NBER WORKING PAPER SERIES

BALANCING COST AND EMISSIONS CERTAINTY: AN ALLOWANCE RESERVE FOR CAP-AND-TRADE

Brian C. Murray
Richard G. Newell
William A. Pizer

Working Paper 14258
http://www.nber.org/papers/w14258

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2008

The research was supported in part by a grant from the Swedish Foundation for Strategic Environmental Research (MISTRA). The authors acknowledge Joseph Aldy, Jon Anda, Jason Grumet, Suzanne Leonard, Tim Profeta, Nicole St. Clair, Robert Stavins, Tracy Terry, an anonymous referee, and participants at the Resources for the Future workshop, "Managing Costs in a U.S. GHG Trading Program," (Pizer and Tatsutani 2008) for their insights and suggestions on this issue. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2008 by Brian C. Murray, Richard G. Newell, and William A. Pizer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Balancing Cost and Emissions Certainty: An Allowance Reserve for Cap-and-Trade
Brian C. Murray, Richard G. Newell, and William A. Pizer
NBER Working Paper No. 14258
August 2008
JEL No. D8,L51,Q54,Q58

ABSTRACT

On efficiency grounds, the economics community has to date tended to emphasize price-based policies to address climate change -- such as taxes or a "safety-valve" price ceiling for cap-and-trade -- while environmental advocates have sought a more clear quantitative limit on emissions. This paper presents a simple modification to the idea of a safety valve: a quantitative limit that we call the allowance reserve. Importantly, this idea may bridge the gap between competing interests and potentially improve efficiency relative to tax or other price-based policies. The last point highlights the deficiencies in several previous studies of price and quantity controls for climate change that do not adequately capture the dynamic opportunities within a cap-and-trade system for allowance banking, borrowing, and intertemporal arbitrage in response to unfolding information.

Brian C. Murray
Duke University
Box 90328
Durham, NC 27708
bcmurray@duke.edu

William A. Pizer
Resources for the Future
1616 P Street, NW
Washington, DC 20036
pizer@rff.org

Richard G. Newell
Nicholas School of the Environment
Duke University
Box 90227
Durham, NC 27708
and NBER
richard.newell@duke.edu

Balancing Cost and Emissions Certainty: An Allowance Reserve for Cap-and-Trade

Brian C. Murray, Richard G. Newell, and William A. Pizer

Introduction

The economic debate over using taxes versus cap-and-trade to control pollution emissions revolves around the relative merits of using prices versus quantities as the policy instrument. A cap-and-trade system fixes the quantity of emissions allowed but leaves the market price of emissions rights uncertain. In contrast, a tax fixes the price of emissions at the tax rate but leaves the quantity of emissions uncertain. This trade-off raises essential questions for policy design: which form of uncertainty is a greater burden to society? What can be done to minimize that burden or maximize net benefits? A sizable economics literature has addressed these questions, dating back to Weitzman (1974) and others.

Taxes and cap-and-trade are, in some sense, extreme examples of the alternative market-based approaches that are available to correct an emissions externality. The government stipulates that emitters must obtain the "right to emit." These rights (typically called allowances or permits) are either supplied with infinite elasticity at a fixed price (the tax) or with zero elasticity at a fixed supply (the cap). A key alternative—initially suggested by Roberts and Spence (1976) and later developed in the context of climate policy by Pizer (2002)—is the idea of a "safety valve," in which a cap-and-trade system is coupled with a price ceiling at which additional allowances can be purchased (in excess of the cap). So long as the allowance price is below the safety-valve price, this hybrid system acts like cap-and-trade, with emissions fixed but the price left to adjust. When the safety-valve price is reached, however, this system behaves like a tax, fixing the price but leaving emissions to adjust. Given the importance attached by many stakeholders and policymakers to containing the costs of any U.S. climate policy, this approach has received considerable attention in the U.S. debate over climate change regulation (e.g., Samuelsohn 2008), and has come to be known as the "cost-containment" issue (Pizer and Tatsutani 2008).

Cap-and-trade with a safety valve represents one of many possible mechanisms that lie between the two extremes of a pure price or a pure quantity instrument. It offers a more malleable supply curve for emissions allowances, containing both vertical and flat segments. This paper discusses a second mechanism that includes features of both price and quantity instruments. We believe this approach, which we call an *allowance reserve*, is particularly

promising. The basic idea goes one step beyond the safety valve: while the safety valve stipulates that an unlimited number of allowances be made available at the specified safety-valve price, the allowance reserve stipulates both a ceiling price at which cost relief is provided and a maximum number of allowances to be issued in exercising that relief. Much like a safety-valve mechanism can mimic either a pure price or pure quantity control, depending on how the cap and safety-valve price are set, an allowance reserve can mimic a pure price, pure quantity, or safety-valve approach, depending on how the ceiling price and volume are set.

Three motivations underlie our interest in this mechanism. The first two are largely practical in nature, while the third hints at a new twist on the conditions underlying optimality, in contrast to the traditional "prices versus quantities" perspective. The first motivation is simple: as we describe below, the safety valve represents a special case of the allowance reserve where the volume of available allowances is very large or unlimited. Thus, an allowance reserve has the capacity to do as well if not better than the safety valve in terms of matching public interest described below as a blend of economic efficiency and political feasibility. That is, political economy conditions suggest that public interest may be better served with an allowance reserve because it is more likely to sustain a coalition that will enable welfare-enhancing policy to be enacted.

Second, the reserve mechanism addresses one problem with a safety valve. Although most cap-and-trade programs permit allowance banking, which can help equilibrate present value prices across different time periods and increase dynamic efficiency, allowance banking coupled with a safety valve creates a dynamic problem. Suppose the cap needs to be tightened and as a result the safety-valve price is expected to increase dramatically at some point in the future. With an ordinary safety valve, an expectation of much higher prices in the future would lead rational firms to buy as many allowances as possible at the current, low safety-valve price in order to save them for use later when prices are high. Absent a mechanism to limit such purchases, they could effectively overwhelm efforts to tighten the future cap, thereby undermining long-term environmental policy goals. An allowance reserve would address this potential problem by placing an upper limit on the available number of extra allowances.

Finally, and most importantly at a fundamental level, most economic analysis of price and quantity controls under uncertainty does not adequately capture the dynamic nature of the regulatory process suggested by the preceding paragraph. In particular, as new information arises—about the benefits, costs, or global commitment to solving the problem of climate change—expectations about the likely long-term emissions level and emissions price will evolve. Therefore, in order to achieve dynamic efficiency, prices need to adjust regularly so that current

prices continue to reflect discounted expected future prices. A cap-and-trade program with banking, borrowing, and eventual adjustment of the cap can achieve that result *if* economic agents have sufficient foresight and capacity to form rational expectations about the longer term (Newell et al. 2005). This factor alone identifies an important advantage of dynamic cap-and-trade with banking and borrowing over other approaches. Nonetheless, these conditions may not hold—or at least are not assured—particularly in the early years of a program when cost uncertainty would be high, a significant bank would not yet have developed, and market actors would still be struggling to understand the new market. An allowance reserve could be used to help the market toward such an equilibrium by anchoring initial prices near or below the ceiling price.

