

Prices versus Quantities versus Bankable Quantities

Harrison Fell*
Ian A. MacKenzie†
William A. Pizer‡

Working Paper EE 12-05
March 2012

*Colorado School of Mines

†Swiss Federal Institute of Technology (ETH), Zurich

‡Sanford School of Public Policy and Nicholas Institute for Environmental Policy
Solutions, Duke University

*The Duke Environmental Economics Working Paper Series provides a forum for Duke
faculty working in environmental and resource economics to disseminate their research.
These working papers have not undergone peer review at the time of posting.*

Prices versus Quantities versus Bankable Quantities

Harrison Fell, Ian A. MacKenzie, and William A. Pizer

Running title: P vs. Q vs. Bankable Q

Abstract

Quantity-based regulation with banking allows regulated firms to shift obligations across time in response to periods of unexpectedly high or low marginal costs. Despite its wide prevalence in existing and proposed emission trading programs, banking has received limited attention in past welfare analyses of policy choice under uncertainty. We address this gap with a model of banking behavior that captures two key constraints: uncertainty about the future from the firm's perspective and a limit on negative bank values (e.g., borrowing). We show conditions where banking provisions reduce price volatility and lower expected costs compared to quantity policies without banking. For plausible parameter values related to U.S. climate change policy, we find that bankable quantities produce behavior quite similar to price policies for about two decades and, during this period, improve welfare by about a \$1 billion per year over fixed quantities.

Key Words: prices, quantities, climate change, allowance banking

Prices versus Quantities versus Bankable Quantities

Harrison Fell, Ian A. MacKenzie, and William A. Pizer*

1. Introduction

Under the presence of uncertainty, welfare comparisons of regulatory instruments have typically focused on price versus temporally fixed quantity instruments. Yet a common trend is for policies to allow regulated quantities to be banked—that is, tradable permits can either be saved for future use or possibly borrowed from future periods. This is true for the majority of existing and proposed tradable permit markets.¹ While the early motivation for banking provisions was to allow firms to smooth the transition to more stringent future limits through early reductions, more recent debates in the 110th Congress focused on accommodating uncertainty and the risk of unexpectedly high costs. Many argued that an expanded banking and borrowing system to regulate U.S. greenhouse gas emissions could alleviate concerns about uncertain costs in a manner similar to a price mechanism.²

But can banking and borrowing provisions really accommodate cost shocks in a way that resembles a price mechanism? To answer this question, we develop a model of optimal firm behavior with a quantitative, periodic emission limit over an arbitrarily long horizon, the flexibility to bank (and possibly borrow within limits) allowances, and cost uncertainty that is

* Colorado School of Mines, Swiss Federal Institute of Technology (ETH) Zurich, and Duke University, respectively.

¹ See <http://www.epa.gov/airmarkets/basic.html>, <http://www.rggi.org/>, <http://ec.europa.eu/environment/climat/emission.htm>, and http://unfccc.int/essential_background/kyoto_protocol/items/1678.php (article 3, paragraph 13) for information on the U.S. SO₂ and NO_x programs, the Regional Greenhouse Gas Initiative, the EU ETS, and the Kyoto Protocol, respectively.

² See Murray et al (2009). The reality of the policy debate in the U.S. is evidenced by the evolution of legislative proposals. The first time cap-and-trade legislation was seriously proposed in 2003, it included a simple banking provision comparable to the sulfur dioxide trading program. By the time legislation passed the U.S. House of Representatives in 2009, it included borrowing provisions, a price floor, and a quantity-limited price ceiling.

potentially correlated over time and never completely resolved. We show banking allows shocks to be spread over time as long as the banking constraint is not expected to bind. When the constraint is expected to bind, banking offers little practical flexibility. This highlights an important connection between banking with and without uncertainty: it is the banking behavior driven by deterministic features – what we could call *deterministic* banking – that creates the window for uncertain shocks to be spread efficiently over time.

In order to quantify these effects and compare welfare outcomes with ordinary price and quantity controls, we must numerically solve for optimal firm banking behavior. For plausible parameters related to U.S. climate change policy, we find bankable quantity regulation improves welfare over a non-bankable system, but does not achieve welfare improvements over a price policy, which represents an upper bound given exceedingly flat marginal benefits. In the early years of the policy, when the banking constraint is not expected to bind for more than two decades, bankable quantities behave like prices in terms of emissions and welfare but gradually transition to behave like non-bankable quantity policies. The welfare improvement over non-bankable quantity policies averages about a billion dollars per year – more than 80 percent of the welfare improvement from prices – during the first twenty years but with little improvement thereafter. While tighter banking constraints and higher discount rates reduce the ability to match the welfare associated with prices, looser banking constraints and lower correlation of cost shocks increase it.

The paper is organized as follows. Section 2 reviews the existing literature, highlighting the gap between discussions of welfare differences among policies and banking behavior with relevant constraints and information assumptions. Section 3 introduces the banking problem and

discusses the solution. Section 4 outlines the benefit model and provides intuition about welfare effects. Section 5 presents the numerical analysis, and Section 6 concludes.

2. Background and Existing Literature

This paper draws from two important strands of the economics literature, one looking at the welfare differences among policy choices that differ under uncertainty about costs but are otherwise equivalent, and one looking at behavior (and sometimes welfare) associated with the ability to bank allowances in a more limited uncertainty environment.³ The first strand has developed a well-known distinction between price versus non-bankable quantity policies (Weitzman, 1974): the net welfare effect of prices over quantities hinges on the relative slopes of marginal costs and benefits.⁴

Papers on banking first emerged after the 1990 Clean Air Act Amendments created the SO₂ trading program with an unlimited banking provision. These studies ignored uncertainty and focused on how banking shifts emissions to the future to equalize otherwise unequal (and higher) discounted marginal costs in the future, lowering total discounted costs (e.g. Cronshaw and Kruse (1996), Rubin (1996), Kling and Rubin (1997), and Leiby and Rubin (2001)).

Analysis of banking with uncertainty followed but has been relatively limited. A number of authors suggested that banking can reduce permit price volatility in the face of uncertain cost shocks (e.g. Godby et al (1997), Jacoby and Ellerman (2004), and Ellerman (2005)). Yates and Cronshaw (2001), Yates (2002), and Feng and Zhao (2006) investigate when banking and

³ Uncertainty in both damage and abatement costs has also been considered. In particular, Stavins (1996) investigated correlated uncertainty between damage and abatement costs and found that in many cases quantity regulation was desirable. However, as our motivating application is global climate change, where the dominant sources of uncertainty are uncorrelated and marginal benefits exceeding flat, we do not focus on this aspect of the problem.

⁴ These results were later extended by Hoel and Karp (2001, 2002) and Newell and Pizer (2003) to consider stock externalities (pollutants).

borrowing improve welfare, with results that support Weitzman's original intuition about the importance of the relative slopes of marginal costs and benefits. However, their models have quite stylized features – two time periods and no borrowing constraints –that make it difficult to assess real-world programs.

Schennach (2000) is the first (and perhaps only) author to build a theoretical model with banking, banking constraints, multiple (even infinite) periods, and uncertainty, in order to study the U.S. SO₂market. She focuses on qualitatively understanding the optimal price and emissions paths (but not welfare) with and without cost uncertainty when emission targets feature a step-wise decline. Our analytical discussion of banking behavior draws heavily from her work.

A study by Williams (2002) mirrors the welfare-focused spirit of our paper. Williams analyzes regulation under prices, quantities and tradable quantities and extends this framework to include banking and borrowing. An important contribution is his intuition that banking and borrowing effectively flatten marginal costs compared to earlier analyses that ignored banking and borrowing (Hoel & Karp, 2002; Newell & Pizer, 2003), consistent with our own results. However, while he goes on to draw a stronger conclusion that banking does not affect the expected benefits from stock pollutants, we relax a number of his assumptions (e.g., no constraints on negative banking) and are more equivocal, a point we return to below.

Currently, there exists no welfare analysis of bankable quantities capturing *both* uncertainty about the future from the firms' perspective as well as the regulators' *and* constraints on banking (e.g., that borrowing is constrained in some way, to zero or to a small, limited amount). Our contribution, therefore, is to provide a more realistic model of banking behavior under uncertainty that lends itself to quantitative analysis, to develop a welfare framework for comparing that policy to fixed quantity and price policies, and to then use these tools to look at

the choice among these policies for the particular problem of controlling greenhouse gas emission to address climate change.

