Emissions trading with profit-neutral permit allocations

Cameron Hepburn,† John K.-H. Quah,‡ and Robert A. Ritz*

Abstract: This paper examines the operation of an emissions trading scheme (ETS) in a Cournot oligopoly. We study the impact of the ETS on industry output, price, costs, emissions, and profits. In particular, we develop formulae for the number of emissions permits that have to be freely allocated to firms in order to neutralize any adverse impact the ETS may have on profits. These formulae tell us that the profit impact of the ETS is usually limited. Indeed, under quite general conditions, industry profits are preserved so long as firms are freely allocated a *fraction* of their total demand for permits, with this fraction being lower than the industry's Herfindahl index.

JEL Classification Numbers: D43, H23, Q58

Keywords: Emissions trading, permit allocation, profit-neutrality, cost pass-through, abatement, grandfathering

---

† School of Enterprise and the Environment, Oxford University. E-mail: cameron.hepburn@economics.ox.ac.uk.

‡ Department of Economics, Oxford University. E-mail: john.quah@economics.ox.ac.uk.

* Department of Economics, Oxford University. Email: robert.ritz@economics.ox.ac.uk

We thank Martin Browning, Simon Cowan, Larry Goulder, Ian Jewitt, Paul Klemperer, Robin Mason, Meg Meyer, Peter Neary, Robin Smale, Bruno Strolovici, and John Vickers for helpful discussions. We are also grateful to seminar and conference audiences at Oxford, Royal Holloway London, RES 2007 (Warwick), EEA 2007 (Budapest), IIOC 2008 (Washington DC), EAERE 2008 (Gothenburg), PET 2008 (Seoul) for valuable comments and feedback.

1 Introduction

There is increasingly broad recognition that greenhouse gas emissions are contributing to changes to Earth's climate and that reducing emissions constitutes an important challenge of economic policy (Stern, 2008). Emerging trading schemes for CO_2 and other greenhouse gases can draw upon considerable experience from other environmental markets, including the acid rain markets in sulfur dioxide (SO_2) and nitrogen oxides (NO_x) created by the 1990 Clean Air Act Amendments,[1] and other trading schemes for water and fishery rights.

The intellectual justification for economic instruments, such as emissions trading and emission taxes, arises from the observation that, under certain assumptions, imposing a common price on emissions equalizes marginal abatement costs across the polluting firms and minimizes the aggregate cost of pollution control.[2] In most cases, this makes economic instruments more efficient than "command-and-control" intervention, which specifies input or output standards or technologies.[3] Cost efficiency is particularly important given the scale of the challenge of reducing greenhouse gas emissions; they are embedded in all aspects of production and consumption. However, there is a significant disadvantage to trading and taxes: while "command-and-control" intervention may impose higher marginal abatement costs, economic instruments tend to create inframarginal wealth transfers, in the form of payments of taxes or for emissions permits, that impose an additional burden on industry. The extent to which this burden can be alleviated affects the magnitude of emissions reductions that are politically feasible.

Policy makers have sought to alleviate this problem by implementing trading schemes where all or some of the emissions permits are given for free. This is often referred to as *grandfathering* since the number of permits freely allocated to a firm

[1] See Schmalensee et al. (1998) and Montero (1999).

[2] See Baumol and Oates (1988) and its references, in particular Montgomery (1972).

[3] Different standards also differ considerably in their impact and efficacy (Helfand, 1991).

is typically related to its past emissions. Grandfathering is the preferred means of winning industry support because it relieves the financial burden of the ETS on industry, without affecting firms' incentives to reduce emissions at the margin. Indeed, the ease with which grandfathering can be coupled with an emissions trading scheme is one reason for the popularity and success of such schemes.[4,5]

For most emissions trading schemes in the US, and also in the early phases of the European Union's ETS for CO_2 (henceforth to be referred to as the EU ETS), almost all permits were freely allocated in this manner. It is clear that not selling permits (at auction, say) entails a significant loss of government revenue which could potentially be more productively employed in other ways.[6] In particular, revenue raised from the sale of permits could allow for the reduction of distortionary taxes imposed on other parts of the economy; this "revenue recycling effect" is important when evaluating the benefits of an ETS (see Fullerton and Metcalf (2001) and Bovenberg, Goulder, and Gurney (2005)). Furthermore, just as a firm's incentive to reduce emissions is unaffected by the free allocation of permits, its incentive to raise prices in response to the higher marginal cost is also unaffected by the free allocation of permits. This raises the possibility that firms will make windfall profits from free permit allocations. For these reasons and others, the question of whether to freely allocate permits, and if so, to what extent, is an important one.

The aim of our paper is to provide a basic theoretical framework in which this

[4]Another method of protecting average profits in an industry is to hold an auction for emissions permits but to then return the revenue back to the firms using some other formula. This was originally proposed by Hahn and Noll (1982); a small fraction of the permits in the Sulfur Allowance Program is allocated through a zero-revenue auction (see Tietenberg (2006, Chapter 6)).

[5]Greenhouse gas emissions trading schemes are adopted more widely than carbon taxes, even though the latter has the same cost efficiency properties as an ETS under certainty and may be even more efficient than an ETS under uncertainty (Weitzman, 1974; Pizer, 2002). See Hepburn (2006) and Stern (2008) for a discussion of the reasons behind this policy 'bias'.

[6]The allocation process can also become the focus of much rent-seeking behavior; for an account of this process in the case of the Acid Rain Program, see Joskow and Schmalensee (1998).

issue and others related to the ETS can be analyzed. We construct a model that we think is a very natural starting point for a theoretical analysis; we show that, in an important set of empirically relevant cases, the number of permits required by an industry for profit-neutrality is a small fraction of its demand for permits, which means that a profit-neutral ETS can still raise substantial revenue for government (through the sale of permits). It is possible that for a particular industry, the profit impact of the ETS is different from that given by this analysis, but it will require the industry to have features that are significantly different from this standard model.

We assume that the industry is an oligopoly in its product market and a price-taker in the market for emissions permits. This is a reasonable assumption since we have in mind a scheme for trading greenhouse gas emissions, like the EU ETS, where permits are traded across many industries in (potentially) many countries while individual industries have oligopolistic structures. To be specific, we assume that the industry is a Cournot oligopoly. We impose two restrictions on the model that, amongst other things, guarantee that the ETS has the desired effect of reducing emissions. First, we assume that firms' emissions intensities are *monotone* in the sense that firms with lower marginal costs also have *weakly* lower emissions intensities (per unit of output).[7] The second assumption is that the industry faces a log-concave demand function. Log-concavity is a commonly-made restriction on the demand function; it is a sufficient (and, in a certain sense, necessary) condition for the Cournot oligopoly to be a game of strategic substitutes (see Section 3.3 for more discussion).

The imposition of a price on emissions will always encourage firms to engage in abatement, thus (weakly) lowering each firm's emissions intensity. But the ETS also leads to changes in output, so that the industry's *average* emissions intensity can increase if it is the dirtier firms that gain market share. This possibility is excluded

[7]Note that this condition is essentially satisfied if firms do not differ significantly in their emissions intensities. Notions of eco-efficiency support a positive correlation between firms' production efficiency and their environmental performance (see footnote 14 for more discussion).

by the two conditions, which jointly ensure that firms with lower marginal costs gain market share and that these firms are *not* more emissions intensive. In this way we guarantee that average emissions intensity is reduced (see Proposition 6). This mechanism also ensures that when the permit price is sufficiently low, the ETS will improve cost efficiency, i.e., the industry's average unit cost of production will *fall*. (We are referring to costs *excluding the cost of permits*.) This may be surprising because the ETS causes firms to substitute away from emissions by using more of other inputs, which tends to raise costs. However, by the envelope theorem, this effect is of second order, so that the only local determinant of average unit costs is the change in market shares caused by the ETS. The ETS lowers average unit cost in the industry because it causes lower cost firms to gain market share (see Proposition 5).

The market gains of lower cost firms is one reason why the adverse profit impact (averaged across the whole industry) of the ETS tends to be limited. We measure the profit impact by looking at the *profit-neutral permit allocation*, i.e., the number of permits that have to be freely allocated to the industry to guarantee that aggregate industry profit before and after the introduction of the ETS are at the same level. Let x be the number of permits required to cover the industry's pre-ETS emissions (had the permits been needed). The profit-neutral permit allocation is always below Hx, where H is the industry's Herfindahl index (see Propositions 10 and 11).[8] For many industries, the Herfindahl index is *much* lower than 0.5. In cases like this, our result says that a large proportion of permits – perhaps more than 50% – can be auctioned without affecting total industry profit.

Our results on profit-neutral permit allocations are obtained by developing a first-order approach that allows us to derive simple expressions for profit-neutral permit allocations at the firm- and industry- level when the permit price is "small". These

[8]Strictly speaking, this result holds only when the Herfindahl index is not too low in some specific sense, but the condition is likely to be satisfied.

formulae give us the bound on profit-neutral permit allocations outlined in the previous paragraph. We then go on to show that our local results remain valid when the permit price is "large". An attractive feature of our formulae is that they involve parameters that can often be estimated with a reasonable degree of accuracy, which makes them potentially amenable to empirical implementation; we illustrate this by applying them to calculate the profit-neutral permit allocation in the U.K. cement industry (which is included the EU ETS).

Our approach inevitably ignores other interesting issues relating to emissions trading, including some that may have an impact on profit-neutral permit allocations.[9] Chief among our assumptions is that permit allocations affect firm profits, but not firm behavior. This will be violated in situations where the market for permits is not significantly broader than the product market; Hahn (1985) and Liski and Montero (2006) consider market power in the emissions market, motivated by the markets for acid rain and particulates.[10] Moreover, we treat the number of firms in the industry as fixed; allowing for potential entry would mean that the manner in which entrants are treated by the permit allocation rules also has strategic and welfare implications. In an intertemporal setting, allocation rules may also have strategic consequences insofar as a firm's actions in one period can affect its allocation in subsequent periods. Finally, in general equilibrium models where there are pre-existing market imperfections, whether permit revenues are used to compensate firms or for some other purpose has efficiency (and not just distributional) consequences (see, e.g., Fullerton and Metcalf (2001) and Bovenberg, Goulder, and Gurney (2005)).

The rest of the paper is organized as follows. Section 2 discusses some general principles underlying profit-neutral permit allocations for general monopoly and oligopoly settings. Section 3 examines the impact of an ETS on a Cournot industry in terms

[9] See also Tietenberg (2006) for a careful summary of these points.

[10] Clearly, the presence of transaction costs also means that initial permit allocations have strategic consequences (Stavins, 1995), although there is some evidence of transaction costs being low in the US sulfur dioxide scheme (see Joskow, Schmalensee, and Bailey, 1998).

of firms' outputs, costs, emissions, and market price. Section 4 presents our results on profit-neutral permit allocations at the firm- and industry- level.