We focus here on the importance for climate policy design of uncertainty in the costs and benefits of greenhouse gas mitigation. There are of course other important design factors to consider, including the degree to which the policy raises revenue (e.g., through taxes or allowance auctions), how those revenues are used (e.g., reducing other taxes, additional spending), and the stringency of the policy (i.e., the cap or tax level). Nonetheless, most of these other elements can be designed largely independently of the instrument choice of cap-and-trade versus a tax.

The remainder of this paper is organized into several sections. The next section provides background on market-based emissions regulation, including the current policy debate about price versus quantity instruments, and discusses the allowance reserve idea in more detail. This is followed by a discussion of the advantages of a reserve-based approach and how it addresses some key practical problems with the current suite of alternatives. We then discuss the issue of optimality in a dynamic context where policies evolve over time, making the case that (1) cap-and-trade with banking and borrowing could approach optimality with sufficient intertemporal flexibility, and (2) absent the institutions or foresight necessary for such optimality, the allowance reserve may be a useful way to help move market outcomes in the correct direction. We end with a discussion of the remaining issues that surround practical implementation of the allowance reserve, including establishing the ceiling price, reserve size, and release mechanisms. We present conclusions in the final section.

Market-Based Emissions Regulation and the Reserve-based Approach

Market-based emissions regulation works by requiring emitters to hold emissions allowances and then establishing a mechanism for supplying those allowances. The two simplest supply mechanisms are (1) a tax that fixes the price associated with purchasing allowances and

(2) cap-and-trade, which establishes a fixed supply of allowances, either auctions or gives them away for free, and then allows trading until the allowances are used to cover emissions. The tax is typically referred to as a price-based approach and cap-and-trade as a quantity-based approach to emissions control.[1]

A key point, highlighted by Weitzman (1974), is that price and quantity controls lead to distinctly different outcomes when there is uncertainty about costs. While emissions are constant under cap-and-trade, price varies; in contrast, under a tax, price is constant but emissions vary. Weitzman (1974) derived conditions under which one or the other policy is preferred in expected efficiency terms based on the relative slopes of the curves for the marginal cost and marginal benefits of emissions control. Since then, many papers have found that for climate change policies, the marginal benefits of mitigation (or marginal damages from emissions) are relatively flat over the relevant range of annual emissions, and, using a somewhat modified Weitzman argument, that price-based policies are therefore preferred in terms of economic efficiency (Kolstad 1996; Pizer 2002; Hoel and Karp 2002; Newell and Pizer 2003). Note that the quantity policy (i.e., cap-and-trade) modeled in these papers corresponds to annual emissions targets without banking or borrowing, a matter we return to below.

Of course, the perfectly inelastic (cap-and-trade) and perfectly elastic (tax) emissions allowance supply curves are the two simplest extremes of a wide range of policies the government could use to provide emissions allowances to the market. Roberts and Spence (1976) examined one alternative: coupling cap-and-trade with a price floor and ceiling. This approach generates three types of outcomes depending on the realized demand: (1) when demand is low, the price is set by the floor, and the quantity of allowances is below the cap; (2) when demand is moderate, the quantity of allowances is determined by the cap, and the price is somewhere between the floor and ceiling; and (3) when demand is high, the price is set by the ceiling, and emissions are above the cap.[2] Depending on the choice of design parameters (i.e., cap, floor, ceiling), the policy also has the ability to mimic either a tax (if the price ceiling or cap level is sufficiently low) or pure cap-and-trade (if the floor is low and the ceiling high). Owing partly to the previously mentioned authors' emphasis on price-based policies and partly to the politics of wanting to have both certainty about prices and stringent emissions limits, there has been a

[1] These two allowance supply approaches are shown in Appendix Figure 1 along with two alternative outcomes for emissions demand.

[2] These outcomes are shown, respectively, as e, f, and g in Appendix Figure 2.

significant emphasis on policy with a relatively low, stringent cap level and low price ceiling. This approach, where the price ceiling is referred to as a "safety valve," has garnered considerable attention and political support over the past five years as climate policy proposals have made their way to Congress (Samuelsohn 2008). The price floor, though it has received less attention in the federal policy debate, is being implemented in the Regional Greenhouse Gas Initiative cap-and-trade program in the Northeastern U.S. states.

Representing the allowance reserve idea requires only a slight adjustment to the Roberts and Spence (1976) supply schedule (see the right panel of Appendix Figure 2). The price ceiling that previously allowed an unlimited volume of allowances to be purchased now also has a quantitative limit, which is the "allowance reserve." Basically, we have simply added another kink in the allowance supply schedule and made it more flexible in its ability to balance price and quantity goals. Indeed, the first-best policy would be to specify an allowance supply schedule that mimicked the marginal damages from higher emissions. In this sense, the allowance reserve offers a well-defined improvement over the alternative policies developed so far, each of which remains a special case. In essence, the reserve can be deployed in a way that reflects something closer to the increasing marginal social cost of emissions.

When implemented, all market-based policies require us to identify a group of regulated entities whose direct emissions or embodied emissions (for upstream regulation of fuels) are measured and reported on a regular basis, typically annually. Under a tax policy, those entities are then required to pay a specified tax ($/ton) applied to the measured amount of emissions. Under cap-and-trade, they are required to acquire and surrender allowances.

A key feature in virtually all proposed greenhouse gas cap-and-trade programs is banking, under which unused allowances in one year can be used in subsequent years. With banking, there can be an incentive to reduce emissions early—particularly during a gradual phasedown of emissions targets—and it is not necessary for the market to meet the target exactly each year. If that were the case, there would be a danger that requiring emissions to match the number of allowances exactly would result in either too few allowances—causing the price to skyrocket— or too many allowances—causing the price to plummet. The former occurred in the California NO_x RECLAIM market; the latter occurred in Phase I of the EU Emissions Trading Scheme (ETS) for greenhouse gases. Both systems significantly restricted banking and borrowing across compliance periods. In the EU ETS, the main culprit was that banking was not allowed between Phase I (pre-Kyoto) and Phase II (Kyoto). That, combined with a generous allocation, eventually led to an excess supply of allowances and drove the price to zero at the end of Phase I. In contrast, systems that have allowed banking (and possibly borrowing) have tended

to have much smoother price behavior as the price at the end of one period tends to match the price at the beginning of the next due to allowance fungibility across periods and market arbitrage.