3. The Banking Problem

We begin by looking at how otherwise equivalent policies – prices, quantities, and quantity controls that allow banking – lead to different optimal emission and marginal cost behavior when uncertainty arises. Ordinary price and quantity controls pose a relatively simple behavioral problem for the firm to solve and the regulator to understand. Consider the quadratic cost function in Newell and Pizer (2003), hereafter NP (in turn, a multi-period extension of Weitzman (1974)):

$$C_t(q_t, \theta_t) = \theta_t(q_t - \bar{q}_t) + \frac{c_t}{2}(q_t - \bar{q}_t)^2 \quad (1)$$

The variable q_t is the chosen emission level, \bar{q}_t is the average, cost-minimizing level of emissions in the absence of regulation (i.e., the expected baseline or “uncontrolled” emissions level), $c_t > 0$ describes the slope of marginal costs, and θ_t is a current-period shock to marginal costs observed by firms (but not by the regulator). Potential changes in c_t and \bar{q}_t allow for cost reductions and growth in uncontrolled emissions over time. We view (1) as an aggregate, representative firm’s cost function arising from a competitive within-period emissions trading market (Rubin, 1996).⁵

How do various policies differ in their emission and marginal cost outcomes? First consider non-banking regulation. In the case of quantity controls, q_t^* , the regulated firm actually

⁵ As shown in Rubin (1996) in a model without uncertainty, the decentralized firm behavior in a cap-and-trade program with banking leads to the same emission path as that of a cost-minimizing social planner. Under the assumption that all regulated firms have the same expectations about future shocks, we can therefore use a representative firm framework to describe aggregate behavior.

faces no choice: it simply emits the regulated volume of emissions, setting $q_t = q_t^*$.⁶ Marginal costs are given by

$$\left. \frac{\partial C_t(q_t, \theta_t)}{\partial q_t} \right|_{q_t=q_t^*} = \theta_t + c_t(q_t - \bar{q}_t) = \theta_t + c_t(q_t^* - \bar{q}_t) \quad (2)$$

Note that as we regulate to reduce emissions, $q_t^* < \bar{q}_t$, so marginal costs of *emissions* are negative – which makes sense because emissions, q_t , are a negative externality. In the case of price controls (e.g., taxes), p_t^* , the firm will want to reduce emissions until marginal *abatement* costs ($-\partial C_t / \partial q_t$) equal the regulated price each period, solving $p_t^* = -\theta_t - c_t(q_t - \bar{q}_t)$ for q_t . Assuming the price policy is set to match the quantity policy in expectation, we can write the emission outcome under a price policy as:

$$q_{t,price}(\theta_t) = \frac{-p_t^*}{c_t} + \bar{q}_t - \frac{\theta_t}{c_t} = q_t^* - \frac{\theta_t}{c_t} \quad (3)$$

The opportunity to bank (and borrow) emission allowances poses a trickier challenge for deriving optimal firm behavior. To be specific, the regulator gives the firm a pre-specified emission allocation each period, which we label y_t . The firm also starts each period with a bank B_t of unused allowances from the end of the previous period. The firm is free to use the sum of these allowances $y_t + B_t$ to cover emissions q_t this period, but any leftover allowances can, in turn, be saved for use in future periods. That is,

$$B_{t+1} = B_t + y_t - q_t \quad (4)$$

becomes the starting point for next period's emission decision.⁷ It may even be possible to leave a negative bank for the next period; generally, the firm is told that the start-of-period $t+1$ bank,

⁶ To be precise, the regulated firm can choose any $q_t \leq q_t^*$. Firms will choose $q_t = q_t^*$ so long as costshocks are small compared to expected marginal costs, which is true in our eventual parameterization. We adopt the general assumption that $q_t = q_t^*$, consistent with Weitzman and NP, as it does not materially affect our results and simplifies exposition by allowing $E[q_t] = q_t^*$.

B_{t+1} , must be equal to or greater than \bar{B}_{t+1} , which is a constraint limiting the minimum level of banking (borrowing). Here, \bar{B}_{t+1} is typically zero but could be negative when a regulation allows limited borrowing. The firm now has the flexibility of choosing emissions q_t anywhere between 0 and $y_t + B_t - \bar{B}_{t+1}$. Finally, the firm needs to make an assumption about the evolution of cost shocks θ_t . For simplicity, we assume a simple autoregressive model.

We can summarize the firm's optimization problem under a bankable quantity policy as an effort to minimize discounted expected costs each period,

$$\min_{q_t} \left\{ E_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\theta_t (q_t - \bar{q}_t) + \frac{c_t}{2} (q_t - \bar{q}_t)^2 \right) \right] \right\} \text{ subject to } \begin{cases} q_t \geq 0 \\ B_{t+1} \geq \bar{B}_{t+1} \\ B_{t+1} = B_t + y_t - q_t \\ \theta_t = \rho \theta_{t-1} + \epsilon_t \end{cases} \quad (5)$$

where $\beta = 1/(1+r)$ is the discount factor and $E_t[\cdot]$ is the expected value given all the information known at time t . As noted, θ_t follows a stationary autoregressive process, with $\epsilon_t \sim N(0, \sigma_0^2)$ and $|\rho| \leq 1$.⁸ Given the ability to pick q_t each period based on new information at time t , it makes sense to rewrite this problem (5) in the form of a Bellman equation,

$$V(B_t, \theta_t) = \min_{q_t, B_{t+1}} \left\{ \theta_t (q_t - \bar{q}_t) + \frac{c_t}{2} (q_t - \bar{q}_t)^2 + E_t[V(B_{t+1}, \theta_{t+1})] \right\} \text{ subject to } \begin{cases} q_t \geq 0 \\ B_{t+1} \geq \bar{B}_{t+1} \\ B_{t+1} = B_t + y_t - q_t \\ \theta_t = \rho \theta_{t-1} + \epsilon_t \end{cases} \quad (6)$$

⁷ A more general banking relationship would allow for non-unitary and even non-uniform trading ratio between periods, so $B_{t+1} = R_t(B_t + y_t - q_t)$ and $R_t(\cdot)$ is some monotonically increasing function. Assuming $R_t(x) = x$, as we do and consistent with actual emission trading programs, implies that whenever banking occurs, prices rise at the interest rate.

⁸ The cost shock variance at time t , $\text{var}(\theta_t)$ is therefore given by $\text{var}(\theta_t) = \sigma_t^2 = \sigma_0^2(1 - \rho^{2t})/(1 - \rho^2)$, assuming that $\theta_{t < 0} = 0$.

The value function $V_t(B_t, \theta_t)$ is defined recursively as the minimized expected net present value of costs, conditional on the current bank B_t and baseline cost shock θ_t , and assuming similarly optimal behavior in every future period. The solution to this problem (6), an optimal decision rule $q_t(B_t, \theta_t)$ and $B_{t+1}(B_t, \theta_t)$, requires a numerical approach, to which we return in Section 5. However, as shown in the appendix, the necessary first-order conditions for the problem (6) are given by:

$$-\theta_t - c_t(q_t - \bar{q}_t) + \mu_t = \pi_t \quad (7)$$

$$\pi_t = \beta E_t[\pi_{t+1}] + \gamma_t \quad (8)$$

where π_t , μ_t , and γ_t , all ≥ 0 , are the Lagrange multipliers on the allowance constraint (4), non-negative emissions constraint, and the minimum bank constraint, respectively.⁹ While not revealing the exact solution path, these first-order conditions provide useful intuition about optimal banking behavior under uncertainty, particularly how banking can smooth out cost shocks and, as we will see in Section 4, affect welfare.

First note that, from Equation (7), π_t is the shadow price of allowances and equals the marginal abatement cost, so long as emissions are strictly positive and $\mu_t = 0$.¹⁰ The Euler equation (8) displays the traditional Hotelling rule that expected allowance prices rise at discount rate $r = (1/\beta) - 1$ each period when the minimum bank constraint is not binding, e.g., when $\gamma_t = 0$.¹¹ That is, firms will try to minimize costs across periods t and $t+1$ by shifting emissions to equalize discounted marginal costs. If they reach $B_{t+1} = \bar{B}_{t+1}$ and discounted expected future

⁹ Note that we associate the Lagrange multiplier γ_t with the constraint $B_{t+1} \geq \bar{B}_{t+1}$. We chose this specification because the bank at the start of next period, B_{t+1} , is fully defined by the current state and the current emission decision in period t , making the Lagrange multiplier on the minimum bank constraint for next period a current period variable.

¹⁰ Compare to Equation (9) in Rubin (1996), where his e_t and A_t are comparable to our q_t and π_t .