2 The impact of the ETS: general principles

This section spells out the major themes that are relevant for analyzing the impact of an emissions trading scheme (ETS) on firm profits. We consider an industry that produces a particular type of emission (e.g., carbon dioxide) that is harmful to the environment. The ETS imposes a cost on all emissions of this type. We assume that the industry is one of many covered by the scheme, so that, although firms have market power in their product market, they are price-takers in the permit market. In this section, we make no substantive assumptions regarding the nature of the strategic interaction in the industry.

2.1 The monopoly case

We begin by considering the case of a monopoly. Although this case is not typical, it has the merit of having a completely general solution and it provides a natural setting to introduce several of our major themes. We assume that the monopolist chooses a production plan that maximizes its profit, given the demand for its output (which may consist of one or several distinct products), its production set, input prices, and the emissions permit price $t \geq 0$. We denote the monopolist's (maximum) profit by $\Pi^*(t)$, and the associated level of emissions by $\zeta^*(t)$. Assuming that one permit is required for each unit of emissions, the profit *before* accounting for the cost of permits is $\underline{\Pi}^*(t) = \Pi^*(t) + t\zeta^*(t)$. (Note that $\Pi^*(0) = \underline{\Pi}^*(0)$.)

The situation before the introduction of the ETS corresponds to the case where $t = 0$, i.e., emissions are unpriced. Therefore, $\Pi^*(0)$ and $\zeta^*(0)$ are the monopolist's initial levels of profits and emissions respectively. To keep our notation simple, we will usually suppress the argument by writing $\Pi^*(0)$ as Π^*, and so forth.

Profit maximization by the monopolist guarantees that, at any $t > 0$,

$$\underline{\Pi}^*(t) \leq \Pi^* \ \text{ and} \tag{1}$$

$$\Pi^*(t) = \underline{\Pi}^*(t) - t\zeta^*(t) \geq \Pi^* - t\zeta^*. \tag{2}$$

Note that (1) follows from the fact that Π^* is the optimal profit at $t = 0$, while the production decision that generates a profit of $\underline{\Pi}^*(t)$ is one that the monopolist *could* have made at $t = 0$, so the latter must be smaller than the former. The right-hand side of (2) is the monopolists' profit if it chooses not to adjust production after the introduction of the ETS—this must be less than $\Pi^*(t)$, which is the *optimal* profit when emissions are priced at t.

Combining (1) and (2) yields two conclusions. First, the introduction of the ETS reduces emissions, since these inequalities only hold simultaneously if $\zeta^*(t) \leq \zeta^*$. Second, the ETS reduces the monopolist's profit, since (1) implies that $\Pi^*(t) = \underline{\Pi}^*(t) - t\zeta^*(t) \leq \Pi^*$.

Consider now the level of free allocation of permits required to compensate the monopolist for the reduction in profits from Π^* to $\Pi^*(t)$. From (2), $\Pi^*(t) + t\zeta^* \geq \Pi^*$, so there is a $0 \leq \gamma(t) \leq 1$ such that

$$\Pi^*(t) + t\left[\gamma(t)\zeta^*\right] = \Pi^*. \tag{3}$$

In other words, $\gamma(t)\zeta^*$ is the number of freely allocated permits—the *profit-neutral allocation* (PNA)—that will leave the monopolist's total profits at the pre-ETS level. Since $\gamma(t) \leq 1$ free permits cover only a *fraction* of the firm's initial emissions. In this case, we say that the profit-neutral allocation is *partial.*

The intuition for this result is as follows. Suppose that the introduction of the ETS is accompanied by a free allocation of permits at the monopolist's original level of emissions; furthermore, suppose that the monopolist chooses *not* to adjust its production plan in response to the introduction of the ETS. Then the increase in her costs would be *exactly* offset by the value of the free allowances. However, the *option* to adjust (e.g., increase price(s) or switch to cleaner inputs) means that the PNA, in general, is partial. It is worth emphasizing that this conclusion is very robust: no restrictions are imposed on the monopolist, except that it is a price-taker in the market for emissions permits.[11]

Finally, suppose that the monopolist indeed receives the PNA of $\gamma(t)\zeta^*$ permits for free. Re-writing (3), we obtain that $\underline{\Pi}^*(t) + t\,[\gamma(t)\zeta^* - \zeta^*(t)] = \Pi^*$. This, together with (1), implies that the monopolist's endowed permits under the PNA, $\gamma(t)\zeta^*$, will *exceed* its requirement $\zeta^*(t)$, so the monopolist will be selling part of its endowment.

The following proposition summarizes our analysis of the monopoly case.

PROPOSITION 1 *Following the introduction of the ETS, a monopolist has lower emissions and lower profit. PNA is partial, i.e.,* $0 \leq \gamma(t) \leq 1$*; with this allocation of permits the monopolist is a net supplier in the market for permits.*

2.2 Characterizing partial PNA

When considering an oligopoly, we can no longer rely solely on the revealed preference arguments that gave us such mileage in the monopoly case. Nevertheless, there are still some general insights we can derive.

Assume that there are $N \geq 2$ firms in an industry that interact with each other strategically. In the interest of generality, we leave the precise manner of their strategic interaction unspecified for now. Retaining our earlier notation, we denote equilibrium industry profits when the permit price is t by $\Pi^*(t)$, the equilibrium (total) emissions by $\zeta^*(t)$, and so on. The corresponding outcomes for firm i are $\Pi_i^*(t)$, $\zeta_i^*(t)$, etc. We assume that these are all smooth functions of the permit price t in some interval $[0, T]$, where $T > 0$. We call this model a *smooth oligopoly.*

By definition, the proportion of free permit allocation needed for profit-neutrality

[11]Furthermore, it is clear that the result holds even if the monopolist is subject to certain regulatory restrictions, such as being prevented from raising prices after the introduction of the ETS.

at the industry-level, $\gamma(t)$, is given by

$$\Pi^*(t) + t\,\gamma(t)\zeta^* = \Pi^*. \tag{4}$$

The next result gives a sufficient (and, as we shall see a bit later, locally necessary) condition for the profit-neutral allocation to be partial.

PROPOSITION 2 *Suppose $\zeta^*(t) < \zeta^*$ and $\underline{\Pi}^*(t) \geq \Pi^*$. Then $\gamma(t) < 1$; with this level of free allocation, the industry has a net demand for permits.*

Proof: Given the assumptions, that there is a $\gamma(t) < 1$ such that $\underline{\Pi}^*(t) - \Pi^* + t\left[\gamma(t)\zeta^* - \zeta^*(t)\right] = 0$. Rearranging this expression and using the fact that $\underline{\Pi}^*(t) - t\zeta^*(t) = \Pi^*(t)$, we obtain (4). Since $\underline{\Pi}^*(t) \geq \Pi^*$, we must have $\gamma(t)\zeta^* - \zeta^*(t) \leq 0$, so the industry has a net demand for permits. QED

This result is quite intuitive. It says that the industry PNA is partial if the introduction of an ETS increases industry profits *before* accounting for emissions costs—in other words, *average PNA is partial if the ETS leads to a more "collusive" equilibrium outcome*. In particular, had the firms in the industry chosen the (same) actions they did upon the introduction of the ETS *before* it was introduced, their total profits (at $\underline{\Pi}^*(t)$) would have exceeded Π^*.[12]

To carry the analysis further, we now concentrate on the behavior of $\gamma(t)$ for low values of t by examining the first-order approximation $\tilde{\gamma} \equiv \lim_{t\to 0} \gamma(t)$. This limit determines the (approximate) proportion of free permit allocation that satisfies (4) by ignoring higher-order terms in t, and allows us to deliver sharp and easily interpretable results. The marginal cost increase due to an ETS is typically small relative to a firm's total marginal costs, so an analysis based on this approach delivers insight without being misleading; in any case, we show in Section 4 that our main insights remain valid even when t is "large". We make three important observations regarding $\tilde{\gamma}$.

[12] For example, in the standard textbook case of symmetric Cournot oligopoly with constant marginal cost, industry profits are lower than for a monopolist. If the ETS leads to a lower industry output that is closer to the monopoly level, then PNA is partial.

(1) If $\tilde{\gamma} < 1$, then for small t, $\gamma(t) < 1$, that is, the industry (on average) requires only partial PNA for profit-neutrality. Moreover, the industry's net demand for permits, assuming it is given this level of free allocation, is $\zeta^*(t) - \gamma(t)\zeta^*$. Since $\lim_{t\to 0} \zeta^*(t) = \zeta^*$, for low values of t, $\zeta^*(t) - \gamma(t)\zeta^* > 0$. In other words, *if* $\tilde{\gamma} < 1$, *the industry's demand for permits under the PNA will exceed its free allocation.* Conversely, if $\tilde{\gamma} > 1$, then the industry will be a net supplier of permits.

Suppose now that there are sufficiently many industries with partial PNA so that (overall) there is a net demand for permits after profit-neutral permit allocations. Then a given permit price of $t > 0$ can only be supported if there is an external party—the government—that meets this net demand of permits. Therefore an ETS with profit-neutral permit allocations raises net revenue for government if it is partial (on average across the industries covered by the ETS).

(2) Recall that if the industry is run by a monopoly then, for all values of t, PNA is partial, i.e., $\gamma(t) \leq 1$, but the monopoly is also a net supplier of permits. Comparing this with our previous observation, we conclude that, *for a monopoly,* $\tilde{\gamma} = 1$; so even though PNA is partial for a monopoly it approaches a full allocation of permits for low permit prices.

(3) Taking the Taylor expansion of $\Pi^*(t)$ around $t = 0$, (4) gives us a simple expression for PNA, namely

$$\tilde{\gamma} \equiv \lim_{t\to 0} \gamma(t) = -\frac{1}{\zeta^*}\frac{d\Pi^*}{dt}(0). \tag{5}$$

To first order, the proportion of free permits required for profit-neutrality is equal to the loss in industry profits per unit of emissions.[13] Since, by definition, $\Pi^*(t) = \underline{\Pi}^*(t) - \zeta^*(t)t$, we can also write

$$\tilde{\gamma} = 1 - \frac{1}{\zeta^*}\frac{d\underline{\Pi}^*}{dt}(0), \tag{6}$$

from which the next proposition follows immediately.

[13] An alternative way of showing that $\tilde{\gamma} = 1$ for a monopolist is to first observe that, by the envelope theorem, $d\Pi^*/dt = -\zeta^*$ and then to apply formula (5).

