Supplemental Appendix: "The Informational Role of Emission Markets: Prices vs Quantities with Dispersed Information About Externalities", Derek Lemoine

Appendix A compares the results to prior work. Appendix B develops graphical intuition. Appendix C contains proofs.

A Comparison to Pizer and Prest (2020)

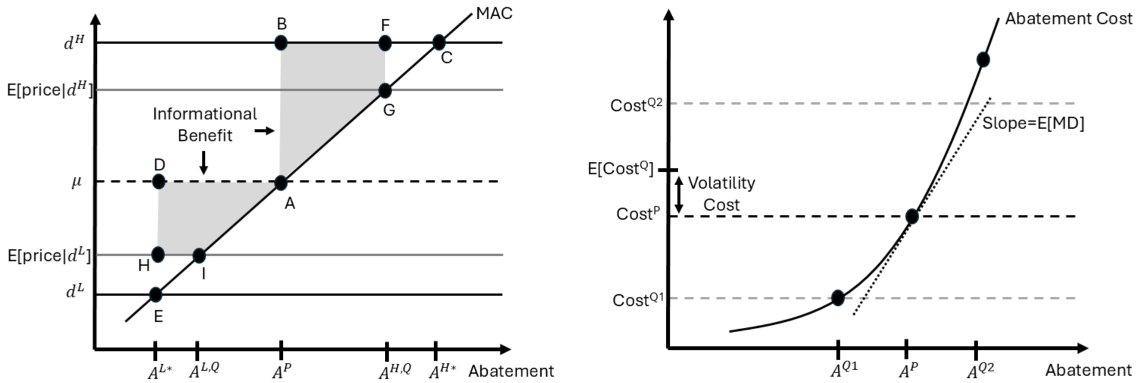
How does equation (5) relate to the most closely related work, Pizer and Prest (2020)? There, private agents learn true social cost d perfectly in period 1, as opposed to needing to learn about d from noisy private signals and equilibrium permit prices. The private signals here become perfectly informative—mimicking their setting—as ν^2 becomes small. In that case, Lemma 1 shows that V approaches zero and the appendix shows that b approaches zero while $b + \beta$ approaches 1.[11] If, in addition, we match the setting of Pizer and Prest (2020) by zeroing out noise traders, then $\lim_{\Theta^2 \to 0} \lim_{\nu^2 \to 0} \Delta = -\frac{\sigma^2}{2\kappa}$. The right-hand side matches the right-hand side of equation (9) in Pizer and Prest (2020) in the special case of a flat marginal benefit curve and known marginal abatement costs. We have recovered their result that a quantity policy always dominates a price policy when private agents learn the truth perfectly.

In their appendix, Pizer and Prest (2020) consider a model in which private agents learn d plus a noise shock. The expression in (5) does not reduce to their expression (which is equivalent to $-\frac{\sigma^2}{2\kappa} + \frac{\nu^2}{\kappa}$) for any combination of parameters. The noise ν^2 favors a price policy less strongly here than in Pizer and Prest (2020).[12] Here learning is an equilibrium process that depends on how emission permit prices encode the private signals that agents throughout the economy receive. When private signals are noisy, private agents learn from the permit market, which introduces terms that are absent from Pizer and Prest (2020). Because the permit market aggregates information, agents' ability to turn to it makes the noise in private signals less costly for a quantity policy than in Pizer and Prest (2020).

In the expressions from Pizer and Prest (2020) given above, Δ becomes arbitrarily large in magnitude as the slope κ of marginal abatement cost becomes small, but here Δ can remain bounded as κ becomes small. The reason is that here traders' risk aversion α and posterior uncertainty V combine to create downward-sloping permit demand even when firms' marginal abatement costs are completely flat. That is why here we need ν^2/σ^2 to also become small for Δ to become unbounded below (and even then Δ is here unbounded only

[11] These limits reflect the Grossman-Stiglitz paradox: from the proof of Lemma 1, prices become perfectly revealing despite every agent ignoring their own private signals.

[12] We can rewrite that term as $\frac{1}{2} \frac{N\kappa}{N\kappa + (1+r)\alpha V} (b+\beta)^2 \frac{\nu^2}{N\kappa + (1+r)\alpha V}$. Comparing to the ν^2/κ from Pizer and Prest (2020), the denominator on ν^2 is here larger and the other terms are all between 0 and 1.