What about more complex policies? The price floor in the Roberts and Spence (1976) hybrid policy could be implemented in two ways. If the allowances associated with the cap are all distributed for free, the only alternative is for the government to agree to buy any allowances that regulated entities are willing to sell at the specified floor price. If, however, some of the allowances are auctioned, the price floor could be implemented by specifying a minimum price in the auction. In this way, allowances only enter the market if the price meets or exceeds the floor; otherwise, less than the full volume of allowances are sold.

The price ceiling, or safety valve, could be implemented by having the government agree to sell additional allowances at the specified ceiling price. However, there has been a wrinkle in such legislative proposals (e.g., S. 1766 in the 110th Congress, the "Bingaman–Specter" bill); that is, unlike ordinary allowances, these additional allowances are not bankable and must be used in the year they are released. This places an implicit limit on the volume of safety-valve allowances that might be sold in any year, namely the total volume of emissions for that year. Thus, under such proposals, one could in principle use safety-valve allowances to meet all of one's current-year emissions obligations and bank ordinary allowances for the future. Another wrinkle in the safety-valve provision of S. 1766 is that the safety valve is only available during one month each year, while firms are doing final balancing of their emissions and allowance holdings. This avoids a potential run on the safety valve while Congress might be debating whether to raise the level or remove the safety valve altogether in the future—a debate that would hopefully be completed during the eleven-month period when the safety valve is unavailable. We return to this issue below, as it is not obvious that such a sequence of events is likely.

The allowance reserve takes the price ceiling idea a step further. As just described, an unlimited nonbankable safety-valve could allow the release of up to one year's worth of emissions at any one time. The allowance reserve, however, could limit the use of this safety valve to a significantly smaller amount. The appropriate size of the reserve will ultimately depend on the stringency of the cap, the ceiling price, and the degree of remaining price volatility that is acceptable. A reserve of perhaps ten to twenty percent of the annual cap would reflect the range of emissions reductions sought by many current proposals over the first decade, coupled with varying assumptions about the price ceiling. The issue of how to choose the reserve size is further addressed later in this paper.

This raises an important question: how does the government allocate the extra allowances from a reserve if demand exceeds reserve supply at the ceiling price? There are several ways to do this. These are outlined in detail below, but perhaps the most compelling is analogous to the price-floor approach, but instead auctions the reserve allowances with a minimum price that is equal to the ceiling price (versus the floor price). The result would be: (1) no sales, (2) sales less than the limit, at the ceiling price, or (3) sales equal to the limit, at or above the ceiling price. Thus, the allowance reserve does not guarantee the ceiling price in the same way as an explicit price ceiling or safety valve. On the other hand, as discussed in the next section, it has several practical and theoretical advantages.

Advantages of an Allowance Reserve

Representing Marginal Damages Across Cumulative Emissions

Based on analyses of marginal damages from emissions, the research cited above finds that the allowance supply schedule for emissions should be roughly flat over the relevant range of annual greenhouse gas emissions. This would seem to suggest that the allowance reserve idea offers no efficiency improvement over either a tax-based or safety-valve approach. Yet that research does not consider the marginal benefit function over cumulative greenhouse gas limits (or in turn the shape of the associated allowance supply schedule) over longer time horizons. Indeed, it seems almost certain that when viewed over many decades of cumulative emissions, the marginal damage of the first ton abated would be higher than the marginal damage of the last. In this case, the additional kink in allowance supply represented by the reserve approach, cumulated over many years, should be able to better represent an upward sloping marginal damage function and deliver an outcome that is more efficient than the tax-based and safety-valve approaches.

Of course, given the tremendous uncertainty and time scale concerning climate change (Weitzman 2008), we must be cautious about economic analyses of the level and shape of the climate mitigation benefit function. Moreover, we believe there are yet other reasons to expect that traditional price and quantity comparisons are problematic in a dynamic setting—an issue we return to in the next section.

Expanding Political-Economic Flexibility

Another important concern is that most environmental advocates have opposed any price-based approach, including the safety-valve variant. In an October 8, 1997, letter to the President

in advance of negotiations on the Kyoto Protocol, seventeen environmental advocacy groups indicated their opposition to a safety valve mechanism.[3] More recently, however, these groups have expressed openness to the idea of a quantity-limited safety valve captured in the allowance reserve approach. Leading environmental advocacy groups, including some of those who signed the 1997 letter opposing a safety valve, supported an amended version of the America's Climate Security Act (S. 3036) in 2008, which included the allowance reserve idea (Eilperin and Mufson 2008). In this way, a simple interpretation of the allowance reserve—that its additional flexibility can better represent public interest—may be the most relevant argument when "public interest" includes not just economic views of optimality but also the perspective of key stakeholders. In other words, one very practical advantage of the allowance reserve idea is that it may be able to bridge differences between environmental advocates seeking a cap on emissions and industrial interests concerned about costs, in much the same way that some viewed the safety valve more than a decade ago (Kopp et al. 1997). Operating under the presumption that failure to enact a climate policy at all would lower social welfare, all else equal, a design element such as an allowance reserve that can break an impasse,can enhance overall efficiency relative to the status quo.

Addressing Concerns over Ability to Achieve Long-term Targets

There is a second practical reason for considering the allowance reserve over the pure safety-valve idea: How would one otherwise deal with evolving expectations of stricter targets and higher prices? That is, despite the attempt to structure current proposals with targets through 2050, it seems almost inevitable that revisions will occur after a decade or so. In anticipation of tightening future caps, current prices would rise assuming that current allowances could be banked for future compliance obligations, which are now anticipated to be more expensive. For

[3] See Samuelsohn (2008) and link to letter at http://www.eenews.net/features/documents/2008/02/21/document_cw_01.pdf. Specifically, they stated "this proposal would weaken, if not eliminate any incentive for private sector innovation and investment in clean technologies." Although one can understand the reluctance of environmental groups to embrace policies allowing greater emissions, the argument against the safety valve based on innovation incentives is flawed. Curtailing the possibility of very high allowances prices would not "eliminate" the incentive for clean technology innovation and adoption, although it would curtail the incentive to do so for very expensive technologies that would only be competitive above the safety valve price. Assuming the safety valve price is set appropriately, however, this is desirable because environmental policies should not, from an economic perspective, seek to promote technology at any cost. Rather policies should induce an *efficient* amount of innovation and adoption, consistent with societal willingness to pay (Kerr and Newell 2003).

example, the U.S. sulfur dioxide trading program was revised in 2005—fifteen years after passage of the 1990 amendments establishing the program—in a way that lowered allowed emissions by fifty percent in 2010 and seventy percent in 2015 (U.S. EPA 2005). In response, as shown in Figure 1, the price of allowances began a significant run-up in 2004 as debate began in earnest over tightening the emissions cap under the program through the Clean Air Interstate Rule. By 2005, the rules were finalized, with a halving of the emissions limit set to begin in 2010. Allowance prices peaked soon after. By 2008, the prices had settled down to roughly double their predebate level, with a May decline in part reflecting legal challenges to the rulemaking (Argus Media 2008) or possibly expectations that climate change regulation will depress future SO_2 allowance prices. All of this has happened years in advance of the actual change in emissions limits. So clearly market participants do act in anticipation of future target stringency.

