we assume that

Fixed typos 8 July 2011

aggregate and firm-level emissions are perfectly observed. Thus we abstract from moral hazard issues that arise when the regulator cannot monitor firms' pollution and abatement outcomes. Furthermore, since abatement costs are deterministic from a firm's perspective, we do not examine their behaviour under state of the world uncertainty. Finally, our discussion in this chapter is static in nature; dynamic issues such as irreversibility, learning over time, and discounting are not addressed. We address these additional topics related to uncertainty in subsequent later chapters.

4.1 Price versus quantities

We begin with a comparison of emission taxes and marketable pollution permits under imperfect information. Weitzman (1974) and Adar and Griffin (1976) present similar research designed to determine the conditions under which price instruments such as emission taxes or quantity instruments such as transferable permits will dominate under different types of uncertainty. We follow convention in this area by first examining damage function uncertainty with known abatement costs, and then cost function uncertainty with a known damage function. We examine the efficiency properties of taxes and permits in these cases and establish some qualitative results. Using the method described by Weitzman, we then add structure to the damage and abatement cost functions, which allows us to derive more precise descriptions of the ramifications of the two types of uncertainty. We close our discussion on prices versus quantities by commenting on how current policy and research continues to be influenced by these findings.

4.1.1 Damage function uncertainty

Continuing our established notation define $D(E)$ as the social damage function from aggregate pollution level E, $C(E)$ as the aggregate abatement cost curve, and $D'(E)$ and $-C'(E)$ as the marginal damage and aggregate marginal abatement cost curves, respectively. Denote the optimal level of pollution by E^* and recall that this is defined as the solution to $-C'(E)=D'(E)$. The functions $D(\cdot)$ and $C(\cdot)$ summarize true schedules that, under different scenarios, may be imperfectly observed by the regulator. Functions that are estimated with error are denoted with '~', so that $\tilde{D}(\cdot)$ and $\tilde{C}(\cdot)$ are the estimated abatement cost and damage functions, and $\tilde{D}'(\cdot)$ and $\tilde{C}'(\cdot)$ are their respective derivatives.

Consider first the case where the regulator knows $-C'(E)$ with certainty, but uses the damage function estimate $\tilde{D}(E)$ to choose a policy emission target. Suppose for illustration that the regulator underestimates the marginal damage function so that

$$\tilde{D}'(E) < D'(E) \quad \forall E, \tag{4.1}$$

and that she chooses a policy target $\tilde{E}$ based on the relationship $-C'(E) = \tilde{D}'(E)$. Two related questions arise from the fact that $\tilde{E} \neq E^*$ in general: what is the welfare (efficiency) loss from targeting $\tilde{E}$, and does the size of the loss depend on the instrument choice? The welfare loss is defined as the difference between total social costs realized at $\tilde{E}$ and those that would occur at E^*:

$$\begin{aligned} WL &= \left[D(\tilde{E}) + C(\tilde{E})\right] - \left[D(E^*) + C(E^*)\right] \\ &= \int_{E^*}^{\tilde{E}} D'(E)dE - \int_{E^*}^{\tilde{E}} -C'(E)dE. \end{aligned} \tag{4.2}$$

The welfare loss for our particular case is illustrated by Panel A of Figure 4.1, shown as area a. Since the emission level is set too high, $\tilde{E} - E^*$ units of pollution having higher marginal damage than marginal abatement cost are emitted.

The figure also shows that, under this scenario, the choice of emission tax versus transferable permits results in the same welfare loss. For the former the regulator sets $\tau(\tilde{E})$, and for the latter she distributes or auctions $L = \tilde{E}$ permits. Since the polluting firms choose emissions such that $\tau = -C'(E)$, both policy instruments result in the same ex post level of pollution, and hence the same welfare loss. That the same result holds when the marginal damage function is overestimated is also clear from the figure. From this we see that damage function uncertainty by itself should not cause the regulator to favour one instrument over the other.

4.1.2 Abatement cost function uncertainty

Suppose now that the regulator knows $D(E)$, but needs to use an estimate of $C(E)$ to set the pollution target. For illustration we assume that

$$-\tilde{C}'(E) < -C'(E) \quad \forall E < \hat{E}, \tag{4.3}$$

Fixed typos 8 July 2011

which is to say that the regulator underestimates the true marginal abatement cost curve for all policy-relevant emission levels. She chooses the emission target $\tilde{E}$ based on the relationship $-\tilde{C}'(E) = D'(E)$. Thus the a priori target implies a level of emissions that is too low relative to the optimum. The ex post emission level and welfare loss, however, will depend on the actual policy chosen.

The difference between the tax and permit policies is shown in Panel B of Figure 4.1. Under the quantity instrument the regulator supplies $L = \tilde{E}$ permits to the polluting firms, and the permit price $\sigma(\tilde{E})$ emerges at the point where permit supply intersects the true aggregate marginal abatement cost curve. Thus the quantity of emissions is fixed by the regulator, but the price is based on firms' behaviour. Under the tax instrument the regulator administratively sets an emission fee $\tau(\tilde{E})$, and firms react by choosing emissions to equate the tax level to their true marginal abatement cost curve. In the figure this causes the ex post aggregate emission level to be $E(\tau)$. In contrast to the quantity instrument, here the emission price is set and the aggregate emission level is determined by firms' behaviour. In a qualitative sense abatement cost uncertainty clearly matters, since the two policies lead to different emission level outcomes. But is there a systematic difference in the efficiency properties of taxes and permits?

In Panel B of Figure 4.1 the welfare loss from the quantity instrument is area *b* and the welfare loss from the tax instrument is area *c*, which are given analytically by

$$
\begin{aligned}
b &= \int_{\tilde{E}}^{E^*} -C'(E)dE - \int_{\tilde{E}}^{E^*} D'(E)dE \\
c &= \int_{E^*}^{E(\tau)} D'(E)dE - \int_{E^*}^{E(\tau)} -C'(E)dE.
\end{aligned} \tag{4.4}
$$

Inspection of the figure and the analytical expressions suggests that we cannot make a general statement about the relative size of the welfare losses – i.e. in some instances $b>c$ and in others $c>b$. We can, however, try to isolate the features of the cost and damage functions that will cause one to be larger than the other. For intuition in this regard consider Figure 4.2. Panels A and B are drawn to reflect the same true and estimated abatement cost functions, as well as the same policy target $\tilde{E}$. The figures differ only in the steepness of the marginal damage curve as it passes through $-\tilde{C}'(\cdot)$ at $\tilde{E}$ (albeit dramatically to make our point). In both cases the regulator distributes $L = \tilde{E}$ to implement a quantity instrument or sets $\tau = \tau(\tilde{E})$ to implement a price instrument. In Panel A the steep, known marginal damage curve restricts the relative gap that can arise between $\tilde{E}$ and E^*, while the unknown marginal abatement cost curve implies the relative size of the gap between $E(\tau)$ and E^* is not bounded. The tax instrument in this case leads to the higher welfare loss – i.e. $b>a$ – due to the larger resulting distance between the optimal and realized emission level. The opposite is true in Panel B. Here the flat marginal damage curve bounds the difference arising between $\tau(\tilde{E})$ and the optimal but unknown emission tax, so that the gap between $E(\tau)$ and E^* is small. In contrast estimation error in the marginal abatement cost curve implies the relative gap between $\tilde{E}$ and E^* is not bounded. With a flat marginal damage curve $c>d$, and a tax policy leads to a smaller welfare loss.

This intuition can be made more formal by writing the marginal damage function to include an explicit parameter for its slope. We rewrite the damage function as $D(E,\eta)$, and assume that $D_{E\eta}(\cdot)>0$. Note that (as implied by Figure 4.2) the optimal pollution level E^* depends on η, which we here denote $E^*(\eta)$. The following proposition summarizes the consequences of abatement cost function uncertainty when the marginal damage function is known:

Proposition 4.1

Suppose the regulator has an estimate of the aggregate marginal abatement cost curve such that $-\tilde{C}'(E) \neq -C'(E)$. Consider a pollution target $\tilde{E}$ defined by $-\tilde{C}'(\tilde{E}) = D_E(\tilde{E},\eta)$, a class of marginal damage functions $D_E(\tilde{E},\eta)$ running through $\tilde{E}$, and a tax level $\tau(\tilde{E})$. For larger values of η – i.e. a steeper marginal damage curve –

i. the gap $|E^*(\eta) - \tilde{E}|$ becomes smaller and the welfare loss under a permit policy shrinks, and
ii. the gap $|E^*(\eta) - E(\tau)|$ becomes larger and the welfare loss under a tax policy increases, where $E(\tau)$ is the emission level that occurs under the tax $\tau = \tau(\tilde{E})$.

This proposition is illustrated by Figure 4.3, for which $\eta_1 < \eta_2$. As the marginal damage curve pivots from $D_E(E,\eta_1)$ to $D_E(E,\eta_2)$, the optimal pollution level moves closer to $\tilde{E}$, and the welfare loss from a permit policy based on $L=\tilde{E}$ shrinks from area *a*+*b*+*c* to area *a*. Simultaneously the pivot in the marginal damage curve causes the optimal emission level to shift further away from $E(\tau)$, and the welfare loss under a tax policy based on $\tau=\tau(\tilde{E})$ increases from area *e* to area *d*+*e*. Thus all else equal, a flatter marginal damage curve favours a tax approach to regulation, and a steeper marginal damage function favours a permit approach.

4.1.3 Weitzman Theorem

Our discussion to this point has been intuitive more than formal in that we have primarily used graphs to highlight the role of relative slopes in determining the performance of tax and permit policies under uncertainty. We have not yet presented a general result summarizing how slope features of both abatement cost and damage functions determine the second best optimal policy when uncertainty is present. For this purpose it is necessary to introduce additional structure to the model. Define the damage function once again by $D(E,\eta)$, where η is now a random variable representing the regulator's uncertainty about the damage function. As usual $D_E(\cdot)>0$ and $D_{EE}(\cdot)>0$; we also assume $D_\eta(\cdot)>0$ and $D_{E\eta}(\cdot)>0$ so that increases in η shift up both the damage and marginal damage functions. Similarly define the aggregate abatement cost function by $C(E,\varepsilon)$, where ε is a random variable representing the regulator's uncertainty about abatement costs. To our familiar assumptions $-C_E(\cdot)>0$ and $-C_{EE}(\cdot)<0$, we add that ε enters the cost function so that total and marginal abatement cost are increasing in ε; that is, $C_\varepsilon(\cdot)>0$ and $-C_{E\varepsilon}(\cdot)>0$.