Proposition 3 *In a smooth oligopoly,*

$$\tilde{\gamma} < 1 \Longleftrightarrow \frac{d\underline{\Pi}^*}{dt}(0) > 0. \tag{7}$$

This result is the local analog of Proposition 2. Indeed, it goes a bit further since it says that, for small t, the condition that $\underline{\Pi}^*(t)$ is increasing with t is both sufficient—and necessary—for partial PNA at the industry level.

2.3 The impact of the ETS on costs

We now consider the impact of the ETS on firm costs, taking into account both the direct effect of the permit price on costs as well as firms' abatement decisions. From this point on, we assume that the industry produces a *single* product using l (costly) inputs, represented by a vector $\bar{x}_i = (\bar{x}_i^1, \bar{x}_i^2, ..., \bar{x}_i^l)$ in R_{++}^l. Production leads to emissions, which we denote by $\bar{z}_i$. Following Baumol and Oates (1988), amongst others, we shall think of emissions as an input in the production process, albeit one that is initially free. Firm i's production function F_i, assumed to exhibit constant returns to scale, maps the input vector $(\bar{x}_i, \bar{z}_i)$ to the output q_i. All inputs (including emissions) are chosen optimally by firms to minimize costs.

The introduction of an ETS typically induces firms to engage in abatement by reducing their emissions and using more of other inputs (whose prices we assume are unchanged). We denote firm i's unit cost at permit price t by $c_i(t)$, its optimal emissions intensity (i.e., emissions per unit of output) by $z_i(t)$, and its unit cost excluding the cost of permits by $\underline{c}_i(t) \equiv c_i(t) - tz_i(t)$.

Standard production theory tells us that, at any $t > 0$, $z_i(t) \leq z_i(0)$ and $a_i(t) \equiv \underline{c}_i(t) - c_i(0) \geq 0$. We can think of $a_i(t)$ as the *abatement cost* incurred by the firm as it reduces emissions intensity from $z_i(0)$ to $z_i(t)$. If the production technology is such that abatement is either non-optimal or simply impossible, then $\underline{c}_i(t) \equiv c_i(0)$. By the envelope theorem,

$$\frac{dc_i}{dt}(t) = z_i(t), \tag{8}$$

which leads to the important observation that

$$\frac{da_i}{dt}(0) = \frac{d\underline{c}_i}{dt}(0) = 0. \tag{9}$$

So even though the ETS leads to increased expenditure on other inputs, this effect is of second order.

Suppose $q_i^*(t)$ is the firm's equilibrium output as a function of the permit price. Then its total cost excluding the cost of permits is $\underline{C}_i^*(t) = \underline{c}_i(t)q_i^*(t)$. Differentiating by t and appealing to (9), we obtain

$$\frac{d\underline{C}_i^*}{dt}(0) = c_i(0)\frac{dq_i^*}{dt}(0). \tag{10}$$

The introduction of the ETS thus has a two-fold impact on $\underline{C}_i^*$. First, it makes the firm switch away from emissions towards other inputs, thus raising expenditure ($\underline{c}_i(t)$) on those inputs. Second, it has an impact on firm i's output via its strategic interaction with other firms. However, by the envelope theorem, the change in $\underline{c}_i(t)$ is of second order, so that, for low values of t, the change in total cost $\underline{C}_i^*(t)$ is simply driven by the change in firm i's output.

An immediate consequence is that the local impact of an ETS on an industry's average cost is *solely* driven by its effect on firms' relative output shares. We denote the industry's total cost, again excluding permits, by $\underline{C}^*(t) = \sum_{i=1}^N \underline{C}_i^*(t)$ and the associated average cost by $\underline{c}^*(t) \equiv \underline{C}^*(t)/Q^*(t)$. Clearly,

$$\underline{c}^*(t) = \sum_{i=1}^N \sigma_i(t)\underline{c}_i(t), \tag{11}$$

where $\sigma_i(t)$ is firm i's market share. Differentiating this expression with respect to the permit price and using (9), we obtain (note that $\underline{c}_i(0) = c_i(0)$)

$$\frac{d\underline{c}^*}{dt}(0) = \sum_{i=1}^N c_i(0)\frac{d\sigma_i}{dt}(0). \tag{12}$$

It is clear from this equation that the industry's average cost (excluding permits) will *fall* with the introduction of an ETS if firms with lower unit costs increase their market share. Obviously, this finding will be important for determining the profit impact of an ETS, and therefore also the industry's PNA.

3 The ETS in a Cournot model

Consider a standard Cournot oligopoly with $N \geq 2$ quantity-setting firms. Without loss of generality, assume that $c_1(0) \leq c_2(0) \leq ... \leq c_N(0)$, so lower indexed firms have lower initial marginal costs. We already know, from the envelope theorem, that $dc_i/dt = z_i$, so the local impact of an ETS on a firm's marginal cost depends (only) on its emissions intensity.

In principle, $z_i > 0$ may vary with i in any possible way, but for the purposes of analysis we shall often focus on two plausible cases. We say that emissions intensity is *monotone* if z_i is weakly increasing with i, so firms with lower marginal costs also tend to pollute less.[14] A special case of monotonicity is uniformity; we say that emissions intensity is *uniform* (across firms) if z_i is equal for all firms (at $t = 0$). We denote by q the vector $(q_i)_{1 \leq i \leq N}$ which gives the output of each firm, the aggregate output associated with q by Q, and the output of all firms except firm i by Q_{-i}. The marginal revenue of firm i at q satisfies $MR_i(q) = P(Q) + q_i P'(Q)$, where $P(Q)$ is the downward-sloping inverse demand curve. Firm i maximizes its profit when marginal revenue equals marginal cost, $MR_i(q) = c_i(t)$.

Before the introduction of the ETS firms are at the Cournot equilibrium $q^* =$

[14] The general notion of eco-efficiency is that reducing waste also reduces costs (Alexander and Buckholz, 1978; Porter and van der Linde, 1995), suggesting that the monotonicity assumption is plausible. Heal (2008) outlines several case studies, such as the internal emissions trading scheme set up by BP, which reduced emissions and also cut costs; Dow Chemicals and Du Pont provide similar evidence. King and Lenox (2001), amongst others, find a positive correlation between environmental and financial performance. Although there is considerable debate on the reasons for this correlation (Konar and Cohen, 2001), for our purposes the nature and direction of causality between environmental and financial performance is irrelevant, and it suffices that they are *correlated*.

$(q_i^*)_{1\leq i\leq N}$ so that total output $Q^* = \sum_{i=1}^N q_i^*$. At this equilibrium,

$$MR_i(q^*) = P(Q^*) + q_i^* P'(Q^*) = c_i(0) \tag{13}$$

for each firm i. Since demand is downward-sloping, $q_1^* \geq q_2^* \geq ... \geq q_N^*$, so equilibrium output varies inversely with marginal cost.

Let $E(Q^*) = -\left[d\log P'(Q)/d\log Q\right]_{Q=Q^*}$ denote the elasticity of the *slope* of inverse demand, evaluated at the initial equilibrium industry output. This can be interpreted as an index of demand curvature. Clearly, $E(Q^*) > 0$ ($E(Q^*) < 0$) if $P''(Q^*) > 0$ ($P''(Q^*) < 0$) and inverse demand is locally convex (concave) at Q^*.

We make only very weak assumptions on demand. The second-order condition for profit-maximization is satisfied for firm i if its marginal revenue is downward-sloping in its own output, $\partial MR_i(q^*)/\partial q_i < 0$, at equilibrium. Using the above, this can be written as $2P'(Q^*) + q_i^* P''(Q^*) < 0$, or equivalently as

$$2 - \sigma_i^* E(Q^*) > 0, \tag{14}$$

where $\sigma_i = q_i^*/Q^*$ is firm i's initial, pre-ETS market share. We also assume that inverse demand is not too convex in the sense that

$$N + 1 - E(Q^*) > 0. \tag{15}$$

In our setting, the main implication of this assumption is that industry output falls when emissions trading is introduced.[15] Another condition which we occasionally refer to (but which is not maintained as a standing assumption) is

$$1 - \sigma_i^* E(Q^*) > 0, \tag{16}$$

It is known that this is a necessary and sufficient condition for the best response curve of firm i to be locally downward sloping at i's equilibrium output (see, for

[15]This condition, if it holds globally, guarantees the uniqueness of the equilibrium industry output in the Cournot model when firms have constant marginal costs (Bergstrom and Varian, 1985).

example, Bulow, Geanakoplos, and Klemperer (1985) and Shapiro's (1989) survey).[16] If it holds for each firm i, then the Cournot oligopoly is (at least) locally a game of strategic substitutes. Obviously, condition (16) is stronger than (14).

For the rest of this section, we shall examine the impact of the ETS on output, the distribution of output across firms, emissions, and costs. Building on this, we examine the impact of the ETS on profits, and thus PNA, in Section 4.

3.1 The impact of the ETS on output

We assume that the Cournot equilibrium varies smoothly with t when the ETS is introduced, i.e., the Cournot oligopoly is a smooth oligopoly in the sense defined in Section 2. The following result shows the impact of the ETS on firm- and industry-level output and is crucial to understanding its impact on costs and firm profits.

Proposition 4 *Output responses at $t = 0$ are given by*

$$\frac{dQ^*}{dt} = \frac{\sum_{i=1}^N z_i}{P'(Q^*)\,[N+1-E(Q^*)]} < 0; \tag{17}$$

$$\frac{dQ^*_{-i}}{dt} = \frac{dQ^*}{dt}\left[[2-\sigma_i E(Q^*)] - [N+1-E(Q^*)]\frac{z_i}{\sum_{i=1}^N z_i}\right]; \tag{18}$$

$$\frac{dq^*_i}{dt} = \frac{dQ^*}{dt}\left[-[1-\sigma_i E(Q^*)] + [N+1-E(Q^*)]\frac{z_i}{\sum_{i=1}^N z_i}\right]. \tag{19}$$

The proofs of Proposition 4 and other results in this section and the next are in Appendix A.

Proposition 4 says that the introduction of the ETS causes industry output to fall in response to firms' increased marginal costs. However, in general, this fall is not shared equally across firms. To isolate the effect of demand curvature on the pattern

[16] Using equation (47), it is straightforward to show that the slope of i's best response curve, $\partial\hat{q}_i/\partial Q_{-i} = -[1-\sigma_i E(Q^*)]/[2-\sigma_i E(Q^*)]$. Given that the denominator of this expression is positive (by (14)), the expression is negative if and only if its numerator is positive, hence condition (16).

of the output response, let us first assume that emissions intensity is uniform across firms, so that $z_i/\sum_{i=1}^{N} z_i = 1/N$. In this case, (19) simplifies to

$$\frac{dq_i^*}{dt} = \frac{dQ^*}{dt}\frac{[1-(1-N\sigma_i)E(Q^*)]}{N}. \tag{20}$$

If the inverse demand function is convex (so $E(Q^*) \geq 0$), dq_i^*/dt increases in i. In other words, larger firms experience larger (absolute) falls in output. Since total output falls, this implies that the largest firm (firm 1, with marginal cost c_1) *must* experience a fall in output. When the inverse demand function in concave, the distribution of the fall in output is reversed: smaller firms bear the brunt of the reduction, and the smallest firm (firm N) must experience a fall in output. This pattern is not surprising: with convex (concave) demand, larger (smaller) firms have a flatter marginal revenue curve, and thus will cut output by more for any given increase in marginal cost.