¹¹ Compare to Rubin (1996) Equation (22) or Schennach (2000) Equation (1).

marginal costs are still lower than current marginal costs, they would prefer to bank even less than \bar{B}_{t+1} for next period, but cannot. In that case, $\gamma_t > 0$ in order to balance Equation (8).

The preceding results have been described by Rubin (1996) and others for the case without uncertainty. In particular, they distinguish those time intervals where the minimum bank constraint is not binding and discounted allowance prices are equalized over the interval, versus those time intervals where the minimum bank constraint is binding and discounted allowance prices are falling. Schennach (2000) looks at this problem with uncertainty and describes a third type of interval that literally connects the other two. That is, uncertainty implies transition periods where it is unknown today whether the constraint is binding or not. Uncertainty tends to smear the exact point envisioned by Rubin where minimum bank constraint begins to bind and γ_t changes from zero to positive. For example, the constraint might bind early if costs are unexpectedly high, or late if costs are unexpectedly low.

These different types of banking intervals turn out to be important for our intuition about emissions, prices, and ultimately welfare outcomes, under different policies. In particular, we can look at how prices respond to shocks θ_t during different banking intervals and contrast it with behavior under price and non-bankable quantity policies. Consider, for example, a perturbation $\delta\theta_t$ to a realized θ_t , and how the perturbation affects prices under different policies. Under a price policy, prices are fixed and there is no price change. Under a non-bankable quantity policy $q_t = q_t^*$, we can work out the perturbation in price based on (2) and the assumption that price is the negative marginal cost of emissions: namely, $\delta p_t = -\delta \left(\frac{\partial C_t}{\partial q_t} \right) = -\delta\theta_t$.

With banking, if we begin in a period where the minimum bank constraint is not binding, and we assume we know with certainty that the bank is constrained by $B_{t+s+1} \geq \bar{B}_{t+s+1}$ beginning s periods in the future, but not before, we show in the appendix that

$$\delta\pi_t = -\delta\theta_t \left(1 + \sum_{j=1}^s \frac{c_t \rho^j}{c_{t+j}}\right) \left(1 + \sum_{j=1}^s \frac{c_t}{\beta^j c_{t+j}}\right)^{-1} \quad (9)$$

Here, $\delta\pi_t$ is the change in allowance price arising from $\delta\theta_t$, ρ is the correlation among cost shocks, c_t is the slope of marginal costs in period t , and β is the discount factor. With no correlation, no discounting, and constant c_t this simplifies to:¹²

$$\delta\pi_t = \frac{-\delta\theta_t}{1+s} \quad (10)$$

In other words, the price increase $\delta\theta_t$ is spread over however many periods the bank is unconstrained. That is, optimal behavior considers this period of unconstrained banking to be a single pool of emission reduction effort, with any required increase or decrease in effort spread evenly throughout the pool. If the unconstrained interval is quite long (e.g., s approaches infinity, as in Newell, Pizer, and Zhang (2005)), the price increase is spread so thinly so as to be imperceptible and behavior will be very similar to a price policy.

On the other hand, with *perfect* auto-correlation and no discounting (9) simplifies to $\delta\pi_t = -\delta\theta_t$. In this case, there is no reason to spread the cost shock because we now expect future cost shocks to be the same as today; behavior will be very similar to a non-bankable quantity policy. This behavioral difference between high and low correlation cost shocks is analogous to the difference between of high and low correlation income shocks for the resulting effect on consumption (Friedman, 1957).

Even without autocorrelation, if s is zero and we are instead in an interval where banking is constrained, there is little flexibility to spread the price increase $\delta\theta_t$ and behavior will be very similar to a non-bankable quantity policy. It will not be *exactly* the same because it is always

¹² This is implied by equation (13) in Schennach (2000) though it is not discussed.

possible that a particularly favorable shock will spontaneously create a brief period of banking (a point ignored in (9-10) by the simplifying assumption that s is known and fixed). That is, a favorable cost shock creates a bounty of unexpected cost savings that can be used today or spread over several periods. When discounted marginal costs are lower in the future, as occurs when banking is constrained, the discounted cost savings from spreading this bounty to the future will generally be less than using it all today. However, if the bounty is large and today's marginal cost is otherwise depressed by a favorable cost shock below the discounted value of future marginal costs, it will make sense to bank some of it for the future.

This leads us to the following summary of banking behavior under uncertainty:

When deterministic trends in baseline emissions, allowance allocation, and costs lead to a period of certain banking activity unrelated to cost shocks and where we do not expect the minimum bank constraint to bind, bankable quantities under uncertainty lead to behavior that is similar to price regulation so long as the end of this period is sufficiently far in the future.

During periods when we do expect the minimum bank constraint is expected to bind, bankable quantities lead to behavior that is similar to non-bankable quantity regulation. As we approach the end of a period of unconstrained banking, outcomes under bankable quantity policies will transition from price-like to non-bankable-quantity-like behavior.

Much of this discussion and the expression (9) assume firms always know the exact future moment $t+s+1$ when the minimum bank constraint becomes binding, ignoring the uncertain transition interval created by cost uncertainty. This creates ambiguity about exactly what is meant by “the end of a period of unconstrained banking” and makes it impossible to derive analytically exactly how prices, π_t , respond to cost shocks. While we expect the aforementioned intuition will continue to hold, we must ultimately turn to numerical analyses to

quantify the results, which we do after rounding out our model of costs and firm behavior with a model of pollution impacts, mitigation benefits, and welfare.

4. Welfare

So far we have focused on mitigation costs and how different price, quantity, and bankable quantity policies lead to different emission and marginal cost outcomes in the face of uncertain cost shocks. To understand the welfare consequences of these outcomes, we need to add a model of mitigation benefits and evaluate the expected net benefits across alternative policies. We continue to use the basic framework in NP to do this.

First, we allow for emissions to accumulate in the environment

$$S_t = (1 - \delta)S_{t-1} + q_t \quad (11)$$

where S_t is the accumulated stock of emissions at time t , which accumulates with decay rate δ .

The decay rate can take on values representing cases ranging from a “pure stock externality” that persists forever ($\delta = 0$) to a “flow externality” ($\delta = 1$) that replicates the traditionally analyzed case. “Benefits” associated with the stock of emissions are then given by

$$B_t(S_t) = -\frac{b_t}{2}(S_t - \bar{S}_t)^2 \quad (12)$$

where \bar{S}_t represents a benefit maximizing level of the stock (possibly zero, possibly a background level) and $b_t \geq 0$. The latter assumption implies that pollution in excess of \bar{S}_t faces increasing marginal damages (e.g., increasingly negative marginal benefits).

The above relations (11-12) and results from the previous section are sufficient to solve the firm’s cost-minimization problem numerically, simulate outcomes under uncertainty, and compute expected net benefits under bankable quantities and alternate policies. However, we can also use the model to derive more refined expressions for comparing welfare that provide

intuition for our results. In the spirit of Weitzman (1974) and NP, we can derive the welfare difference Δ_t between alternate policies and a benchmark of non-bankable quantity controls in a given period t . That is, we can derive an expression for

$$\Delta_t^x = E_0[\text{NB}_{t,\text{policy } x}] - E[\text{NB}_{t,\text{quantity}}], \quad (13)$$

where $E[\text{NB}_{t,\text{policy } x}]$ is the expected net benefits under policy x – price controls or a system of bankable permits in our case – in period t , and $E[\text{NB}_{t,\text{quantity}}]$ is the expected net benefits of a (non-bankable) quantity policy in period t . Here, the quantity control is set at the same emission level as the expected emissions under policy x .¹³

As we show in the appendix, we can use our model of costs and benefits to write

$$\Delta_t^x = \frac{\sigma_t^2}{2c_t^2} \left(c_t \left(1 - \frac{\text{var}(p_t)}{\sigma_t^2} \right) - b_t \left(\frac{\text{var}(S_t)}{\sigma_t^2/c_t^2} \right) \right) \quad (14)$$

where $\text{var}(p_t)$ and $\text{var}(S_t)$ are the variance in allowance prices and the accumulated emission stock, respectively, under policy x . To understand this expression, first note that aside from the time subscripts and factors multiplying c_t and b_t , Equation (14) appears remarkably similar to Weitzman's (1974) expression comparing welfare under alternate price and quantity controls in a simpler one-period model,

$$\Delta = \frac{\sigma^2}{2c^2} (c - b) \quad (15)$$

The intuition in his simpler model and policy scenarios is straightforward: prices offer a cost savings over quantities, given by $(1/2)(\sigma^2/c^2)c$, by trading off high marginal cost emissions for low marginal cost emissions in different states of the world. This goes along with a loss in

¹³Note that in a comparison of bankable to non-bankable quantity policies, fixing non-bankable quantities at the expected emission level $E_0[q_t]$ will not be the same as fixing them at the allocation level y_t . Fixing emissions at the allocation level would conflate advantages of banking unrelated to uncertainty.

expected benefits associated with prices, $-(1/2)(\sigma^2/c^2)b$, stemming from the induced variance in quantities coupled with concave benefits.