Faced with uncertain damage and abatement cost functions, the regulator's task when designing a quantity instrument is to find the emission target that minimizes the expected sum of abatement costs and damages:

$$\min_E EV\left[C(E,\varepsilon)+D(E,\eta)\right], \tag{4.5}$$

where EV is the expected value operator. The emission target $\tilde{E}$ satisfies the first order condition

$$EV\left[-C_E(E,\varepsilon)\right]=EV\left[D_E(E,\eta)\right]. \tag{4.6}$$

To implement the quantity instrument, the regulator issues or auctions $L=\tilde{E}$ permits to the polluting industry. Under such a quantity instrument the expected social cost at pollution level

$\tilde{E}$ is

$$ESC^Q(\tilde{E}) = EV\left[C(\tilde{E},\varepsilon)\right] + EV\left[D(\tilde{E},\eta)\right]. \tag{4.7}$$

Things are somewhat more complicated under the tax policy. Behaviour among polluting firms determines the aggregate emission level under a tax policy according to $-C_E(E,\varepsilon)=\tau$. We can solve this for $E=E(\tau,\varepsilon)$, which is the ex post emission level that arises given firms' actual costs and a particular tax rate τ. Knowing the response function $E(\cdot)$ but not the value of ε, a rational regulator's objective is to choose τ to solve the problem

$$\min_{\tau} EV\left[C\left(E(\tau,\varepsilon),\varepsilon\right) + D\left(E(\tau,\varepsilon),\eta\right)\right], \tag{4.8}$$

which arises from substituting $E=E(\tau,\varepsilon)$ into (4.5). er choice of tax instrument $\tilde{\tau}$ satisfies the first order condition

$$EV\left[-C_E\left(E(\tau,\varepsilon),\varepsilon\right)\times E_\tau(\tau,\varepsilon)\right] = EV\left[D_E\left(E(\tau,\varepsilon),\eta\right)\times E_\tau(\tau,\varepsilon)\right], \tag{4.9}$$

where $E_\tau(\cdot)$ is the partial derivative with respect to τ. Using the identity $-C_E\left(E(\tau,\varepsilon),\varepsilon\right)\equiv\tau$ and substituting into (4.9) we obtain

$$EV\left[\tau\times E_\tau(\tau,\varepsilon)\right] = \tau\times EV\left[E_\tau(\tau,\varepsilon)\right] = EV\left[D_E\left(E(\tau,\varepsilon),\eta\right)\times E_\tau(\tau,\varepsilon)\right], \tag{4.10}$$

from which we can implicitly define the target tax rate $\tilde{\tau}$ by

$$\tilde{\tau} = \frac{EV\left[D_E\left(E(\tilde{\tau},\varepsilon),\eta\right)\times E_\tau(\tilde{\tau},\varepsilon)\right]}{EV\left[E_\tau(\tilde{\tau},\varepsilon)\right]}. \tag{4.11}$$

The expected social cost from pollution under this price instrument is

$$ESC^P(\tilde{\tau}) = EV\left[C\left(E(\tilde{\tau},\varepsilon),\varepsilon\right)\right] + EV\left[D\left(E(\tilde{\tau},\varepsilon),\eta\right)\right]. \tag{4.12}$$

Ex post, it is generally the case that $\tilde{E}\neq E(\tilde{\tau},\varepsilon)$, and so the preferred second best policy depends on the relative sizes of ESC^Q and ESC^P.

To gain traction on comparing the two expected social cost levels, we approximate the true abatement cost and damage functions using a second order Taylor series expansion around $\tilde{E}$. Define

$$C(E,\varepsilon) \approx c(\varepsilon) + \left(\mathrm{K}_1 + \varepsilon\right) \times (E - \tilde{E}) + \frac{\mathrm{K}_2}{2} \times (E - \tilde{E})^2$$
$$D(E,\eta) \approx d(\eta) + \left(\Delta_1 + \eta\right) \times (E - \tilde{E}) + \frac{\Delta_2}{2} \times (E - \tilde{E})^2, \tag{4.13}$$

where ε and η are zero mean random variables, $c(\varepsilon)$ and $d(\eta)$ are random functions, and (K_1, K_2, Δ_1, Δ_2) are constants. Differentiating with respect to E, the implied marginal damage and marginal abatement cost functions from these approximations are

$$-C_E(E,\varepsilon) \approx -\left(\mathrm{K}_1 + \varepsilon\right) - \mathrm{K}_2 \times (E - \tilde{E})$$
$$D_E(E,\eta) \approx \Delta_1 + \eta + \Delta_2 \times (E - \tilde{E}). \tag{4.14}$$

A few things about these functions are noteworthy. First, ε and η summarize the regulator's uncertainty about the marginal damage and cost curves, and they serve to shift up and down the marginal damage and abatement cost curves. Second, if we evaluate the functions at $E = \tilde{E}$, from (4.14) we can see that

$$EV\left[-C_E(\tilde{E},\varepsilon)\right] \approx \mathrm{K}_1 > 0$$
$$EV\left[D_E(\tilde{E},\eta)\right] \approx \Delta_1 > 0, \tag{4.15}$$

since $EV(\varepsilon)$=$EV(\eta)$=0 and the last terms drop out when $E = \tilde{E}$. Finally, differentiating the marginal damage and abatement cost approximations with respect to E leads to

$$\mathrm{K}_2 \approx C_{EE}(E,\varepsilon) > 0$$
$$\Delta_2 \approx D_{EE}(E,\eta) > 0. \tag{4.16}$$

The Taylor series approximation allows us to derive a statement for $ESC^P - ESC^Q$ that depends only on terms with known signs. We summarize this finding in the following proposition:

Proposition 4.2 (Weitzman Theorem)

For approximations of the abatement cost and damage functions given by equation (4.13), and for ε and η independently distributed with expected value of zero, the comparative advantage of a price instrument over a quantity instrument is

$$ESC^P(\tilde{\tau}) - ESC^Q(\tilde{E}) = \frac{\sigma^2(\Delta_2 - \mathrm{K}_2)}{2\left(\mathrm{K}_2\right)^2}, \tag{4.17}$$

where

$$\sigma^2 = EV\left[\left(-C_E(E,\varepsilon)+EV\left[C_E(E,\varepsilon)\right]\right)^2\right] = EV\left[\varepsilon^2\right] > 0. \qquad (4.18)$$

We provide a derivation of equation (4.17) in the appendix to this chapter, and give an opportunity to consider the case when ε and η are correlated in Exercise 4.1.

This result shows that the welfare loss difference depends on the relative steepness (in absolute value) of the marginal abatement cost and marginal damage functions. Extreme examples can help to illustrate this finding. If the marginal damage function is a horizontal line (i.e. all units of pollution contribute equally to total damages), then Δ_2=0 and $ESC^P(\tilde{\tau}) - ESC^Q(\tilde{E}) < 0$, suggesting a tax instrument will be the preferred choice. In contrast, if the marginal abatement cost is constant, then K_2=0 and $ESC^P(\tilde{\tau}) - ESC^Q(\tilde{E}) > 0$. This implies a system of pollution permits will be preferred. The more general case is illustrated by Figure 4.4, where Panel A shows a comparatively steeper marginal damage curve, and Panel B shows a comparatively steeper marginal abatement cost curve. The efficiency loss from tax and permit policies under both scenarios are shown as areas *a*, *b*, *c*, and *d*. Note that in each case uncertainty in the marginal damage function (or a shift in its estimate) affects the size of the inefficiencies but not their ranking. Consistent with the Weitzman Theorem, the characteristics of the functions in Panel A favour a quantity instrument, while those in Panel B favour a tax instrument.

4.1.4 Contemporary policy and research relevance

The theoretical results from the 1970s comparing price and quantity instruments continue to have relevance in both policy and research circles to this day. The policy relevance comes from the fact that, while damage and abatement cost functions are almost always unknown, it is often possible to develop intuition on their shapes over the policy relevant range of emissions. Climate change is a good case in point. As discussed by Nordhaus (2007, p.37) and others, the marginal cost of greenhouse gas abatement is related to the current level of emissions and is therefore sensitive to the degree of reduction, as is the case for most common pollutants. In contrast, the damages from greenhouse gas emissions occur based on the cumulative stock of gasses in the atmosphere; as such the marginal damage from an additional unit of CO_2 emission, for example, is largely independent of how much CO_2 is currently being emitted. This intuition suggests the marginal damage of a particular ton of CO_2 emitted is the same, whether it was the tenth or ten thousandth ton emitted in a month. These features suggest that, for climate change, the marginal abatement cost function is likely to be steeper than the marginal damage function, and all else being equal a tax approach to climate change policy is preferred to a permit-based scheme. Several authors (Metcalf, 2008, is a good example) have used this argument in policy debates to support their preference for carbon taxes rather than a system of tradable carbon emission permits.

The research agenda launched by the original prices versus quantities analysis has focused on examining how specific features of policy design that were left out of the base model may interact with abatement cost and damage function uncertainty to alter the original conclusions. A non-exhaustive list of specific areas that have been examined include correlated uncertainty (Stavins, 1996), modelling a stock externality (Newell and Pizer, 2003), and dealing with imperfect enforcement (Montero, 2002).

4.2 Hybrid Policies

Our analysis thus far has considered only linear tax schemes. However, non-linear tax schemes may be preferred when the regulator has imperfect information about firms' abatement cost curves. Non-linear tax schemes are common in other contexts, the best example being progressive income taxes in both Europe and the United States, where the marginal tax rate depends on a person's income. In principle something similar could be used for environmental policy. Consider again the case of a known damage function $D(E)$ but unknown abatement cost function, and define a pollution tax schedule $\tau(E)$ based on total emissions such that $\tau(E)=D'(E)$. Firm j's pollution-related total cost under this type of policy is

$$TC_j = C_j(e_j) + \tau(E) \times e_j, \tag{4.19}$$

where $C_j(\cdot)$ and e_j are the abatement cost function and emission level, respectively. If we assume that firms cannot influence the tax schedule (i.e. each behaves as a price taker regarding the tax system), the first order condition for cost minimization results in $-C'_j(e_j)=\tau(E)=D'(E)$ for all firms, suggesting the condition for optimal pollution is met. Intuitively, the regulator and the firms know the marginal damage function, but only firms know their abatement cost function. The regulator sets the tax rate ex post to the level of marginal damage that occurs based on the observed pollution level. If the industry emits $E>E^*$, a comparatively high emission fee (i.e. far up the $D'(E)$ curve) is assessed ex post, and firms must pay a marginal tax rate that is higher than their marginal abatement cost at E. If the industry emits $E<E^*$ a low emission fee is assessed, but firms' marginal abatement cost level at E is higher than the marginal tax rate. In either case firms would have been better off at E^*, since it is only at this emission level that the ex post tax will be equal to the marginal abatement cost level.

Though attractive in concept, a non-linear tax scheme as described here is fraught with practical difficulties. Since firms do not know the actual price when making emission decisions, this logic only holds when firms have knowledge of their own *and* their competitors' cost structures. In addition, tax bills based on collective rather than individual behaviour tend to be infeasible from a moral and fairness perspective. For this reason research has focused on investigating policies that mimic features of a non-linear tax, but do so without similar barriers to their use. By and large these have taken the form of hybrid policies that combine elements of tradable permit and emission tax instruments. In this section we review theoretical developments in this area.