To isolate the influence of emissions intensity on the relative impact of the ETS, consider the case where emissions intensity is monotone and demand is linear, so $E(Q^*) = 0$. Then (19) simplifies to

$$\frac{dq_i^*}{dt} = \frac{dQ^*}{dt}\left[-1+\frac{(N+1)}{N}z_i\right]. \tag{21}$$

Clearly, dq_i^*/dt now decreases with i. The cost impact of emissions trading is greatest on the small (and highest polluting) firms, so they experience the largest fall in output.

In summary, when demand is concave ($E(Q^*) \leq 0$) and emissions intensity is monotone, then the demand curvature effect and the emissions intensity effect reinforce each other, so that dq_i^*/dt decreases with i and smaller firms experience greater decreases in output. When demand is convex, however, the two effects work against each other and (without further assumptions) it is ambiguous whether smaller or larger firms cut output by more.

3.2 The impact of the ETS on price

We know from (17) that equilibrium output falls after the introduction of the ETS, which means that the equilibrium price of output must increase. To be specific, (17) gives us

$$\frac{dP^*}{dt} = P'(Q^*)\frac{dQ^*}{dt} = \frac{\sum_{i=1}^{N} z_i}{[N+1-E(Q^*)]}. \tag{22}$$

This formula is remarkably simple in one respect: the price increase depends on only the unweighted average of the emissions intensities and on no other feature of its distribution. So a change in emissions intensities that leaves its unweighted average unchanged does not modify the price impact of the ETS.

To have a better understanding of (22), consider the hypothetical situation where $z_i = 1$ for all i. Then

$$\frac{dP^*}{dt} = \frac{N}{[N+1-E(Q^*)]} \equiv \kappa. \tag{23}$$

The term κ is known as the *rate of cost pass-through* since it measures the change in the equilibrium price following a common increase in the marginal cost of every firm in the oligopoly. Loosely speaking, if marginal cost increases by a dollar at every firm, then the equilibrium price rises by κ dollars. Denoting the unweighted average marginal cost in the industry by $\hat{c}(t) = (\sum_{i=1}^{N} c_i(t))/N$, it follows from (8) that

$$\frac{d\hat{c}}{dt}(0) = \frac{\sum_{i=1}^{N} z_i}{N}, \tag{24}$$

so (22) may be rewritten as

$$\frac{dP^*}{dt} = \kappa\,\frac{d\hat{c}}{dt}. \tag{25}$$

In finite terms, $\Delta P^* \approx \kappa\,\Delta\hat{c}$, so the price increase following the introduction of the ETS is approximately proportional to the rise in the unweighted marginal cost, with the cost pass-through κ as the proportionality constant.

3.3 The impact of the ETS on costs

The standard justification for the use of an ETS is the cost minimization theorem, which says that it represents the cheapest way of achieving an emissions target subject to a given set of output constraints. We now give a formal statement of this result in the context of a Cournot model.

Let r denote the input price vector of all inputs except emissions and let t be the unit price of emissions. Keeping r fixed, we denote firm i's optimal input vector (for producing a single unit of output) by $(x_i(t), z_i(t))$, where $x_i(t)$ is the vector of inputs excluding emissions and $z_i(t)$ is the emissions intensity. Thus, the unit cost excluding the cost of permits is $\underline{c}_i(t) = r \cdot x_i(t)$. As above, firm i's output at the Cournot equilibrium is $q_i^*(t)$. The problem of minimizing costs subject to achieving an emissions target can be stated as:

$$\min r \cdot \left[\sum_{i=1}^{N} \hat{x}_i\right]$$

$$\begin{aligned} \text{subject to } (i)\ F_i(\hat{x}_i, \hat{z}_i) &= q_i^*(t) \text{ for all } i, \\ \text{and } (ii)\ \sum_{i=1}^{N} \hat{z}_i &= \sum_{i=1}^{N} z_i(t) q_i^*(t), \end{aligned}$$

where F_i is firm i's production function, so $F_i(\hat{x}_i, \hat{z}_i)$ is the firm's output at the input vector $(\hat{x}_i, \hat{z}_i)$. This optimization determines the cheapest way of achieving an emissions target of $\sum_{i=1}^{N} z_i(t)q_i^*(t)$ subject to firm i producing $q_i^*(t)$. By the cost minimization theorem (see Baumol and Oates (1988)), the solution to this problem is $\hat{x}_i = x_i(t)q_i^*(t)$ and $\hat{z}_i = z_i(t)q_i^*(t)$. In short, the solution coincides with that achieved by an ETS with permits priced at t.

This result is difficult to interpret in a Cournot setting: since firms' outputs are perfect substitutes, there is no normative reason for imposing condition (i). However, if we replace it with the condition that *total* output equals $\sum_{i=1}^{N} q_i^*(t)$, then it is clear that neither the initial Cournot equilibrium nor that after the introduction of the ETS are cost efficient. Cost minimization in both instances would require all output to be produced by the firm with lowest marginal cost.

A more useful criterion in evaluating the impact of the ETS in this setting is to ask whether it makes the industry more (or less) efficient in its use of resources (other than emissions). Formally, we wish to study the impact of emissions trading on the industry's average cost, $\underline{c}^*(t)$. It follows from (12) that the local impact of the ETS on average cost is *solely* driven by its impact on relative output shares. Therefore, if firms with lower unit costs (lower i) increase their market share, then cost efficiency improves. Our next result identifies conditions under which $d\sigma_i/dt$ is indeed decreasing in i. This implies that the output responses $\{d\sigma_i/dt\}_{1\leq i\leq N}$ obey the *single crossing property*, and there is k such that for all $i \leq k$, $d\sigma_i/dt \geq 0$ and for all $i > k$, $d\sigma_i/dt < 0$. In other words, firms may be partitioned into two sets: the firms in $\mathcal{J} = \{1, 2, ..., k\}$ (weakly) increase their market share while all other firms lose market share.

PROPOSITION 5 *Suppose that emissions intensity is monotone and* $E(Q^*) \leq 1$.
(i) Then $d\sigma_i/dt$ *is decreasing in* i. *Furthermore, either* (a) $z_j \leq (\sum_{i=1}^{N} z_i)/N$ *for every firm* j *in* $\mathcal{J}$ *or* (b) $\sigma_j \geq 1/N$ *for every firm* j *in* $\mathcal{J}$ *(with* $\mathcal{J}$ *as defined above).*
(ii) The industry's average cost $\underline{c}^*$ *obeys*

$$\frac{d\underline{c}^*}{dt}(0) \leq 0.$$

Proposition 5(i) guarantees the single crossing property and also identifies a common feature of the firms that gain market share: before the introduction of the ETS, either they all have lower than average emissions intensities or they all have higher than average market shares.[17] Proposition 5(ii) shows that the introduction of an ETS can actually *reduce* an industry's average cost of production (recalling that marginal costs and market shares are inversely related).

Both parts of this proposition rely on the condition that $E(Q^*) \leq 1$, which is just another way of saying that the demand curve is locally log-concave (at Q^*). It follows immediately from (23) that this condition is equivalent to *partial* cost pass-through, i.e., $\kappa \leq 1$ (so a dollar increase in the marginal cost of every firm raises price by less than a dollar). It is also clear from (16) that this condition is sufficient to guarantee that each firm's best response function is locally downward-sloping. Indeed,

[17] It is also straightforward to modify the proof of Proposition 5(i) to show that all firms that lose market share either all have higher than average emissions or lower than average market shares.

$E(Q^*) \leq 1$ is a *necessary* condition for (16) to hold for any distribution of market shares. For this reason and others, log-concavity is a common standing assumption in Cournot analysis.[18,19]

Even when demand is not log-concave, it is clear that sufficiently strong monotonicity of emissions intensities will ensure that cost efficiency improves.[20] So we conclude that while it is possible for cost efficiency to worsen with the introduction of the ETS,[21] an improvement in cost efficiency is the far more likely scenario (for small t). This serves as an additional basic justification for the ETS in the context of an oligopolistic industry producing a homogeneous good.

Of course, the result that the industry's average cost will fall is driven by the fact that firm-level cost increases are of second order, which means that for large t these increases could still be significant enough to raise average costs. Nonetheless, even for large t, log-concavity of demand implies that the market share of relatively low cost firms increase, which helps moderate the impact of any increase in costs (see Section 4.4 and Lemma A2 in Appendix A).

3.4 The impact of the ETS on emissions

We already know that the ETS lowers total industry output, so total emissions will fall if the average emissions intensity of firms falls. Letting $z^*(t)$ denote average

[18]See, e.g., Farrell and Shapiro (1990) and Shapiro (1989). Log-concavity also guarantees local stability of the Cournot equilibrium and, when it holds globally, the global uniqueness of the Cournot equilibrium, see again Shapiro (1989).

[19]For products that are consumed only as a single unit or none at all, market demand at price p is proportional to $\bar{F}(p) = 1 - F(p)$, where F is the distribution of the reservation prices. A sufficient condition for $\bar{F}$ to be log-concave is for F to be generated by a log-concave density function. Many commonly used density functions have this property, see Bagnoli and Bergstrom (2005).

[20]To be precise, see (50) in the proof of Proposition 5 in Appendix A.

[21]For example, if $E(Q^*) \geq 1$ and emissions intensity is uniform, then $d\underline{c}^*/dt \geq 0$. (See the proof of Proposition 6.) In this case, the rate of cost pass-through κ exceeds 100%, so firms' profit margins increase and smaller, less efficient firms gain market share.

emissions intensity and noting that $z^*(t) = \sum_{i=1}^N \sigma(t) z_i(t)$, we obtain

$$\frac{dz^*}{dt}(0) = \sum_{i=1}^N z_i(0)\frac{d\sigma_i}{dt}(0) + \sum_{i=1}^N \frac{dz_i}{dt}(0)\sigma_i(0). \tag{26}$$

Standard production theory tells us that $dz_i/dt < 0$ (in other words, firms make abatement decisions), so the second term on the right of this equation is always negative. If emissions intensity is uniform across firms, the first term on the right equals zero since $\sum_{i=1}^N d\sigma_i/dt = 0$ and we conclude that average emissions intensity must fall.

