Prices or Quantities Dominate Banking and Borrowing

Martin L. Weitzman*

January 7, 2018 version.
Comments Appreciated.

Abstract

The possibility of intertemporal banking and borrowing of tradeable permits is often viewed as tilting the various policy debates about optimal pollution control instruments toward favoring such time-flexible quantities. The present paper shows that this view is misleading, at least for the simplest dynamic extension of the original 'prices vs. quantities' information structure. The model of this paper allows the firms to know and act upon the realization of uncertain future costs two full periods ahead of the regulators. For any given circumstance, the paper shows that either a fixed price or a fixed quantity is superior in expected welfare to time-flexible banking and borrowing. Furthermore, the standard original formula for the comparative advantage of prices over quantities contains sufficient information to completely characterize the regulatory role of intertemporal banking and borrowing. The logic and implications of these results are analyzed and discussed.

JEL Codes: Q50, Q51, Q52, Q54, Q58

Keywords: prices, quantities, prices versus quantities, regulatory instruments, pollution, climate change

---

*Department of Economics, Harvard University (mweitzman@harvard.edu). For helpful constructive comments, but without implicating them in errors, omissions, or interpretations, I am indebted to: Reyer Gerlagh, Kenneth Gillingham, Ben Groom, Garth Heutel, Ryan Kellogg, Frank Krysiak, Derek Lemoine, Jing Li, Jussi Lintunen, Ian MacKenzie, Gilbert Metcalf, Guy Meunier, Ian Parry, Cedric Philibert, Simon Quemin, Philippe Quirion, Mar Reguant, Till Requate, Richard Revesz, Andrew Schein, Richard Schmalansee, Joseph Shapiro, Jean Tirole, and Yi-Ming Wei.

1 Introduction

Choosing the best instrument for controlling pollution has been a long-standing central issue in environmental economics. This paper is primarily concerned with the theory of environmental policy in terms of the basic theoretical foundations of the instrument choice issue among the three alternatives of: fixed prices, fixed quantities, and time-flexible quantities from intertemporal banking and borrowing.[1]

Pollution is a negative externality. Pigou (1920) introduced and subsequently popularized the central concept of placing a price-charge on pollution (since called a 'Pigouvian tax') as an efficient way to correct a pollution externality. This Pigouvian-tax approach dominated economic thinking about the pollution externality problem for about the next half-century.

Dales (1968) introduced the idea of creating property rights in the form of tradeable pollution permits (aka 'allowances') as an efficient alternative to a Pigouvian tax. Montgomery (1972) proved rigorously the formal equivalence between a price on pollution emissions and a dual quantity representing the total allotment of tradeable permits. Henceforth, it became widely accepted that there is a fundamental isomorphism between a Pigouvian tax on pollution and the total quantity of allotted caps in a cap-and-trade system that ends up with all permits trading at the same competitive-equilibrium market price as the Pigouvian tax. For every given Pigouvian tax, there is a total quantity of tradeable permits allotted whose competitive-equilibrium market price equals the Pigouvian tax. And for every given total quantity of tradeable permits allotted, there is a competitive-equilibrium market price that would yield the same result if it were imposed as a Pigouvian tax. Thus far all of the analysis took place in a deterministic context where everything was known with certainty.

Weitzman (1974) showed that with uncertainty in cost and benefit functions there is no longer an isomorphism between price and quantity instruments. The key distinction under uncertainty is that setting a fixed price stabilizes marginal cost while leaving the total quantity variable, whereas setting a fixed total quantity of tradeable permits stabilizes total quantity while leaving price (or marginal cost) variable. The question then becomes: which instrument is better under which circumstances?

The Weitzman (1974) paper 'Prices vs. Quantities' derived a relatively simple formula for the 'comparative advantage of prices over quantities', denoted in the paper as Δ. The sign

[1] I sprinkle some comments on 'practicality' throughout the paper, but the main focus is on the pure economics of instrument choice. Thus, the broader dimensions of real-world political, ideological, legal, social, historical, administrative, motivational, informational, timing, lobbying, monitoring, enforcing, or other non-purely-economic issues are beyond the scope of this paper. For a balanced overall discussion of both theoretical and practical issues involved in choosing between carbon taxes and cap-and-trade, see the review article of Goulder and Schein (2013) and the many further references that they cite.

of Δ depends on the relative slopes of the marginal abatement-cost curve and the marginal abatement-benefit curve. The sign of Δ is positive (prices are favored over quantities) when the marginal benefit curve is flatter than the marginal cost curve. Conversely, the sign of Δ is negative (quantities are preferred over prices) when the marginal benefit curve is steeper than the marginal cost curve.[2]

There subsequently developed a sizable literature on the optimal choice of price vs. quantity policy instruments under uncertainty.[3] Adar and Griffin (1976), Fishelson (1976), and Roberts and Spence (1976) analyzed seemingly alternative (but ultimately similar) forms of uncertainty. Weitzman (1978), Yohe (1978), Kaplow and Shavell (2002), and Kelly (2005) extended the basic model to cover various aspects of nonlinear marginal benefits and nonlinear marginal costs. Yohe (1978) and Stavins (1995) analyzed a situation where uncertain marginal costs are correlated with uncertain marginal benefits. Chao and Wilson (1993), and Zhao (2003) incorporated investment behavior into the basic framework of instrument choice under uncertainty. In these extensions, the results seemed by and large to preserve the earlier insight that, other things being equal, flatter marginal benefits or steeper marginal costs tend to favor prices while steeper marginal benefits or flatter marginal costs tend to favor quantities.[4]

Extensions to cover stock externalities in a dynamic multi-period context were made by Hoel and Karp (2002), Pizer (2002), Newell and Pizer (2003), and Fell, MacKenzie and Pizer (2012), among others. These dynamic stock externality extensions appeared to keep much of the 'flavor' of the original Δ story, which was phrased in terms of emission flows throughout a regulatory period (after which comes a new regulatory period with new decision-relevant parameters). In particular, for the case of climate change from accumulated stocks of atmospheric carbon dioxide (CO_2), this stock-based literature concludes that, throughout the relevant regulatory period, Pigouvian taxes are strongly favored over cap-and-trade. This is because the flow of CO_2 emissions throughout a realistic regulatory period is but a tiny fraction of the total stock of atmospheric CO_2 (which actually does the damage), and therefore the corresponding marginal flow benefits of CO_2 abatement within, say, a five or ten year regulatory period, are very flat, implying that prices have a strong comparative

[2] This same 'prices vs. quantities' condition for the sign of Δ will later be shown to be a sufficient statistic for completely characterizing the welfare-maximizing regulatory role of intertemporal banking and borrowing.

[3] This literature is too large to cover here all published papers concerning various aspects of 'prices vs. quantities'. Instead, I have included here only a subset that I subjectively judged to be most relevant to this paper.

[4] But note that combinations of instruments, such as a fixed price with a floor and ceiling on quantities, must supersede in expected welfare both a pure price and a pure quantity because both of these pure instruments are special cases of such combinations of instruments. This insight traces back to Roberts and Spence (1976).

advantage over quantities.

The role of intertemporal banking and borrowing of emissions permits in a dynamic multi-period context was analyzed, or at least touched upon, in a methodologically heterogeneous series of papers[5] including Rubin (1995), Yates and Cronshaw (2001), Williams (2002), Newell, Pizer and Zhang (2005), Feng and Zhao (2006), Murray, Newell, and Pizer (2009), Fell, MacKenzie, and Pizer (2012), and Pizer and Prest (2016). Conclusions of this heterogeneous collection of banking-and-borrowing papers are mixed, but several are favorable to intertemporal banking and borrowing of permits (or allowances).

The papers cited above do not fully and completely confront the main themes of this paper.[6] The two-period model of this paper features symmetric linear marginal-cost and marginal-benefit functions in each period, along with an entirely general joint probability specification of the two marginal-cost-shifting random variables of each period. This paper shows exactly how the *quantitative* formulas for comparative advantage (between each different pairwise combination of the three regulatory instruments being analyzed) depend upon the slopes of the marginal benefit and marginal cost functions, and also upon the probability distribution of two-period uncertainty, as well as other parameters of the model. Such quantitative formulas could be used to examine, for example, the magnitude of the pairwise welfare implications of the wrong policy, a task that is left for future research. Instead, the analysis of this paper is focused primarily on providing a complete *qualitative* welfare ordering among the three instruments of fixed prices, fixed quantities, and intertemporal banking and borrowing – which welfare ordering depends only on the relative slopes of the marginal benefit and marginal cost functions and is independent of other parameters of the model or the joint probability distribution of two-period uncertainty.