4.2.1 Safety valves

Roberts and Spence (1976) suggest a mixed system in which a polluting industry receives an allocation of transferable permits to cover emissions, as well as the option to pay a tax on emissions that are not covered by a permit or receive a subsidy for permits that are not used. By providing a safety valve in the form of the announced prices, the regulator guards against the more extreme consequences of improperly estimating E^* due abatement cost uncertainty. To see how the mechanism works, consider the following additions to our model. The regulator distributes or auctions $L=\tilde{E}$ permits based on the known damage function and her estimate of the aggregate marginal abatement cost function. Denote the price of an emission permit by σ, and the quantity of permits that firm j holds after the permit market clears by $\bar{e}_j$. The regulator announces a tax τ that firms must pay for all emissions in excess of their permit holdings, so that the tax bill is $\tau(e_j-\bar{e}_j)$ if $e_j>\bar{e}_j$. She also announces a subsidy ζ that firms receive for each permit they hold but do not use. For $e_j<\bar{e}_j$ the subsidy amount is given by $\zeta(\bar{e}_j-\bar{e}_j)$.

Under this type of policy firm j chooses e_j and $\bar{e}_j$ to minimize its pollution related expenses according to

$$\min_{e_j,\bar{e}_j} TC_j = \begin{cases} C_j(e_j)+\sigma\bar{e}_j+\tau(e_j-\bar{e}_j) & e_j \geq \bar{e}_j \\ C_j(e_j)+\sigma\bar{e}_j-\zeta(\bar{e}_j-e_j) & e_j \leq \bar{e}_j \end{cases}. \tag{4.20}$$

Before examining the firm's optimal behaviour, consider the following argument for why, in equilibrium, the permit price must be bounded from above by the tax rate and from below by the subsidy. To see this, suppose instead that $\sigma>\tau$. In this case firms will always choose to pay the emission tax rather than covering emissions with the more expensive permits. Firms will sell their permits (or not bid in the case of auctioned permits), and the permit price will fall at least until $\sigma\leq\tau$. If $\sigma<\zeta$, firms could earn profit by buying permits at σ and turning them back in for the larger amount ζ. This type of arbitrage opportunity will increase the demand for permits, and push the price up at least until $\zeta\leq\sigma$. So in equilibrium, it must be the case that $\zeta\leq\sigma\leq\tau$.

This condition suggests there are three types of firm-level outcomes that we can observe. If $\zeta<\sigma<\tau$, firms set $e_j=\bar{e}_j$, since it is cheaper to cover emission with a permit than to pay the tax, and holding extra permits to receive ζ results in a loss per permit. In this case firms operate such that $-C'_j(e_j)=\sigma$. If instead $\sigma=\tau$, firms operate such that $-C'_j(e_j)=\tau$, and emission levels are not constrained by the supply of permits that was original distributed. Finally, if $\sigma=\zeta$ firms operate such that $-C'_j(e_j)=\zeta$, and there will be an excess supply of permits, resulting in less pollution than the regulator's target. The Roberts and Spence mechanism is a hybrid in that in some instances it functions like a price instrument, and in others like a quantity instrument. The tax rate provides a safety valve to polluting firms, in that it assures that emission rights will always be available at a fixed and known price, regardless of the position of the actual or estimated marginal abatement cost curve. The subsidy rate provides a safety valve for the environment, in that emission reductions beyond that implied by L can still be achieved when marginal abatement costs are overestimated by the regulator.

A hybrid policy of this type provides advantages over pure price or quantity instruments when the regulator is uncertain about abatement costs. If she underestimates the aggregate marginal abatement cost curve (and therefore the price of permits), the tax rate serves to cap how high firms' marginal abatement cost levels can climb. If the regulator overestimates the aggregate marginal abatement costs and (and therefore the price of permits), the subsidy assures there are still incentives for emissions reductions. These advantages are illustrated in Figure 4.5. Panel A shows three different aggregate marginal abatement cost curves and policy parameters corresponding to $L = \tilde{E}$, τ and ζ. By specifying the prices along with the quantity the regulator can achieve an efficient outcome (zero welfare loss) for three possible marginal abatement cost curves. For the cost function $C^1(E)$, the relevant policy parameter is ζ, and firms receive subsidies for reducing emissions below $\tilde{E}$. For $C^2(E)$ the market price of permits lies between τ and ζ, suggesting the quantity parameter $L = \tilde{E}$ is in play. Finally, for $C^3(E)$ the tax level determines aggregate emissions in that firms pay τ for each unit of pollution beyond $\tilde{E}$.

The more practical advantages of the hybrid system are shown in Panels B and C. Panel B shows a case in which the regulator has underestimated the aggregate marginal abatement cost curve. To implement a pure quantity instrument she would issue $L = \tilde{E}$ and expect a price of $\tilde{\sigma}$ to arise. Once in place, such a policy would result in an actual price of σ and a welfare loss given by area a+b. If she instead implements a mixed instrument that also includes a safety valve tax τ, the realized level of emissions is $E(\tau)$ and the welfare loss reduces to area c. The parallel case in which the regulator overestimates the marginal abatement cost function is shown in Panel C. If $L = \tilde{E}$ permits were distributed without an accompanying subsidy price ζ, there would be a gap of size $\tilde{E} - E^*$ between realized and optimal emissions, leading to a welfare loss shown by area d+e. With the subsidy option included, firms accept payment to reduce emissions below the number of permits distributed, leading to aggregate emissions $E(\zeta)$ and the smaller welfare loss shown as area d. So while the Roberts and Spence mechanism cannot prevent welfare losses from occurring, it does serve to limit their magnitude relative to pure quantity or price approaches.

BOXED EXAMPLE

Discussions over the inclusion of safety-valve like provisions are common in almost all of the ongoing debates regarding the formation of permit trading schemes for controlling CO_2 emissions. The world’s largest existing carbon market is in the European Union, where the member states have agreed to limit carbon emissions through the European Union Emissions Trading Scheme (*EU ETS*). For example, the *EU ETS* has a compliance design feature that stipulates a 100€ fine per ton of CO_2 emitted without surrendering an allowance. This is not a pure safety valve, however, in that a covering allowance (a European Union Allowance, or *EUA*) from the next year's allotment must subsequently be surrendered. From the market's founding in 2005 until the time of this writing, the price of an *EUA* has generally not exceeded 30€, and fines have not been assessed for uncovered emissions. In the United States a number of carbon trading schemes have been proposed and debated, though not implemented as of this writing. An example carbon control bill from the US 110th Congress – S.1766, the 'Bingaman-Specter bill – stipulated a safety valve price on carbon emissions that would be available one month a year, during the time when emissions and allowances are being balanced. See Murray, Newell, and Pizer (2009) and the citations therein for additional discussion.

END BOXED EXAMPLE

4.2.2 Approximating the marginal damage curve

In a coarse sense the Roberts and Spence instrument traces out the marginal damage curve using step-wise constant functions. This is most apparent in Panel A of Figure 4.5, where the price line firms face begins at ζ, is vertical at $\tilde{E}$, and then levels off at τ. Recall that the non-linear tax scheme discussed above defines the tax schedule to be the marginal damage curve. Thus any policy that connects the effective price firms pay for emission rights to the marginal damage function will emulate the non-linear tax, and reduce the welfare loss that occurs due to uncertainty about abatement costs. This realization has led researchers to consider generalizations of the price/quantity hybrid instruments that provide finer approximations to the marginal damage function, while avoiding the difficulties of non-linear taxes.

In an appendix to their article Roberts and Spence themselves suggest a generalized version of their instrument in which a portfolio of permit types $L_1,...,L_N$ is announced, along with corresponding tax levels $\tau_1,...,\tau_N$ and subsidy levels $\zeta_1,...,\zeta_N$. As market prices $\sigma_1,...,\sigma_N$ for each permit type emerge, the regulator obtains a stepwise approximation to the aggregate marginal abatement cost curve. From this she determines the optimal (i.e. welfare loss minimizing) permit supply, denoted L_n, from among the N that were announced. Through a system of complicated side payments and purchases, she then adjusts the system so that only type n permits are held by the firms and $\zeta_n < \sigma_n < \tau_n$. Thus a flexible permit supply system can, in principle, used to reduce the degree of inefficiency compared to the simple hybrid model. In practice the Roberts and Spence mechanism is cumbersome to explain, and likely impossible to implement. Nonetheless the idea is worth pursuing, and so we proceed by first presenting a simpler approach proposed by Henry (1989), which allows us to develop the intuition of approximating the marginal damage curve with multiple permit types. We then discuss a related mechanism suggested by Unold and Requate (2001) that, given its reliance on familiar financial mechanisms, is more feasible to implement.

Like Roberts and Spence, Henry's (1989) approach uses a flexible supply menu of permits $L_1 < L_2 ... < L_N$, where L_1 and L_N are the lowest and highest conceivable emission levels, respectively. For each interval $(L_{n-1}, L_n]$, an upper threshold price $\bar{\sigma}_n$ and a lower threshold price $\underline{\sigma}_n$ are fixed and announced, where $\underline{\sigma}_{n+1} = \bar{\sigma}_n$. The regulator uses the permit levels and threshold prices in the following algorithm to implement the instrument:

a) Issue an initial amount of permits L_n based on the condition $-\tilde{C}'(L_n) = D'(L_n)$, and observe the permit market price σ that emerges. If $\underline{\sigma}_n < \sigma < \bar{\sigma}_n$, take no further action.
b) If $\sigma > \bar{\sigma}_n$, issue an additional $L_{n+1} - L_n$ permits and observe the new price.
    i. If $\underline{\sigma}_{n+1} < \sigma < \bar{\sigma}_{n+1}$, take no further action. If $\sigma > \bar{\sigma}_{n+1}$ an additional $L_{n+2} - L_{n+1}$ permits are issued and a new price is observed. This continues a total of K times until either $\underline{\sigma}_{n+K} < \sigma < \bar{\sigma}_{n+K}$, or $\underline{\sigma}_{n+K} > \sigma$.
    ii. If at iteration K the permit price emerges below the lower threshold such that $\underline{\sigma}_{n+K} > \sigma$, buy back a quantity of permits $(L_{n+K} - L_{n+K-1})/2$ and observe the permit price. Continue buying back permits in smaller increments until $\underline{\sigma}_{n+K} < \sigma < \bar{\sigma}_{n+K}$.
c) If $\sigma < \underline{\sigma}_n$, buy back $L_n - L_{n-1}$ permits and observe the new price.
    i. If $\underline{\sigma}_{n-1} < \sigma < \bar{\sigma}_{n-1}$, take no further action. If $\sigma < \underline{\sigma}_{n-1}$, buy an additional $L_{n-1} - L_{n-2}$ back and observe the market price. This continues K times until either $\underline{\sigma}_{n-K} < \sigma < \bar{\sigma}_{n-K}$, or $\bar{\sigma}_{n-K} < \sigma$.
    ii. If at iteration K the permit price emerges above the upper threshold such that $\bar{\sigma}_{n-K} < \sigma$, issue an additional quantity of permits $(L_{n-K} - L_{n-K-1})/2$ and observe the permit price. Continue issuing additional permits in smaller increments until

$$\underline{\sigma}_{n-K} < \sigma < \bar{\sigma}_{n-K}.$$

Henry's algorithm allows the regulator to incrementally buy and/or sell permits at fixed and known prices, in order to adjust the permit supply level to correct errors based on her initial estimate of the industry's costs. Figure 4.6 displays a stylized example of how the algorithm unfolds. Panel A shows how the permit levels and threshold prices trace out the known marginal damage curve in a step-wise fashion. The behaviour of firms and the reactions of the regulator assure that the price firms ultimately face lies within the step through which the true but unknown marginal abatement cost curve runs. This is shown in Panel B. Two permit levels L_1 and L_2 are drawn along with their respective threshold prices. Based on her best estimate $-\tilde{C}'(E)$, the regulator initially issues L_1 permits. The firms' reaction, based on the true cost curve $-C'(E)$, causes the initial permit price to be $\sigma(L_1)$, where $\sigma(L_1) > \bar{\sigma}_1$. Seeing this, the regulator issues an additional L_2-L_1 permits; this causes the market price for permits to fall to $\sigma(L_2) < \underline{\sigma}_2$. In this example the regulator has over-adjusted, and she now needs to purchase permits back from the industry. The figure shows a buyback of L_2-L permits and a final price of $\sigma(L)$, where $\underline{\sigma}_2 < \sigma(L) < \bar{\sigma}_2$. With this buyback the regulator has reached her stopping point, and the final pollution/permit level occurs at $E=L$. As the figure is drawn a small welfare loss still occurs relative to the optimal pollution level, but the iterative algorithm has allowed the regulator to come arbitrarily close to the optimum – without ex ante knowledge of the industry's abatement cost function.