If emissions intensity is *not* uniform, the sign of the first term on the right of (26)—and thus the sign of dz^*/dt—cannot be guaranteed without further assumptions. This reflects the fact that while the ETS induces each firm to lower its emissions intensity, it is possible for this effect to be negated in part or in whole by strategic effects. If the ETS causes firms with (initially) low emissions to gain market share, then it has a doubly beneficial effect. On the other hand, if these firms lose market share, this diminishes its ability to lower emissions.[22] The latter possibility is excluded if firms' output responses obey the single crossing property, which in turn follows from the log-concavity of demand.

PROPOSITION 6 *Average emissions intensity* z^* *and total emissions* ζ^* *satisfy*

$$\frac{dz^*}{dt}(0) \leq 0 \text{ and } \frac{d\zeta^*}{dt}(0) \leq 0$$

if either (a) *emissions intensities are uniform or* (b) *emissions intensities are monotone and* $E(Q^*) \leq 1$.

[22] The possibility of such perverse effects have also been noted in Levin's (1985) study of taxation in a Cournot model.

4 Profit-neutral permit allocations

Having established the impact of the ETS on output, price, costs and emissions, we now turn to examine its impact on profits. In particular, we develop formulae that determine the level of free permit allocations required to ensure profit-neutrality at the level of the firm and of the industry. We use these formulae to show that, under reasonable conditions, average PNA in the industry is not just partial, but *low*. We also perform some illustrative calculations on the level of PNA for the UK cement industry.

4.1 PNA for an individual firm

By definition, the proportion of free permit allocation, $\gamma_i(t)$, needed to preserve firm i's profit satisfies

$$\Pi_i^*(t) + t\,\gamma_i(t)\zeta_i^* = \Pi_i^*. \tag{27}$$

Taking the Taylor expansion of $\Pi_i^*(t)$ at $t = 0$ yields the approximate PNA

$$\tilde{\gamma}_i \equiv \lim_{t\to 0} \gamma(t) = -\frac{1}{\zeta_i^*}\frac{d\Pi_i^*}{dt}(0). \tag{28}$$

The free allocation required to ensure profit-neutrality (to first order) at firm i is equal to the profit lost per unit of emissions.

In a Cournot setting, one can naturally think of the operating profit of firm i as a function of its own output, q_i, all other firms' output, Q_{-i}, and the market price of emissions permits, t. More formally, $\Pi_i(q_i, Q_{-i}, t) = q_i P(q_i + Q_{-i}) - c_i(t)q_i$. The equilibrium profit $\Pi_i^*(t) \equiv \Pi_i(t, q_i^*(t), Q_{-i}^*(t))$ then varies with t according to

$$\begin{aligned} \frac{d\Pi_i^*}{dt} &= \frac{\partial \Pi_i}{\partial t} + \frac{\partial \Pi_i}{\partial q_i}\frac{dq_i^*}{dt} + \frac{\partial \Pi_i}{\partial Q_{-i}}\frac{dQ_{-i}^*}{dt} \\ &= -z_i q_i^* + q_i^* P'(Q^*)\frac{dQ_{-i}^*}{dt}, \end{aligned} \tag{29}$$

where $\zeta_i^* = z_i q_i^*$ and the second equality relies on the first-order condition for profit-maximization that $\partial\Pi_i/\partial q_i = 0$. Using this last result in (28), we obtain a simple expression for PNA:

$$\tilde{\gamma}_i = 1 - \frac{P'(Q^*)}{z_i} \frac{dQ^*_{-i}}{dt}. \tag{30}$$

Our next result follows immediately.

Proposition 7 *The first-order PNA for firm i has the following property:*

$$\tilde{\gamma}_i < 1 \Longleftrightarrow \frac{dQ^*_{-i}}{dt} < 0.$$

This result says that, for small t, PNA for firm i is partial if (and only if) the total output of all *other* firms, $Q^*_{-i}(t)$, falls in response to the introduction of emissions trading. Like a monopolist (for which recall $\tilde{\gamma} = 1$), an individual oligopolist faces an increase in marginal cost from the ETS, but, in addition, it also faces a change in its residual demand curve $p = P(q_i + Q^*_{-i}(t))$. Therefore, the PNA for an individual firm $\tilde{\gamma}_i$ is less than unity if its residual demand curve becomes more favorable, $Q^*_{-i}(t) < Q^*_{-i}$, and vice versa.[23] Therefore, it is clear that PNA for each firm *must* be partial if firms in an industry are sufficiently symmetric (in that they all cut output).

Equations (17) and (18) from Proposition 5, together with (30), now give us an explicit formula for PNA at the firm level.

Proposition 8 *The first-order PNA for firm i,*

$$\tilde{\gamma}_i = 2 - \frac{[2 - \sigma_i E(Q^*)]}{[N + 1 - E(Q^*)]} \frac{\sum_{i=1}^N z_i}{z_i}. \tag{31}$$

It is clear from this formula that PNA (measured as a proportion of the firm's initial emissions) will typically not be the same across firms. Almost inevitably, a "one-size-fits-all" allocation policy, in which every firm receives the same proportion of freely allocated permits, will lead to overcompensation for some firms and a degree of undercompensation for others. One special case where this proportion *is* constant across firms is when demand is linear and firms have the same emissions intensity

[23] Note that this holds independently of emissions intensities (although the change in residual demand itself is, of course, a function thereof).

(even if they have different costs, and hence market shares). In this case, $\tilde{\gamma}_i = 2/(N+1) \in (0, 2/3]$ for all i, so PNA is positive but partial.

When the firms are symmetric, i.e., have identical costs and emissions intensities, then $\tilde{\gamma}_i = (2 - E(Q^*))/(N + 1 - E(Q^*)) < 1$ for all i, thus confirming our remarks following Proposition 7. More generally, it is not hard to check that $\gamma_i < 1$ if emissions are uniform and firm i is sufficiently 'typical' in the sense that $\sigma_i \in [1/N, 2/(N+1)]$ (for any value of $E(Q^*)$). When firm i's market share$_i$ is outside that region, it is possible for $\gamma_i > 1$ so firm i's PNA is *larger* than its initial emissions.[24] (A specific example of this is provided in the next subsection, following Proposition 9.)

Differing emissions intensities are another source of differences in PNA across firms. As one would expect, $\tilde{\gamma}_i$ is decreasing in the emissions of other firms and increasing in its own emissions (see (31), (14), and (15)). For example, assuming that demand is linear (so $E(Q^*) = 0$),

$$\tilde{\gamma}_i = 2\left[1 - \frac{\sum_{i=1}^{N} z_i}{(N+1)z_i}\right]. \tag{32}$$

Clearly, $\tilde{\gamma}_i > 1$ if $z_i/(\sum_{i=1}^{N} z_i)$ is sufficiently close to 1. At the other extreme, if $z_i/(\sum_{i=1}^{N} z_i)$ is sufficiently close to zero, then $\tilde{\gamma}_i < 0$ since the scheme has a greater impact on firm i's rivals and its strategic position improves to the extent that it actually makes a *higher* profit after the introduction of the ETS.

4.2 Average PNA for an industry

We now examine the level of PNA needed for profit-neutrality for an industry as a whole. This number, rather than firm-specific PNA, is likely to be more policy-relevant in terms of deciding how many permits to freely allocate to firms (and conversely how many to sell or auction).

Recalling the definition of industry-level PNA that $\Pi^*(t) + t\,\gamma(t)\zeta^* = \Pi^*$, and

[24]However, it is also clear from (31) and the assumptions (14) and (15), that $\tilde{\gamma}_i$ never exceeds 2.

since $\Pi^*(t) = \sum_{i=1}^N \Pi_i^*(t)$, we obtain (after Taylor expansion) the first-order PNA

$$\tilde{\gamma} = -\frac{1}{\zeta^*} \sum_{i=1}^{N} \frac{d\Pi_i^*}{dt}(0). \quad (33)$$

Now, since from (28) $d\Pi_i^*/dt = -z_i^* q_i^* \tilde{\gamma}_i$, we can rewrite this as

$$\tilde{\gamma} = \frac{\sum_{i=1}^N z_i \sigma_i \tilde{\gamma}_i}{\sum_{i=1}^N z_i \sigma_i}. \quad (34)$$

In other words, PNA for an industry is an average of PNAs for individual firms, weighted by market shares and emissions intensities. Denoting the industry's Herfindahl index by $H = \sum_{i=1}^N \sigma_i^2$, where $1 > H \geq 1/N$, equations (34) and (31) yield an explicit formula for PNA for an industry.

Proposition 9 *The first-order PNA for an industry,*

$$\tilde{\gamma} = 2 - \frac{[2 - HE(Q^*)] \sum_{i=1}^N z_i}{[N + 1 - E(Q^*)] \sum_{i=1}^N z_i \sigma_i}. \quad (35)$$

In principle, $\tilde{\gamma}$ can take on a wide range of values, both positive and negative. For example, it is known (see Kimmel (1992)) that in a symmetric Cournot oligopoly, a common increase in marginal cost *raises* total profit (in our notation, $\Pi^*(t) > \Pi^*$) if and only if $E(Q^*) > 2$. We can recover this result using (35); at a symmetric equilibrium and assuming uniform emissions intensities, $\tilde{\gamma} < 0$ if and only if $E(Q^*) > 2$. In this case, industry profits increase with the introduction of the ETS, so the industry is (at least weakly) better off even if it has to buy all the permits it needs at the market price. It is also clear from our earlier discussion of firm-level PNA that if firms (and their emissions intensities) are roughly symmetric, then PNA for each firm – and thus for the industry as a whole – is partial.

However, with sufficiently asymmetric firms, it is possible for $\tilde{\gamma}$ to exceed unity. For example, consider a duopoly with uniform emissions intensity that faces a unit-elastic demand curve $P(Q) = K/Q$ (so industry revenue is constant at K), so $N = 2$ and $E = 2$. It is easily checked that $\tilde{\gamma}_1 = 2(\sigma_1 - \sigma_2)$ and hence that $\tilde{\gamma}_1 = -\tilde{\gamma}_2$. With symmetric firms, therefore, PNA is zero for both firms (and for the industry as well), but if $\sigma_1 > \frac{3}{4}$, then $\tilde{\gamma}_1 > 1$ and $\tilde{\gamma}_2 < -1$. The average PNA $\tilde{\gamma} = \sigma_1\tilde{\gamma}_1 + \sigma_2\tilde{\gamma}_2 = 2(\sigma_1 - \sigma_2)^2$ exceeds unity if $\sigma_1 > (\sqrt{2}+1)/2\sqrt{2} \approx 85\%$.