All of this points to a potential problem with the ordinary safety valve when it is coupled with banking and evolving expectations of stricter targets. Under these circumstances, as firms and individuals become convinced that future prices will be well above the current safety valve, they will want to make use of the safety valve as much as possible, acquiring emissions allowances cheaply now that will quickly become more valuable in the future. Or, if safety-valve allowances cannot be banked, will allow regulated entities to preserve more valuable ordinary allowances for the future. That is, even without the ability to bank safety-valve allowances, there is a real possibility of accumulating multiple years' worth of allowances if people become convinced of the impending change many years in advance. The SO_2 trading program, for example, saw more than a year's worth of allowances accumulated early in the program without a safety valve, owing to the relatively easy targets from 1995 through 1999 and anticipation of stricter targets legislated for 2000.

The accumulation of a large bank of allowances—perhaps more than an entire year's worth of allowances—poses two related problems. The first is superficial: from an appearance standpoint, people may see a run on the safety valve, and a large accumulation of allowances from it, as a systemic failure. The second is related, but more substantive: a particularly large bank could begin to thwart efforts to cut emissions in the future. This is not an issue in the SO_2 program because emissions reductions are relatively large compared with historic emissions—fifty percent in the 1990 amendments, starting in 2000, and fifty percent again in 2010 under the Clean Air Interstate Rule. One year's worth of banked allowances would be used up in two years following a fifty percent cut (were facilities to try avoiding their fifty percent cut in emissions). In contrast, CO_2 emissions reductions are anticipated to occur more slowly as entirely new

technologies cutting across many sectors must be brought into use. A relatively tough target might mean a ten to twenty percent reduction from baseline within the first decade, in which case a bank on the order of one year of allowances could delay such a change for five to ten years without reducing emissions. We emphasize only that this could (but not necessarily would) be a problem because, even in the worst case, the tougher target could be designed with the bank in mind, in much the same way that programs with offset credits from uncapped sources often seek a tougher target than would be practical if those offset opportunities did not exist. Further, there is little evidence concerning how large of an allowance bank firms might accumulate (it could, in fact, be much larger than one year's worth of allowances), how fast they might spend it down, and in turn how much this might affect any future tightening of the cap.

The allowance reserve tackles both potential problems head on by simply limiting the volume of extra allowances entering the market and therefore limiting the potential for these extra allowances to contribute to an excessively large bank. As noted above, existing legislative proposals for a safety valve limit the released volume to the annual emissions level. With emissions reductions of perhaps ten to twenty percent per decade, this seems far more than is necessary to deal with anything except the desire to bank. In this case, an annual allowance reserve limit of about ten to twenty percent of the cap should be sufficient to address short-term uncertainty while leaving longer-term expectations free to drive near-term prices.

Optimal Policy in a Dynamic Setting

Most of the literature comparing price and quantity policies has ignored the aforementioned dynamic feature: that policies will inevitably be revised as new information arises and policymakers revisit the issue—what we might call dynamic price and quantity policies (i.e., policies that are updated over time). Newell et al (2005) emphasize that such revisions can be used to make a dynamic quantity policy mimic an unadjusted price policy. Here we suggest that a dynamic quantity policy might do better, even when the price policy is dynamic as well, particularly in a world where future damages depend only on cumulative emissions and not on their time path, as is roughly true for greenhouse gases. The key is intertemporal flexibility coupled with foresight about these revisions. As we elaborate below, the allowance reserve may in turn help foresight to drive near-term prices—in this case a desirable, even optimal, feature.

To illustrate this point that dynamic quantity policies may do better than dynamic prices, let's imagine a simple world with three time periods: when current policy is set (period 0), when that policy takes effect and firms respond (period 1), and some period in the future when policy

can be revised (period 2). Importantly, improved information on costs, benefits, and participation is arriving each period, so that there is a better notion of the optimal policy in period 1 (when no policy adjustment is possible) and an even better notion in period 2. For simplicity, we could assume that period 2 involves complete knowledge of costs and benefits and is also the last period of relevant activities. In any case, with better information in period 2, one can revise either a price or quantity policy to deliver improved outcomes in period 2 because there will be a better sense of how to balance costs and benefits compared to period 0.

Assuming revised price and quantity controls are equally efficient in period 2, the question of comparing various policies hinges on what happens in period 1 when firms respond to policies set in period 0, but with improved knowledge about costs and benefits as well as foresight about period 2. Consider the first-best outcome. Based on the working assumption that damages depend only on cumulative emissions, efficiency would lead us to minimize the expected present value of the total emissions abatement costs associated with achieving the cumulative emissions limit decided in period 2. This leads to a simple efficiency condition that the marginal cost (i.e., emissions price) each period should equal the present value of expected long-run marginal costs (see the Technical Appendix for a mathematical formulation of this first-order condition and the arguments that follow.) That is, it would be optimal to choose period 1 emissions such that marginal costs in period 1 are equal to the (discounted) expected marginal costs of meeting the cumulative target through the revised period 2 cap. Given the limited information available when period 1 emissions must be chosen, the cumulative cap will not be known exactly, but with additional information relative to period 0, expectations of the period 2 cap should be revised from the expectations in period 0 when the policy is set. Now that we understand the first-best outcome conditional on available information, we can examine how dynamic price and quantity policies compare in period 1.

Specifically, consider two policies set in period 0 and revised in period 2: a tax and a cap-and-trade program, where the cap-and-trade program allows banking (and borrowing, if necessary).

Performance of a Tax Program

An optimizing government that is setting the period 1 tax in period 0 would choose a tax level that equates the present value of expected marginal costs across the periods given the information it has at the time, thereby minimizing expected total costs as seen in period 0. The important point is that in period 1, firms would then choose to emit an amount such that their marginal costs—given the resolution of cost uncertainty in period 1—are equal to the tax which

was set in period 0. Firms will not match their marginal costs to the expected period 2 marginal costs (updated with new information on both costs and mitigation benefits in period 1) because there is no incentive to do so. Specifically, there is no ability to shift compliance obligations from the period with high (expected, discounted) costs to the one with low (expected, discounted) costs in a tax-based system. The emissions outcome in the first period would therefore not generally satisfy the previously mentioned efficiency condition because expectations about period 2 marginal costs will have changed between periods 0 and 1, but no responsive action will be taken by the affected parties.