(a) Informational Benefit of a Cap

(b) Volatility Cost of a Cap

Figure A-1: Illustrating the informational benefit and volatility cost of a quantity policy (i.e., a cap), relative to a price policy (i.e., an emission tax).

below, not above).

B Graphical Intuition

Figure A-1 provides graphical intuition for the informational benefit (left panel) and volatility cost (right panel) of a cap.

In the left panel, the upward-sloping black line is marginal abatement cost (MAC), with slope κ. The horizontal solid black lines, labeled d^H and d^L, illustrate an equiprobable two-point distribution over marginal damage d. The dashed black line is μ, the prior mean of marginal damage.

A price policy sets the period 1 emission tax to μ. Equilibrium abatement in period 1 is A^P. If the true d is d^H, then socially optimal abatement is A^{H*} and the deadweight loss is triangle ABC. If the true d is d^L, then socially optimal abatement is A^{L*} and the deadweight loss is triangle ADE.

Under a quantity policy, private agents receive noisy signals of d and condition their demand for emission permits on that information. If the true d is d^H, then, on average, the equilibrium permit price may be at a line like the upper solid gray line, labeled $E[price|d^H]$. It moves partway, but not all the way, towards the upper solid black line from the dashed black line. Equilibrium abatement is $A^{H,Q}$. There is still a deadweight loss, in triangle CFG. Relative to a price policy, the quantity policy reduces deadweight loss by the difference between triangles ABC and CFG, which is the gray polygon ABFG. If the true d is d^L, then the reduction in deadweight loss relative to a price policy is the gray polygon ADHI. The informational benefit of a cap is the average of the two gray polygons. The benefit is larger when σ^2 is larger because there is more to learn, with d^H and d^L further from μ.

However, the equilibrium permit price is itself noisy, even conditional on d. The right panel depicts this noise for an illustrative case in which, ex post, $d = \mu$. The curved solid line depicts total abatement cost. The price policy sets abatement A^P to the point whose tangent line (in light dots) has slope equal to expected marginal damage (i.e., has slope μ), for total cost $Cost^P$. Abatement under the quantity policy is noisy due to the ϵ_i and θ. A^{Q1} and A^{Q2} here illustrate two equiprobable outcomes, which correspond to costs $Cost^{Q1}$ and $Cost^{Q2}$. Due to Jensen's inequality, expected cost under the quantity policy ($E[Cost^Q]$) is greater than the cost at expected abatement ($Cost^P$). The volatility cost of the cap is that additional cost, indicated by the double arrow. The volatility cost is larger when the abatement cost curve is more convex (i.e., when κ is larger) and the permit price under the cap is noisier (i.e., when ν^2 and Θ^2 are larger).

Which instrument is preferred depends on whether the informational benefit of a cap dominates the volatility cost of a cap. Because the volatility cost of a cap vanishes as κ becomes small, the quantity policy always dominates the price policy for small κ. If the volatility cost of a cap ever dominates its informational benefit, then it must do so for large κ, where Jensen's inequality has a lot of bite. However, if σ^2 is large, then the informational benefit can be so large that the quantity policy dominates for all κ. Alternately, if ν^2 is small, abatement is not volatile because traders do not need to learn from the permit price, in which case the quantity policy again dominates for all κ.

C Proofs

C.1 Derivation of Equation (2)

Problem (1) is equivalent to

$$\max_{X_i} - \exp\left\{-\alpha\left((1+r)(w_i - X_i p_1) + (1+r)X_i E[p_2|s_i,p_1]\right) + \frac{1}{2}[\alpha(1+r)X_i]^2 Var[p_2|s_i,p_1]\right\}.$$

The first-order condition is:

$$X_i = \frac{E[p_2|s_i,p_1] - p_1}{\alpha(1+r)Var[p_2|s_i,p_1]}.$$

By normal-normal updating, $Var[p_2|s_i,p_1]$ is independent of the value of s_i. Then, because of identical precisions and priors across types i, $Var[p_2|s_i,p_1]$ is independent of i. Denote that constant, exogenous variance as V. We then have equation (2).