While the Henry algorithm is intuitive, the institutional structure needed to actually implement such a policy is substantial. In particular, multiple points of interaction between the regulator and the polluting industry are needed, and examples of such iterative approaches to environmental policy almost nonexistent. With this as motivation, Unold and Requate (2001) describe a policy that is similar in spirit, but relies instead on financial institutions to adjust the supply of pollution permits to the near-optimal level. To see their logic, consider the aggregate abatement cost curve function $C(E,\varepsilon)$, where ε again denotes the regulator's uncertainty, and it serves to shift the total and marginal abatement cost curves such that $C_\varepsilon(E,\varepsilon)>0$ and $-C_{E\varepsilon}(E,\varepsilon)>0$. Suppose, however, that the regulator observes the upper and lower bounds on abatement costs such that $-C_E(E,\underline{\varepsilon})$ and $-C_E(E,\overline{\varepsilon})$ are known, where $\underline{\varepsilon}$ and $\overline{\varepsilon}$ are the lowest and highest values ε can take, respectively. Define E_0^* as the solution to $D'(E_0^*) = -C_E(E_0^*,\underline{\varepsilon})$, and let E_N^* be similarly defined by $D'(E_N^*) = -C_E(E_N^*,\underline{\varepsilon})$. Note that by assumption E_0^* and E_N^* are known by the regulator, so long as she can observe the damage function.

With this as setup, the Unold and Requate instrument proceeds as follows. First, the regulator formulates a menu of potential pollution levels $E_0,...,E_N$, where $E_0 = E_0^*$ and $E_N = E_N^*$. She then distributes or auctions L_0=E_0 pollution permits to the industry, and simultaneously announces a system of call options that allows firms to purchase additional permits at set prices, should they choose to do so. Denote the quantity available for each of N options by $co_1,...,co_N$, and their corresponding striking prices as $s_1,...,s_N$. Furthermore, let co_n and s_n be a priori set by the regulator so that

$$\begin{aligned} co_n &= E_n - E_{n-1} \\ s_n &= D'\left(\textstyle\sum_{k=1}^{n} co_k\right). \end{aligned} \tag{4.21}$$

Thus the striking prices are ordered such that $s_1<s_2<...<s_N$, and they serve to approximate the marginal damage curve at N discrete points in the same manner as the Roberts and Spence and Henry procedures. If the polluting firms decide to do so, they can obtain additional pollution rights beyond L_0 by exercising the options at the announced striking prices. The first option to be exercised will be co_1, since its striking price s_1 is the lowest. If there is still excess permit demand when co_1 is exhausted, firms will exercise the next lowest priced option co_2. This continues until there is no longer excess permit demand at the striking prices for the remaining options. In this sense the supply of permits is flexible, and the price that firms must pay as more permits are consumed, and correspondingly more pollution is emitted, is tied to the marginal damage curve.

This is best seen through an example. Suppose that upon the regulator's release of L_0 permits a (spot market) permit price of $\sigma(L_0)$ would emerge, absent any call options. With the call options, firms will find it optimal to exercise the first option, so long as $s_1<\sigma(L_0)$. This puts downward pressure on σ, which continues so long as there are options available that cost less than σ to exercise. This leads to an equilibrium in which σ=s_n, where s_n is the striking price for the last call option exercised by the industry. The final emission level is determined by the familiar condition $-C'(E,\varepsilon)=\sigma$, and the ultimate supply of permits at large among the regulated firms is

$$\begin{aligned} L &= L_0 + \sum_{k=1}^{n-1} co_k + \beta co_n \\ &= E, \end{aligned} \tag{4.22}$$

where β is the fraction of the quantity co_n sold at s_n. Figure 4.7 displays the intuition of this process. Here, the regulator distributes L_0 permits and makes available four types of call options with quantities co_1, co_2, co_3, co_4 and striking prices s_1, s_2, s_3, s_4. The initial reference price is $\sigma(L_0)$, but this cannot be an equilibrium since firms can exercise options and obtain pollution rights at the lower striking prices. The options with striking prices s_1 and s_2 are exercised by individual firms, so that the total emission permit availability corresponds to E_2. At a price of s_2, however, the marginal abatement cost curve shows there is still excess demand for emission allowances. This leads some firms to obtain permits through the third call option, at a price of s_3. Equilibrium is reached when some, but not all, of these options are exercised such that permits corresponding to emissions $L=\check{E}$ are held by the industry, and the permit price settles at $-C_E(\check{E},\varepsilon)=s_3=\sigma(\check{E})$. In this example the fourth call option is not exercised by the industry. Firms' behaviour in the options market results in a pollution level that minimizes ex post inefficiency, based on the size and number of the call options.

Compared to Henry's flexible supply approach, the advantage here is that, once the regulator has issued the permits and options, she does not need to take further action, since firms' decisions on exercising options determine the optimal outcome. In this sense a system of flexible supply via call options is more feasible institutionally, and one can imagine such a policy being put in place.

4.3 Mechanism design

Our discussion thus far has focused on the design of policies that minimize welfare losses arising from the regulator's uncertainty about firms' abatement cost functions. We have taken uncertainty as given and studied how taxes, permits, and hybrid instruments are best designed conditional on imperfect information. From this perspective the main objective has not been to learn about firms' abatement costs ex ante (though in some instances this information became available as the policy unfolded), and so we generally did not expect to obtain the ex post social optimum. A good policy was judged by how close it came to the social optimum, and how feasibly it could be implemented.

An alternative perspective for studying the imperfect information problem is mechanism design. Mechanism design is an area of research that grew out of principle-agent problems in which asymmetric and incomplete information among players (e.g. workers, firms, regulatory agencies) with different incentives leads to sub-optimal outcomes. The classic example is the employee/employer relationship in which workers do not always have incentive to act in the best interest of the firm. The objective in these situations is to design 'mechanisms' – contracts or regulatory schemes – that align the incentives of all players such that their self-interested actions lead to the preferred group outcome. In many instances this requires revelation of private information by the agents to the principle.

In our case polluting firms are agents who hold private information about their abatement cost functions, without which the regulator cannot design efficient environmental policy. The regulator must effectively ask firms about their abatement costs, and depending on the circumstances, firms may find it optimal to misrepresent or truthfully reveal their cost structure. Kwerel (1976) describes how this may play out in the case of simple permit or emission tax schemes. For the former, firms would like to see a generous supply of emission rights, and so they have incentive to *overstate* abatement costs. As is clear from the models developed earlier in the chapter, the regulator's resulting overestimate of marginal abatement costs causes her to issue a number of permits that is greater than the true social optimum. For tax schemes, firms would like to see an emission fee that is as low as possible, and so they have incentive to *understate* abatement costs. Through her underestimate of marginal abatement costs, the regulator sets a tax level that is too low relative to what is needed to obtain the true social optimum. Thus, in neither pure tax nor pure permit schemes do firms have incentive to fully cooperate.

To obtain the information needed to implement an efficient policy the regulator needs to set a mechanism that manipulates firms' payoff functions to be incentive compatible and individually rational. Incentive compatibility means a firm's best strategy is to reveal its true cost structure, given that other firms do as well, and individually rational means the firm will find it best to participate in the mechanism. A relatively small number of authors have considered environmental policy from the perspective of mechanism design. Examples include Kwerel (1977), Dasgupta, Hammond and Maskin (1980), and Montero (2008). In this section we briefly outline Kwerel's approach for motivation, and then describe Montero's mechanism in greater detail.

4.3.1 Kwerel's Mechanism

Kwerel (1977) proposes a hybrid instrument in which the regulator announces a quantity of permits L and a subsidy ζ, which the firm receives for each pollution permit it possess but does not use. If the market for permits is competitive, a particular firm j minimizes pollution related costs by choosing emissions and permit holdings to minimize

$$\begin{aligned} TC_j &= C_j(e_j) - \zeta(\bar{e}_j - e_j) + \sigma\bar{e}_j \\ &= C_j(e_j) + (\sigma - \zeta)\bar{e}_j + \zeta e_j, \end{aligned} \tag{4.23}$$

where as usual $C_j(\cdot)$ is the abatement cost function, $\bar{e}_j$ is the stock of permits firm j holds, and σ is the permit price. The regulator sets L and ζ based on information she receives from the firms about their abatement costs. In particular, each firm reports $-\tilde{C}_j'(e_j)$ to the regulator, which in general may not be equal to the firm's true cost function. Information from all firms in the industry is used to construct an estimated aggregate marginal abatement cost function $-\tilde{C}'(E)$, which the regulator uses to set the policy parameters L and ζ so that

$$D'(L) = \tilde{C}'(L) = \zeta. \tag{4.24}$$

Kwerel's claim is that the combination of a permit market and subsidy so structured will lead firms to report their true cost function such that $-\tilde{C}_j'(e_j) = -C_j'(e_j)$, as long as firms believe all their counterparts will also truthfully report. This allows the regulator to choose the socially optimal permit and subsidy levels $L=E^*$ and ζ^*, respectively.