Such examples notwithstanding, the following result shows that reasonable restrictions on the industry guarantee that industry PNA is partial, and indeed low.

PROPOSITION 10 *Provided emissions intensity is monotone*

$$\tilde{\gamma} \leq \tilde{\beta} \equiv 2 - \frac{N\left[2 - HE(Q^*)\right]}{\left[N + 1 - E(Q^*))\right]}. \tag{36}$$

Furthermore, (i) if $H \leq 2/(N+1)$, then

$$\tilde{\gamma} \leq \tilde{\beta} \leq 2 - NH \tag{37}$$

and (ii) if $H \geq 2/(N+1)$, then

$$\tilde{\gamma} \leq \tilde{\beta} \leq \max\{1, E(Q^*)\}\, H. \tag{38}$$

Note that the bounds on $\tilde{\gamma}$ given in Proposition 10 are all independent of emissions intensities (subject to them being monotone). This result allows us to draw three important conclusions about average PNA.

(1) *PNA is partial for any industry that is sufficiently fragmented in the sense that its Herfindahl index is sufficiently low* (with $H \leq 2/(N+1)$). This conclusion follows immediately from Proposition 10(i) since H is always bounded below by $1/N$, so $\tilde{\gamma} \leq 2 - NH \leq 1$. This result does not depend on the demand curvature $E(Q^*)$.[25]

(2) *PNA is partial for any industry in which firms compete in strategic substitutes (i.e., all firms have downward-sloping best responses).* To see this, recall that firm i's best response is downward-sloping if (16) holds; if (16) holds for all i, we obtain

[25] At $H = 2/(N+1)$ and with uniform emissions, formula (35) gives $\tilde{\gamma} = 2/(N+1)$, so the bound in (i) is tight at this value of H. With uniform emissions and symmetric market shares (so $H = 1/N$), the formula (35) gives us $\tilde{\gamma} = (2 - E(Q^*))/(N + 1 - E(Q^*))$; this expression approaches 1 as $E(Q^*)$ approaches $-\infty$, so the bound of 1 given by (37) is indeed tight.

$\sum_{i=1}^{N} \sigma_i(1 - E(Q^*)\sigma_i) > 0$, which may be rewritten as $1 > HE(Q^*)$. Proposition 10(ii) tells us that $\tilde{\gamma} \leq \max\{1, E(Q^*)\}\, H < 1$ if $H \geq 2/(N+1)$. For the case of $H \leq 2/(N+1)$, we know from Proposition 10(i) that PNA is partial (for any value of $E(Q^*)$).

It follows from this result that in the example we considered after Proposition 9, $\tilde{\gamma} > 1$ only when either firm has an *upward*-sloping best response curve; indeed, one can check directly that this holds for Firm 1.

(3) *PNA is bounded above by the Herfindahl index for any industry that has log-concave demand and is sufficiently concentrated* ($H \geq 2/(N+1)$). This is clear since if $E(Q^*) \leq 1$, Proposition 10(ii) says that $\tilde{\gamma} \leq H$. Note that the required degree of concentration ($H \geq 2/(N+1)$) is very low and likely to hold in many cases. This result gives a far tighter bound on PNA since the Herfindahl index is usually below 50%, and *much* below this level for many industries. (The U.S. Department of Justice considers a Herfindahl index between 0.1 and 0.18 as 'moderately concentrated' and anything above 0.18 as 'concentrated'.)

4.3 Calculating PNA: an example

An attractive feature of our formulae for PNA is that they involve parameters that are familiar and can often be estimated with a reasonable degree of accuracy, so that we can readily use them to make indicative calculations of PNA.

We illustrate this with an application to the U.K. cement industry, which is part of the EU ETS. In this example, the relevant market definition is at the UK level, and it is reasonable to set the Herfindahl index $H = 0.28$ and the number of firms $N = 8$.[26] We do not have detailed information on emissions intensities across firms, but we are not aware of any reason to believe that it departs significantly from monotonicity, so we shall make this assumption.

[26]See Appendix B for a justification of these values and any other data used in this example.

The main complication when using our PNA formula is that demand curvature $E(Q^*)$ is not directly observable. If we assume that demand is locally log-concave, then $E(Q^*) \leq 1$, and so (by Proposition 10(ii)) PNA is bounded above by the industry's Herfindahl index, so $\tilde{\gamma} \leq H = 0.28$. (It is trivial to check that this industry obeys the condition $H \geq 2(N+1)$.) In other words, to preserve the industry's profits, it requires a free allocation of permits covering just 30% of its pre-ETS emissions.

If we are not content with making this assumption, there is another, less stringent, way of bounding $E(Q^*)$, which we call the *elasticity approach*. Note that we can write

$$E(Q^*) = \left[1 + \frac{1}{\eta(Q)} + \frac{d \log \eta(Q)}{d \log Q}\right]_{Q=Q^*}, \tag{39}$$

where $\eta(Q) = |P(Q)/QP'(Q)|$ is the industry price elasticity of demand. Imposing the commonly-made and reasonable assumption that demand elasticity is non-decreasing in price (which implies that $\partial\eta(Q)/\partial Q \leq 0$), we obtain an upper bound on demand curvature $E(Q^*) \leq 1 + 1/\eta(Q^*) \equiv \bar{E}$, where $\bar{E} > 1$. If demand has constant elasticity $E(Q^*) = \bar{E}$, but otherwise $\bar{E}$ may be a significant overestimate of the true demand curvature. Calculating $\bar{E}$ is usually straightforward, as it is relatively easy to find estimates of price elasticity $\eta(Q^*)$ for many industries from previous empirical work. For the U.K. cement industry, $\eta = 0.8$ is our 'best guess,' but we also use a low estimate of 0.5 and a high estimate of 2.0 to check robustness of our results.

Since $\tilde{\beta}$ in Proposition 10 is increasing in $E(Q^*)$,[27] we can write

$$\tilde{\gamma} \leq \tilde{\beta} \leq \bar{\beta} \equiv 2 - N\frac{(2 - H\bar{E})}{(N + 1 - \bar{E})}. \tag{40}$$

Table 1 displays values for the upper bound $\bar{E}$ on demand curvature, as well as of the upper bound $\bar{\beta}$ on PNA for our range of elasticity estimates. As is consistent with our earlier analysis, PNA (as bounded by $\bar{\beta}$) is well below unity for all of these estimates, and indeed is always below 50%. We also repeat these calculations for a larger number of firms in the industry (to account for any potential ambiguity over

[27] Note that $\tilde{\beta} = \tilde{\gamma}$ if emissions is uniform rather than just monotone.

any very small firms not captured in our industry data—since $\bar{\beta}$ also increases with N). The estimates of PNA remain well below 100% for these cases, even as we let $N \to \infty$ and so $\bar{\beta} \to \bar{E}H$.

Table 1: PNA ($\tilde{\gamma} \leq \bar{\beta}$) and price elasticity ($\eta$)

Elasticity (η)	$\bar{E}$	$\bar{\beta}$ ($N = 8$)	$\bar{\beta}$ ($N = 10$)	$\bar{\beta}$ ($N = 12$)	$\bar{E}H$
0.5 (low estimate)	3.00	0.45	0.55	0.61	0.84
0.8 (best guess)	2.25	0.38	0.43	0.47	0.63
2.0 (high estimate)	1.50	0.31	0.34	0.35	0.42

We can also double-check our results on PNA using a *cost pass-through approach.* Recall from (23) that the rate of cost pass-through

$$\kappa = \frac{N}{N + 1 - E(Q^*)}. \tag{41}$$

This formula shows that it is possible to back out the value of $E(Q^*)$ from information on cost pass-through. We do not have any data on pass-through for the cement industry, and therefore report estimates of PNA (as well as the implied $E(Q^*)$) for a very wide range of rates (between 25% and 200%) in Table 2 (for $N = 8$ and $H = 0.28$).[28]

This approach covers a much larger range of values for $E(Q^*)$ than our previous elasticity approach, however, PNA (as bounded above by $\tilde{\beta}$; see (36)) remains significantly below 100% in all cases. Notice also that when demand is log-concave ($E(Q^*) \leq 1$; equivalently, when cost pass-through is below 100%), PNA is bounded above by the industry's Herfindahl index $\tilde{\gamma} \leq \bar{\beta} \leq H = 0.28$, as expected from Proposition 10(ii). Moreover, when cost pass-through is below around 40% (and demand is rather concave), then PNA is certainly negative as $\tilde{\beta} \leq 0$. In these cases, the

[28]The literature on tax incidence finds empirical evidence for cost pass-through both above 100% ('overshifting') and below 100% ('undershifting') in markets such as cigarettes, gasoline and groceries, but there is little evidence of rates of pass-through above 200%. See Fullerton and Metcalf (2002) for an overview of this literature.

ETS *raises* average industry profit, so that no free allocation of permits is needed for profit-neutrality.

Table 2: PNA ($\tilde{\gamma} \leq \tilde{\beta}$) and cost pass-through (κ)

κ	25%	50%	75%	100%	125%	150%	175%	200%
$E(Q^*)$	–23.0	–7.0	–1.7	1.0	2.6	3.7	4.4	5.0
$\tilde{\beta}$	–0.11	0.02	0.15	0.28	0.41	0.54	0.67	0.80

4.4 The ETS when t is large

Our analysis of the impact of an ETS in a Cournot model has thus far focused on its first-order impact (for small t). We now show that our main results are preserved when the permit price is not small. In particular, when demand is log-concave, PNA is bounded by the Herfindhal index and thus likely to be lower (perhaps even *much* lower) than 50%.

We assume that the permit price is $T > 0$, and for all t in $[0, T]$, the Cournot equilibrium exists and varies smoothly with t. It follows from (4) that the industry PNA may be written as

$$\gamma(T) = \frac{1}{T\zeta^*(0)} \left[\Pi^*(0) - \Pi^*(T) \right]. \tag{42}$$

Since $\Pi^*(T) = \Pi^*(0) + \int_{t=0}^{T} [d\Pi^*(t)/dt] \, dt$ and defining $\lambda(t) = -\left[d\Pi^*(t)/dt\right]/\zeta^*(t)$, we obtain from (42)

$$\gamma(T) = \frac{1}{T\zeta^*(0)} \int_{t=0}^{T} \lambda(t)\zeta^*(t)dt \tag{43}$$

Using essentially the same arguments that led to the formula for $\tilde{\gamma}$ (see Proposition 9), we find that

$$\lambda(t) = 2 - \frac{[2 - H(t)E(Q^*(t))]}{[N + 1 - E(Q^*(t))]} \frac{\sum_{i=1}^{N} z_i(t)}{\sum_{i=1}^{N} z_i(t)\sigma_i(t)}. \tag{44}$$

It follows that for any positive scalar K, $\gamma(T) < K$ provided (i) $\lambda(t) < K$ and (ii) $\zeta^*(t)$ is decreasing in t for all t in $[0, T]$. It is possible to bound $\lambda(t)$—and thus $\gamma(T)$—by mimicking the arguments of the last sub-section (and, in particular, Proposition

10). Instead of covering all the possible scenarios, we here simply give a flavor of how the argument works by focusing on two important cases that guarantee partial PNA.