Second, our more general model reduces to Weitzman's result when policy x represents price regulation and we consider a flow pollutant where $S_t = q_t$. In this case by construction $\text{var}(p_t) = 0$ and, from (3) and the assumption that $S_t = q_t$, we have $\text{var}(S_t) = \text{var}(q_t)$ and $\text{var}(q_t) = \sigma_t^2/c_t^2$. That is, the Weitzman price-versus-quantity welfare comparison is a special case of (14).

More generally, these factors multiplying c_t and b_t in Equation (14) provide intuition about how expected costs and benefits change based on a more general set of policies. The first factor shows that cost savings relative to a non-bankable quantity policy has everything to do with reducing price (or marginal cost) variance under a given emission response function. Price variance represents missed opportunities to trade off high price emission reductions for low price emission reductions while maintaining the same expected emission level. Under a quantity policy, $\text{var}(p_t) = \sigma_t^2$ based on (2) and there is no cost savings by construction. Under a price policy, $\text{var}(p_t) = 0$ and cost savings is maximized. Other reasonable policies will generally fall somewhere in between.

The second factor shows that benefit losses relative to a non-bankable quantity policy are related to variance in the accumulated pollution stock, which is in turn driven by variance in emissions. Under a quantity policy, $\text{var}(S_t) = \text{var}(q_t) = 0$ and there is no benefit loss by construction. Under a price policy, $\text{var}(q_t) = \sigma_t^2/c_t^2$ based on (3). However, except in the case of a flow pollutant, where $\text{var}(S_t) = \text{var}(q_t)$, translating this emission variation into accumulated stock variation is non-trivial. Calculating $\text{var}(S_t)$ is even more problematic under bankable quantities, where emissions are path-, as well as shock-, dependent and must be numerically derived. While we know there is a benefit loss from policies that introduce emissions

uncertainty compared to a regulation that fixes the emission path, ultimately a numerical analysis is required to quantify this loss.

We note these results generally are consistent with Williams (2002), who concludes that banking and borrowing effectively reduce the slope of marginal costs when comparing prices and quantities and offer welfare improvements over non-bankable quantity controls for stock pollutants. His interpretation of flatter marginal costs when comparing prices to bankable rather than non-bankable quantities is consistent with our factor (<1) multiplying c_t term in (14).¹⁴ Going further, certain assumptions in his analysis eliminate (or minimize) the loss of expected benefits under bankable versus non-bankable quantities for a stock pollutant. In our approach, we relax his assumptions (e.g., no banking constraint, a fixed banking horizon, no correlation of shocks, full knowledge of all shocks in all periods) and leave to a numerical analysis a determination about the magnitude of expected benefit loss, a task to which we now turn.

5. Numerical Analysis

To parameterize the problem, we consider sets of values related to an emissions control policy implemented in the U.S. This allows us to answer the question of whether banking significantly reduce price volatility, and consequently expected costs, in an immediately relevant context.¹⁵ Table 1 presents our benchmark parameter values, where cost-related parameters are based largely on recent estimates of U.S. compliance costs with H.R. 2454 (Waxman-Markey) given in (EPA, 2009) and benefit-related parameters derive from NP. Additional detail on the rationale and data sources can be found in the appendix.

¹⁴ The c_t term in a price-to-bankable-quantity comparison, $\Delta_t^{\text{price}} - \Delta_t^{\text{bankable } q}$, would be $(\text{var}(p_t)/\sigma^2)c_t$, where $\text{var}(p_t) < \sigma^2$ would be measured for bankable quantities. In that sense, the effective marginal cost slope, e.g., the entire c_t term including the pre-multiplying factor, is flatter for the price-to-bankable-quantity comparison versus the effective marginal cost slope in a price-to-non-bankable quantity comparison (c_t).

¹⁵ For reference, we also compare our results to those in NP who consider a global policy in footnote 20.

The numerical approach is relatively straightforward: we take a finite-horizon approximation to (5), discretize the state space in (6), and solve backwards from a final period for the value function $V_t(B_t, \theta_t)$ and optimal decision rule $q_t(B_t, \theta_t)$ or $B_{t+1}(B_t, \theta_t)$ each period.¹⁶ For this application, we allow both the shock state variable θ_t and the bank state variable B_t to take on 201 discrete values. Thus, a 201×201 grid can fully describe the range of possible states at any given period. Given the discretized shocks and the assumed autoregressive process for the shocks, we can also derive a transition probability matrix, P , that describes the probability of moving from any of the 201 θ_t values to each θ_{t+1} value. We consider a terminal period at $T = 51$, allowing us to consider behavior through 2050 without endpoint effects.¹⁷

With a chosen horizon T , we can easily find the optimal rule for $q_T(B_T, \theta_T)$ and value function $V_T(B_T, \theta_T)$ in the final period T , for all possible 201×201 states. We assume the terminal conditions that $B_{T+1} \geq 0$, which implies that the representative firm must settle all outstanding allowance debts by the end of the finite horizon, and that $V_{T+1}(B_{T+1}, \theta_{T+1}) = 0$. For a given bank B_T at the beginning of period T , the problem in period T then simplifies to a fixed quantity problem with an allocation of $B_T + y_T$ that the firm can either use in period T or discard. Solving this problem yields the function $V_T(B_T, \theta_T)$; we can then recursively solve the problem of finding the optimal $q(B_t, \theta_t)$ and $V(B_t, \theta_t)$ functions going from period $T - 1$ back to 0 using (6).

The end result of this backward recursion process gives us T , 201×201 matrices of optimal q_t values for each possible (B_t, θ_t) state in each time period (i.e., the optimal policy rules). Given the optimal q_t matrices and starting bank $B_0 (=0)$, we can then run simulations given a set of generated θ values. Each iteration provides a different realization of the shock series $\theta_1, \dots, \theta_T$,

¹⁶ Discretizing the bank state variable implicitly discretizes the control variable, or vice versa, given the bank transition equation (4).

¹⁷ None of our results through 2050 change when instead use $T = 101$ or when we switch to a gridsize of 301×301 .

which we use to simulate the emission level q_t (from our optimal q matrices), bank B_{t+1} (based on (4)), and accumulated stock S_t (based on (11)), period by period. Assuming emission allowance prices equal marginal abatement costs (i.e., a competitive allowance market), prices are calculated using equation (2) evaluated at the chosen emission level q_t with banking.

We use this procedure to forward simulate banking behavior over 10,000 random realizations of the cost shocks. This yields a distribution of 10,000 emission levels, bank levels, prices, and pollution stocks in each period, which we can use to compute $\text{var}(S_t)$ and $\text{var}(p_t)$. We then use these measures to calculate welfare differences for banking compared to non-bankable quantities given in (14).¹⁸Welfare differences Δ_t for multiple years t are then discounted and added together to calculate expected welfare differences over various horizons.

The results are summarized in Table 2.¹⁹From earlier work by NP and others, we know that marginal benefits are exceedingly flat for greenhouse gas emissions, so prices are close to the first-best optimum. We therefore start with prices versus non-bankable quantities in column 2 as a benchmark. That column closely replicates NP and reminds us the prices are unambiguously better than non-bankable quantities.²⁰The third column provides our central estimates for the welfare gains from banking versus non-bankable quantities, which range from under a billion dollars for one year to over \$15 billion for 20years. Column 4 expresses this gain

¹⁸ A similar procedure can be used to simulate the welfare advantage of price policies compared to non-bankable quantities, replacing the recursively solved function $q_t(B_t, \theta_t)$ with equation (3) in the simulation procedure.

¹⁹ These welfare comparisons are based on price and non-bankable quantity policies determined by the expected price and emission paths associated with the bankable quantity policy. In this way, the comparison is among policies with the same expected outcomes.

²⁰ NP report welfare gains for prices versus fixed quantities of 0.52, 4.6, 11, 21, and 35 billion dollars, respectively, over the 1, 5, 10, 20, and 40 year horizons. While our estimates for prices versus fixed quantities are quite similar, they reflect cost shocks that are roughly one-third smaller and a marginal cost slope that is roughly ten times larger – differences that cancel out in Equation (1). These differences reflect not only global versus U.S. cost functions but also different data sources. For both marginal cost slopes and cost shock magnitudes, we use EPA (2009) while NP use Weyant and Hill (1999).

as a share of the (essentially first-best) welfare gain from prices, and we see that bankable quantities in the U.S. achieve almost all of those welfare gains in the early periods of the regulation. However, the relative advantage of bankable quantities stops growing after 20 years and prices begin to look much better. Why?