C.2 Proof of Lemma 1

Conjecture that

$$E[p_2|s_i,p_1] = a\,\mu + b\,(s_i - \mu) + B\,(p_1 - E[p_1]). \tag{A-1}$$

The coefficients do not vary with i because traders are symmetric ex ante. Substituting from (2) into period 1 market-clearing and using this conjecture and A_1^*,

$$Q_1 = Z - \frac{p_1 - \chi}{\kappa} + \theta + \sum_i \frac{a\,\mu + b\,(s_i - \mu) + B\,(p_1 - E[p_1]) - p_1}{(1+r)\alpha V}.$$

Rearrange to solve for p_1:

$$p_1 = \frac{(1+r)\alpha V\,[\kappa(Q_1 - Z - \theta) - \chi] - N\kappa a\mu - \kappa b \sum_i (s_i - \mu) + N\kappa B E[p_1]}{N\kappa[B-1] - (1+r)\alpha V}. \qquad \text{(A-2)}$$

Under the common prior, the unconditional mean is

$$E[p_1] = \frac{(1+r)\alpha V\,[\kappa(Z - Q_1) + \chi] + N\kappa a\mu}{N\kappa + (1+r)\alpha V} \qquad \text{(A-3)}$$

and the unconditional variance is

$$Var[p_1] = \frac{\kappa^2 b^2 (N^2\sigma^2 + N\nu^2) + [(1+r)\alpha V\kappa]^2\Theta^2}{\left[N\kappa[B-1] - (1+r)\alpha V\right]^2}. \qquad \text{(A-4)}$$

Using $E[p_1]$ in (A-2), we find:

$$p_1 = \frac{-(1+r)\alpha V\kappa\theta - \kappa b \sum_i (s_i - \mu)}{N\kappa[B-1] - (1+r)\alpha V} + \frac{(1+r)\alpha V\,[\kappa(Q_1 - Z) - \chi] - N\kappa a\mu}{-N\kappa - (1+r)\alpha V}. \qquad \text{(A-5)}$$

Create the 3×1 random vector

$$S_i = \begin{bmatrix} p_2 \\ p_1 \\ s_i \end{bmatrix},$$

where $p_2 = d$. This random vector is multivariate normal. Using (A-3), it has unconditional mean

$$E[S_i] = \begin{bmatrix} \mu \\ \frac{(1+r)\alpha V[\kappa(Q_1 - Z) - \chi] - N\kappa a\mu}{-N\kappa - (1+r)\alpha V} \\ \mu \end{bmatrix},$$

and using (A-5), it has covariance matrix

$$Cov[S_i] = \begin{bmatrix} \sigma^2 & -N\sigma^2 \frac{\kappa b}{N\kappa[B-1]-(1+r)\alpha V} & \sigma^2 \\ -N\sigma^2 \frac{\kappa b}{N\kappa[B-1]-(1+r)\alpha V} & Var[p_1] & \frac{-\kappa b}{N\kappa[B-1]-(1+r)\alpha V}(N\sigma^2+\nu^2) \\ \sigma^2 & \frac{-\kappa b}{N\kappa[B-1]-(1+r)\alpha V}(N\sigma^2+\nu^2) & \sigma^2+\nu^2 \end{bmatrix}.$$

From the projection theorem and (A-4),

$$\begin{aligned} E[p_2|p_1, s_i] =&\mu + \frac{-(N-1)\nu^2\sigma^2\kappa b\big[N\kappa[B-1]-(1+r)\alpha V\big]}{\kappa^2 b^2(N\sigma^2+\nu^2)(N-1)\nu^2+(\sigma^2+\nu^2)[(1+r)\alpha V\kappa]^2\Theta^2}(p_1 - E[p_1]) \\ &+ \frac{\sigma^2[(1+r)\alpha V\kappa]^2\Theta^2}{\kappa^2 b^2(N\sigma^2+\nu^2)(N-1)\nu^2+(\sigma^2+\nu^2)[(1+r)\alpha V\kappa]^2\Theta^2}(s_i-\mu). \end{aligned}$$