Condition (i) (for $K = 1$) and condition (ii) hold in the following cases:

Case A. The demand function obeys $[N + 1 - E(Q)] > 0$ for any $Q > 0$. All firms operate the same technology, so (in symmetric equilibrium) each firm produces at the same output level using an identical input (including emissions) mix.

Case B. The demand function is globally log-concave, i.e., $E(Q) < 1$ for $Q > 0$. Market concentration before the introduction of the ETS obeys $H(0) \geq 2/(N + 1)$. For any t in $[0, T]$, $c_i(t)$ increases with i, and emissions intensity is monotone in the sense that $z_i(t)$ also increases (weakly) with i.

Solving Case A is straightforward. Equilibrium output $Q^*(t)$ falls with t since $[N + 1 - E(Q)] > 0$ (which can be proved using essentially the same arguments that led to (17)). Standard revealed preference arguments guarantee that the optimal emissions intensity at firm i, $z_i(t)$, decreases with t. Since firms are identical, $z_i(t)$ also equals the average emissions intensity, $z^*(t)$. It follows that total emissions $\zeta^*(t) = z^*(t)Q^*(t)$ are decreasing in t, so condition (ii) is satisfied. Furthermore, with a symmetric equilibrium, $(\sum_{i=1}^{N} z_i(t))/(\sum_{i=1}^{N} z_i(t)\sigma_i(t)) \equiv 1$ and $H(t) \equiv 1/N$, so

$$\lambda(t) = \frac{2 - E(Q^*(t))}{[N + 1 - E(Q^*(t))]} \leq \frac{2 - E_m}{[N + 1 - E_m]} \tag{45}$$

where $E_m = \min_{0 \leq t \leq T} E(Q^*(t))$. Note that the bound $M \equiv (2 - E_m)/(N + 1 - E_m)$ is strictly less than unity. Furthermore, it follows from (43) that $\gamma(T) \leq M$. We conclude that, *under the assumptions of Case A, the introduction of the ETS reduces output, emissions intensity, and emissions; PNA is partial and bounded by M.*

A similar conclusion holds for Case B, though the proof is less straightforward.

Proposition 11 *For Case B, the introduction of the ETS lowers industry output, emissions intensity, and total emissions; formally, $Q^*(T) < Q^*(0)$, $z^*(T) \leq z^*(0)$,*

$\zeta^*(T) < \zeta^*(0)$. *Industry PNA,* $\gamma(T) \leq H(T)$*, where* $H(T)$ *is the Herfindahl index at the post-ETS equilibrium. In particular, industry PNA is partial.*

We complete the proof of this result in Appendix A, but it is instructive to sketch out its major elements here. Since $E(Q) < 1$, the inequality $N + 1 - E(Q) > 0$ holds and guarantees that equilibrium output $Q^*(t)$ falls with t. Revealed preference arguments guarantee that $z_i(t)$ decreases with t. Thus, *average* emissions intensity $z^*(t)$ falls with t if firms with (weakly) lower emissions intensity gain market share with increasing t. This follows from the fact that, with monotonicity, firms with lower emissions also have lower marginal costs and these firms gain market share when t increases and demand is log-concave.

To obtain the bound on industry PNA, first observe that the Herfindahl index $H(t)$ rises with t. This is because the firms with lower marginal cost have larger market shares, and gain further market share as t increases. Essentially the same argument that guaranteed (38) in Proposition 10(ii) can be used to show that $\lambda(t) \leq H(t)$. (This argument requires $H(t) \geq 2/(N+1)$, which is true since $H(t)$ increases with t and we assume that $H(0) \geq 2/(N+1)$.) Therefore, $\lambda(t) \leq H(T)$ for all t in $[0, T]$. Using (43), we obtain $\gamma(T) \leq H(T)$.

In this paper we have constructed a basic model in which the impact of the ETS – on profits, costs, emissions, and output – can be analyzed. We have shown that, under quite natural conditions, the adverse profit impact of the ETS, as reflected in the level of PNA, is limited. We hope that this analysis will inform the public discussion of such schemes as they are implemented in different parts of the world. Our results could also provide a starting point for further theoretical and empirical studies.

Appendix A

Proof of Proposition 4: Let $\hat{q}_i(Q_{-i}, t)$ be the best response of firm i when the other firms are producing Q_{-i} and its marginal cost is $c_i(t)$. Abusing notation, let $MR_i(q_i, Q_{-i})$ denote firm i's marginal revenue when its output is q_i and the other firms are producing Q_{-i}. The first-order condition guarantees that $MR_i(\hat{q}_i, Q_{-i}) = c_i(t)$. Differentiating this equation by t and evaluating it at $t = 0$, we obtain

$$\frac{\partial \hat{q}_i}{\partial t} = \frac{z_i}{\partial MR_i/\partial q_i} = \frac{z_i}{2P' + q_i P''} \tag{46}$$

Differentiating the same equation by Q_{-i}, we obtain

$$\frac{\partial \hat{q}_i}{\partial Q_{-i}} = -\frac{\partial MR_i/\partial Q_{-i}}{\partial MR_i/\partial q_i} = -1 + \frac{P'}{2P' + q_i P''}. \tag{47}$$

At equilibrium, $\hat{q}_i(Q^*_{-i}, t) + Q^*_{-i} \equiv Q^*$. Differentiating this with respect to t and using (46) and (47), we obtain

$$\frac{dQ^*_{-i}}{dt} = \frac{dQ^*}{dt}\left[2 - \sigma_i E(Q^*)\right] - \frac{z_i}{P'}. \tag{48}$$

Summing this equation across firms gives us

$$(N-1)\frac{dQ^*}{dt} = \frac{dQ^*}{dt}[2N - E(Q^*)] - \frac{\sum_{i=1}^N z_i}{P'}. \tag{49}$$

Rearranging now yields (17). Using (17) to substitute for P' in (48), we obtain (18). Since $dq^*_i/dt = dQ^*/dt - dQ^*_{-i}/dt$,(19) may be derived from (17) and (18). QED

Proof of Proposition 5: Appealing to (17) and (19), we can write

$$\frac{d\sigma_i}{dt}(0) = \frac{1}{Q^*(0)}\frac{dQ^*}{dt}\left[-1 + \sigma_i\left[E(Q^*) - 1\right] + \left[N + 1 - E(Q^*)\right]\frac{z_i}{\sum_{i=1}^N z_i}\right]. \tag{50}$$

Since σ_i decreases with i, it is clear that $d\sigma_i/dt$ decreases with i if $E(Q^*) < 1$ and z_i increases with i. This implies that $\{d\sigma_i/dt\}_{1\le i\le N}$ has the single crossing property.

By definition, for every firm j in $\mathcal{J} = \{1, 2, ..., k\}$ we have $d\sigma_j/dt \geq 0$. In particular, $d\sigma_k/dt \geq 0$, which is possible if and only if

$$-1 + \sigma_k\left[E(Q^*) - 1\right] + \left[N + 1 - E(Q^*)\right]\frac{z_k}{\sum_{i=1}^N z_i} \leq 0 \tag{51}$$

(since $dQ^*/dt < 0$). If $\sigma_k < 1/N$, (51) implies that

$$-1 + \frac{1}{N}\left[E(Q^*) - 1\right] + \left[N + 1 - E(Q^*)\right] \frac{z_k}{\sum_{i=1}^{N} z_i} \leq 0. \tag{52}$$

Dividing this inequality by $[N + 1 - E(Q^*)] > 0$, we obtain $z_k \leq \sum_{i=1}^{N} z_i/N$ and, by monotonicity, $z_j \leq \sum_{i=1}^{N} z_i/N$ for all j in $\mathcal{J}$. If this property does not hold, then $\sigma_k \geq 1/N$, which implies $\sigma_j \geq 1/N$ for all firms in $\mathcal{J}$. This completes the proof of part (i).

Part (ii) of the proposition relies on the following lemma.[29]

Lemma A1: *Suppose that $\{b_i\}_{1\leq i\leq N}$ obeys the single crossing property with the crossing at $i = k$, i.e., $b_k \geq 0$ and $b_{k+1} < 0$. Let $\{a_i\}_{1\leq i\leq N}$ be a collection of positive numbers that are increasing in i. Then*

$$\sum_{i=1}^{N} a_i b_i \leq a_k \left[\sum_{i=1}^{N} b_i\right] \tag{53}$$

Proof: Since $b_i \geq 0$ and $a_i \leq a_k$ for $i \leq k$, we must have $\sum_{i=1}^{k} a_i b_i \leq a_k \sum_{i=1}^{k} b_i$. Similarly, for $i > k$, $b_i < 0$ and $a_i > a_k$, so we obtain $\sum_{i=k}^{N} a_i b_i \leq a_k \sum_{i=k}^{N} b_i$. Adding up these two inequalities give us (53). QED

Returning to the proof of Proposition 5(ii), we can now apply Lemma A1 since $\{d\sigma_i/dt\}_{1\leq i\leq N}$ obeys the single crossing property and $c_i(0)$ increases with i. Then

$$\frac{d\underline{c}^*}{dt}(0) \leq \underline{c}_k(0) \left[\sum_{i=1}^{N} \frac{d\sigma_i}{dt}\right] = 0, \tag{54}$$

where the second equality follows from the fact that $\sum_{i=1}^{N} \sigma_i(t) \equiv 1$. QED

Proof of Proposition 6: Since $dQ^*/dt < 0$ (see (17)), $d\zeta^*/dt \leq 0$ if $dz^*/dt \leq 0$. For case (a), we have already established that $dz^*/dt \leq 0$ in the main part of the paper. So we turn to case (b). From Proposition 5 we know that the output response obeys

[29] In one guise or another, Lemma A1 and its simple proof are well-known.

the single crossing property. So

$$\frac{dz^*}{dt}(0) \le \sum_{i=1}^{N} z_i(0)\frac{d\sigma_i}{dt}(0) \le z_K(0)\left[\sum_{i=1}^{N}\frac{d\sigma_i}{dt}\right] = 0, \tag{55}$$

where the first inequality follows from (26) and the second from Lemma A1 (with $a_i = z_i(0)$). QED