As described above, when deterministic trends in baseline emissions, allocation, and costs create a long period of certain banking activity, bankable quantity policies lead to behavior much like a price policy when the banking constraint remains far in the future; e.g., based on (10) and a large horizon s , cost shocks $\delta\theta_t$ do not translate into significant price changes $\delta\pi_t$. However, as time passes the expected horizon to the banking constraints falls; in cases with more adverse shocks it falls more quickly and in cases with more favorable shocks it falls more slowly. With a small s in (10), costs shocks translate more directly into price changes and a bankable-quantity policy looks more like a non-bankable quantity policy.

To better understand the relationship between the banking constraint and welfare, it is useful to look at the path of relevant variables over time. Figure 1 and Figure 2 show how the banking constraints become binding over time. Figure 1 shows the banking path (both the median value and 95 percent simulation intervals) while Figure 2 shows the probability that the banking constraint is binding. Both reflect a large amount of certain banking through 2030 driven by the deterministic trends in baseline emissions, allocations, and costs, which would arise even without uncertainty. By 2030, both figures indicate that the banking constraint begins to bind with an observable probability even though it still does not bind with certainty by the end of our reporting horizon in 2050. This wide range of possible endpoints for the initial banking period sharply contrasts with the Williams (2002) assumption that the endpoint is known with

certainty and is qualitatively different from Schennach (2000), where the transition interval is relatively short.

Of particular interest to us, the effectiveness of banking as a cost control (and welfare improving) mechanism begins to diminish long before the initial probability of a banking constraint arises in 2030. As the horizon to 2030 becomes too short to adequately smooth out shocks based on (9), particularly given the high autocorrelation in the benchmark scenario ($\rho = 0.8$), price variance increases. This can be seen clearly in Figure 3, showing the path of allowances prices (both the mean and 95 percent simulation intervals). Allowance price volatility under bankable quantities starts very small but increases steadily through 2035, by which point it matches the variability under non-bankable quantities. Thus, expected gains from a bankable quantity policy relative to a non-bankable quantity policy diminish as we look over longer horizons because in later periods the bankable quantity policy begins to resemble the non-bankable quantity policy. This explains the results from Table 2, where most welfare gains arise in the first 10 years and none arise after 20 years.

Sensitivity analysis in Table 3 confirms this intuition and provides additional insight into what happens under a banking policy. The first row repeats the 20-year results from Table 2, labeled “benchmark,” followed by simulation results with the noted parameter changes. All other parameters are otherwise the same as the benchmark. The second row of the table, labeled “double shocks”, shows that doubling the standard deviation of cost shocks roughly quadruples the welfare gain of both prices and banking over non-bankable quantities (as we would expect from (14)) with little relative change between prices and banking. For the “high discount rate” case, where the discount rate is increased to 7%, there is a significant drop in the welfare advantage of both prices and banking, as the net present value of the same Δ_t^x each period

declines. However, the banking advantage over non-bankable quantities declines more (indicated by the lower ratio of banking gain / price gain). Higher discount rates reduce deterministic banking because the future is valued less. This means that the horizon for unconstrained banking shrinks, and, consequently, the flexibility and ultimate welfare advantage gained through a bankable versus non-bankable quantity policy diminishes.

In the “flat allocation path” scenario we completely eliminate deterministic banking by setting $i_t = 2.84$ billion tons for all t (which has the same cumulative allocation and emissions as the benchmark; compare to the benchmark initial allocations of 6 billion tons declining by -3.4 percent, given in Table 1). With such a low initial allocation, there is no banking absent uncertainty; prices start at over \$35 per ton and then rise at less than 1 percent per year.²¹ The firm would desperately prefer to do less abatement now and more in the future, but cannot do so because of the borrowing constraint. Even with uncertainty, there is little banking. Favorable shocks have to drive down prices by more than 20 percent in order to encourage spontaneous banking, which is unlikely because the shocks’ standard deviation is barely 10 percent of the initial price.²² Correspondingly, the bankable-quantity policy offers little benefit over a fixed-quantity policy.

While the high discount rate and flat allocation cases show why banking performance might decline, the last two highlight why it might improve. In the case where we relax the

²¹ Initial abatement is 3.16 billion tons and the initial price is \$36 per ton. Prices then grow at the growth rate of abatement. Given abatement equals $\bar{q}_t - y_t$, y_t is constant, and \bar{q}_t grows at 0.35 percent, the growth rate of abatement is 0.35 percent divided by abatement, which begins at slightly more than 50 percent and grows. Hence prices start growing at 0.70 percent and decline over time to 0.35 percent.

²² A favorable shock that drove down current prices by 20 percent would also drive down expected prices next period by 16 percent, based on the correlation of 0.8. The improved expected return to banking of $20 - 16 = 4$ percent, on top of the original expected return of about 1 percent, would just match the discount rate. More favorable shocks would offer a return to banking in excess of the discount rate, and therefore banking would occur. This would improve welfare compared to fixed quantities, but only in these limited states where large favorable shocks occur.

borrowing constraint to $\bar{B}_{t+1} = -4y_{t+1}$, four times the annual allowance allocation, the probability of a binding borrowing constraint decreases and the bankable-quantity policy mimics the price policy for a much longer period of time. The advantage of a bankable-quantity policy under this relaxed borrowing constraint therefore increases relative to the benchmark almost matching the price policy and similar to the results discussed in Newell, Pfizer, and Zhang (2005). In the last row, we consider a no correlation case ($\rho = 0$) and see that bankable quantities now nearly match prices. Absent correlation, it is considerably easier to accommodate shocks with banking because shocks are unlikely to build on one another. This was pointed out analytically in equation (9) where lower correlation reduces deviations in prices for a given cost shock and time horizon.

Summarizing, our results and sensitivity analysis support the intuition provided at the end of Section 3: banking (and borrowing) provides the flexibility to spread cost shocks over time, to emit less and save allowances when costs are low, consequently emitting more and using banked allowance when costs are high. However, the value of this flexibility depends on the horizon over which banking is already occurring unrelated to uncertainty, and shifting emissions over time is costless. Such banking arises due to expected trends in baseline emissions, allocations, and costs – what we call deterministic banking. Without deterministic banking, when the banking constraint is binding, it is impossible to borrow allowances in the face of an adverse shock and, even with a favorable shock, the amount of cost savings today typically does not make it prudent to save allowances for the future (unless the shock is particularly large). Higher discounting and lower initial allocations all bring forward the likely time when deterministic banking will end and the banking constraint will bind, lowering the benefits of banking relative to fixed quantities. Looser borrowing constraints clearly delay the onset of the

banking constraint, while lower correlation makes it easier to accommodate shocks for a given banking horizon.

6. Conclusion

Welfare comparisons between otherwise equivalent price and quantity instruments have long emphasized a basic difference. Price instruments accommodate cost shocks and maintain fixed prices; quantity instruments provide no such accommodation and maintain fixed quantities. Yet most real quantity-based policies provide flexibility not accounted for in this dichotomy, namely the ability to bank (and sometimes borrow) against a multi-period quantitative constraint. For example, the ability to over-comply with a tradable permit system and bank unused allowances for future use is a central part of most observed emission trading systems. Intuitively, this banking flexibility should lead otherwise fixed quantities to behave more like prices.

While many papers have looked at banking behavior as a way to smooth deterministic costs over time, few have looked at banking behavior as a way to smooth uncertain costs. Those papers that have looked at banking under uncertainty did not focus on welfare, did not consider uncertainty about the future from the vantage point of the firm, and/or did not consider the typical banking constraint (that borrowing is limited). These features are important if we are interested in a normative analysis of realistic policies over a horizon where firms as well as regulators have incomplete information.

Our results confirm the intuition that optimizing firms would use banking to spread cost shocks over time, lowering costs compared to fixed quantities. However, this tendency is limited to periods where deterministic banking arises and by the horizon before which the banking constraint is likely to bind. Once the banking constraint is binding, adverse shocks cannot be

accommodated and even favorable shocks are unlikely to lead firms to save allowances unless the shock is sufficient to reduce today's marginal cost below expected future discounted marginal costs (which are typically much lower than today's marginal costs when banking is constrained). Based on the exceedingly flat marginal benefits associated with climate change policies, we know that price policies represent a limit on how much costs can be lowered from this kind of emission shifting, and the question remains: how much can banking match the performance of price policies?