For any given set of circumstances, the model of this paper shows that *either* a fixed price *or* a fixed quantity is superior in expected welfare to time-flexible banking and borrowing. (For example, in instances where banking and borrowing dominates fixed quantities in expected welfare, then fixed prices dominate both.) Furthermore, the paper shows that the standard original formula for the comparative advantage of prices over quantities Δ contains sufficient information to completely characterize the welfare-maximizing regulatory role of intertemporal banking and borrowing relative to fixed prices or fixed quantities. The logic and implications of these seemingly counterintuitive results are analyzed and discussed.

[5] The literature concerning intertemporal banking and borrowing is too large to cover here all papers related to this subject. Instead, I have included here only a subset that I subjectively judged to be most relevant to this paper.

[6] There is a partial overlap in the methodology and conclusions of some of the above-cited papers with this paper. Williams (2002), in particular, overlaps with this paper in several of his conclusions, although his methodology is very different by depending upon an insightful analogy between spatial and temporal formulations.

2 The Model

The emphasis in the model of this paper is on clarity of exposition and the appealing simplicity of clean crisp analytical results. Hopefully the model preserves enough of 'reality' to give some useful insights, if only at a fairly high level of abstraction.

For the sake of preserving a unified familiar notation, we follow the standard convention that goods are good. This means that rather than dealing with polluted air, we deal instead with its negative – clean air. And instead of dealing with pollution emissions, we deal instead with its negative – emissions reduction (or abatement). This standard convention implies, under the usual assumptions, that the marginal benefit curve is familiarly downward sloping and the marginal cost curve is familiarly upward sloping.

We analyze a *regulatory cycle* of fixed length, at the very beginning of which *the regulator* chooses instrument values that *the firm* will respond to during the regulatory cycle. 'The regulator' is shorthand for some government regulatory agency that sets policy instruments to maximize overall social welfare.[7] 'The firm' refers to an *aggregate* of all emitting firms in the economy, all of which are united, within the same period, by facing an identical price on emissions (whether as the outcome of a Pigouvian tax, or as the outcome of the competitive market equilibrium of a cap-and-trade system, or as the shadow-price outcome of intertemporal banking and borrowing). Whatever the origin, charging a single uniform price causes the marginal costs of all firms to be equal. But this is exactly the condition for horizontal summation (of quantities across one parametric price) of all of the different firms' marginal cost (or supply) curves into one overall aggregate marginal cost (or supply) curve. This aspect justifies the rigorous aggregation of all firms into one big economy-wide as-if firm with an aggregate cost curve as just described (and where the various cost uncertainties of the various firms are aggregated into the single cost uncertainty of the single as-if firm). Henceforth, for notational and analytical convenience, this as-if aggregate firm will simply be referred to as *the* firm, with the understanding that its marginal cost function has a rigorous aggregation-theoretic underpinning in a situation where all firms face the same uniform price on pollution.

This paper analyzes and compares the expected welfare of three different regulatory instruments: prices, quantities, and time-flexible quantities. In the nomenclature of this paper, a 'price' instrument is synonymous with a fixed 'Pigouvian tax', a 'quantity' instrument is synonymous with fixed one-period 'total allocated tradeable permits', and a 'time-flexible quantity' instrument is synonymous with 'intertemporal banking and borrowing'. In each

[7] Only the efficiency side is covered here. The distribution side is outside the scope of this paper and depends, among other things, on how revenues are collected and spent.

case the regulator acts like a Stackelberg leader, setting policy instruments for the regulatory cycle given the firm's response as a function of the particular value of the regulatory instrument that has been chosen and the realization of uncertainty.

At the end of one regulatory cycle, a new regulatory cycle begins with new cost and benefit functions along with newly-updated specifications of uncertainty, and the process proceeds anew – thus allowing us, at least in principle, to analyze regulatory cycles in isolation.[8] Here we introduce a critical change in the more standard one-period structure of uncertainty, information, and decision-making in a regulatory cycle.

Any analysis of intertemporal banking and borrowing requires at least two periods, which motivates the approach taken here. The regulatory cycle is divided into two periods of equal duration. Period one is henceforth denoted period**1**, while period two is henceforth denoted period**2**. Costs are uncertain in both periods. The regulator sets policy instruments at the beginning of the regulatory cycle just before period**1** when it knows only the distribution of the cost-shifting random variables in period**1** and period**2**, and must then live with the consequences of the realizations of uncertainty throughout the entire two-period regulatory cycle.[9] By contrast, the time-flexible firm here enjoys a situation of knowing already at the very start of period**1** the realizations of the cost-shifting random variables in both period**1** and period**2** – and can intertemporally bank or borrow accordingly beginning in period**1**.[10]

What follows in this section is the specification of the overarching two-period model framework that will subsequently be applied to the analysis of all three regulatory instruments of fixed prices, fixed quantities, and, most critically, time-flexible quantities.

Turning first to period**1**, let q_1 be any value of the quantity (here period**1** pollution abatement). Consider a quadratic specification of period**1** costs of the form

$$C_1(q_1;\theta_1) = f_1(\theta_1) + (\gamma + \theta_1)(q_1 - \overline{q}) + \frac{c}{2}(q_1 - \overline{q})^2, \tag{1}$$

[8] I can only hope in this paper that, as is often the case in economic theory, an analytically-tractable flow model, standing in for a more complicated stock-flow situation, is capable of offering some useful insights. I have in mind here a regulatory cycle of maybe five to ten years or so, which is short enough to justify the model specification here, yet is long enough to allow banking and borrowing. The model of this paper abstracts away from highly durable capital investments, as well as long-term innovation and stock externalities. These could in principle be covered by a fully-dynamic multi-period stock-flow model, but the lack of analytical solutions for such a complicated formulation is a big drawback and I opt here for a surrogate flow model.

[9] Thus, importantly throughout the formulation here, the regulator knows the *structure* of this model of uncertainty, if not the *realization* of the uncertainty. It is undoubtedly more accurate to say that the regulator knows how the firm would react to various realizations of uncertainty than to say that the regulator can write down the structure of the uncertain cost functions, even though, at least in principle, the two concepts are logically equivalent.

[10] An alternative interpretation of the banking-and-borrowing setup here is that both the regulator and the firm learn the realization of two-period cost uncertainty at the same time, but there is a regulatory lag for the regulator, whereas, hypothetically, perfect futures markets with perfect information in principle would allow the firm to realize any price consequences immediately.

where the known coefficients γ and c are positive, while the known value of $\overline{q}$ will be explained later, and θ_1 is a random shift-variable normalized so that $E[\theta_1] = 0$.[11] The stochastic function $f_1(\theta_1)$ represents a pure vertical displacement of the period**1** total-cost curve. The period**1** marginal cost corresponding to (1) is the linear form

$$MC_1(q_1; \theta_1) = \gamma + \theta_1 + c\,(q_1 - \overline{q}), \tag{2}$$

which means that the random variable θ_1 with $E[\theta_1] = 0$ represents a pure neutral shift of the period**1** marginal cost function.

Turning next to the period**2** cost function, for analytical convenience we assume the same structure as period**1** except that the period**2** random variable θ_2 replaces the period**1** random variable θ_1. Thus, for quantity q_2 (here period**2** pollution abatement) the quadratic specification of period**2** costs is of the form

$$C_2(q_2; \theta_2) = f_2(\theta_2) + (\gamma + \theta_2)(q_2 - \overline{q}) + \frac{c}{2}(q_2 - \overline{q})^2, \tag{3}$$

where the known coefficients γ and c are positive, while the known value of $\overline{q}$ will be explained later, and θ_2 is a random shift-variable normalized so that $E[\theta_2] = 0$. The stochastic function $f_2(\theta_2)$ represents a pure vertical displacement of the period**2** total-cost curve. The period**2** marginal cost corresponding to (3) is the linear form

$$MC_2(q_2; \theta_2) = \gamma + \theta_2 + c\,(q_2 - \overline{q}), \tag{4}$$

which means that the random variable θ_2 with $E[\theta_2] = 0$ represents a pure neutral shift of the period**2** marginal cost function.

We specify here a fully general relationship between the random variables θ_1 and θ_2. The joint probability distribution of (θ_1, θ_2) is completely unrestricted (except for the normalization $E[\theta_1] = E[\theta_2] = 0$).[12] Some special cases of this generality include: all uncertainty concentrated in period**2** ($\theta_1 \equiv 0$), all uncertainty concentrated in period**1** ($\theta_2 \equiv 0$), and (θ_2, θ_1) following an AR(1) stochastic process.

For any value of the quantity q_1, the quadratic specification of benefits in period**1** is taken to be of the deterministic form:

$$B_1(q_1) = g_1 + \beta(q_1 - \overline{q}) - \frac{b}{2}(q_1 - \overline{q})^2, \tag{5}$$

[11] For any random variable z, the notation $E[z]$ represents the expected value of z.