This type of result is inherent in the classic Weitzman framework where policies are fixed prior to uncertainty being revealed. Neither a price nor (a nonbankable) quantity policy is optimal ex post because neither exactly matches realized (or updated expectations about) marginal costs and marginal benefits. Both instruments are generally inefficient in such a setting, so the issue becomes one of choosing the instrument with the lowest deadweight loss. Even when period 1 brings about expected changes in period 2 tax rates, there is virtually no incentive to deviate from the otherwise standard behavior setting period 1 marginal costs equal to the fixed-in-period-0 tax. The only possible incentive to deviate arises if changed expectations about future tax rates affects investment in long-lived emissions abatement capital that would be subject to the future tax.

Performance of a Cap-and-Trade System

The question is can cap-and-trade in a dynamic setting with banking and borrowing do any better? To find out, let's imagine that instead of a tax the government sets a period 1 cap in period 0, firms decide how much to emit during period 1 and bank or borrow until period 2, and everyone expects that in period 2 the government will set a period 2 cap to deliver the ultimate, optimized-with-complete-information-in-period-2, cumulative emissions target (or, more generally, a cap based on better information in period 2). Note that in period 2 the government can enforce any emissions level by accommodating or absorbing the bank or allowance debt acquired by firms in period 1. With this information in period 1, a cost-minimizing firm would form its best expectation of period 2 marginal costs and choose period 1 emissions such that marginal costs in period 1 would equal the discounted value of the expected period 2 marginal cost, regardless of the first period cap.

Thus, a cap-and-trade system with banking, borrowing, and an expectation of eventual adjustment of the emissions target can achieve the best possible outcome given the information that is known in period 1 even though policy is set in period 0. What drives this result?

Why Dynamic Cap-and-Trade Can Deliver a Better Outcome

The key is intertemporal flexibility and foresight. Through dynamic market arbitrage, whereby firms equate (present value) prices in periods 1 and 2 for the perfectly fungible allowances, the cap-and-trade system allows the information revealed about benefits, costs, and future expected targets to be transmitted to markets today. That is, knowing that new information on costs and benefits gained during the first period of the policy will lead to adjustment of future caps, firms have an incentive to adjust emissions during period 1 so that they can bank (or borrow) more (or less) now in order to equate marginal costs over the two periods. The existence of an intertemporal market for emissions allowances—something that is absent with a tax—provides the vehicle for doing this. Note that in terms of the efficiency condition, benefit information is transmitted through expectations about the cumulative target, while cost information is transmitted through both the cost function itself and expectations about the cumulative target.

The tax instrument, in contrast, only provides market incentives for adjusting emissions in response to information revealed about period 1 costs in a simple way that keeps marginal costs equal to the fixed-in-period-0 tax (and does not respond at all to changes in expectations about benefits or future targets). With a tax instrument, even if firms correctly anticipate a higher marginal cost or tax in the future, they cannot arbitrage against this outcome by overcomplying now and banking residual allowances for use in the future. This undermines their ability to efficiently manage costs. Taxes (like the cap) can of course be adjusted over time, but during the period between adjustments there will be inefficiently high or low levels of abatement and costs.

Interestingly, this incentive structure differs in a fundamental way from the classic Weitzman setting. In that static setting, only the tax (or price) policy provides incentives for firms to change behavior, only in response to new cost information, and only in a simple way that keeps marginal costs constant. The quantity policy in that case does not transmit any new information—firms must simply meet the target and have no flexibility to adjust by banking or borrowing. Neither policy transmits any new information about benefits or future targets. In contrast, in the dynamic cap-and-trade setting that is relevant here, firms do have an incentive to adjust under the quantity policy in response to both new cost and new benefit information because of adjusted expectations about future targets and marginal costs. While both policies can eventually be adjusted to achieve the desired target, the dynamic cap-and-trade policy provides a mechanism for firms to respond during the first period, when policy is fixed, while the tax does not.

All of this suggests that for a cumulative emissions problem like greenhouse gases, a cap-and-trade program with sufficient banking and borrowing can in principle deliver a better outcome than taxing emissions. This conclusion has been recognized to some degree for some time (Jacoby and Ellerman 2003). Extending prior research on optimal banking and borrowing (Rubin 1996, Kling and Rubin 1997) to a stochastic instrument choice context, Newell et al. (2005) rigorously showed how intertemporal banking and borrowing would allow firms to smooth abatement costs across time, thereby offsetting the traditional disadvantage of cap-and-trade relative to taxes. They also suggested several practical mechanisms for implementing such an approach, including an allowance reserve. What is new here, we believe, is that this is the first time conventional economics has suggested cap-and-trade can be better than tax-based approaches based on Weitzman-like efficiency grounds, with appropriate dynamic modifications. The key, as discussed above, is that most previous analyses have either ignored or underappreciated both the evolution of information and the dynamic nature of policymaking that are core features of a long-term problem like climate change—as well as the common feature of banking in most trading programs.

How Can an Allowance Reserve Enhance Efficiency?

The discussion and results above raise the question: why do we need an allowance reserve at all if cap-and-trade with sufficient banking and borrowing can be optimal (given available information)? There are at least three reasons. First, the allowance reserve does nothing to upset this result. Indeed, the main point is that the period 1 cap does not really matter so long as there are rational expectations about future caps. Second (and somewhat countering the first point), it may be important for the government to send signals concerning its current expectations about the long-term cap and expected price. This means that not only is the period 1 cap important; so are expectations (in period 0) of future marginal costs and allowance prices in period 2, which also depend on future targets and benefits. The ceiling price in the allowance reserve mechanism is one way the government can signal an initial expectation about the correct current and future prices.

A third and important reason for considering an allowance reserve is the concern that borrowing—a key mechanism for dealing with unexpectedly high costs in the short-term—may not work as we have assumed. Borrowing may not be implemented or it may be constrained in ways that limit its usefulness. To date, market-based policies have included only limited borrowing mechanisms. For example, the corporate average fuel economy (CAFE) program for light-duty vehicles allows a firm to undercomply in a given model year if it repays the borrowed

credits within the subsequent three model years. Meanwhile, there are examples of exceptionally high prices early in a borrowing-constrained cap-and-trade program as market participants anticipated or experienced a shortage of allowances. These include both the NO_x State Implementation Plan (SIP) Call in the United States and the EU ETS. In the context of an emissions phasedown of the type discussed for greenhouse gas policy, a well-designed allowance reserve would change the market dynamics so that high prices tap the reserve and alter the market from tending to borrow allowances in the short term to either meeting demand or potentially banking allowances.

Implementation Issues

We turn next to a number of important practical issues surrounding the implementation of an allowance reserve. Most immediate are determining the appropriate ceiling price at which the reserve can be drawn down and the size of the reserve. Additional issues include whether the reserve expands or attempts to maintain the cumulative cap, how reserve allowances are introduced to the market, and whether the reserve design parameters would be managed by an executive board or decided through legislation.