Parameterizing our model based on recent U.S. cap-and-trade proposals, we find banking improves welfare by about \$1 billion per year compared to fixed quantities over a 20-year horizon, but subsequently banking constraints imply no further gains. In relative terms, banking achieves over 95 percent of the price advantage over the initial decade, and over 80 percent over the first two decades. More expansive borrowing provisions can extend the banking gains, as would lower correlation among cost shocks. Higher discount rates and lower initial allocations (or more restrictive borrowing), however, would all reduce the banking gains.

Given the observation that such policies frequently remain fixed for a decade or two at the most, limiting the relevant horizon for welfare comparisons, these results provide strong evidence that bankable quantity policies to address climate change can provide the bulk of the welfare advantages previously afforded to price policies. Still, a number of interesting areas remain for further work. Importantly, this analysis has not examined how the likelihood of policy adjustments might affect both banking behavior and the welfare analysis. Recent changes in the regulation of sulfur dioxide, for example, suggest potential advantages and disadvantages from banking in the face of policy changes. In addition, while we have sought to understand

how banking *ought* to proceed, it remains for future work to more carefully compare these predictions to observed behavior.

References

- Cronshaw, M. B., & Kruse, J. (1996). Regulated Firms in Pollution Permit Markets with Banking. *Journal of Regulatory Economics*, 9(2), 179-189.
- Ellerman, A. D. (2005). US Experience with Emissions Trading: Lessons for CO2 Emissions Trading". In B. Hansjgens (Ed.), *Emissions Trading for Climate Policy: US and European Perspectives*. Cambridge: Cambridge University Press.
- Feng, H., & Zhao, J. (2006). Alternative Intertemporal Permit Trading Regimes with Stochastic Abatement Costs. *Resource and Energy Economics*, 28(1), 24-40.
- Friedman, M. (1957). *A Theory of Consumption Function*. Princeton, NJ: Princeton University Press.
- Godby, R. W., Mestelman, S., Muller, R. A., & Welland, J. D. (1997). Emissions Trading with Shares and Coupons when Control over Discharge Is Uncertain. *Journal of Environmental Economics and Management*, 32(3), 359-381.
- Hoel, M., & Karp, L. (2001). Taxes Versus Quotas for a Stock Pollutant with Multiplicative Uncertainty. *Journal of Public Economics*, 82, 91-114.
- Hoel, M., & Karp, L. (2002). Taxes Versus Quantities for a Stock Pollutant. *Resource and Energy Economics*, 24(4), 367-384.
- Jacoby, H. J., & Ellerman, A. D. (2004). The Safety Valve and Climate Policy. *Energy Policy*, 32(4), 481-491.
- Kling, C., & Rubin, J. (1997). Bankable Permits for the Control of Environment Pollution. *Journal of Public Economics*, 64(1), 101-115.
- Leiby, P., & Rubin, J. (2001). Intertemporal Permit Trading for the Control of Greenhouse Gas Emissions. *Environment and Resource Economics*, 19(3), 229-256.
- Murray, B. C., Newell, R. G., & Pizer, W. A. (2009). Balancing Cost and Emissions Certainty: An Allowance Reserve for Cap-and-Trade. *Review of Environmental Economics and Policy*, 3(1), 84-103.
- Newell, R. G., & Pizer, W. A. (2003). Regulating Stock Externalities Under Uncertainty. *Journal of Environmental Economics and Management*, 45(2), 416-432.
- Rubin, J. (1996). A Model of Intertemporal Emission Trading, Banking and Borrowing. *Journal of Environmental Economics and Management*, 31(3), 269-286.
- Schennach, S. M. (2000). The Economics of Pollution Permit Banking in the Context of Title IV of the 1990 Clean Air Act Amendments. *Journal of Environmental Economics and Management*, 40(3), 189-210.
- Stavins, R. N. (1996). Correlated Uncertainty and Policy Instrument Choice. *Journal of Environmental Economics and Management*, 30(2), 218-232.
- U.S. Environmental Protection Agency (EPA). (2009). *The United States Environmental Protection Agency's Analysis of H.R. 2454 in the 111th Congress, The Waxman-Markey American Clean Energy and Security Act of 2009*. Environmental Protection Agency. Washington, DC.
- Weitzman, M. (1974). Prices vs. Quantities. *Review of Economic Studies*, 41(4), 477-491.
- Williams, R. C. (2002). *Prices vs. Quantities vs. Tradable Quantities*. NBER Working Papers 9283. National Bureau of Economic Research, Inc.
- Yates, A. J. (2002). Decentralization in Pollution Permit Markets. *Journal of Public Economics*, 4(4), 641-660.

Yates, A. J., & Cronshaw, M. B. (2001). Pollution Permit Markets with Intertemporal Trading and Asymmetric Information. *Journal of Environmental Economics and Management*, 42, 104-118.

Tables and Figures

Table 1: Parameter Values

<i>Parameter</i>	<i>Value</i>
Discount rate (r)	5.0%
Marginal cost slope (c_t)	12×10^{-9} \$/ton ²
Cost uncertainty (σ_0)	4 \$/ton
Correlation of shocks (ρ)	0.8
Initial emissions (\bar{q}_1)	6 billion tons
Baseline emissions growth ($g_{\bar{q}}$)	0.35%
Initial allocation (y_1)	6 billion tons
Allocation growth (g_y)	-3.4 %
Borrowing Limit (\bar{B}_{t+1})	$-y_{t+1}$
Marginal benefit slope (b_1)	5.6×10^{-13} \$/ton ²
Benefit growth rate (g_b)	2.5%

Note: \$ refers to 2005 US dollars and tons refers to metric tons of carbon dioxide.

Table 2: Relative Welfare Advantages over Fixed Quantity Policy

Policy Horizon	Price	Bankable Q	Banking gain/Price gain
1	0.67	0.66	0.98
5	5.7	5.6	0.98
10	12	11	0.96
20	21	17	0.83
40	29	17	0.58

Notes: Welfare advantages are the discounted expected cost differences between the policy listed and a pure quantity policy over given policy horizon. Welfare advantages are in billions of 2005 US dollars. Banking results are based on 10,000 simulations.

Table 3: 20-year Relative Welfare Advantages over Fixed Quantity – Sensitivity Analysis

Sensitivity	Price	Bankable Q	Banking gain/Price gain
Benchmark	21	17	0.83
Double shocks	83	67	0.80
High Discount Rate	17	12	0.71
Flat Allocation Path	21	2.6	0.12
Unlimited Borrowing	21	19	0.93
No Correlation	8.6	8.4	0.97

Notes: Welfare advantages are given at a 20-year policy horizon. Welfare advantages are in billions of 2005 US dollars. Banking results based on 10,000 simulations. All parameters at benchmark levels unless otherwise noted. Double shocks sets the cost uncertainty (σ_0) equal to \$8/ton. High Discount Rate sets $r = 0.075$. Flat allocation path sets $y_t = 2.84$ billion tons for all t . Unlimited borrowing sets $\bar{B}_t = -4y_t$. No correlation sets $\rho = 0$.