[12] Note that while the *unconditional* expectation of θ_2 (namely $E[\theta_2]$) is normalized to be zero, the expectation of θ_2 *conditional on* θ_1 (namely $E[\theta_2 \mid \theta_1]$) is *not* restricted to be zero.

where all coefficients are known (with β and b positive, while $\overline{q}$ will be explained later). For any value of q_1, the deterministic period**1** marginal benefit corresponding to (5) is of the linear form

$$MB_1(q_1) = \beta - b\,(q_1 - \overline{q}). \tag{6}$$

Analogously, for any value of the quantity q_2, the quadratic specification of benefits in period**2** is taken to be of the deterministic form:

$$B_2(q_2) = g_2 + \beta\,(q_2 - \overline{q}) - \frac{b}{2}(q_2 - \overline{q})^2, \tag{7}$$

where all coefficients are known (with β and b positive, while $\overline{q}$ will be explained later). For any value of q_2, the deterministic period**2** marginal benefit corresponding to (7) is of the linear form

$$MB_2(q_2) = \beta - b\,(q_2 - \overline{q}). \tag{8}$$

We could readily consider an extension of the benefit functions to include uncertainty, where the random variable η_1 with $E[\eta_1] = 0$ represents a pure neutral shift of the period**1** marginal benefit function, while the random variable η_2 with $E[\eta_2] = 0$ represents a pure neutral shift of the period**2** marginal benefit function. Since the unknown factors connecting q_1 with C_1 and q_2 with C_2 are likely to be quite different from those linking q_1 to B_1 and q_2 to B_2, it seems plausible to assume as a base case that the random variables η_1 and η_2 are independently distributed from the random variables θ_1 and θ_2.[13] So long as η_1 and η_2 are independent of θ_1 and θ_2, careful, if tedious, calculations show that the influence of η_1 and η_2 has zero effect on all of the comparative-advantage formulas to be derived throughout this paper. This zero influence of η_1 and η_2 on all relevant formulas occurs here because uncertainty in η_1 and η_2 that is independent of θ_1 and θ_2 affects all regulatory instruments equally adversely, which is a well-known effect throughout the 'prices vs. quantities' literature. Thus, with independent uncertainty between marginal benefit and marginal cost functions, there is no loss of generality (in making pairwise welfare comparisons between binary instrument choices) by considering benefits here to be deterministic.

The value $\overline{q}$ is intended to represent both the period**1** quantity where period**1** marginal benefit equals period**1** expected marginal cost, or

$$MB_1(\overline{q}) = E[MC_1(\overline{q}; \theta_1)], \tag{9}$$

[13] The condition $E[\theta\eta] = 0$ seems like an especially relevant abstraction for the case of climate change. The situation where θ and η have non-zero correlation in the standard one-period model is considered in the footnote on page 485 of Weitzman (1974) and explicated in more detail in Yohe (1978) and Stavins (1995).

and also, simultaneously, the period**2** quantity where period**2** marginal benefit equals period**2** expected marginal cost, or

$$MB_2(\overline{q}) = E[MC_2(\overline{q};\theta_2)]. \tag{10}$$

With the normalization of $\overline{q}$ implied by (9) and (10), and making use of formulas (6), (2), (8), and (4) (along with $E[\theta_1] = E[\theta_2] = 0$), then (9) and (10) both imply

$$\beta = \gamma. \tag{11}$$

Henceforth we substitute γ for β in all formulas of this paper.

The above description of $\overline{q}$ allows an important elucidation of cost and benefit functions in both period**1** and period**2**. The value $\overline{q}$ is a natural focal point or point of departure where marginal benefits equal expected marginal costs in both periods. The cost functions $C_1(q_1;\theta_1)$ of (1) and $C_2(q_2;\theta_2)$ of (3) then represent rigorously a quadratic approximation of costs in a relevant neighborhood of $\overline{q}$. Analogously, the benefit functions $B_1(q_1)$ of (5) and $B_2(q_2)$ of (7) then represent rigorously a quadratic approximation of benefits in a relevant neighborhood of $\overline{q}$.

This concludes our description of the basic model.[14] Now we are ready to apply this framework to evaluate and compare the expected welfare of the three different regulatory control instruments analyzed in this paper.

3 Expected Welfare of Time-Flexible Quantities

We begin this section with a reminder that, for a two-period version of a fixed-price or fixed-quantity instrument, the firm is not required to foresee the resolution of period**2** uncertain costs at any time before period**2** because it makes no difference for the analysis whether the firm foresees or not. Period**2** adjustments by the firm to a fixed price or to a fixed quantity, which are imposed by the regulator at the beginning of the regulatory cycle just before period**1**, need occur only during period**2** (after the firm learns the realization of the random variable θ_2), not earlier. By way of sharp contrast, to make the strongest case for banking and borrowing the model is seemingly most favorable when the firm foresees the resolution of period**2** uncertain costs (as well as the resolution of period**1** uncertain costs)

[14] It goes without saying that we are imposing a tremendous amount of structure on the basic model (including linear-quadratic costs and benefits, a two-period analysis, independence of regulatory cycles, additively-separable costs and benefits each having essentially the same structure in both periods, additive marginal-cost shocks, statistical independence of costs and benefits, a highly simplified information and decision structure, a purely static flow analysis, and so forth and so on). Absent such kind of structure it is difficult to address simply and analytically the basic issues of this paper.

instantaneously at the very beginning of period**1**.

By time-flexible quantities, we mean here the ability of the firm to intertemporally bank and borrow freely between the two periods within the regulatory cycle on the basis of perfect-foresight knowledge of costs in both periods. After the end of one regulatory cycle, there begins a new regulatory cycle with new regulations based on new coefficients of cost and benefit functions along with new joint probability distributions, and the process proceeds anew.[15] The regulator sets its control instruments just before the beginning of the regulatory cycle, which here means just before the start of period**1**. At that initial instant, the regulator comprehends θ_1 and θ_2 just as probability *distributions* that will make their realizations known only during period**1** (for θ_1) and period**2** (for θ_2), well after regulatory policy has already been set. The time-flexible firm, by distinction, here is allowed to have perfect foresight throughout by knowing already at the very beginning of period**1** the *realizations* of the random variables θ_1 in period**1** and θ_2 in period**2**. By striking contrast with this model of banking and borrowing, both a fixed-price instrument and a fixed-quantity instrument are far less informationally demanding of the firm by *not* requiring perfect-foresight knowledge of the period**2** resolution of uncertainty θ_2 at any earlier time before period**2** arrives.

In this model of time-flexible quantities the firm can freely bank or borrow intertemporally between period**1** and period**2** as it sees fit on the basis of its super-prescient knowledge of θ_1 and θ_2. This strong informational asymmetry would appear to stack the odds in favor of the expected welfare of intertemporal banking and borrowing. Time-flexible quantities are made to appear more attractive than they actually are in practice because, in reality, firms have at best imperfect foresight and, while they are able to bank permits in many regulatory situations, they only rarely can borrow. The artificially-enhanced attractiveness of perfect-foresight intertemporal banking and borrowing in this model multiplies the force of the two basic counterintuitive results from this paper that (for any given circumstance) either fixed prices or fixed quantities are superior in expected welfare to intertemporal banking and borrowing, and that the standard original formula for the comparative advantage of prices over quantities Δ contains sufficient information to completely characterize the appropriate regulatory role of time-flexible quantities.

[15] Unfettered banking and borrowing is allowed here *within* a regulatory cycle but not *across* regulatory cycles, because in a new regulatory cycle there are new regulations based on new information. Pizer and Prest (2016) analyze a model with policy updating based on banking and borrowing across regulatory cycles, in which the regulator and the firm are both effectively locked together within a mutual rational-expectations-type equilibrium where each is acting contingent upon a self-fulfilling expectation of how the other is acting. This forceful assumption of a joint rational-expectations equilibrium represents a significantly higher order of abstraction about 'rational' human behavior than the model here, and leads to quite different and, typical of rational expectations, very strong policy implications. By contrast, in this paper the regulator is merely assumed to be a Stackelberg leader within the regulatory cycle.

At the very instant before the regulatory cycle, when the regulator comprehends θ_1 and θ_2 just as probability distributions that will make their realizations known only during period**1** (for θ_1) and period**2** (for θ_2), let the regulator here set a *total* quantity of permits Q covering both periods. With unlimited banking and borrowing by the super-prescient firm, the form in which Q is assigned is inessential. For example, the regulator might specify an assigned number of permits q_1' for period**1** and an assigned number of permits q_2' for period**2**. With the ability to bank or borrow unlimited amounts under perfect foresight, the firm cares only about the two-period total quantity of permits $Q = q_1' + q_2'$. Henceforth we deal only with Q.