Ceiling Price and Reserve Size

The most challenging implementation questions are the ceiling price at which the reserve can be tapped and the size of the reserve. In principle, the ceiling price should be related to the marginal benefit of emissions reduction to ensure that allowance prices—an indicator of marginal abatement costs—stay in line with marginal benefits. As noted early on, however, the marginal benefits of greenhouse gas reduction are not likely to be well-defined and are affected by some factors beyond policymakers' control, including the extent to which other countries undertake emissions reductions. Thus policymakers may be more likely to focus on choosing a target ceiling price that is simply "not too high," meaning that it does not create seemingly excessive hardship for the overall economy. Or, if there is a range of likely allowance prices and economic impacts associated with a chosen emissions limit, the ceiling price might be set at the upper end of the predicted range, assuming policymakers and stakeholders are comfortable with both the cap and the price range.

If the allowance reserve is intended to credibly meet near-term demand at the ceiling price, then the ceiling price and the size of the reserve are inter-related. A low ceiling price will require a larger reserve to credibly deliver that price. A greater number of near-term events (e.g., weather, economic fluctuations) would be likely to come up against a low ceiling price and

therefore require a larger reserve to meet that near-term demand. Alternatively, if the ceiling price is set high, the reserve plays a lesser role and can be smaller, as the circumstances under which it is likely to be used become more rare. The size of the reserve essentially determines its power to keep the allowance price at or below a given ceiling price; a larger reserve is necessary to ensure lower prices. A distinct issue—not directly addressed here—is whether the allowance reserve might be capped not only annually, but also cumulatively over time and/or phased out.

One might argue that if the reserve is going to work as advertised—by providing strong and reliable relief against a run-up in prices driven by near-term events—it should be large enough to meet demand at the specified ceiling price under most foreseeable circumstances. The reserve could be set up to accommodate, for example, all conceivable demand shifts and still maintain the ceiling price by providing enough additional allowances to meet demand at that price. This could be informed by a convincingly large number of modeling scenarios that exogenously set the allowance price equal to the candidate ceiling price. The possible shortfall of allowances at that price (the difference between the emissions projected at that price and the proposed cap) would provide an estimate of the reserve size necessary to maintain that price and cover potential shortfalls.

Another possibility—if the program is expected to lean heavily on offsets (i.e., emissions reductions from outside capped sources) to achieve the cap—is to size the reserve to match the expected offset supply in the event that such offsets fail to materialize. The Regional Greenhouse Gas Initiative, for example, allows additional use of offsets at certain allowance price thresholds. However, the availability of cost-effective offset opportunities is only one source of cost uncertainty, so its importance would have to be evaluated alongside other sources of uncertainty.

Finally, in determining an upper bound for the size of the allowance reserve, it would not make sense to have the allowance reserve be larger than the difference between the target and the highest business-as-usual emissions forecast. An indefinite reserve of that size would be capable of lowering allowance prices to zero under the most pessimistic conditions, and therefore in practice it would go underused since reserve allowances would only be released at a price at or above the ceiling price.

Maintaining the Cumulative Cap vs. Establishing a Range

Because uncertainty about costs and allowance demand is likely to be highest at the beginning of a cap-and-trade program, we presume the reserve would be in place from the program's inception. Before trading can begin, the government must allocate allowances to

regulated entities either through free allocation or an auction. The existence of a reserve means that a separate allocation must also be made to a reserve account. There are two options for creating this reserve account: (1) create it from future allocations that, if never used, go back to the future allocation, which would maintain the cumulative cap over time; or (2) create it from allowances that, if never used, would vanish, which would establish a range of possible cumulative emissions outcomes that depend on the degree to which the reserve is tapped. It is also possible to construct a combination of these two options, with some reserve allowances drawn from the future and others not. If the cap-and-trade policy includes a price floor, as suggested above, reserve allowances could also come from any allowances that remained unused in prior periods. And, revenue from the sales of reserve allowances could finance offset purchases. Both of these options are variants that lie somewhere between maintaining the cumulative cap and creating a range of possible cumulative emissions outcomes.

Based on recent policy proposals, a U.S. cap-and-trade program would likely establish an allowance cap that starts in the near term with allowance quantities that are perhaps five to ten percent below current emissions levels. The cap would then be scheduled to decline over several decades until a substantial reduction in annual emissions is achieved. Recent proposals have called for reductions on the order of fifty to eighty percent below current levels by 2050, or about ten to twenty percent per decade. It is unlikely that all of the allowances over an almost forty-year period would be allocated up front. Therefore, the unallocated future allowances could serve as a source of reserve allowances. Again, if the objective is to maintain the long-term cap, and if these reserve allowances are in fact drawn down, this implies that future caps (unless modified in the future) will be that much tighter.

A policy that seeks to maintain the same cumulative cap, even as the allowance reserve is tapped, would likely create expectations of higher future prices if the reserve allowances are used now to lower current prices (rather than banked to comply with the now tighter future cap). Such behavior would make sense if current prices are high compared to long-term expectations because borrowing—which would be desired to arbitrage long-term low prices against short-term high prices—is either constrained or unpalatable. However, most recent economic modeling of cap-and-trade proposals shows a strong tendency toward allowance banking in the early years of a program (EIA 2008, EPA 2008, Murray and Ross 2007, Paltsev et al 2007). In this case, allowances from the reserve should not be necessary to offset near-term shocks (which the banked allowances can address) and, if the reserve is tapped, it should not depress current prices. With allowances already being banked to reflect future scarcity, any allowances moved from the future to the present via the reserve would tend to be added to the current bank and

returned to the future. It is only in the situation when firms are constrained in some way that is not well-captured by the referenced modeling results—particularly by a near-term shock before a bank can be built coupled with an explicit or implicit limit on individual borrowing—that system-wide borrowing from future allocations would represent a relaxation of that constraint, thereby lowering prices and containing costs.

An alternative approach would be to establish the reserve with allowances that, if unused, would vanish. Here, the cumulative cap is a range and tapping the reserve more clearly loosens the emissions constraint. The lower bound of the cumulative cap defines the aspirational target of the policy if the reserve is never tapped and the price remains below the ceiling price. The upper end of the range, defined by the cumulative effect of tapping the reserve, reflects the maximum allowable cumulative emissions. Based on the earlier discussion of how one would set the size of the reserve, this should be sufficient to maintain the ceiling price unless future expectations drive prices higher.

Just as the approach of system-wide borrowing from future allocations may make more sense if there is strong societal commitment to a specific cumulative cap (and a willingness to accept the cost consequences), the cap-range approach may make more sense if there is strong societal commitment to maintaining incremental costs below the ceiling price (and a willingness to accept the emissions consequences). Of course, in either case the long-term cap will undoubtedly be adjusted in the future; the main issue here is how the specification of a default cumulative cap (be it larger or smaller) may affect future expectations and indeed future action. Both approaches address short-term constraints with an appropriately chosen ceiling price and reserve size. However, the future borrowing approach, which maintains a predetermined cumulative cap, may create higher future price expectations and induce more mitigation than the range approach with the same aspirational cap. On the other hand, the caps are not exogenous to the choice of design; a range approach where the aspirational cap is significantly more aggressive than the cap under the future borrowing approach could create even higher price expectations.