Figure 1: Bank Path

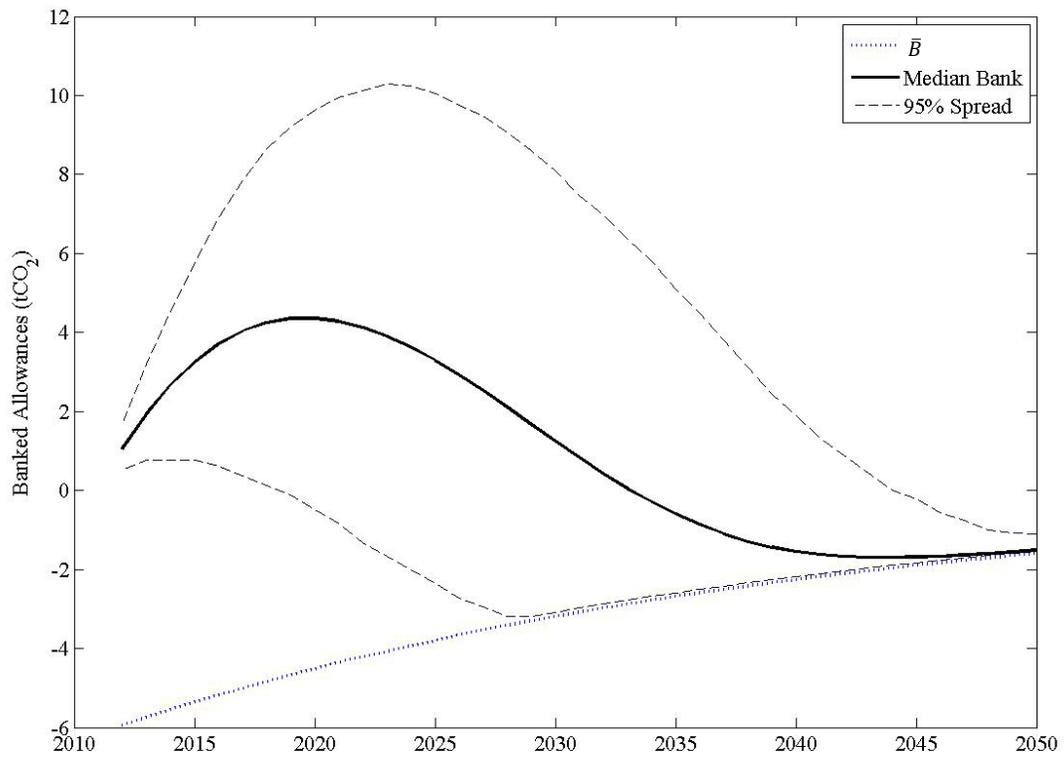


Figure 2: Probability of a Binding Bank Constraint in Each Year

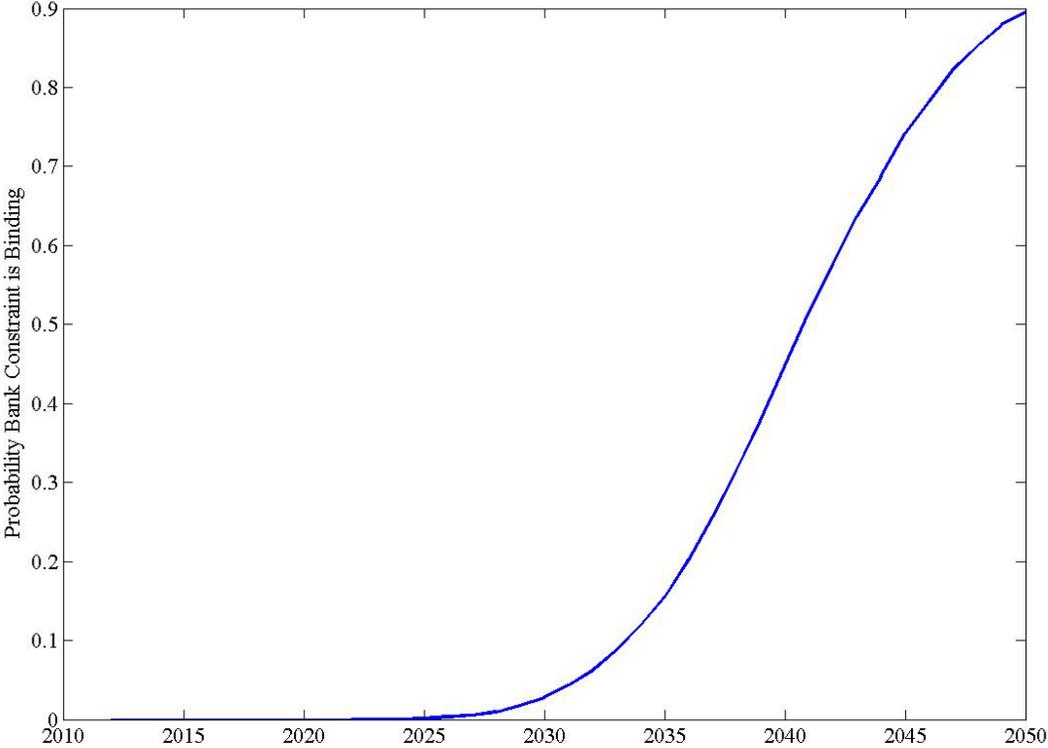
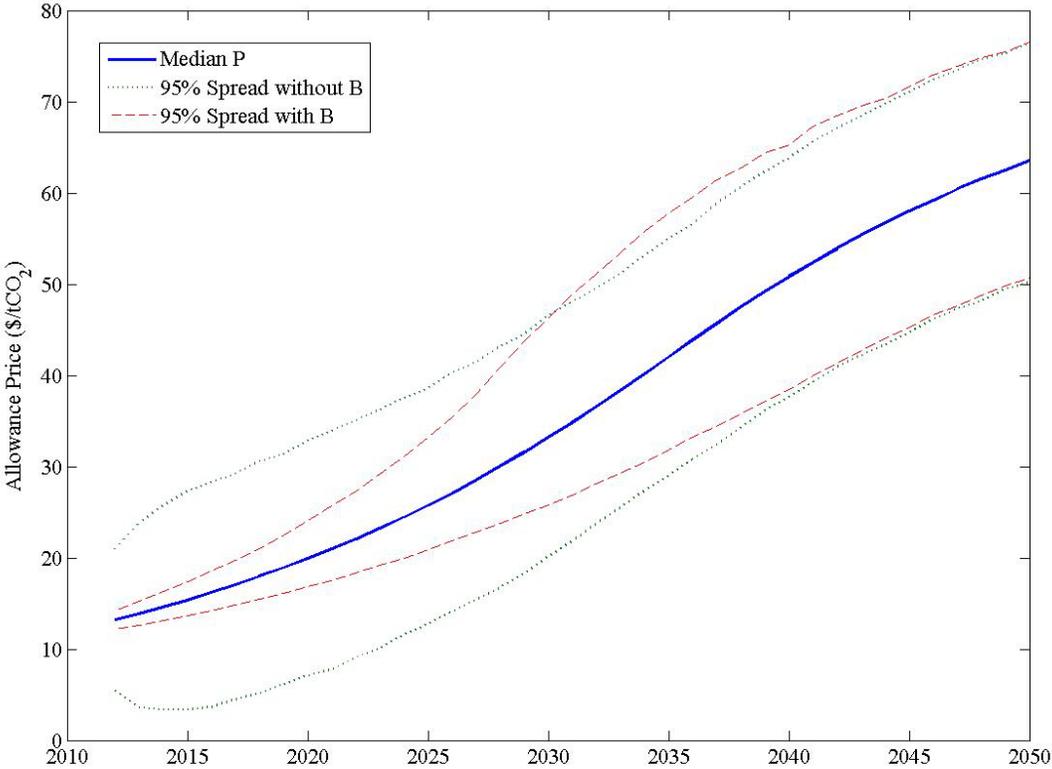


Figure 3: Price Path with and without Banking



Appendix

First-order conditions

We start by augmenting the Bellman equation (6) with the specified optimization constraints:

$$V_t(B_t, \theta_t) = \min_{q_t, B_{t+1}} \left\{ \begin{array}{l} \theta_t(q_t - \bar{q}_t) + \frac{c_t}{2}(q_t - \bar{q}_t)^2 + \beta E_t[V_{t+1}(B_{t+1}, \theta_{t+1})] \\ -\pi_t(B_t + y_t - q_t - B_{t+1}) - \mu_t q_t - \gamma_t(B_{t+1} - \bar{B}_{t+1}) \end{array} \right\} \quad (\text{A.1})$$

Here, $\beta = 1/(1+r)$ is the discount factor, E_t is the expectations operator given information at time t , and π_t , μ_t , and γ_t , all ≥ 0 , are Lagrange multipliers on the allowance constraint (4), non-negative emissions constraint, and the next-period minimum bank constraint, respectively.²³ The first-order conditions are then:

$$-\theta_t - c_t(q_t - \bar{q}_t) + \mu_t = \pi_t \quad (\text{A.2})$$

$$-\beta E_t \left[\frac{\partial V_{t+1}(B_{t+1}, \theta_{t+1})}{\partial B_{t+1}} \right] + \gamma_t = \pi_t \quad (\text{A.3})$$

In addition the constraints noted in (6), we have the complementarity conditions $\mu_t q_t = 0$ and $\gamma_t(B_{t+1} - \bar{B}_{t+1}) = 0$. Given the envelope condition result that $\partial V_t / \partial B_t = -\pi_t$, Equation (A.3) can be rewritten as the following Euler equation:

$$\pi_t = \beta E_t[\pi_{t+1}] + \gamma_t \quad (\text{A.4})$$

Derivation of price response to $\delta\theta_t$ with active banking for s periods

With banking active for s periods, but not longer, small quantity adjustments δq_{t+j} , $0 \leq j \leq s$, can be made over the next s periods but

$$\sum_{j=0}^s \delta q_{t+j} = 0$$

²³ See Section 3 and Appendix D of (Schennach, 2000) for further discussion. Her model is similar to ours (compare her Equation (16) with m and λ to our Equation (A.4) with π and γ) but her discussion is focused on the expected path rather than uncertainty and welfare.