Given Q and the firm's perfect-foresight super-prescient knowledge of the realizations of the random variables θ_1 and θ_2 already at the very beginning of period**1**, the firm (postulated as having the unlimited ability to bank or borrow) faces the problem of optimally determining $\widehat{q}_1(Q;\theta_1,\theta_2)$ in period**1** and $\widehat{q}_2(Q;\theta_1,\theta_2) = Q - \widehat{q}_1(Q;\theta_1,\theta_2)$ in period**2**. The firm therefore seeks to minimize the total costs of both periods, conditional on knowing the realizations of θ_1 and θ_2, which means that it minimizes over $\widehat{q}_1$ the expression[16]

$$C_1(\widehat{q}_1;\theta_1) + C_2((Q - \widehat{q}_1);\theta_2). \tag{12}$$

Let the solution of the firm minimizing over $\widehat{q}_1$ the total-cost expression (12) be denoted $\widehat{q}_1^*(Q;\theta_1,\theta_2)$. Then $\widehat{q}_1^*(Q;\theta_1,\theta_2)$ must satisfy the first order condition[17]

$$MC_1(\widehat{q}_1^*(Q;\theta_1,\theta_2);\theta_1) = MC_2((Q - \widehat{q}_1^*(Q;\theta_1,\theta_2));\theta_2). \tag{13}$$

Substituting $\widehat{q}_1^*(Q;\theta_1,\theta_2)$ into the MC_1 formula (2) yields

$$MC_1(\widehat{q}_1^*(Q;\theta_1,\theta_2);\theta_1) = \gamma + \theta_1 + c\,(\widehat{q}_1^*(Q;\theta_1,\theta_2) - \overline{q}). \tag{14}$$

Substituting $\widehat{q}_2^*(Q;\theta_1,\theta_2) = Q - \widehat{q}_1^*(Q;\theta_1,\theta_2)$ into the MC_2 formula (4) yields

$$MC_2((Q - \widehat{q}_1^*(Q;\theta_1,\theta_2));\theta_2) = \gamma + \theta_2 + c\,(Q - \widehat{q}_1^*(Q;\theta_1,\theta_2) - \overline{q}). \tag{15}$$

Next plug (14) into the left hand side of equation (13) and plug (15) into the right hand side of (13). Then, after carefully cancelling, consolidating, and rearranging terms, the

[16]For simplicity, we assume throughout that there is no discounting by the firm within the two-period regulatory cycle. Thus, the intertemporal trading ratio $\theta = 0$ in the notation of Feng and Zhao (2006). More generally, if the intertemporal trading ratio equals the regulator's discount factor, then all of the results of the paper go through.

[17]As if there is a single shadow price for tradeable permits covering emissions from both periods.

substituted version of (13) reduces to

$$\widehat{q}_1^*(Q;\theta_1,\theta_2) = \frac{Q}{2} + \frac{\theta_2-\theta_1}{2c}, \tag{16}$$

which implies

$$\widehat{q}_2^*(Q;\theta_1,\theta_2) = \frac{Q}{2} - \frac{\theta_2-\theta_1}{2c}. \tag{17}$$

Thus far the regulator imposes a value of Q at the instant just before period**1** (when it knows θ_1 and θ_2 only as probability distributions) and the super-prescient firm reacts with (16) and (17). Next, the welfare-maximizing regulator seeks to maximize over Q the welfare expression[18]

$$E[B_1(\widehat{q}_1^*(Q;\theta_1,\theta_2))]-E[C_1(\widehat{q}_1^*(Q;\theta_1,\theta_2);\theta_1)]+E\left[B_2\left(\widehat{q}_2^*(Q;\theta_1,\theta_2)\right)\right]-E\left[C_2(\widehat{q}_2^*(Q;\theta_1,\theta_2);\theta_2)\right]. \tag{18}$$

Let the value of Q that maximizes (18) be denoted Q^*. Then Q^* must satisfy the first-order condition

$$\begin{aligned}&(E[MB_1(\widehat{q}_1^*(Q^*;\theta_1,\theta_2))] - E[MC_1(\widehat{q}_1^*(Q^*;\theta_1,\theta_2);\theta_1)])\left(\frac{\partial \widehat{q}_1^*}{\partial Q}\right)\\ &+(E\left[MB_2\left(\widehat{q}_2^*(Q^*;\theta_1,\theta_2)\right)\right] - E\left[MC_2(\widehat{q}_2^*(Q^*;\theta_1,\theta_2);\theta_2)\right])\left(\frac{\partial \widehat{q}_2^*}{\partial Q}\right) = 0.\end{aligned} \tag{19}$$

Note from (16) and (17) that

$$\frac{\partial \widehat{q}_1^*}{\partial Q} = \frac{\partial \widehat{q}_2^*}{\partial Q}\left(=\frac{1}{2}\right), \tag{20}$$

and therefore the two terms in (20) cancel out of equation (19).

Next, use the results (16) and (17). Plug (16) into formulas (6) and (2), and plug (17) into formulas (8) and (4) (paying attention to condition (11) when applicable); thereafter plug the resulting four pre-expectation terms into equation (19). Take expected values, remembering that $E[\theta_1] = E[\theta_2] = 0$. Rearrange terms to derive the basic consequence that

$$Q^* = 2\overline{q}. \tag{21}$$

[18] For simplicity, we assume throughout this paper that there is no discounting by the regulator within the two-period regulatory cycle. Thus, the discount rate $r = 0$ in the notation of Feng and Zhao (2006). More generally, if the intertemporal trading ratio equals the regulator's discount factor, then all of the results of the paper go through.

With (21) holding, (16) becomes

$$\widehat{q}_1^*(Q^*;\theta_1,\theta_2) = \overline{q} + \frac{\theta_2-\theta_1}{2c}, \tag{22}$$

while (17) becomes[19]

$$\widehat{q}_2^*(Q^*;\theta_1,\theta_2) = \overline{q} - \frac{\theta_2-\theta_1}{2c}. \tag{23}$$

Let the expected welfare of optimized time-flexible quantities with banking and borrowing here be denoted $EW(\widehat{q})$. Making use of formulas (22) and (23), we then have

$$EW(\widehat{q}) = E\left[B_1\left(\overline{q}+\frac{\theta_2-\theta_1}{2c}\right) - C_1\left(\overline{q}+\frac{\theta_2-\theta_1}{2c};\theta_1\right) + B_2\left(\overline{q}-\frac{\theta_2-\theta_1}{2c}\right) - C_2\left(\overline{q}-\frac{\theta_2-\theta_1}{2c};\theta_2\right)\right]. \tag{24}$$

Substitute carefully the four basic pre-expectation terms of the right-hand side of (24) into the corresponding equations (5) for $B_1(\overline{q}+\frac{\theta_2-\theta_1}{2c})$, (1) for $C_1((\overline{q}+\frac{\theta_2-\theta_1}{2c});\theta_1)$, (7) for $B_2\left(\overline{q}-\frac{\theta_2-\theta_1}{2c}\right)$, and (3) for $C_2((\overline{q}-\frac{\theta_2-\theta_1}{2c});\theta_2)$ (paying attention to (11) when applicable). Take the expected values of the resulting substituted-for version of (24). Then use $E[\theta_1] = E[\theta_2] = 0$ to consolidate and eliminate terms. After much careful algebra, cancellation, and rearrangement, we obtain finally the basic result

$$EW(\widehat{q}) = k + \left(\frac{E[(\theta_2-\theta_1)^2]}{4c^2}\right) \times (c-b), \tag{25}$$

where the constant k is defined by the equation

$$k = g_1 + g_2 - E[f_1(\theta_1)] - E[f_2(\theta_2)]. \tag{26}$$

4 Expected Welfare of a Fixed Quantity

In this section, assume that the regulator imposes the same fixed quantity q on output (here emissions abatement) in both period**1** and period**2**.[20] We now seek to know the regulator's expected welfare in this two-period fixed-quantity setting, which will be indispensable later for addressing, understanding, and comparing the three basic regulatory instruments being analyzed in this paper.

[19] If a reader wishes to take (22) and (23) on faith, it can be excused. However, I heartily believe that at least skimming the intermediate arguments leading up to (22) and (23) is vital for understanding what underlying logic these two expressions rest upon.

[20] It makes no difference if the regulator is allowed to set different quantities in the two different periods, since in an optimal quantity solution here they will both end up being the same anyway.