Showing that this is the case is relatively straightforward, and proceeds as follows. First, if $\sigma>\zeta$ there is no advantage to holding an extra permit since the cost of doing so exceeds the revenue from returning it unused. So $e_j=\bar{e}_j$ for all firms, and the cost minimization problem in (4.23) implies that firm j chooses emissions such that $-C'_j(e_j)=\sigma$; and aggregate behaviour leads to

$-C'(L)=\sigma$. If instead $\sigma=\zeta$, equation (4.23) reduces to

$$TC_j = C_j(e_j) + \zeta e_j, \tag{4.25}$$

and individual and aggregate firm behaviour implies $-C'_j(e_j)=\zeta=-C'(L)$. Finally, the case where $\sigma<\zeta$ cannot exist in equilibrium, since the demand for permits is infinite when such arbitrage profits are available, causing the price of permits to rise to $\sigma=\zeta$. These three cases imply that the equilibrium pollution permit price under policy parameters (L, ζ) is $\sigma=\max\{\zeta, -C'(L)\}$. Since the regulator has chosen $\zeta=D'(L)$, we can further note that

$$\sigma = \max\{D'(L), -C'(L)\}. \tag{4.26}$$

Thus firms face a price determined by their true abatement cost functions and the damage function, and they know this to be the case when announcing their cost type to the regulator.

To close the argument, consider now the social optimum $L=E^*$ defined, as always, by $-C'(E^*)=D'(E^*)$. Note that for any permit level $L<E^*$ or $L>E^*$, (4.26) implies the permit price σ will be *higher* than for $L=E^*$. This follows formally from the assumptions $D''(\cdot)>0$ and $-C''(\cdot)<0$, but it is most easily seen by inspecting any of the figures shown above that contain $C'(\cdot)$, $D'(\cdot)$, and E^* (e.g. Figure 4.1 Panel A). Because firms' pollution-related costs are an increasing function of σ, costs will be at a minimum when $\sigma=D'(E^*)$. Thus under the assumptions of the model, firms can do no better than to reveal their true abatement cost type to the regulator, so the mechanism is incentive compatible. More to the point, the regulator can achieve an ex post optimal level of pollution without knowing firms' abatement cost functions ex ante.

4.3.2 Montero's Mechanism

In recent research, Montero (2008) proposes a mechanism that is more general than Kwerel's, and as such is applicable in a wider range of settings, including when the pollution permit market is not perfectly competitive. In his model the regulator first auctions a fixed number of permits based on firms' self-reported abatement cost functions, and then she reimburses firms a fraction of the revenue. The auction rules assure that the outcome – i.e. the number of permits made available – is ex post efficient. Furthermore, firms pay a net amount that is equal to their contribution to pollution damages; in this sense it is also ex post equitable from a polluter pays perspective. The combination of generality, efficiency and equity properties, and ease of implementation, makes this mechanism deserving of further study.

For intuition, we first consider the case of a single firm with abatement cost function $C(E)$, where we have dropped subscripts since the single firm is equivalent to the industry. To implement the policy, the regulator informs the firm about the following steps, which are then executed:

a) The firm submits a marginal abatement cost schedule $-\tilde{C}'(E)$, which in general need not be its true cost structure.
b) The regulator sells L permits at a price of σ, where L and σ are determined by $-\tilde{C}'(L) = D'(L) = \sigma$. The firm pays the regulator σL and receives L pollution permits.
c) The firm receives a fraction of the auction revenues $\beta(L)$, so that an amount $\beta(L)\cdot\sigma L$ is returned after the auction.

Under these rules the firm decides what form of marginal abatement cost schedule to submit in order to minimize total pollution-related costs, which are

$$TC = C(L) + \sigma L - \beta(L)\cdot\sigma\cdot L. \qquad (4.27)$$

Note that the last term in the total cost expression is the revenue returned following the permit auction, and that if we substitute out $\sigma = D'(L)$ based on the mechanism rules, the firm's objective function is

$$\min_L C(L) + D'(L)\cdot L - \beta(L)\cdot D'(L)\cdot L. \qquad (4.28)$$

Denote the solution to (4.28) – i.e. the level of pollution permits/pollution that minimizes total costs under the mechanism rules – by $\tilde{L}$, and the firm's marginal abatement cost at this level of pollution by $-C'(\tilde{L})$. Given this, the firm will announce a marginal abatement cost schedule running through $-C'(\tilde{L})$, so that at $\tilde{L}$, it is the case that the true and announced functions overlap – i.e. $-\tilde{C}'(\tilde{L}) = -C'(\tilde{L})$. The rest of the announced schedule may or may not overlap with the true schedule.

Consider now how $\tilde{L}$ is determined. Differentiating (4.28) with respect to L leads to the first order condition

$$C'(L)+D'(L)+D''(L)\cdot L \\ -\beta'(L)\cdot D'(L)\cdot L-\beta(L)\cdot\left[D''(L)\cdot L+D'(L)\right]=0. \tag{4.29}$$

The regulator's task is to announce $\beta(L)$ in a way that makes the condition in (4.29) match the condition for a social optimum, which is defined as the level of L that makes $C'(L)+D'(L)=0$. For this, $\beta(L)$ needs to be set so that it satisfies

$$\beta'(L)+\beta(L)\frac{D''(L)\cdot L+D'(L)}{D'(L)\cdot L}=\frac{D''(L)}{D'(L)}, \tag{4.30}$$

since this is the condition that needs to hold if we want (4.29) to collapse to $C'(L)+D'(L)=0$. Equation (4.30) can be viewed as a differential equation in L, which we can solve for $\beta(L)$ to obtain the correct payback function. The solution is (see Montero, p. 515)

$$\beta(L)=1-\frac{D(L)}{D'(L)\cdot L}, \tag{4.31}$$

which lies in the unit interval by the weak convexity of the damage function. If we substitute (4.31) into (4.28) we see that the firm's pollution-related total cost function is $C(L)+D(L)$, which matches the regulator's objective function. This form of $\beta(L)$ will therefore induce the firm to set its emissions to the social optimum, and announce a marginal abatement cost schedule in the auction-relevant range that allows it to do so. Furthermore, the firm bears both the cost of pollution abatement, and the damages from any remaining emissions.

Examination of the range of values β might take in special cases helps further clarify the mechanism. Note that when the damage function is linear (i.e. $D(E)=d\cdot E$) it is optimal for the regulator to set $\beta=0$. In this case the mechanism is effectively an emissions tax set at the constant marginal damage level, which from our previous analysis we know to be optimal regardless of the abatement cost function. More generally, it will be optimal for the regulator to keep a proportion, but not all, of the revenue she raises through the auction. Figure 4.8 helps show why this is so. Panel A displays the case in which $\beta=1$. If the firm truthfully reveals its marginal abatement cost curve to be $-C'(E)$, the regulator issues L^* permits at a price of σ^*. The firm pays area $a+b$ in abatement costs, and the permit expenditures $\sigma^*\cdot L^*$ are fully refunded back to the firm. If instead the firm reports its marginal abatement cost curve to be $-\tilde{C}'(E)$, the regulator auctions $\tilde{L}$ permits at a price of $\tilde{\sigma}$. The firm now pays only area a in abatement costs,

Fixed typos 8 July 2011

and once again expenditures on permits are fully refunded. Thus, when auction revenues are fully refunded, the firm spends less on regulatory compliance when it exaggerates its abatement costs. The opposite case for β=0 is shown in Panel B. If the firm truthfully reports its abatement cost function, the regulator auctions L^* permits at a price of σ^*. The firm pays area *c*+*d* in abatement costs, and spends *e*+*f*+*g*+*i*+*j*+*k*+*l* to purchase permits, none of which is refunded. If instead the firm reports $-\tilde{C}'(E)$, the regulator sets policy parameters $\tilde{L}$ and $\tilde{\sigma}$. Abatement expenditures are now (the larger) area *c*+*d*+*e*+*f*+*g*+*h*, but the payment for permits falls to area *k*+*l*. So long as *j*+*i*>*h*, the firm pays less in pollution-related costs if it underreports its abatement cost structure. These two extreme cases echo Kwerel's intuition about firms' incentives under pure price and quantity instruments, and help illustrate why it is optimal for the regulator to balance the misreporting incentives by refunding some, but not all, of the auction revenue.

Montero's mechanism for multiple firms proceeds similarly to the single firm case. The regulator informs the firms about the auction rules, and then the following steps are taken.

a) Each firm *j* submits a marginal abatement cost schedule $-\tilde{C}'_j(e_j)$.

b) For each firm *j*, the regulator sums the collection of submissions $\{-\tilde{C}'_k(e_k)\forall k \neq j\}$ to obtain $-\tilde{C}'_{-j}(E_{-j})$, which is the aggregate marginal abatement cost function absent firm *j*, where $E_{-j}=\Sigma e_k$ for all $k\neq j$.

c) For each firm *j* the regulator uses $-\tilde{C}'_{-j}(E_{-j})$ to compute a residual marginal damage function for firm *j*, defined as

$$D'_j(e_j) = D'(E) - \tilde{C}'_{-j}(E_{-j}). \tag{4.32}$$

d) The regulator clears the auction for each firm by determining the number of permits l_j and the price σ_j for each bidder according to the rule

$$-\tilde{C}'_j(l_j) = D'_j(l_j) = \sigma_j. \tag{4.33}$$

Firm *j* spends $\sigma_j \cdot l_j$ and receives l_j pollution permits.

e) Each firm *j* receives back a fraction of its expenditure on permits according to the rule

$$\beta_j(l_j) = 1 - \frac{D_j(l_j)}{D'_j(l_j)\cdot l_j}, \tag{4.34}$$

where $D_j(l_j)$ is the integral of $D'_j(l)$ between 0 and l_j. The firm receives a rebate of $\beta_j \cdot \sigma_j \cdot l_j$.

Fixed typos 8 July 2011

Montero's claim is that this mechanism is incentive compatible – i.e. firms will find it optimal to submit their true marginal abatement cost function – regardless of what other firms do. Showing this analytically involves replacing $D'(L)$ in the single firm case with $D'_j(l_j)$ for each firm in the general case, so that the objective function for firm *j* is

$$\min_{l_j} C_j(l_j) + D'_j(l_j) \cdot l_j - \beta_j(l_j) \cdot D'_j(l_j) \cdot l_j. \tag{4.35}$$

It follows analogously from the single firm derivation that plugging (4.34) into (4.35) results in total pollution related costs of $C_j(l_j)+D_j(l_j)$, and the firm will announce its marginal abatement cost function based on the condition $-C'_j(l_j)=D'_j(l_j)$. The rules of the mechanism therefore result in each firms' behaviour being determined by

$$-\tilde{C}'_j(l_j) = -C'_j(l_j) = D'_j(l_j) = \sigma_j. \tag{4.36}$$

Montero (p. 504) explains this condition as '...basically informing the firm that, whatever (abatement cost function) it chooses to submit to the regulator, that report, together with the other firms, will be used efficiently.' Importantly, equation (4.36) suggests it is optimal for firm *j* to reveal its true cost function, regardless of what other firms do. If other firms for some reason misrepresent their costs, $D'_j(\cdot)$ will be incorrect from an efficiency perspective, but firm *j* can still do no better than to announce its costs according to (4.36). The mechanism therefore eliminates the role of expectations of other firms' actions, and knowledge of competitors' cost structures, in the decisions of any single firm.