Matching coefficients to (A-1),

$$B = \frac{(N-1)\nu^2\sigma^2\kappa b}{\kappa^2 b^2(N\sigma^2+\nu^2)(N-1)\nu^2+(\sigma^2+\nu^2)[(1+r)\alpha V\kappa]^2\Theta^2}\big[N\kappa[1-B]+(1+r)\alpha V\big], \tag{A-6}$$

$$b = \frac{\sigma^2[(1+r)\alpha V]^2\Theta^2}{\sigma^2[(1+r)\alpha V]^2\Theta^2+\big[(N-1)(N\sigma^2+\nu^2)b^2+[(1+r)\alpha V]^2\Theta^2\big]\nu^2} \in (0,1), \tag{A-7}$$

$$a = 1. \tag{A-8}$$

Note that, from (A-6):

$$Nb\kappa B = \beta\big[N\kappa[1-B]+(1+r)\alpha V\big] \tag{A-9}$$

for

$$\beta \triangleq \frac{N(N-1)\nu^2\sigma^2 b^2}{b^2(N\sigma^2+\nu^2)(N-1)\nu^2+(\sigma^2+\nu^2)[(1+r)\alpha V]^2\Theta^2} \in (0,1). \tag{A-10}$$

Observe that $\beta + b \in (0,1)$. Solving for B in (A-9),

$$B = \frac{\beta}{b+\beta}\frac{N\kappa+(1+r)\alpha V}{N\kappa}.$$

Using that,

$$\big[N\kappa[1-B]+(1+r)\alpha V\big] = \frac{b}{\beta+b}\left[N\kappa+(1+r)\alpha V\right]. \tag{A-11}$$

Substitute from (A-11) and (A-8) into (A-5):

$$p_1 = \frac{N\kappa}{N\kappa+(1+r)\alpha V}\left[\mu + \frac{b+\beta}{N}\sum_i(s_i-\mu)\right] + \frac{(1+r)\alpha V}{N\kappa+(1+r)\alpha V}\left[\frac{b+\beta}{b}\kappa\theta+\chi-\kappa(Q_1-Z)\right]. \tag{A-12}$$

We saw in the main text that the regulator sets the period 1 cap Q_1 so that $E[p_1] = \mu$. Use (A-3) to solve for the Q_1 that sets $E[p_1] = \mu$, then substitute into (A-12) to obtain equation (3).

To finish proving the lemma, it remains to establish bounds on V, independence from κ, and existence of b. From the random vector S_i, the definition of V as the posterior variance of p_2, and the projection theorem,

$$V = \frac{(N-1)\nu^2 + \left(\frac{(1+r)\alpha V}{b}\right)^2 \Theta^2}{(N\sigma^2+\nu^2)(N-1)\nu^2 + (\sigma^2+\nu^2)\left(\frac{(1+r)\alpha V}{b}\right)^2 \Theta^2}\nu^2\sigma^2. \tag{A-13}$$

From (A-7):

$$\left(\frac{(1+r)\alpha V}{b}\right)^2 \Theta^2 = \frac{(N\sigma^2+\nu^2)(N-1)\nu^2 b}{\sigma^2-(\sigma^2+\nu^2)b}.$$

Substitute into (A-13):

$$V = \frac{\sigma^2\nu^2}{N\sigma^2+\nu^2}[1+(N-1)b]. \tag{A-14}$$

The bounds on V follow from $b \in (0,1)$ and V being monotonic in b.

Substitute V from (A-14) into (A-7) and simplify:

$$b = \frac{\sigma^2\nu^2\left((1+r)\alpha\frac{\sigma^2}{N\sigma^2+\nu^2}\right)^2 \Theta^2[1+(N-1)b]^2}{b^2(N\sigma^2+\nu^2)(N-1) + \nu^2(\sigma^2+\nu^2)\left((1+r)\alpha\frac{\sigma^2}{N\sigma^2+\nu^2}\right)^2 \Theta^2[1+(N-1)b]^2}. \tag{A-15}$$

The left-hand side of (A-15) is just b and the right-hand side is in $(0,1)$. The derivative of the right-hand side with respect to b is proportional to $-2b(N\sigma^2+\nu^2)(N-1)$. So the right-hand side decreases in b for $b \in (0,1)$. The limit of the right-hand side as b goes to 0 from above is $\frac{\sigma^2}{\sigma^2+\nu^2}$, which is greater than the left-hand side at $b=0$. The limit of the right-hand side as b goes to 1 from below is strictly less than 1 and thus is less than the left-hand side. Therefore there is one and only one intersection between the left- and right-hand sides, so the b that solves (A-15) is unique.