The regulator seeks to maximize over q the welfare expression

$$B_1(q) - E[C_1(q;\theta_1)] + B_2(q) - E\left[C_2\left(q;\theta_2\right)\right], \tag{27}$$

implying that the solution q^* must satisfy the first-order condition

$$MB_1(q^*) - E[MC_1(q^*;\theta_1)] + MB_2(q^*) - E\left[MC_2\left(q^*;\theta_2\right)\right] = 0. \tag{28}$$

For $q = q^*$, plug (6), (2), (8), and (4) into the left hand side of equation (28), being aware of (11). Take expected values where indicated, and make use of the condition $E[\theta_1] = E[\theta_2] = 0$ to infer (after eliminating and consolidating terms) that equation (28) then implies

$$q^* = \overline{q}. \tag{29}$$

Let $EW(\overline{q})$ stand for expected welfare in the optimally-fixed-quantity setup of this section where the basic equation (29) holds. Then $EW(\overline{q})$ equals benefits minus expected costs in period**1** plus benefits minus expected costs in period**2**. More formally, substituting (29) into (27) for $q = \overline{q}$ and taking expectations where indicated, we then have

$$EW(\overline{q}) = B_1(\overline{q}) - E[C_1(\overline{q};\theta_1)] + B_2(\overline{q}) - E\left[C_2\left(\overline{q};\theta_2\right)\right]. \tag{30}$$

Substitute the four basic terms of the right-hand side of (30) into the corresponding equations (5) for $B_1(\overline{q})$, (1) for $C_1(\overline{q};\theta_1)$, (7) for $B_2(\overline{q})$, and (3) for $C_2\left(\overline{q};\theta_2\right)$ (paying attention to (11)). Take expected values where indicated. Use $E[\theta_1] = E[\theta_2] = 0$. We then immediately obtain the basic result

$$EW(\overline{q}) = k, \tag{31}$$

where the constant k is given by equation (26).

5 Expected Welfare of a Fixed Price

Assume that the regulator imposes the same fixed price on output (here emissions abatement) in period**1** and period**2**.[21] We now seek to know the regulator's expected welfare in this two-period fixed-price setting, which will be indispensable later for addressing, understanding, and comparing the three basic regulatory instruments being analyzed in this paper.

[21] It makes no difference if the regulator is allowed to set different prices in the two different periods, since in an optimal price solution here they will both end up being the same anyway.

If the regulator imposes the price p in period**1**, the firm's quantity response is then to react with quantity $\widetilde{q}_1(p;\theta_1)$ satisfying $MC_1(\widetilde{q}_1(p;\theta_1);\theta_1) = p$. Substituting $\widetilde{q}_1(p;\theta_1)$ into (2) and inverting to solve for $\widetilde{q}_1(p;\theta_1)$ yields the period**1** quantity-reaction or supply function

$$\widetilde{q}_1(p;\theta_1) = \overline{q} + \frac{p-\gamma-\theta_1}{c}. \tag{32}$$

Likewise if the regulator imposes the price p in period**2**, the firm's quantity response is then to react with quantity $\widetilde{q}_2(p;\theta_2)$ satisfying $MC_2(\widetilde{q}_2(p;\theta_2);\theta_2) = p$. Substituting $\widetilde{q}_2(p;\theta_2)$ into (4) and inverting to solve for $\widetilde{q}_2(p;\theta_2)$ yields the period**2** quantity-reaction or supply function

$$\widetilde{q}_2(p;\theta_2) = \overline{q} + \frac{p-\gamma-\theta_2}{c}. \tag{33}$$

The regulator seeks to maximize over p the welfare expression

$$E[B_1(\widetilde{q}_1(p;\theta_1))] - E[C_1(\widetilde{q}_1(p;\theta_1);\theta_1)] + E[B_2(\widetilde{q}_2(p;\theta_2))] - E\left[C_2\left(\widetilde{q}_2(p;\theta_2);\theta_2\right)\right], \tag{34}$$

implying that the solution p^* must satisfy the first-order condition[22]

$$E[MB_1(\widetilde{q}_1(p^*;\theta_1))] - E[MC_1(\widetilde{q}_1(p^*;\theta_1);\theta_1)] + E[MB_2(\widetilde{q}_2(p^*;\theta_2))] - E\left[MC_2\left(\widetilde{q}_2(p^*;\theta_2);\theta_2\right)\right] = 0. \tag{35}$$

Plug (32) into (6) and (2). Plug (33) into (8) and (4). Then substitute the results into equation (35), being aware of (11). Take expected values of the resulting substituted-for terms in the left-hand side of (35). Make use of the conditions $E[\theta_1] = E[\theta_2] = 0$ to infer (after eliminating and consolidating terms) that equation (35) then implies

$$p^* = \gamma, \tag{36}$$

in which case (32) and (33) reduce to

$$\widetilde{q}_1(p^*;\theta_1) = \overline{q} - \frac{\theta_1}{c} \tag{37}$$

and

$$\widetilde{q}_2(p^*;\theta_2) = \overline{q} - \frac{\theta_2}{c}. \tag{38}$$

Let $EW(p)$ stand for expected welfare in the optimally-fixed-price setup of this section where the basic equations (37) and (38) both hold. Then $EW(p)$ equals expected benefits

[22] Note from (32) and (33) that $\frac{\partial \widetilde{q}_1}{\partial p} = \frac{\partial \widetilde{q}_2}{\partial p}$ $(= \frac{1}{c})$, so that these two terms have already been cancelled out of equation (35).

minus expected costs in period**1** plus expected benefits minus expected costs in period**2**. More formally, making use of (37) and (38), we have here:

$$EW(p) = E\left[B_1\left(\overline{q} - \frac{\theta_1}{c}\right)\right] - E\left[C_1\left(\overline{q} - \frac{\theta_1}{c}; \theta_1\right)\right] + E\left[B_2\left(\overline{q} - \frac{\theta_2}{c}\right)\right] - E\left[C_2\left(\overline{q} - \frac{\theta_2}{c}; \theta_2\right)\right]. \tag{39}$$

Substitute the four basic pre-expectation terms of the right-hand side of (39) into the corresponding equations (5) for $B_1(\overline{q} - \frac{\theta_1}{c})$, (1) for $C_1((\overline{q} - \frac{\theta_1}{c}); \theta_1)$, (7) for $B_2(\overline{q} - \frac{\theta_2}{c})$, and (3) for $C_2((\overline{q} - \frac{\theta_2}{c}); \theta_2)$ (paying attention to (11) whenever applicable). Take expected values in the substituted-for version of (39). Then use $E[\theta_1] = E[\theta_2] = 0$ to consolidate and eliminate terms. After much careful algebra, cancellation, and rearrangement, we obtain finally the basic result

$$EW(p) = k + \left(\frac{E[(\theta_1)^2] + E[(\theta_2)^2]}{2c^2}\right) \times (c - b), \tag{40}$$

where the constant k is given by equation (26).

6 Comparative Advantage of Fixed Prices over Fixed Quantities

In all of the comparative advantage calculations to be undertaken throughout the rest of this paper, we will employ the following useful notation. Let x be a real number. The sign of x will be denoted $sign[x]$.[23] If x and y are two real numbers with the same sign, we will write $sign[x] = sign[y]$.

Within the two-period framework of this paper, denote the comparative advantage of fixed prices over fixed quantities as $\Delta^p_{\overline{q}}$. In symbols,[24]

$$\Delta^p_{\overline{q}} = EW(p) - EW(\overline{q}). \tag{41}$$

Substituting from formulas (40) and (31) into equation (41), we then have the basic result that

$$\Delta^p_{\overline{q}} = \left(\frac{E[(\theta_1)^2] + E[(\theta_2)^2]}{2c^2}\right) \times (c - b). \tag{42}$$

Since the term $E[(\theta_1)^2] + E[(\theta_2)^2]$ of (42) is strictly positive, we are entitled here to write

$$sign\left[\Delta^p_{\overline{q}}\right] = sign\left[(c - b)\right]. \tag{43}$$

[23] The sign operator $sign[x]$ takes on one of three values: $x > 0 \Leftrightarrow sign[x] = +$; $x < 0 \Leftrightarrow sign[x] = -$; $x = 0 \Leftrightarrow sign[x] = 0$.

[24] Note that $\Delta^y_x = -\Delta^x_y$, which will be used throughout the rest of this paper.

From (43), the analysis of the two-period comparative advantage of prices over quantities, here $\Delta_{\overline{q}}^{p}$, essentially duplicates the traditional standard analysis of the one-period comparative advantage of prices over quantities Δ. The sign of $\Delta_{\overline{q}}^{p}$ is positive (prices are favored over quantities) when the marginal benefit curve is flatter than the marginal cost curve ($b < c$). Conversely, the sign of $\Delta_{\overline{q}}^{p}$ is negative (quantities are preferred over prices) when the marginal benefit curve is steeper than the marginal cost curve ($b > c$).