Figure 4.9 illustrates the mechanism and the efficiency consequences of its truth telling property. The curves $-C'(E)$ and $-C'_j(E_{-j})$ show the industry marginal abatement cost curves with and without firm *j*, respectively. Note that E is the efficient pollution level and σ is the corresponding efficient price, based on the intersection of $-C'(E)$ and $D'(E)$. When $D'_j(e_j)$ is the true contribution by firm *j* to marginal damages and $-C'_j(e_j)$ is its true marginal abatement cost curve, the two curves by construction must intersect at σ, and all firms face the same permit price in the regulator's auction. In this example firm *j* emits e_j units and, based on its pre-regulation emission level $\hat{e}_j$, the firm abates $\hat{e}_j - e_j$ units. Thus the firm bears abatement costs traced out by the points $e_j a \hat{e}_j$, and at the auction pays the area $\sigma a e_j 0$ in permit costs. The reimbursement mechanism, however, implies that firm *j* receives a refund that results in its total regulation-related expenditures are given by the area $ba\hat{e}_j0$. This is just equal to the sum of the firm's abatement costs, and its contributions to total damages.

4.4 Summary

Our objective in this chapter has been to introduce uncertainty in abatement cost and damage functions, and examine how this generalization changes the efficiency properties of the policy instruments introduced in the previous chapter. By limiting attention to a specific type of uncertainty – the regulator's inability to fully observe the damage function and/or firms' private abatement cost functions – three fairly robust conclusions emerge.

First, the ex post efficiency properties of emission taxes and marketable permits generally diverge when uncertainty is introduced. This is in contrast to the certainty case, for which the full menu of economic incentive approaches provided similarly efficient outcomes, and varied only in their distributional impacts. The specifics of how taxes and permits (i.e. price and quantities) diverge under uncertainty have been well established in the three plus decades since Weitzman's (1974) seminal article. When the uncertainty in the damage function and the cost function are uncorrelated it is the case that

- damage function uncertainty does not affect the relative performance of price versus quantity instruments, and
- abatement cost function uncertainty affects the relative performance based on the fact that ex post pollution levels will differ under price and quantity instruments due to the unknown position of the aggregate marginal abatement cost curve.

The preferred instrument under the latter type of uncertainty depends systematically on the curvature properties of the aggregate damage and abatement cost functions. A flat marginal damage curve relative to the marginal abatement cost favours a tax instrument, while a steep marginal damage curve favours a permit instrument, all else equal. Using linear approximations for the damage and abatement cost functions, Weitzman provides a simple formula relating the efficiency loss under both instruments to the slope parameters for the two functions, and the variance of abatement cost function. From this we learn that the efficiency loss due to uncertainty is absolutely larger when the estimates for the functions are less precise, but that relative efficiency loss between emission taxes and permits depends on the relative slopes of the marginal damage and abatement cost functions. If the (absolute value) of the slope of the marginal abatement cost curve is greater than the (absolute value) of the slope of the marginal damage curve, a tax approach minimizes the ex post welfare loss. If the opposite holds – i.e. the magnitude of the slope of the marginal damage function is greater – a permit approach is preferred.

A defining characteristic of the Weitzman theorem and related research is that ex post inefficiency remains regardless of whether the correct price or quantity instrument is chosen ex ante. Thus related research has examined the potential for hybrid instruments that mix elements of price and quantity controls to improve on the performance of either alone. Our second major conclusion regarding uncertainty emerges from our review of Spence and Robert's (1976) hybrid policy proposal, and subsequent research. Specifically, combining a permit trading scheme with floor/ceiling limits (i.e. safety valves) on the permit price guards against the extreme consequences of a poorly estimated marginal abatement cost function. If abatement costs are higher than expected by the regulator, firms can pay an emission fee to obtain pollution rights beyond the initial permit distribution. If abatement costs are lower than expected, the regulator can secure additional pollution reduction through an effective buyback of permits. In either case, the efficiency loss is smaller than would occur under a permit or tax only approach. More generally, the ex post efficiency loss can be made arbitrarily small if a flexible system of permit supply can be used to provide price incentives that trace out the marginal damage curve. Unold and Requate's (2001) proposal in which a polluting industry receives a minimum supply of permits, as well as a series of call options to obtain additional pollution rights as needed, can in principle provide the correct price signals to the industry within a realistic institutional structure.

Finally, our examination of mechanism design suggests that environmental policy can be formulated in a way that leads to ex post efficiency (as opposed to minimized efficiency loss) if polluting firms have incentive to truthfully reveal their abatement cost structure to the regulator. Montero's (2008) mechanism is the best example of how a specific set of pollution permit auction rules – in his case payment combined with partial reimbursement – can be used to obtain the first best pollution outcome under fairly general circumstances, and with favourable equity and institutional features.

4.5 Further reading

Our discussion in this chapter has followed fairly carefully the original articles on this topic. In the prices versus quantities discussion our analytical approach is drawn primarily from Weitzman (1974), though our graphical analysis has been motivated by Adar and Griffin (1976). Our discussion of hybrid policies draws on many sources, but is fundamentally based on Roberts and Spence (1977) and extensions thereof. Montero's (2008) discussion of mechanism design as related to environmental economics is particularly useful, and we have borrowed heavily from his original analysis in explaining his models.

Exercises

4.1 Our derivation of the Weitzman theorem has proceeded by assuming that the uncertainty in the abatement cost and damage functions (ε and η, respectively) are uncorrelated. If we generalize the problem and allow $COV(\varepsilon,\eta)\neq 0$, a slight different expression for Δ as shown in equation (4.17) arises. Consider the case of correlated uncertainty to

a. Derive an explicit expression for Δ in the general case. See Weitzman, p. 485, footnote 1 for guidance on how our derivation in the Appendix must be modified.

b. Describe how positive or negative correlation in damages and abatement costs augments the purely slope-based arguments from the independence case.

c. Stavins (1996) argues that the type of correlation we might expect to see in many situations will lead us to favour quantity instruments, all else equal. Summarize his arguments, and comment on the extent to which you agree.

4.2 Weitzman's research has motivated several subsequent papers examining how features of policy design in specific contexts interact with benefit or cost uncertainty to alter the original model's conclusions. Examples include Newell and Pizer (2003), Montero (2002), and Quirion (2004). Consider the following for these papers:

a. Explain in words the generalization examined relative to Weitzman's baseline analysis, and how the modelling framework is altered to accommodate the generalization.

b. Explain how the conclusions regarding the regulator's appropriate choice of a quantity or price instrument are altered by the generalization.

c. Summarize any specific examples discussed in the papers in which the generalization is thought to matter for a particular policy or environmental issue.

Appendix 4.1 Derivation of Weitzman theorem result:

Recall from (4.15) that

$$\begin{aligned} \mathrm{K}_1 &\equiv EV\left[-C_E(\tilde{E},\varepsilon)\right] \\ \Delta_1 &\equiv EV\left[D_E(\tilde{E},\eta)\right], \end{aligned} \qquad \text{(A4.1)}$$

and that for the quantity instrument the regulator chooses $\tilde{E}$ such that

$$EV\left[-C_E(\tilde{E},\varepsilon)\right] = EV\left[D_E(\tilde{E},\eta)\right], \qquad \text{(A4.2)}$$

and $\mathrm{K}_1 = \Delta_1$. Also, for subsequent use, we can obtain the variance of the marginal abatement cost function by

$$\begin{aligned} \sigma^2 &= EV\left[\left(-C_E(E,\varepsilon) - EV\left[-C_E(E,\varepsilon)\right]\right)^2\right] \\ &= EV\left[\left(-\mathrm{K}_1 - \varepsilon - \mathrm{K}_2(E-\tilde{E}) + \mathrm{K}_1 + \mathrm{K}_2(E-\tilde{E})\right)^2\right] \\ &= EV\left[(-\varepsilon)^2\right] = EV(\varepsilon^2). \end{aligned} \qquad \text{(A4.3)}$$

Recall that for the price instrument, firms' behaviour results in their equating the marginal abatement cost curve with the tax rate such that

$$\begin{aligned} \tau &= -\mathrm{K}_1 - \varepsilon - \mathrm{K}_2(E-\tilde{E}) \\ &= -\mathrm{K}_1 - \varepsilon - \mathrm{K}_2\left[E(\tau,\varepsilon) - \tilde{E}\right], \end{aligned} \qquad \text{(A4.4)}$$

where the marginal abatement cost curve is from (4.14), and we have substituted the behavioural relationship $E=E(\tau,\varepsilon)$ into the equation. Solving for $E(\tau,\varepsilon)$ we obtain

$$E(\tau,\varepsilon) = \frac{\mathrm{K}_1 - \tau + \varepsilon}{\mathrm{K}_2} + \tilde{E}. \qquad \text{(A4.5)}$$

Differentiating (A4.5) with respect to τ leads to

$$E_\tau(\tau,\varepsilon) = -\frac{1}{K_2}. \qquad \text{(A4.6)}$$

By using this expression for $h_\tau(\tau,\varepsilon)$, we can write the optimal tax rate from equation (4.11) as

$$\tilde{\tau} = \frac{EV\left[D_E\left(E(\tilde{\tau},\varepsilon),\eta\right)\right] \times (-1/\mathrm{K}_2)}{-1/\mathrm{K}_2} = EV\left[D_E\left(E(\tilde{\tau},\varepsilon),\eta\right)\right]. \qquad \text{(A4.7)}$$

Furthermore, we can substitute (A4.5) into the formula for the marginal damage function in (4.14) and take expectations to simplify the expression

$$\tilde{\tau} = \Delta_1 + \frac{\Delta_2}{\mathrm{K}_2}(\mathrm{K}_1 - \tilde{\tau}) = \Delta_1. \tag{A4.8}$$

This, along with $\mathrm{K}_1 = \Delta_1$, allows us to simplify (A4.5) to

$$E(\tau, \varepsilon) = \tilde{E} + \frac{\varepsilon}{\mathrm{K}_2}. \tag{A4.9}$$

Recall from equation (4.7) that the expected social cost from the quantity instrument is