Independence of b, β, and V from κ follows by inspection of (A-15), (A-10), and (A-14). □

C.3 Proof of Proposition 1

In period 2, A_2 is the same under a tax or a cap. So $\Delta = E\big[TC_1^Q - TC_1^P\big]$. Under the price instrument, abatement in period 1 is not stochastic: $A_1^P = (\mu-\chi)/\kappa$ and $E[TC_1^P] = C(A_1^P) + \mu\{Z - A_1^P\}$. Second-order approximate $C(A_1^Q)$ around $A_1^Q = A_1^P$ and substitute into (4) (and observe that the approximation is exact under the given functional forms):

$$E[TC_1^Q] = E[TC_1^P] + E\left[C'(A_1^P)\left[A_1^Q - A_1^P\right] + \frac{1}{2}C''(A_1^P)\left[A_1^Q - A_1^P\right]^2\right] - E\left[d[A_1^Q - A_1^P]\right]. \tag{A-16}$$

Using (3) in A_1^* under a cap,

$$A_1^Q = \frac{\mu - \chi}{\kappa} + \frac{N(b+\beta)}{N\kappa + (1+r)\alpha V}\left[d - \mu + \frac{1}{N}\sum_i \epsilon_i\right] + \frac{(1+r)\alpha V}{N\kappa + (1+r)\alpha V}\frac{b+\beta}{b}\theta$$

and thus:

$$\begin{aligned}
E\left[A_1^Q\right] &= \frac{\mu - \chi}{\kappa} = A_1^P, \\
E\left[A_1^Q - A_1^P\right]^2 &= Var[A_1^Q] = \left(\frac{b+\beta}{N\kappa + (1+r)\alpha V}\right)^2 \left[\frac{[(1+r)\alpha V]^2}{b^2}\Theta^2 + N^2\sigma^2 + N\nu^2\right], \\
E\left[d\{A_1^Q - A_1^P\}\right] &= Cov[d, A_1^Q] = \frac{b+\beta}{N\kappa + (1+r)\alpha V}N\sigma^2.
\end{aligned}$$

The proposition follows from substituting into (A-16) and then substituting that result into $\Delta = E\left[TC_1^Q - TC_1^P\right]$.

□

C.4 Proof of Proposition 2

Part 1 of the proposition follows from inspection of (5), recalling from Proposition 1 that b, β, and V are independent of κ.

Rewrite (5) as:

$$\Delta = -\frac{b+\beta}{N\kappa + (1+r)\alpha V}\left\{N\sigma^2 - \frac{1}{2}\kappa\left(\frac{b+\beta}{N\kappa + (1+r)\alpha V}\right)\left[\frac{[(1+r)\alpha V]^2}{b^2}\Theta^2 + N^2\sigma^2 + N\nu^2\right]\right\}. \tag{A-17}$$

For Δ to be strictly positive at some κ, we need the expression in curly braces to hit zero at some $\hat{\kappa} > 0$. Again using that b, β, and V are independent of κ, we have the unique $\hat{\kappa}$:

$$\hat{\kappa} = \frac{N\sigma^2(1+r)\alpha V}{\frac{1}{2}(b+\beta)\left[\frac{[(1+r)\alpha V]^2}{b^2}\Theta^2 + N^2\sigma^2 + N\nu^2\right] - N^2\sigma^2}.$$

For $\hat{\kappa}$ to be strictly positive, the denominator must be strictly positive, which in turn means, using (A-14),

$$\frac{1}{2}(b+\beta)\left[\left(\frac{(1+r)\alpha\sigma^2\nu^2[1+(N-1)b]}{N(\sigma^2+\nu^2)b}\right)^2\Theta^2+N\nu^2\right]>\left[1-\frac{1}{2}(b+\beta)\right]N^2\sigma^2. \quad \text{(A-18)}$$