7 Comparative Advantage of Fixed Quantities over Time-Flexible Quantities

Within the two-period framework of this paper, denote the comparative advantage of fixed quantities over time-flexible banked-and-borrowed quantities as $\Delta_{\widehat{q}}^{\overline{q}}$. In symbols,

$$\Delta_{\widehat{q}}^{\overline{q}} = EW(\overline{q}) - EW(\widehat{q}). \tag{44}$$

Substituting from formulas (31) and (25) into equation (44), we then have the basic result that

$$\Delta_{\widehat{q}}^{\overline{q}} = -\left(\frac{E[(\theta_2 - \theta_1)^2]}{4c^2}\right) \times (c - b), \tag{45}$$

which, multiplying by -1, can be trivially rearranged into an equivalent form more useful for the purposes of this paper:

$$\Delta_{\widehat{q}}^{\overline{q}} = \left(\frac{E[(\theta_2 - \theta_1)^2]}{4c^2}\right) \times (b - c). \tag{46}$$

Since the term $E[(\theta_2 - \theta_1)^2]$ of expression (46) is strictly positive (unless $\theta_2 \equiv \theta_1$), we can write

$$sign\left[\Delta_{\widehat{q}}^{\overline{q}}\right] = sign\left[(b - c)\right]. \tag{47}$$

Condition (47) signifies that the fixed-quantity $\overline{q}$ regulatory system is preferred to the time-flexible banking-and-borrowing $\widehat{q}$ regulatory system when the slope of the marginal benefit curve is steeper than the slope of the marginal cost curve ($b > c$). Conversely, the time-flexible banking-and-borrowing $\widehat{q}$ regulatory system is preferred to the fixed-quantity $\overline{q}$ regulatory system when the slope of the marginal benefit curve is flatter than the slope of the marginal cost curve ($b < c$).

A prime question, which is begging to be answered, is why the fixed-quantity $\overline{q}$ regulatory system is *ever* preferred to the time-flexible banking-and-borrowing $\widehat{q}$ regulatory system. After all, with the *same* total output across both periods ($= 2\overline{q}$), the time-flexible $\widehat{q}$ system is ex-post intertemporally cost-efficient ($MC_1^{\widehat{q}} = MC_2^{\widehat{q}}$), whereas the fixed-quantity $\overline{q}$ system is ex-post intertemporally cost-inefficient ($MC_1^{\overline{q}} \neq MC_2^{\overline{q}}$).[25] This cost-efficiency aspect is, perhaps naturally, what most economists have instinctively focused upon in intuitively preferring time-flexible $\widehat{q}$ to fixed-quantity $\overline{q}$. Why, then, is the sign of $\Delta_{\widehat{q}}^{\overline{q}}$ not unambiguously negative, thereby indicating the regulator's dominant preference for the time-flexible banking-and-borrowing $\widehat{q}$ system over the fixed-quantity $\overline{q}$ system?[26]

The answer is that in the choice of control instruments for this model we must pay attention to the benefit side (as well as the cost side). Marginal costs in the time-flexible $\widehat{q}$ system are intertemporally equal, but marginal benefits are intertemporally unequal. Conversely, marginal benefits in the fixed-quantity $\overline{q}$ system are intertemporally equal, but marginal costs are intertemporally unequal.

While the time-flexible $\widehat{q}$ system is intertemporally cost-efficient, it increases the variability of $\widehat{q}_1^*$ in equation (22) and $\widehat{q}_2^*$ in equation (23) relative to the fixed quantity assignments of $\overline{q}_1 = \overline{q}_2 = \overline{q}$. The risk-aversion to quantity variability here is captured by the b coefficient in the quadratic total-benefit formulas (5) and (7) when $q_1 = \widehat{q}_1^*$ is given by (22) and $q_2 = \widehat{q}_2^*$ is given by (23). The more that b increases in formula (46) (or (45)), the more are the expected drawbacks of quantity variability diminishing the expected benefits minus expected costs (summed over the two periods) of variable time-flexible cost-efficient quantities relative to stable cost-inefficient fixed quantities. In other words, all else being equal, the relative advantage of stable (but intertemporally cost-inefficient) fixed quantities over variable (but intertemporally cost-efficient) time-flexible quantities increases in the risk-aversion coefficient b, and translates into a positive comparative advantage when $b > c$.

The logic of (47) is somewhat analogous to the logic of the corresponding comparative advantage formula from a 'prices vs. quantities' perspective. Time-flexible banked-and-borrowed quantities have a figurative 'price like' aspect in the sense that there is one shadow price across both periods due to intertemporal arbitrage by the firms, so that quantities adjust between periods to efficiently make all marginal costs equal to the single shadow price. In this paper, let the comparative advantage of time-flexible quantities over fixed quantities be denoted $\Delta_{\overline{q}}^{\widehat{q}}$. Then $\Delta_{\overline{q}}^{\widehat{q}} = -\Delta_{\widehat{q}}^{\overline{q}}$, implying from (47) that $sign\left[\Delta_{\overline{q}}^{\widehat{q}}\right] = sign\left[(c-b)\right]$. The corresponding formula here for the comparative advantage of fixed prices over fixed quantities $\Delta_{\overline{q}}^{p}$ from (42) implies by (43) that $sign\left[\Delta_{\overline{q}}^{p}\right] = sign\left[(c-b)\right]$. In this admittedly incomplete

[25] More precisely, $MC_1^{\overline{q}} \neq MC_2^{\overline{q}}$ except on a set of measure zero.

[26] Note that if $B(q_1, q_2) = B_{1,2}(q_1 + q_2)$ (instead of the paper's assumed form $B(q_1, q_2) = B_1(q_1) + B_2(q_2)$), then $\Delta_{\widehat{q}}^{\overline{q}} < 0$ for any values of b or c ($b \neq c$), thereby seemingly justifying the economist's first instinct. But the case $B(q_1, q_2) = B_{1,2}(q_1 + q_2)$ is presumably trying to capture stock effects, in which situation the corresponding one-period marginal flow-benefit curve is relatively flat and a price instrument dominates both fixed and time-flexible quantity instruments anyway. This issue is discussed further in footnote 29.

story, the sign of $\Delta_{q}^{\widehat{q}}$ is the same as the sign of Δ_{q}^{p} because the underlying considerations are roughly analogous – namely adjusting variable quantities between each period to have the same marginal cost, as if equal to a single price (thereby guaranteeing intertemporal cost efficiency), versus the stability of fixed quantities in each period. We elaborate on this theme later in the paper.

8 Comparative Advantage of Fixed Prices over Time-Flexible Quantities

Within the two-period framework of this paper, denote the comparative advantage of fixed prices over time-flexible banked-and-borrowed quantities as $\Delta_{\widehat{q}}^{p}$. In symbols,

$$\Delta_{\widehat{q}}^{p} = EW(p) - EW(\widehat{q}). \tag{48}$$

Substituting from formulas (40) and (25) into equation (48), we then have, after rewriting, the result

$$\Delta_{\widehat{q}}^{p} = \frac{1}{4c^2} \left\{ 2E[(\theta_1)^2] + 2E[(\theta_2)^2] - E[(\theta_2 - \theta_1)^2 \right\} \times (c - b). \tag{49}$$

Collecting and consolidating terms within the curly brackets of (49) yields

$$2E[(\theta_1)^2] + 2E[(\theta_2)^2] - E[(\theta_2 - \theta_1)^2] = E[(\theta_1)^2] + E[(\theta_2)^2] + 2E[\theta_1\theta_2] = E[(\theta_2 + \theta_1)^2], \tag{50}$$

thereby allowing (49) to be expressed in the form of a basic result that

$$\Delta_{\widehat{q}}^{p} = \left(\frac{E[(\theta_2 + \theta_1)^2]}{4c^2} \right) \times (c - b) \, . \tag{51}$$

Since the term $E[(\theta_2 + \theta_1)^2]$ in expression (51) is strictly positive (unless $\theta_2 \equiv -\theta_1$), we can write

$$sign \left[\Delta_{\widehat{q}}^{p} \right] = sign \left[(c - b) \right] . \tag{52}$$

From (52), fixed prices are preferred to time-flexible quantities with intertemporal banking and borrowing when the slope of the marginal benefit curve is flatter than the slope of the marginal cost curve ($b < c$). Conversely, time-flexible quantities with intertemporal banking and borrowing are preferred to fixed prices when the slope of the marginal benefit curve is steeper than the slope of the marginal cost curve ($b > c$).