That is, the cumulative change must be zero. From Equation (2) we know that

$\delta\pi_t = -\delta\left(\frac{\partial c_t}{\partial q_t}\right) = -\delta\theta_t - c_t\delta q_t$ and $E_t[\delta\pi_{t+j}] = -\rho^j\delta\theta_t - c_{t+j}\delta q_{t+j}$ for $s \geq j \geq 1$. Solving for δq_{t+j} and substituting into the above expression,

$$\delta\theta_t \left(\frac{1}{c_t} + \sum_{j=1}^s \frac{\rho^j}{c_{t+j}} \right) + \sum_{j=0}^s \frac{E_t[\delta\pi_{t+j}]}{c_{t+j}} = 0 \quad (\text{A.5})$$

Using equation (A.4), we can recursively solve for $\delta\pi_t = \beta^j E_t[\delta\pi_{t+j}]$ or $E_t[\delta\pi_{t+j}] = \beta^{-j} \delta\pi_t$.

Substituting again into the above expression,

$$\delta\theta_t \left(\frac{1}{c_t} + \sum_{j=1}^s \frac{\rho^j}{c_{t+j}} \right) + \delta\pi_t \sum_{j=0}^s \frac{\beta^{-j}}{c_{t+j}} = 0 \quad (\text{A.6})$$

Rearranging,

$$\delta\pi_t = -\delta\theta_t \left(1 + \sum_{j=1}^s \frac{c_t \rho^j}{c_{t+j}} \right) \left(1 + \sum_{j=1}^s \frac{c_t}{\beta^j c_{t+j}} \right)^{-1} \quad (\text{A.7})$$

Derivation of the relative advantage expression Δ_t^x in (14)

As noted in the text, the general policy x -quantity welfare difference Δ_t^x can be decomposed into effects on expected costs and effects on expected benefits. That is, Equation (13) can be rewritten:

$$\Delta_t^x = E[\text{NB}_{t,\text{policy}x}] - E[\text{NB}_{t,\text{quantity}}] = (B_{t,\text{policy}x} - B_{t,\text{quantity}}) - (C_{t,\text{policy}x} - C_{t,\text{quantity}}) \quad (\text{A.8})$$

Starting with the change in expected costs, if we rewrite the cost function by completing the square, we can express the cost function as

$$C_t(q_t, \theta_t) = \frac{c_t}{2} \left(q_t - \bar{q}_t + \frac{\theta_t}{c_t} \right)^2 - \frac{1}{2c_t} \theta_t^2 \quad (\text{A.9})$$

$$= \frac{1}{2c_t} p_t^2(q_t, \theta_t) - \frac{1}{2c_t} \theta_t^2 \quad (\text{A.10})$$

where equilibrium price p_t is defined by marginal abatement costs, $-\partial C_t / \partial q_t$, and the second line recognizes the squared expression in the first line as $-p_t/c_t$. Taking expectations and subtracting the case where $q_t = q_t^*$ to determine cost savings relative to non-bankable quantities in (A.8), we have

$$E_0[C_{t,\text{policy } x} - C_{t,\text{quantity}}] = E_0 \left[\frac{1}{2c_t} p_t^2(q_t, \theta_t) \right] - E_0 \left[\frac{1}{2c_t} \theta_t^2 \right] - \frac{c_t}{2} (q_t^* - \bar{q}_t)^2 \quad (\text{A.11})$$

$$= \frac{1}{2c_t} (E_0[p_t]^2 + \text{var}(p_t)) - \frac{1}{2c_t} \sigma_t^2 - \frac{c_t}{2} (q_t^* - \bar{q}_t)^2 \quad (\text{A.12})$$

Here, $\sigma_t^2 = \text{var}(\theta_t)$ is the variation in prices under non-bankable quantities, and $\text{var}(p_t)$ is the variation in prices observed under a mean preserving alternative, policy x .²⁴ Note that by constructing a mean preserving alternative x , $E_0[p_t] = E_0[-c_t(q_t - \bar{q}_t)] = E_0[-c_t(q_t^* - \bar{q}_t)]$, a result we used to simplify the final result:

$$E_0[C_{t,\text{policy } x} - C_{t,\text{quantity}}] = \frac{\sigma_t^2}{2c_t^2} c_t \left(\frac{\text{var}(p_t)}{\sigma_t^2} - 1 \right) \quad (\text{A.13})$$

Turning to the benefit expression in (A.8), we note that $E_0[S_t]$ under policy x equals $E_0[S_t^*]$, based on the linear stock accumulation (11) coupled with the assumption that policy x is a mean preserving alternative to q_t^* . Therefore,

²⁴ Note that the abatement cost function does not include adjustment costs. The inclusion of such costs would reduce the abatement cost advantage flexible quantity policies have over fixed quantity policies.

$$\begin{aligned}
E_0[B_{t,\text{policy}} - B_{t,\text{quantity}}] &= E_0 \left[\frac{b_t}{2} S_t^2 \right] - E_0 \left[\frac{b_t}{2} (S_t^*)^2 \right] = \frac{b_t}{2} \text{var}(S_t^2) \\
&= \frac{\sigma_t^2}{2c_t^2} b_t \frac{\text{var}(S_t^2)}{\sigma_t^2/c_t^2}
\end{aligned} \tag{A.14}$$

All of this leads to the generalized form given in (14):

$$\frac{\sigma_t^2}{2c_t^2} \left(c_t \left(1 - \frac{\text{var}(p_t)}{\sigma_t^2} \right) - b_t \left(\frac{\text{var}(S_t)}{\sigma_t^2/c_t^2} \right) \right) \tag{A.15}$$

where $\text{var}(p_t)$ and $\text{var}(q_t)$ are measured for any mean-preserving alternative to non-bankable quantities and which simplifies back to (15) when $p_t = -c_t(q_t^* - \bar{q}_t)$, $q_t = q_t^* - (\theta_t / c_t)$, $\delta = 1$, and $S_t = q_t$ under a price policy for a flow pollutant.

Parameterization

Our cost parameterization is based on recent estimates of U.S. compliance costs with H.R. 2454 (Waxman-Markey) given in (U.S. Environmental Protection Agency (EPA), 2009). In particular, we assume initial baseline emissions, \bar{q}_1 , of 6 billion tons CO₂e and initial allocation of allowances, y_1 , of 6 billion tons, with \bar{q}_t growing at 0.35% per year ($g_{\bar{q}} = 0.0035$) and y_t declining at 3.4% per year ($g_y = -0.034$). This parameterization is roughly in line with regulated abatement path of H.R. 2454 considered in EPA (2009). Given this allocation path and baseline emissions path, we parameterize c_1 such that $P_1 \approx$ \$13 per ton CO₂ underbankable quantities, the initial price estimate given in EPA (2009).²⁵ This leads to $c_1 =$ \$12 per ton per billion tons, or $\$12 \times 10^{-9}$ \$/ton². We assume an initial standard error for cost shocks, σ_0 , of \$4 per ton CO₂ based on the reported variation in EPA scenarios,²⁶ and a correlation of 0.8 based on

²⁵ See slide 14 of (U.S. Environmental Protection Agency (EPA), 2009) for initial allowance price estimates.

²⁶ See slide 12 of EPA (2009) for sensitivity analyses.

NP. Consistent with Waxman-Markey, we assume that firms can costlessly borrow from the next period's allocation, thus we set $\bar{B}_{t+1} = -y_{t+1}$.²⁷

On the benefit side, we start with the NP estimate of b_1 , 8.7×10^{-13} $\$/\text{ton}^2$ and multiply by 0.64 to convert from a price of *retained* CO₂ to a price of *emitted* CO₂.²⁸ This allows easier comparison with conventional cost analyses based on emissions, such as EPA (2009), but requires an adjustment to compare with the original NP results. As in NP, we assume these benefits grow at 2.5% and that the discount rate is 5%. Table 1 summarizes these benchmark parameter values.

²⁷ The Waxman-Markey legislation allows for borrowing from allocations beyond the next period's allocation, but an eight percent interest charge is assessed to these borrowed allowances. Such a steep borrowing charge makes borrowing from allocations beyond next year's allocation highly unlikely (see the "flat allocation path" sensitivity analysis). Thus, we omit this type of borrowing from our analysis.

²⁸ Roughly 36 percent of CO₂ emissions decay within one year; see appendix B in Newell and Pizer (1999).