Proof of Proposition 10: Since σ_i is decreasing in i while z_i is increasing in i (by monotonicity), $\sum_{i=1}^{N} z_i\sigma_i \le (\sum_{i=1}^{N} z_i)/N$. Furthermore, (14) guarantees that $2 - HE(Q^*) = \sum_{i=1}^{N}\sigma_i(2-\sigma_i E(Q^*)) > 0$. Therefore, using(35), we obtain (36). We may rewrite $\tilde{\beta}$ as

$$2 - NH - N(N+1)\frac{[2/(N+1) - H]}{[N+1-E(Q^*)]}. \tag{56}$$

If $H \le 2/(N+1)$, this term is always less than $2 - NH$, so we obtain (i). (Note that $N + 1 - E(Q^*) > 0$ by (15).) If $H \ge 2/(N+1)$, this term increases with $E(Q^*)$, so we may replace $E(Q^*)$ with $\widehat{E} = \max\{1, E(Q^*)\}$ to obtain

$$\tilde{\gamma} \le \tilde{\beta} \le \hat{\beta} \equiv 2 - \frac{N\left[2 - H\widehat{E}\right]}{\left[N+1-\widehat{E}\right]}. \tag{57}$$

Since $E(Q^*)$ obeys (15), we also have $N + 1 - \widehat{E} > 0$ and since it obeys (14) we obtain $2 - H\widehat{E} = \sum_{i=1}^{N}\sigma_i(2-\sigma_i\widehat{E}) > 0$. Given these and the fact that $\widehat{E} \ge 1$ by construction, it is easy to check that $\hat{\beta}$ (as defined in (57)) is increasing in N and has a supremum of $H\widehat{E}$. QED

The proof of Proposition 11 requires the following lemma, which gives the different senses in which larger firms gain market share when market demand is log-concave.

LEMMA A2: *Under Case B, the following hold for t and t' in $[0, T]$ with $t < t'$: (i) $\sigma_i(t') - \sigma_i(t)$ is decreasing in i; (ii) for any k, we have $\sum_{i=1}^{k}\sigma_i(t') \ge \sum_{i=1}^{k}\sigma_i(t)$; and (iii) $H(t') \ge H(t)$.*

Proof: (i) is equivalent to having $d\sigma_i/dt$ decreasing in i (at every t). To see that the latter is true, we need only adapt the argument used to prove Proposition 5. To prove (ii), we apply Lemma A1, with $a_i = -1$ for $i \leq k$ and $a_i = 0$ for $i > k$. This gives us

$$-\frac{d}{dt}\left[\sum_{i=1}^{k}\sigma_i(t)\right] \leq 0, \tag{58}$$

from which we obtain (ii). To prove (iii), note that

$$-\frac{dH}{dt}(t) = -\sum_{i=1}^{N} 2\sigma_i(t)\,\frac{d\sigma_i}{dt}(t). \tag{59}$$

Set $a_i = -2\sigma_i(t)$ and apply Lemma A1 to obtain $dH/dt \geq 0$. QED

Proof of Proposition 11: It suffices to show that $\zeta^*(t)$ decreases with t and $H(t)$ increases with t. The latter claim is Lemma A2(iii). Lemma A2(i) guarantees that at any t, $d\sigma_i/dt$ is decreasing in i, which implies that it obeys the single crossing property. It is now straightforward to adapt the argument used in Proposition 6 to obtain $d\zeta^*/dt \leq 0$ at t in $[0, T]$. QED

Appendix B

There are five certified types of cement—Portland cement, Portland blast furnace cement, sulphate-resisting cement, masonry cement, and Portland pulverized fuel ash cement—which we group together because they are manufactured with essentially the same process (Environment Agency, 2005). The UK cement market is dominated by the four members of the British Cement Association: Lafarge Cement UK (previously Blue Circle), Castle Cement (owned by Heidelberg Cement), Cemex (previously Rugby Cement) and Buxton Lime Industries. These four firms collectively produce around 90% of the cement sold in the UK, with approximate market shares of 40%, 25%, 20% and 5% (Environment Agency, 2005). Imports from four other firms (all manufacturing within the EU and subject to the EU ETS) supply the remainder. This gives a Herfindahl index of around H =0.28 and with a number of firms $N = 8$.

Estimates of the price elasticity of demand for cement in the UK do not seem to be readily available. Jans and Rosenbaum (1997) find an average elasticity of demand of 0.80 for cement industry in the U.S. More recently, Ryan (2005) finds an elasticity of 2.95 from US market-level data on prices and quantities. While noting this is a rather high estimate, he argues that it is consistent with data on profit margins and plant costs. Finally, Roller and Steen (2005) find a short-run elasticity of 0.46 and a corresponding long-run elasticity of 1.47 for the Norwegian market. For our calculations, we employ price elasticities of 0.5 (low), 0.8 (best guess) and 2.0 (high).

References

[1] Alexander, Gordon and Buchholz, Rogene (1978). Corporate Social Responsibility and Stock Market Performance. *Academy of Management Journal*, 21(3) pp 479-486.

[2] Baumol, William and Oates, Wallace (1988). *Theory of Environmental Policy, 2nd edition.*, Cambridge: Cambridge University Press.

[3] Bagnoli, Mark and Bergstrom, Ted (2005). Log-concave probability and its applications. *Economic Theory* 26, 445469.

[4] Bergstrom, Ted and Varian, Hal (1985). When Are Nash Equilibria Independent of the Distribution of Agents' Characteristics? *Review of Economics Studies*, 52(4), 715-718.

[5] Bovenberg, Lans; Goulder, Lawrence; and Gurney, Derek (2005). Efficiency Costs of Meeting Industry-Distributional Constraints Under Environmental Permits and Taxes. *RAND Journal of Economics*, 36, 950–970.

[6] Bulow, Jeremy; Geanakoplos, John; and Klemperer, Paul (1985). Multimarket Oligopoly: Strategic Substitutes and Complements. *Journal of Political Economy*, 93(3), 488-511.

[7] Environment Agency (2005). Measuring Environmental Performance: Sector Report for the Cement Industry. Bristol.

[8] Farrell, Joseph and Shapiro, Carl (1990). Horizontal Mergers: An equilibrium analysis. *American Economic Review*, 80, 107-126.

[9] Fullerton, Don and Metcalf, Gilbert (2001). Environmental Controls, Scarcity Rents, and Pre-Existing Distortions. *Journal of Public Economics*, 80, 249–267.

[10] Fullerton, Don and Metcalf, Gilbert (2002). Tax Incidence. In: A. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*. Amsterdam: Elsevier, 1787–1872.

[11] Hahn, Robert (1985). Market power and transferable property rights. *Quarterly Journal of Economics*, 99, 753–765.

[12] Hahn, Robert and Noll, Roger (1982). Designing an efficient permits market. In *Implementing tradeable permits for sulfur oxide emissions* (vol. II), Cass, Glen et al. (eds.) Pasadena: CALTECH, 102-134.

[13] Heal, Geoffrey (2008). *When Principles Pay: Corporate Social Responsibility and the Bottom Line.* New York: Columbia University Press.

[14] Helfand, Gloria (1991). Standards vs standards: the effects of different pollution restrictions. *American Economic Review*, 81(3), 622–634.

[15] Hepburn, Cameron (2006). Regulating by prices, quantities or both: an update and an overview, *Oxford Review of Economic Policy*, 22:2, 226–247.

[16] Jans, Ivette and Rosenbaum, David (1997). Multimarket Contact and Pricing: Evidence from the U.S. Cement Industry. *International Journal of Industrial Organization*, 15, 391–412.

[17] Joskow, Paul and Schmalensee, Richard. (1998). The Political Economy of Market-Based Environmental Policy: The U.S. Acid Rain Programme. *Journal of Law and Economics*, 41(1), 37–84.

[18] Joskow, Paul; Schmalensee, Richard; and Bailey, Elizabeth (1998). The Market for Sulfur Dioxide Emissions, *American Economic Review*, 88(4), 669–685.

[19] Kimmel, Sheldon (1992). Effects of Cost Changes on Oligopolists' Profits. *Journal of Industrial Economics*, 40, 441–449.

[20] King, Andrew and Lenox, Michael (2001). Does it Really Pay to be Green? *Journal of Industrial Ecology*, 5(1), 105-116.

[21] Konar, Shameek and Cohen, Mark (2001). Does the Market Value Environmental Performance? *Review of Economics and Statistics*, 83(2), 281-289.

[22] Levin, Dan (1985). Taxation Within Cournot Oligopoly. *Journal of Public Economics*, 27, 81-290.

[23] Liski, Matti and Montero, Juan-Pablo (2006). On pollution permit banking and market power, *Journal of Regulatory Economics*, 29, 283–302.

[24] Montero, Juan-Pablo (1999). Voluntary Compliance with Market-Based Environmental Policy: Evidence from the U.S. Acid Raid Program. *Journal of Political Economy*, 107(5), 998–1033.

[25] Montgomery, David (1972). Markets in licences and efficient pollution control programs. *Journal of Economic Theory*, 5, 395–418.

[26] Pizer, William (2002). Combining price and quantity controls to mitigate global climate change, *Journal of Public Economics*, 85, 409-434.

[27] Porter, Michael and van der Linde, Claas (1995). Green and Competitive: Ending the Stalemate. *Harvard Business Review*, 73(5), 120-134.

[28] Roller, Lars-Hendrik and Steen, Frode (2005). On the Workings of a Cartel: Evidence from the Norwegian Cement Industry. *American Economic Review*, 96, 321–338.

[29] Ryan, Stephen (2005). The Costs of Environmental Regulation in a Concentrated Industry. MIT Center for Energy and Environmental Policy Research, 05-010.

[30] Schmalensee, Richard; Joskow, Paul; Ellerman, Denny; Montero, Juan-Pablo; and Bailey, Elizabeth (1998). An Interim Evaluation of Sulfur Dioxide Trading. *Journal of Economic Perspectives*, 12(3), 53–68.

[31] Shapiro, Carl (1989). Theories of Oligopoly Behavior, in *The Handbook of Industrial Organization Vol. I*, Schmalensee, Richard and Willig, Robert (eds.). Elsevier.

[32] Stavins, Robert (1995). Transaction Costs and Tradeable Permits. *Journal of Environmental Economics and Management* 29, 133–148.

[33] Stern, Nicholas (2008). The Economics of Climate Change, *American Economic Review: Papers and Proceedings*, 98(2), 1–37.

[34] Tietenberg, Thomas (2006). *Emissions Trading: Principles and Practice*, 2nd ed. Washington, DC: Resources for the Future.

[35] Weitzman, Martin (1974). Prices vs. Quantities. *Review of Economic Studies*, 41(4), 477–491.