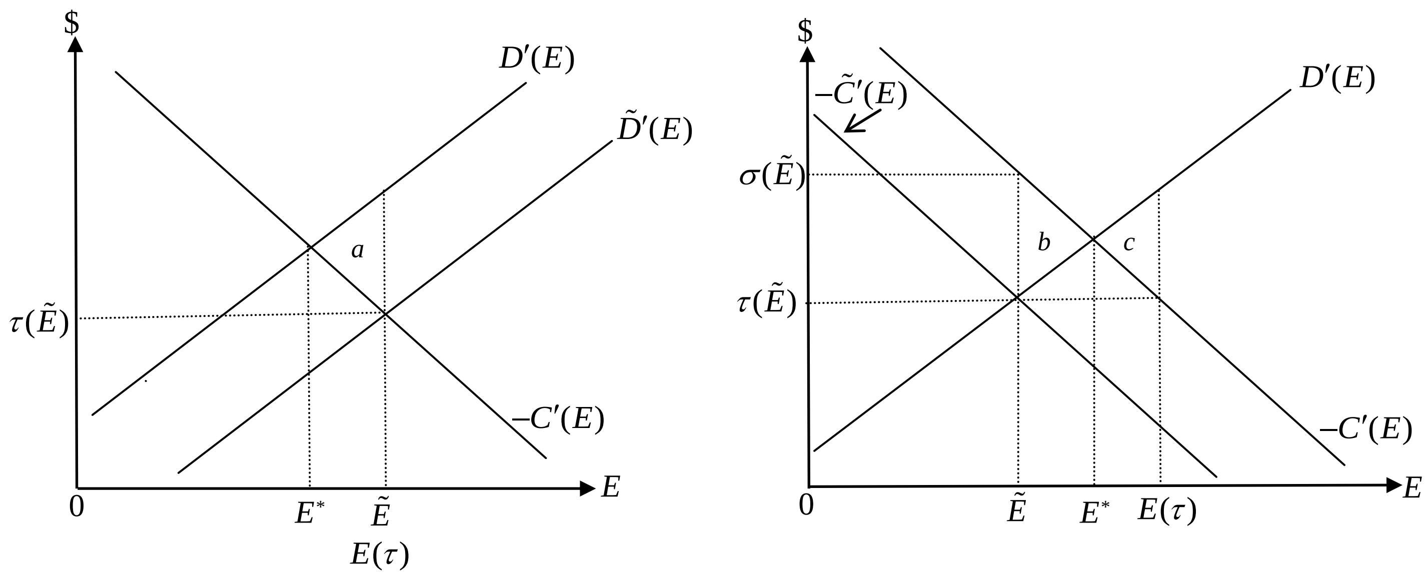
$$\begin{aligned} ESC^Q(\tilde{E}) &= EV\left[C(\tilde{E}, \varepsilon)\right] + EV\left[D(\tilde{E}, \eta)\right] \\ &= EV\left[c(\varepsilon)\right] + EV\left[d(\eta)\right] + \frac{\sigma^2}{\mathrm{K}_2}. \end{aligned} \tag{A4.10}$$

Likewise the expected social cost from the price instrument is

$$\begin{aligned} ESC^P &= EV\left[C\left(E(\tilde{\tau}, \varepsilon), \varepsilon\right)\right] + EV\left[D\left(E(\tilde{\tau}, \varepsilon), \eta\right)\right] \\ &= EV\left[c(\varepsilon) + \mathrm{K}_1\left(\frac{\varepsilon}{\mathrm{K}_2}\right) + \frac{\varepsilon^2}{\mathrm{K}_2} + \frac{\mathrm{K}_2}{2\left(\mathrm{K}_2\right)^2}\varepsilon^2\right] \\ &\quad + EV\left[d(\eta) + \Delta_1\left(\frac{\varepsilon}{\mathrm{K}_2}\right) + \frac{\eta \times \varepsilon}{\mathrm{K}_2} + \frac{\Delta_2}{2\left(\mathrm{K}_2\right)^2}\varepsilon^2\right] \\ &= EV\left[c(\varepsilon) + d(\eta)\right] + \frac{\sigma^2}{\mathrm{K}_2} + \frac{\sigma^2}{2\left(\mathrm{K}_2\right)^2} + \frac{\Delta_2\sigma^2}{2\left(\mathrm{K}_2\right)^2}, \end{aligned} \tag{A4.11}$$

where the last step uses that fact that ε and η are independent so that $EV(\varepsilon \times \eta) = 0$, and $EV(\varepsilon^2) = \sigma^2$ is the variance of ε. Taking the difference $ESC^P - ESP^Q$ matches equation (4.17).