Introducing Reserve Allowances to the Market

Given the structure of the allowance reserve approach, use of the reserve must involve, at a minimum, payment to the government of the ceiling price for any tapped reserve allowances. Otherwise there can be no assurance that the cap only expands when the ceiling price is reached as there is no other way to ensure that the market is really willing to pay that much. More generally, there are a variety of ways to increase the cap in response to high prices. Newell et al.

(2005) mention several approaches, including the announcement of an allowance split that makes each outstanding allowance worth more than one ton. However, this and other approaches that do not require payment to the government of the ceiling price have trouble maintaining the ceiling price without issuing too many or too few allowances. We therefore focus on two other options, which are somewhat equivalent.

The first approach, described briefly above, introduces reserve allowances into the market via a supplemental reserve auction prior to the end of the period when firms must balance their emissions and allowances (i.e., the true-up period). Here the government offers a fixed number of allowances (i.e., the reserve size) to the market via an auction with a minimum price equal to the ceiling price. If, at the time of the supplemental auction, the market expects the ordinary allowances to meet demand at a market price below the ceiling price, then presumably the allowances would remain in the reserve unsold. If, on the other hand, allowance demand is sufficient to push prices up to or above the ceiling price, then there should be some willingness to purchase reserve allowances at the ceiling price. If the reserve size is sufficient to meet demand at the ceiling price, then there should be enough allowances for both the allowance market price and the reserve auction price to equilibrate at the ceiling price.

However, if the demand for additional allowances at the ceiling price exceeds the reserve auction quantity, the auction process would lead to prices being bid up until the market clears at some price above the ceiling price. In this case, the allowance reserve does not guarantee a ceiling price in the same way as an explicit price ceiling or an unlimited safety valve. It puts only so much weight on addressing cost concerns, leaving some guaranteed maximum level of emissions intact. As noted earlier, the case when the ceiling price is exceeded should correspond to a situation when (discounted) long-term price expectations exceed the ceiling price, not when there is only a near-term disruption or shortage (unless the reserve size is too small).

Another approach, based on well-known financial instruments, is to have the government provide financial contracts (call options) that would give the holder the right, but not the obligation, to buy a certain quantity of allowances at the ceiling price (i.e., the strike price) during the true-up period each year.[4] In fact, as pointed out by Unold and Requate (2001), a

[4] If the execution date is not constrained in this way, it would create a very important difference: the effective annual reserve could accumulate over time if options accumulate, unexercised, year- after- year. If options can be executed well before the true-up period, reserve allowances could enter the system based on early expectations of high prices which, by the time the true up period arrives, have been revised.

series of such options—of different size and with different strike prices—could be used to replicate any known marginal damage function or desired allowance supply function. Such options could be auctioned (like ordinary allowances) or they could come attached to ordinary allowances on a pro rata basis.[5] The option's value, whether auctioned or given away, would be determined by the ceiling (strike) price and expectations of whether, when, and with what eventual market price it would be exercised. In the event that allowance prices exceeded the ceiling (strike) price level, options holders would begin to exercise their option rights, but would stop once prices fell back below the ceiling price level. One substantive difference with the reserve auction discussed above is the timing of the allocation of reserve allowances or options for reserve allowances, with the allocation of options most likely occurring sooner. Whether this would be an advantage or disadvantage requires further analysis. Another difference is that under the option approach, the difference between the market and ceiling price—if there is one—goes to the option holder, and options could be either auctioned or allocated for free. This feature suggests that options could be allocated in a way to help ensure that legislation passes, but can also create wasteful rent-seeking behavior.

An Allowance Reserve Board or Legislative Specification?

While some envision a reserve or other cost-containment program with key design parameters specified in legislation, an alternative is to delegate that responsibility to an independent executive board. Specific reserve design elements might be better managed by an independent executive board because, over time, there would be a clear need to update policy in response to new information and Congress may not respond in a timely manner. Indeed, part of the motivation for the reserve in the first place is a recognition that the drivers of long-term prices will evolve over time, that policymakers will be slow to adjust parameters, and that borrowing may not be a fully effective or implemented element, which is required for dynamic efficiency. While a suitably designed mechanism would, in principle, allow the market to operate for long periods of time without revision (driven by the expectation of an eventual revision), it is certainly possible that Congressional inaction might challenge that capacity. Therefore, governance by an independent board may be useful.

---

[5] See presentation by Jon A. Anda at the Carbon Market Insights Americas Conference. October 29–31, 2007, New York, NY. Available at: www.pointcarbon.com/events/recentevents/cmiamericas07.

As discussed in Newell et al. (2005), an important remaining issue would be the precise governing mandate for such a board, the tools available to it, and the degree to which it operated subject to legislated rules versus having complete discretion. There has been a tendency to draw an analogy with the U.S. Board of Governors of the Federal Reserve, along with parallels between its dual mandate of managing growth versus inflation and the dual objectives of climate protection versus containing costs. However, there are a variety of differences (Pizer and Tatsutani 2008), and this remains an active area of discussion.

Conclusions

While much of the debate in the literature on the economics of climate change regulation has focused on comparing pure price and pure quantity mechanisms—i.e., taxes versus cap-and-trade—these policies are increasingly being viewed as too extreme to meet both practical and political needs. This paper has presented recent and perhaps provocative new arguments suggesting that a sufficiently flexible cap-and-trade system can in theory do at least as well as and potentially better than a tax (despite previous literature pointing the other direction). However, it is unlikely that the required flexibility to borrow allowances from the future and the associated requirement for rational expectations in dynamic allowance markets would be ensured in practice. All of this recommends a hybrid mechanism. Roberts and Spence (1976) first suggested the idea of a cap-and-trade system with both a floor and ceiling price. We have taken their idea one step further and suggest that the ceiling price could come with a quantitative limit: what we call the allowance reserve.

We have argued that the allowance reserve addresses certain shortcomings of the Roberts and Spence idea, including the need for more flexibility in the elements of policy design to balance competing political interests. It also solves a possibly thorny technical problem that arises when the Roberts and Spence idea is applied in a dynamic world that includes banking and a need to update policies—which has the potential to lead to a run on the price ceiling. But perhaps most fundamentally, it supports the idea of a flexible cap-and-trade system that seeks to achieve an intertemporal optimum.

A number of additional details remain to be resolved, most notably setting the ceiling price and reserve size and how an allowance reserve would be institutionalized. Other issues tend to be primarily cosmetic or of a more general nature applicable to any market-based policy. In summary, the allowance reserve may help solve several previously insurmountable challenges in the current debate over climate policy design. This paper demonstrates that the notion of cap-and-trade with an allowance reserve is more than simply a political solution. Rather given the

considerable uncertainties we face now over the costs and benefits of greenhouse gas mitigation, the institutional difficulties faced by firm-level borrowing mechanisms, and the need for market-based institutions that will react to the unfolding of new information over time, a cap-and-trade system with an allowance reserve is well supported by an economic view of efficient long-term climate policy.