From (A-10), (A-14), and (A-15) we have:

$$b+\beta=\frac{N(N-1)\sigma^2b^2+\sigma^2\nu^2\left((1+r)\alpha\frac{\sigma^2}{N\sigma^2+\nu^2}\right)^2\Theta^2[1+(N-1)b]^2}{b^2(N\sigma^2+\nu^2)(N-1)+\nu^2(\sigma^2+\nu^2)\left((1+r)\alpha\frac{\sigma^2}{N\sigma^2+\nu^2}\right)^2\Theta^2[1+(N-1)b]^2}. \quad \text{(A-19)}$$

We see that $\lim_{\sigma^2\to\infty}(b+\beta)=1$. The right-hand side of inequality (A-18) approaches infinity as σ^2 approaches infinity, in which case the inequality does not hold. From (A-15),

$$\lim_{\nu^2\to 0} b=\frac{\sigma^2\left((1+r)\alpha\frac{1}{N}\right)^2\Theta^2\lim_{\nu^2\to 0}\left(\nu^2[1+(N-1)b]^2\right)}{N\sigma^2(N-1)\lim_{\nu^2\to 0}b^2+\sigma^2\left((1+r)\alpha\frac{1}{N}\right)^2\Theta^2\lim_{\nu^2\to 0}\left(\nu^2[1+(N-1)b]^2\right)}.$$

We have a contradiction if we assume $\lim_{\nu^2\to 0} b>0$, so the limit must be zero, which can be the case only if b^2 goes to zero more slowly than does ν^2. In that case, $\lim_{\nu^2\to 0}(b+\beta)=1$. The left-hand side of inequality (A-18) approaches zero as ν^2 approaches 0, and the right-hand side of inequality (A-18) approaches $0.5N^2\sigma^2$. Inequality (A-18) then does not hold. We have established the second part of the proposition.

If $\hat{\kappa}>0$, then, from (A-17) and b, β, and V being independent of κ, $\Delta>0$ for all $\kappa>\hat{\kappa}$ and $\Delta\leq 0$ for all $\kappa\leq\hat{\kappa}$. We have established the third part of the proposition.

As κ becomes large, Δ approaches 0, from (A-17) and b, β, and V being independent of κ. From (A-15),

$$\lim_{\sigma^2\to 0} b=\frac{\nu^2\left((1+r)\alpha\frac{1}{\nu^2}\right)^2\Theta^2\lim_{\sigma^2\to 0}\left([\sigma^2]^3[1+(N-1)b]^2\right)}{\nu^2(N-1)\lim_{\sigma^2\to 0}b^2+[\nu^2]^2\left((1+r)\alpha\frac{1}{\nu^2}\right)^2\Theta^2\lim_{\sigma^2\to 0}\left([\sigma^2]^2[1+(N-1)b]^2\right)}.$$

We have a contradiction if we assume $\lim_{\sigma^2\to 0} b>0$, so the limit must be zero, which can be the case only if b^2 goes to zero more slowly than does $[\sigma^2]^3$. Then, from (A-19), $\lim_{\sigma^2\to 0}(\beta+b)=0$, because $[\sigma^2]^3$ in the numerator goes to zero faster than does $[\sigma^2]^2$ in the denominator and σ^2b^2 in the numerator goes to zero faster than does b^2 in the denominator. Then, from (A-17), Δ goes to 0 as σ^2 goes to 0. We have established the fourth part of the proposition.

If $\Delta>0$, then boundedness of $b+\beta$ and V and $\lim_{\kappa\to\infty}\Delta=0$ imply with (A-17) that Δ is finite. Using b, β, and V independent of κ, Δ is minimized as κ becomes small. Using (A-14),

$$\lim_{\kappa\to 0}\Delta=-N\left(\frac{N\sigma^2}{\nu^2}+1\right)\frac{b+\beta}{(1+r)\alpha[1+(N-1)b]}. \quad \text{(A-20)}$$

We saw above that $\lim_{\sigma^2\to\infty}(b+\beta)=1$ and $\lim_{\nu^2\to 0}(b+\beta)=1$. Therefore the right-hand side of (A-20) approaches negative infinity as σ^2/ν^2 approaches infinity. We have established the final part of the proposition.

□