Figure 4.1: Damage and Abatement Cost Curve Uncertainty

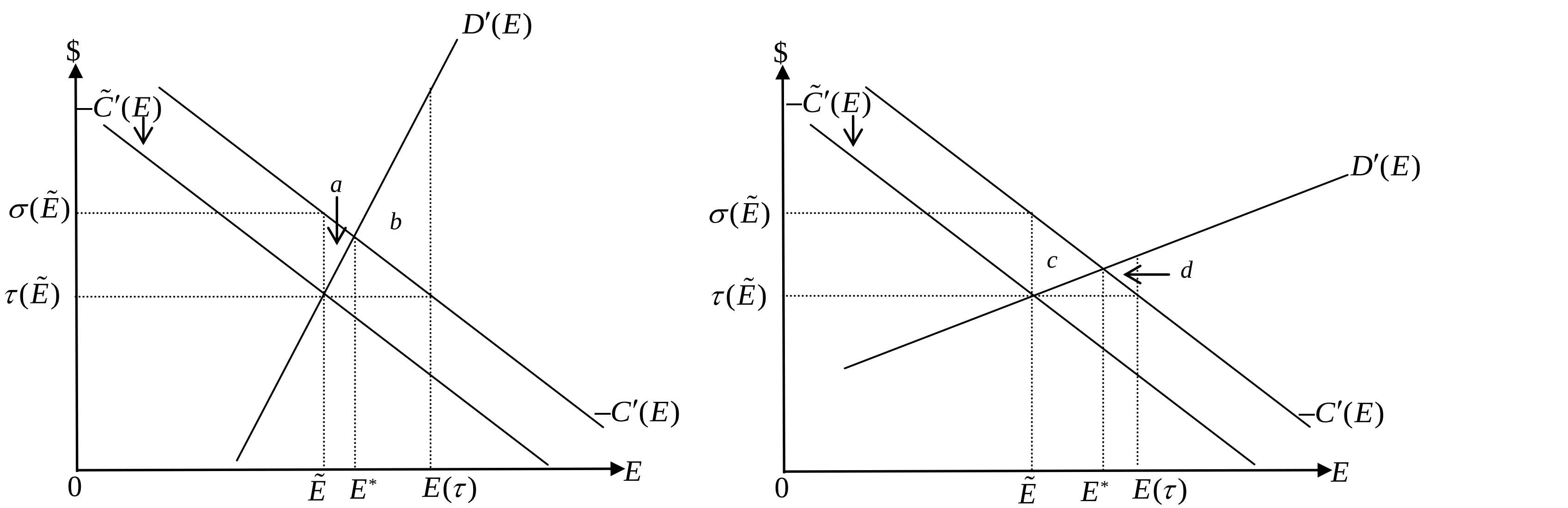


Panel A

Panel B

Fixed typos 8 July 2011

Figure 4.2: Welfare Loss with Steep and Flat Marginal Damage Curves

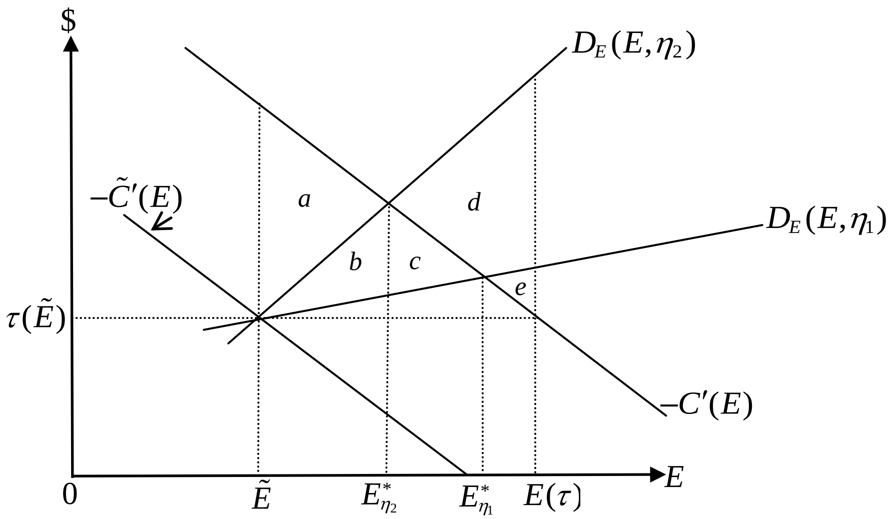


Panel A

Panel B

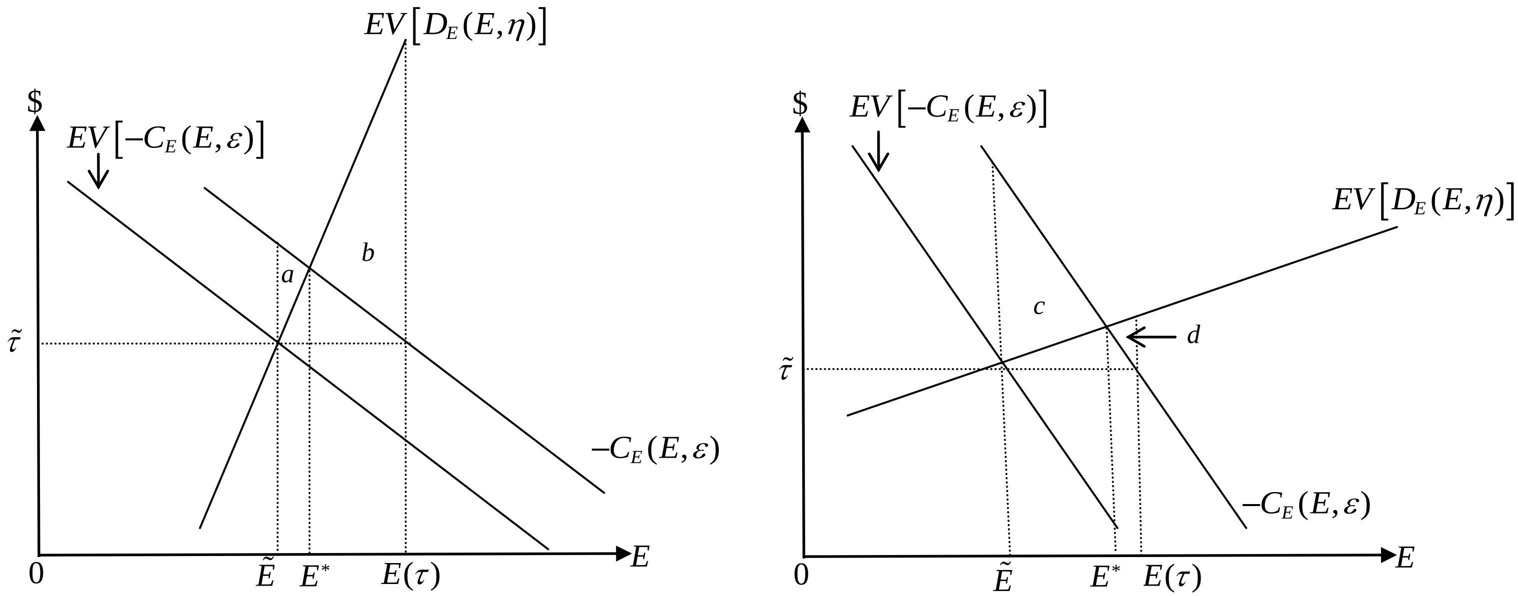
Fixed typos 8 July 2011

Figure 4.3: Illustration of Proposition 4.1



Fixed typos 8 July 2011

Figure 4.4: Illustration of the Weitzman Theorem

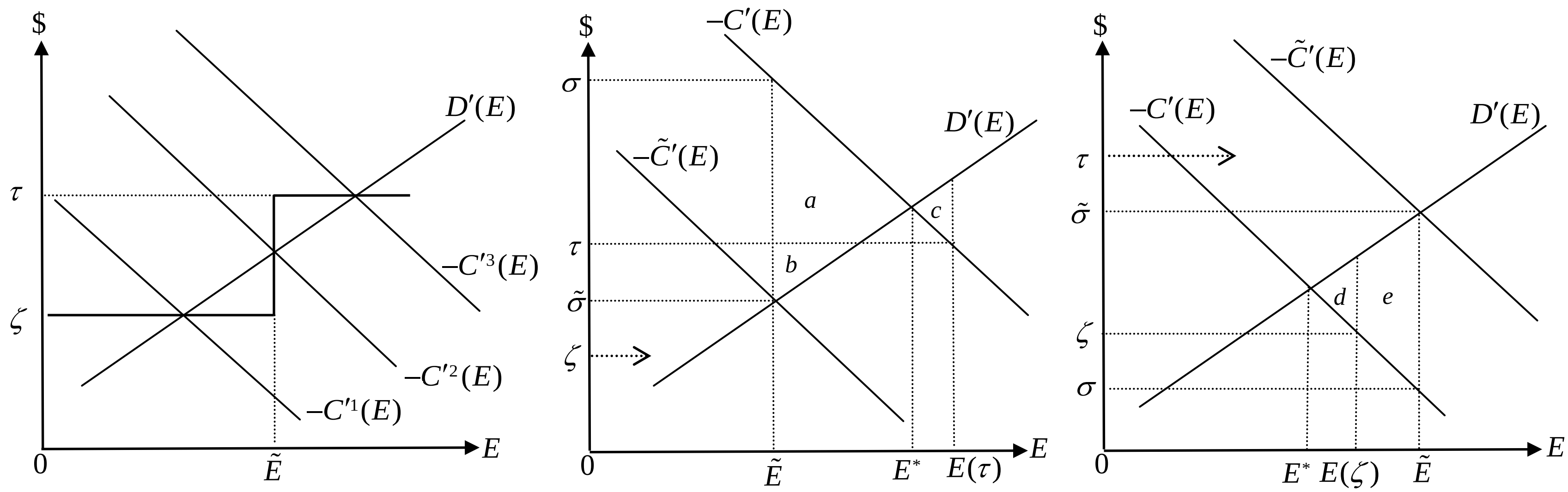


Panel A

Panel B

Fixed typos 8 July 2011

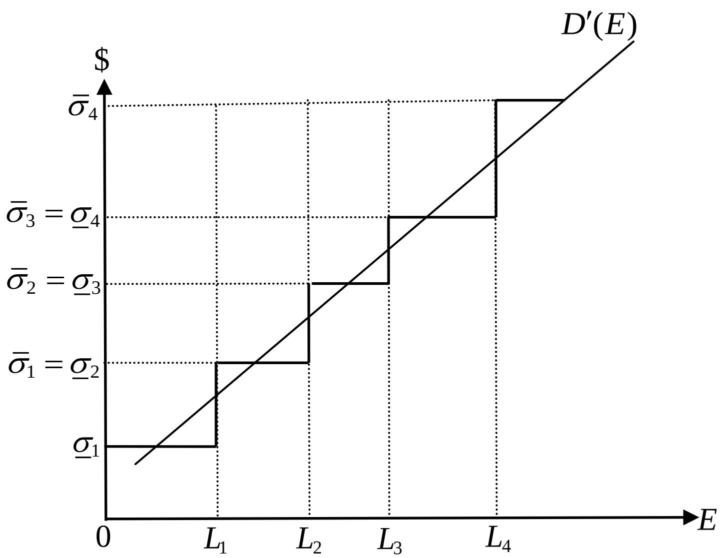
Figure 4.5: Roberts and Spence Instrument



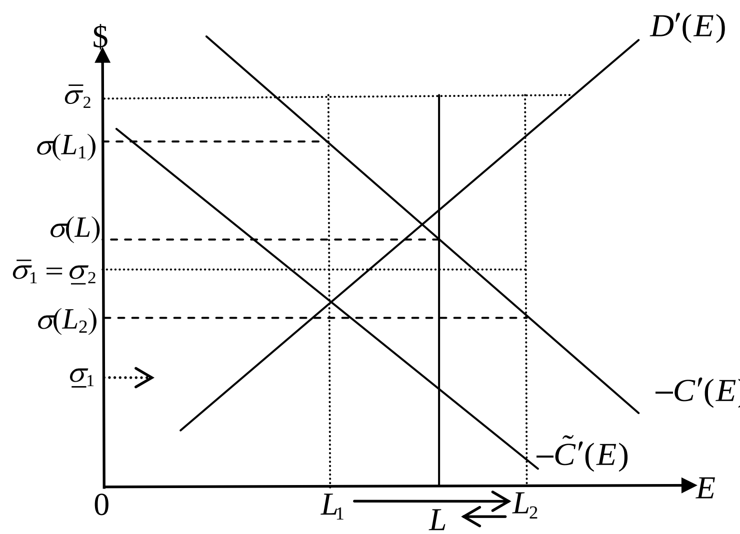
Panel A Panel B Panel C

Fixed typos 8 July 2011

Figure 4.6: Henry's Instrument

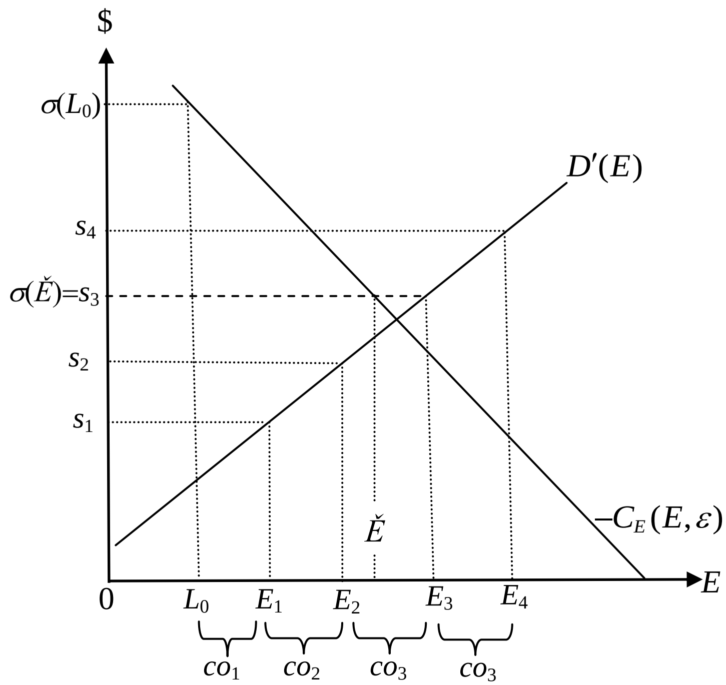


Panel A



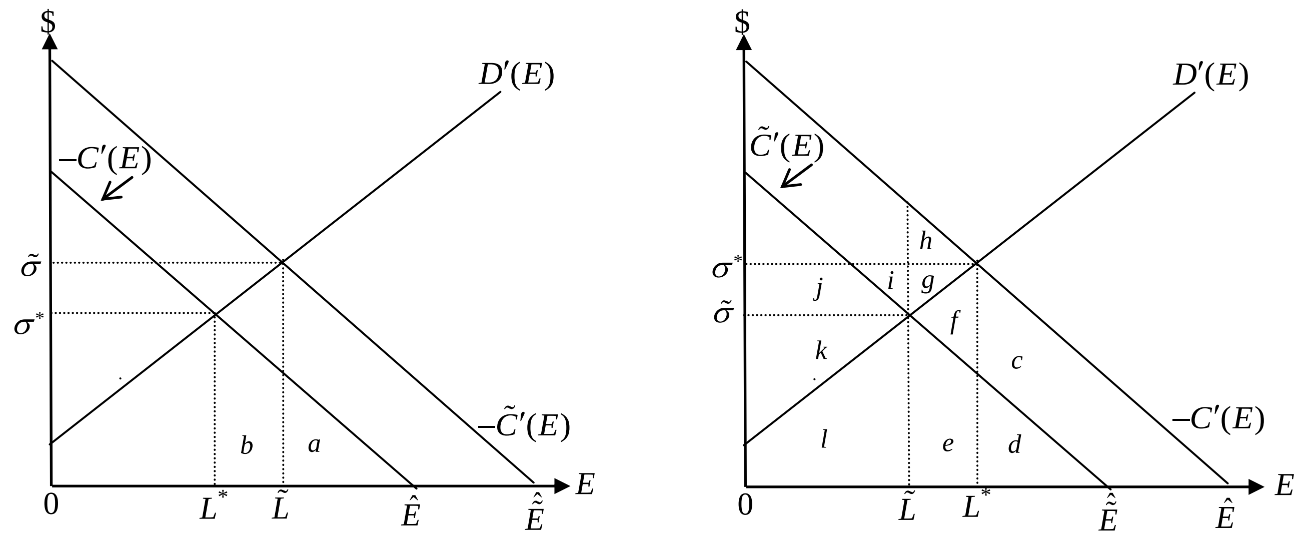
Panel B

Figure 4.7: Unold and Requate Instrument



Fixed typos 8 July 2011

Figure 4.8: Montero Mechanism with β=0 or β=1

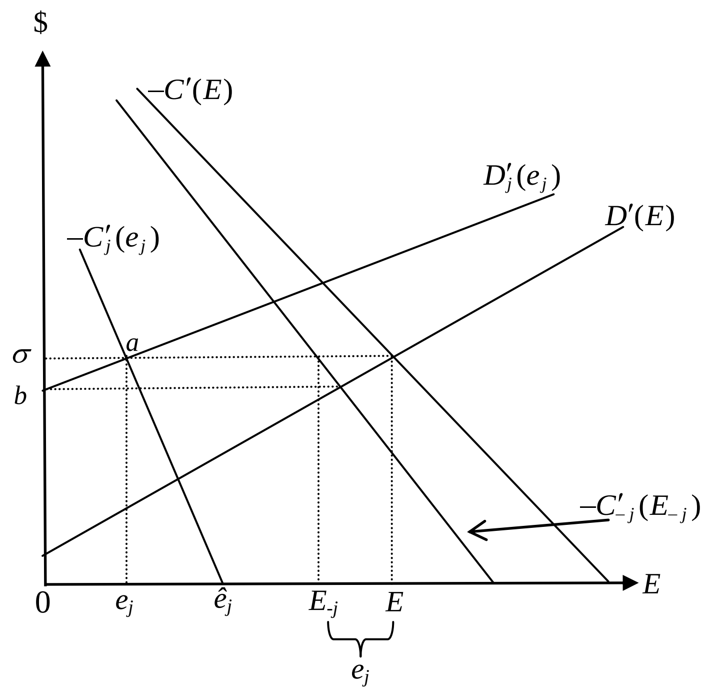


Panel A

Panel B

Fixed typos 8 July 2011

Figure 4.9: Montero Mechanism for firm *j*



Fixed typos 8 July 2011