The logic of (52) is somewhat analogous to the logic of the corresponding comparative advantage formula from a 'prices vs. quantities' perspective. Time-flexible banked-and-borrowed quantities here have a figurative 'quantity like' aspect reminiscent of fixed quantities in the sense that total output across both periods is fixed at the same total quantity of $2\overline{q}$ (whereas total output in a fixed-price system is variable). The formula here for the comparative advantage of fixed prices over fixed quantities $\Delta_{\overline{q}}^{p}$ from (42) implies by (43) that $sign\left[\Delta_{\overline{q}}^{p}\right] = sign[(c-b)]$, to be compared with the comparative advantage of fixed prices over time-flexible quantities $\Delta_{\widehat{q}}^{p}$ from (51), which implies by (52) that $sign\left[\Delta_{\widehat{q}}^{p}\right] = sign[(c-b)]$. In this admittedly incomplete story, the sign of $\Delta_{\widehat{q}}^{p}$ (from (52)) here is the same as the sign of $\Delta_{\overline{q}}^{p}$ (from (43)) because the underlying considerations are roughly analogous – namely variable total quantities (being adjusted within each period to make each period's marginal cost equal to the single fixed price of γ) versus, by comparison, the relative stability of having total quantities across both periods fixed at $2\overline{q}$. We elaborate on this theme in the next section.

9 Prices or Quantities Dominate Banking and Borrowing

The most striking result of this paper is that, for *any* given values of the model's primitive parameters $(b, c, \gamma, \overline{q}, k)$ and for *any* joint probability distribution of (θ_1, θ_2), there is a complete welfare ordering such that *either*

$$b < c \Longrightarrow EW(p) > EW(\widehat{q}) > EW(\overline{q}), \tag{53}$$

or[27]

$$b > c \Longrightarrow EW(\overline{q}) > EW(\widehat{q}) > EW(p). \tag{54}$$

The proof of conditions (53) and (54) simply relies on systematically examining the signs of the relevant comparative advantage formulas $\Delta_{\widehat{q}}^{p}$ and $\Delta_{\overline{q}}^{\widehat{q}}$ for $b < c$ in (53), along with $\Delta_{\widehat{q}}^{\overline{q}}$ and $\Delta_{p}^{\widehat{q}}$ for $b > c$ in (54). When $b < c$, then $sign[\Delta_{\widehat{q}}^{p}] > 0$ from (52), and also when $b < c$, then $sign[\Delta_{\overline{q}}^{\widehat{q}}] = -sign[\Delta_{\widehat{q}}^{\overline{q}}] > 0$ from (47) – this explains condition (53). When $b > c$, then $sign[\Delta_{\widehat{q}}^{\overline{q}}] > 0$ from (47), and also when $b > c$, then $sign[\Delta_{p}^{\widehat{q}}] = -sign\left[\Delta_{\widehat{q}}^{p}\right] > 0$ from (52) – this explains condition (54).

Conditions (53) and (54) immediately imply that *either* a fixed price *or* a fixed quantity is superior in expected welfare to intertemporal banking and borrowing. Such an outcome appears all the more surprising because the two-period model of this paper has seemingly

[27] I ignore here and throughout the razor's edge case where $b = c$, in which zero-measure situation all three instruments have the same expected welfare: $EW(p) = EW(\widehat{q}) = EW(\overline{q})$.

biased the odds in favor of intertemporal banking and borrowing by allowing the firm to know, and immediately act upon, the perfect-foresight realizations of both of the uncertain future costs a full two periods ahead of the regulator (and also, although less importantly, by allowing unlimited borrowing).[28] By stark contrast, it is important to appreciate fully that both in the fixed-price regime and, more trivially, in the fixed-quantity regime, the firm is *not* required (or even permitted) to foresee period**2** costs at any time before period**2**.

Such a prominently counterintuitive result as the dominance of prices or quantities over banking and borrowing is a mathematical surprise of the model, so that it cannot be too obvious why it is occurring. A heuristic story might be attempted along the rough lines that a time-flexible quantity acts like a sort of a metaphorical intermediate hybrid between a fixed price and a fixed quantity. Intertemporal banking and borrowing is somewhat related to a price instrument here insofar as time-flexible quantities adjust across the two periods as if by reacting to a single shadow price, thereby inducing intertemporal cost efficiency via setting marginal costs in both periods equal to the shadow price (much as a fixed price of γ induces intertemporal cost efficiency). Simultaneously, intertemporal banking and borrowing is also somewhat related to a quantity instrument here insofar as total output across both periods is fixed (at $2\overline{q}$). The mutual exclusivity of conditions (53) and (54), with $EW(\widehat{q})$ in the middle of both inequalities, also hints at an intermediate-hybrid interpretation of banking and borrowing as a kind of a metaphorical cross between (fixed) prices and (fixed) quantities.

If $b < c$, the 'prices vs. quantities' tradition suggests figuratively that the regulator should want to move away from a quantity instrument in favor of moving toward a price instrument. But then, in this heuristic story, rather than settling for the mongrel half-price hybrid banking-and-borrowing instrument, which has the other half tainted by its similarity to a here-inferior quantity-hybrid half, the regulator would prefer to go all the way to a purebred price instrument. When $b < c$, the part of the banking-and-borrowing instrument that is attractive to the regulator is the price component, while the quantity component is unattractive. Here the full-price is preferred to the half-price-half-quantity is preferred to the full-quantity. Hence the ordering (53).

Conversely, if $b > c$, the 'prices vs. quantities' tradition suggests figuratively that the regulator should want to move away from a price instrument in favor of moving toward a quantity instrument. But then, in this heuristic story, rather than settling for the mongrel half-quantity hybrid banking-and-borrowing instrument, which has the other half tainted by its similarity to a here-inferior price-hybrid half, the regulator would prefer to go all the way

[28] In reality, as has been pointed out, firms have at best imperfect foresight about the resolution of future cost uncertainty and, while they are able to bank permits in many regulatory situations, they only rarely can borrow.

to a purebred quantity instrument. When $b > c$, the part of the banking-and-borrowing instrument that is attractive to the regulator is the quantity component, while the price component is unattractive. Here the full-quantity is preferred to the half-quantity-half-price is preferred to the full-price. Hence the ordering (54).

For both cases of $b < c$ or $b > c$, instead of a mongrel hybrid policy instrument mixing attractive with unattractive components, the regulator prefers a purebred policy instrument that is fully attractive through and through. This heuristic story is not rock-solid, but perhaps it is the best tale that can be told in these circumstances to try to intuit what is, after all, a genuine mathematical surprise.

Note from conditions (53) and (54) that if the first-best instrument of either a price or a quantity is ruled out, then banking and borrowing might have a genuine, if limited, role to play as a second-best option. For example in the case of CO_2, since the marginal benefit curve within a regulatory period is very flat, so that $b \approx 0 << c$, the theory strongly advises a fixed price as the optimal regulatory instrument; however, if policy makers override the adviser and insist on a cap-and-trade system, then condition (53) informs us that banking and borrowing is then some-ways superior to fixed quantities and should therefore he allowed. Another possible example is when the coefficients b or c are themselves genuinely unknown, so that it is unclear whether $b < c$ or $b > c$, in which case the intermediate banking-and-borrowing instrument might represent a compromise strategy that avoids worst-case outcomes.[29]

10 Prices vs. Quantities, Again

From the literature on 'prices vs. quantities', the sign of the traditional formula for the comparative advantage of prices over quantities in a one-period setting is given by the condition

$$sign[\Delta] = sign\left[(c - b)\right]. \tag{55}$$

Comparing (43) with (55) indicates that $sign\left[\Delta^p_q\right] = sign[\Delta]$, thus allowing the familiar

[29] If $B(q_1, q_2) = B_{1,2}(q_1 + q_2)$ (instead of the paper's assumed form $B(q_1, q_2) = B_1(q_1) + B_2(q_2)$), then intertemporal banking and borrowing necessarily dominates fixed quantities for *any* values of b or c (with $b \neq c$). For the corresponding linear-quadratic setup of this paper applied to the specification $B(q_1, q_2) = B_{1,2}(q_1 + q_2)$, we have $b < c \Rightarrow EW(p) > EW(\widehat{q}) > EW(\overline{q})$, and $b > c \Rightarrow EW(\widehat{q}) > EW(\overline{q}) > EW(p)$. However, the specification $B_{1,2}(q_1+q_2)$ is presumably attempting to capture cumulative stock effects, in which case (at least for CO_2 in climate change) $b \approx 0$ within a regulatory cycle of five to ten years or so. But then we are effectively in preference-ordering situation (53) anyway, and the insight that $EW(\widehat{q}) > EW(\overline{q})$ for the specification $B_{1,2}(q_1+q_2)$ is not challenging the ordering prescribed by (53). If $B(q_1, q_2) = B_{1,2}(q_1+q_2)$ is *not* representing cumulative stock effects, then the additively-separable specification $B(q_1, q_2) = B_1(q_1) + B_2(q_2)$ of this paper seems appropriate, or at least seems defensible.

one-period formula (55) for the sign of Δ to effectively 'stand in' for the two-period formula (43) for the sign of Δ^{p}_{q}.

If $b < c$ then, from (53), a (fixed) price instrument is welfare-superior to *both* a (fixed) quantity instrument *and* a time-flexible quantity instrument. If $b > c$ then, from (54), a (fixed) quantity instrument is welfare-superior to *both* a (fixed) price instrument *and* a time-flexible quantity instrument.

What the above observations mean for policy is that, in principle, the regulator here need not ever consult the sign formula (52) for $\Delta^{p}_{\widehat{q}}$ or the sign formula (47) for $\Delta^{\overline{q}}_{\widehat{q}}$. A useful conceptual-informational shortcut is available because $sign[\Delta]$ is a sufficient statistic for *both* $sign\left[\Delta^{p}_{\widehat{q}}\right]$ *and* $sign\left[\Delta^{\overline{q}}_{\widehat{q}}\right]$. To determine whether p dominates $\widehat{q}$ in the two-period sign formula (52) for $\Delta^{p}_{\widehat{q}}$, it suffices to check whether p dominates q in the traditional one-period formula (55) for the sign of Δ. To determine whether $\overline{q}$ dominates $\widehat{q}$ in the two-period sign formula (47) for $\Delta^{\overline{q}}_{\widehat{q}}$, it suffices to check whether q dominates p in the traditional one-period sign formula (55) for Δ. In other words,

$$sign[\Delta] > 0 \iff sign\left[\Delta^{p}_{\widehat{q}}\right] > 0, \tag{56}$$

while

$$sign[\Delta] < 0 \iff sign\left[\Delta^{\overline{q}}_{\widehat{q}}\right] > 0. \tag{57}$$

Thus, the standard original formula for the comparative advantage of prices over quantities Δ contains enough information to completely characterize the appropriate regulatory role of time-flexible quantities. In this sense (fixed) prices vs. time-flexible quantities or (fixed) quantities vs. time-flexible quantities present no new issues relative to the standard insights offered by the original formula for Δ – at least within the confines of the model of this paper. Nothing is otherwise gained by, or added to, the economic analysis by conceptualizing the regulatory role of intertemporal banking and borrowing as somehow fundamentally different from the usual analysis in terms of 'prices vs. quantities'. Therefore, far from undoing the standard 'prices vs. quantities' message about the formula Δ for the comparative advantage of prices over quantities, the introduction of intertemporal banking and borrowing seems to reinforce the robust scope of the original basic message.

11 Concluding Remarks

There is no question that the model of this paper is highly stylized with a great many restrictive assumptions (including linear-quadratic costs and benefits, a two-period analysis, independence of regulatory cycles, additively-separable costs and benefits each having essentially the same structure in both periods, additive marginal-cost shocks, statistical independence of costs and benefits, a highly simplified information and decision structure, a purely static flow analysis, and so forth and so on). The problem confronting a modeler here is that the issues addressed in the model of this paper tend to become analytically intractable without making strong structural assumptions. The model of this paper might be appended and extended, but largely at the detriment of adding complexity. As a consequence of the model's sparsely-crisp simplicity, the results of this paper are applicable for providing insight and rough guidance only at a high level of abstraction. At the very least, I would hope that the ideas expressed in this paper might be found to be thought-provoking about an important set of issues.

The most striking insight of the paper is the result that, for any given parameter values, either a fixed price or a fixed quantity is superior in expected welfare to intertemporal banking and borrowing. Of course, this strong result can be criticized as stemming from unrealistic assumptions. However, we must also keep in mind that the dominance of prices or quantities emerges in spite of the model here seemingly being biased in favor of artificially boosting the welfare status of intertemporal banking and borrowing by allowing the firm to know, and immediately act upon, the perfect-foresight realizations of both of the uncertain future costs at the very beginning of the regulatory cycle.[30] This unrealistic assumption about the exaggerated power of time-flexible quantities (over fixed prices or fixed quantities, which do *not* rely on perfect foresight) would have to be weighed against other unrealistic assumptions in evaluating the overall real-world relevance of the idea that prices or quantities dominate banking and borrowing.

Furthermore, the paper shows that the addition of intertemporal banking and borrowing does not alter the fundamental insights gained from the traditional framework for analyzing 'prices vs. quantities'. The customary formula for the comparative advantage of prices over quantities is sufficiently robust for the regulators to make correct economic decisions even when intertemporal banking and borrowing is included. I believe this insight too is useful, and parts of it might even survive under more realistic assumptions.

References

[1] Adar, Zvi, and James M. Griffin (1976). "Uncertainty and the Choice of Pollution Control Instruments." *Journal of Environmental Economics and Management*, 3: 178-188.

[30] And also, but less importantly, by allowing unlimited borrowing.

[2] Chao, Hung-Po, and Robert Wilson (1993). "Option Value of Emission Allowances." *Journal of Regulatory Economics*, 5: 233-249.

[3] Dales, John H. (1968). *Pollution, Property, and Prices: An Essay on Policy-making and Economics.* Toronto: University of Toronto Press.

[4] Fell, Harrison, Ian A. MacKenzie, and William A. Pizer (2012). "Prices versus Quantities versus Bankable Quantities." *Resource and Energy Economics*, 34(4): 607-623.

[5] Feng, Hongli, and Jinhua Zhao (2006). "Alternative intertemporal trading regimes with stochastic abatement costs." *Resource and Energy Economics*, 28(1): 24-40.

[6] Fishelson, Gideon (1976). "Emission Control Policies under Uncertainty." *Journal of Environmental Economics and Management*, 3: 189–197.

[7] Goulder, Lawrence H., and Andrew R. Schein (2013). "Carbon Taxes vs. Cap and Trade: A Critical Review." *Climate Change Economics,* 4(3): 1-28.

[8] Hoel, Michael, and Larry Karp (2002). "Taxes versus Quotas for a Stock Pollutant." *Resource and Energy Economics*, 24(4): 367-384.

[9] Kaplow, Louis, and Steven Shavell (2002). "On the superiority of corrective taxes to quantity regulation." *American Law and Economics Review*, 4(1): 1-17.

[10] Kelly, David L. (2005). "Price and quantity regulation in general equilibrium." *Journal of Economic Theory*, 125: 36-60.

[11] Montgomery, W. David (1972). "Markets in Licenses and Efficient Pollution Control Programs." *Journal of Economic Theory*, 5(3): 395-418.

[12] Murray, Brian C., Richard G. Newell, and William A. Pizer (2009). "Balancing cost and emission certainty: An allowance reserve for cap-and-trade." *Review of Environmental Economics and Policy*, 3: 84-103.

[13] Newell, Richard, and William Pizer (2003). "Regulating stock externalities under uncertainty." *Journal of Environmental Economics and Management*, 45: 416-432.

[14] Newell, Richard, William Pizer, and Jiangfeng Zhang (2005). "Managing permit markets to stabilize prices." *Environment and Resource Economics*, 31: 133-157.

[15] Pigou, Arthur C. (1920). *The Economics of Welfare.* London: Macmillan.

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

[17] Pizer, William A. and Brian Prest (2016). "Prices versus Quantities with Policy Updating." *NBER Working Paper No. 22379.*

[18] Roberts, Mark J., and A Michael Spence (1976). "Effluent charges and licenses under uncertainty." *Journal of Public Economics*, 5: 193-208.

[19] Rubin, Jonathan D. (1995). "A Model of Intertemporal Emission Trading, Banking, and Borrowing." *Journal of Environmental Economics and Management*, 31: 269-286.

[20] Weitzman, Martin L. (1974). "Prices *vs.* Quantities." *Review of Economic Studies*, 41(4): 477-491.

[21] Weitzman, Martin L. (1978). "Optimal rewards for economic regulation." *American Economic Review*, 68: 683-691.

[22] Williams, Roberton C. III (2002). "Prices vs. Quantities vs. Tradeable Permits." *NBER Working Paper No. 9283.*

[23] Yates, Andrew J., and Mark B. Cronshaw (2001). "Pollution permit markets with intertemporal trading and asymmetric information." *Journal of Environmental Economics and Management*, 42: 104-118.

[24] Yohe, Gary W. (1978). "Towards a general comparison of price controls and quantity controls under uncertainty." *Review of Economic Studies*, 45: 229-238.

[25] Zhao, Jinhua (2003). "Irreversible abatement investment under cost uncertainties: tradable emission permits and emissions charges." *Journal of Public Economics*, 87: 2765-2789.