References

Argus Media. 2008. Markets Slowed by CAIR Concerns. *Argus Air Daily*, May 30, 15 (104).

Eilperin, Juliet, and Steven Mufson. 2008. Climate Bill Underlines Obstacles to Capping Greenhouse Gases. *Washington Post*, June 1, A12.

Energy Information Administration (EIA). 2008. Energy Market and Economic Impacts of S. 2191, the Lieberman–Warner Climate Security Act of 2007. April. Report SR/OIAF/2008-01. Washington, DC: EIA.

Hoel, Michael, and Larry Karp. 2002. Taxes versus Quotas for a Stock Pollutant. *Resource and Energy Economics* 24: 367–384.

Jacoby, H.J., and A.D. Ellerman. 2003. The Safety Valve and Climate Policy. *Energy Policy* 32(4): 481–491.

Kerr, Suzi and Richard G. Newell. 2003. Policy-Induced Technology Adoption: Evidence from the U.S, Lead Phasedown. *Journal of Industrial Economics* 51(3):271-343.

Kling, C., and J. Rubin. 1997. Bankable Permits for the Control of Environmental Pollution. *Journal of Public Economics* 64(1): 101–15.

Kolstad, C.D. 1996. Learning and Stock Effects in Environmental Regulation: The Case of Greenhouse Gas Emissions. *Journal of Environmental Economics and Management* 31(1): 1–18.

Kopp, Raymond, Richard Morgenstern, and William Pizer. 1997. Something for Everyone: A Climate Policy that Both Environmentalists and Industry Can Live With.*Weathervane.* Washington, DC: Resources for the Future. http://www.weathervane.rff.org/SomethingForEveryone.pdf.

Murray, Brian, and Martin Ross. 2007. The Lieberman–Warner America's Climate Security Act: A Preliminary Assessment of Potential Economic Impacts. Policy Brief NI PB 0704. Durham, NC: Nicholas Institute for Environmental Policy Solutions, Duke University.

Newell, Richard G., and William A. Pizer. 2003. Regulating Stock Externalities under Uncertainty. *Journal of Environmental Economics and Management* 45: 416–432.

Newell, Richard G., William A. Pizer, and Jiangfeng Zhang. 2005. Managing Permit Markets to Stabilize Prices. *Environmental and Resource Economics* 31: 133–157.

Paltsev, Sergey, John M. Reilly, Henry D. Jacoby, Angelo C. Gurgel, Gilbert E. Metcalf, Andrei P. Sokolov, and Jennifer F. Holak. 2007. Assessment of U.S. Cap-and-Trade Proposals. Report 146. Cambridge, MA: MIT Joint Program on the Science and Policy of Global Change.

Pizer, William A. 2002. Combining Price and Quantity Controls to Mitigate Global Climate Change. *Journal of Public Economics* 85(3): 409–434.

Pizer, William A. and Marika Tatsutani. 2008. Managing Costs in a U.S. Greenhouse Gas Trading Program: A Workshop Summary. Discussion paper 08-23. Washington, DC: Resources for the Future.

Roberts, M.J., and M. Spence. 1976. Effluent Charges and Licenses under Uncertainty. *Journal of Public Economics* 5(3–4): 193–208.

Rubin, J.D. 1996. A Model of Intertemporal Emission Trading, Banking, and Borrowing. *Journal of Environmental Economics and Management* 31: 269–286.

Samuelsohn, Darren. 2008. Behind "Safety Valve" Debate Resides 30+ Years of History. *ClimateWire*, March 11.

U.S. Environmental Protection Agency. 2005. Rule to Reduce Interstate Transport of Fine Particulate Matter and Ozone (Clean Air Interstate Rule); Revisions to Acid Rain Program; Revisions to the NOX SIP Call. *Federal Register* 70(91): 25162–25405.

U.S. Environmental Protection Agency. 2008. EPA Analysis of the Lieberman–Warner Climate Security Act of 2008: S. 2191 in 110th Congress. http://www.epa.gov/climatechange/downloads/s2191_EPA_Analysis.pdf.

Unold, Wolfram, and Till Requate. 2001. Pollution Control by Options Trading. *Economics Letters* 73: 353–358.

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

Weitzman, Martin L. 2008. On Modeling and Interpreting the Economics of Catastrophic Climate Change. Forthcoming *Review of Economics and Statistics*.

Technical Appendix

Consider an effort to seek: $\min C_1(q_1) + C_2(Q - q_1)$, where q_t is the emissions each period t, Q is the ultimate cumulative emissions goal (unknown until period 2), C_t is the cost of emissions level q_t (which, unknown until period t, is positive if q_t is below some baseline level and zero otherwise), and the cost functions include adjustment for discounting to the present.

Given the ultimate resolution of uncertainty in period 2, we know costs in period 2 are $C_2(Q - q_1)$ and it would therefore be optimal to choose q_1 such that $C_1'(q_1) = C_2'(Q - q_1)$. With the information available when q_1 has to be chosen, the best practical outcome would be $C_1'(q_1) = E_1\left[C_2'(Q - q_1)\right]$, where E_1 reflects the expectation formed in period 1 about the costs and target set in period 2.

An optimizing government setting a period 1 tax, t_1, in period 0 would choose $t_1 = E_0\left[C_1'(\hat{q}_1)\right]$ where $\hat{q}_1$ satisfies $E_0\left[C_1'(\hat{q}_1)\right] = E_0\left[C_2'(Q - \hat{q}_1)\right]$, thereby minimizing expected costs as seen in period 0. Firms would then choose to emit q_1 such that $t_1 = C_1'(q_1)$, given the resolution of cost uncertainty in period 1. This outcome for q_1 would not generally satisfy the efficiency condition, $C_1'(q_1) = E_1\left[C_2'(Q - q_1)\right]$because $E_1[C_2'(Q - q_1)]$ will not generally equal $E_0[C_2'(Q - q_1)] = t_1 = C_1'(q_1)$ under the tax.

Now imagine the government instead sets a cap $\overline{q}_1$ in period 0 and a second period cap $\overline{q}_2$ to deliver the ultimate objective $Q = \overline{q}_1 + \overline{q}_2$. Note that with second period cap $\overline{q}_2$, if firms have banked $\overline{q}_1 - q_1$ at the end of period 1, in period 2 emissions would be $Q - q_1$ and marginal costs would be $C_2'(Q - q_1)$. In this setting, a cost-minimizing firm would choose q_1 such that $C_1'(q_1) = E_1\left[C_2'(Q - q_1)\right]$ regardless of the first period cap.

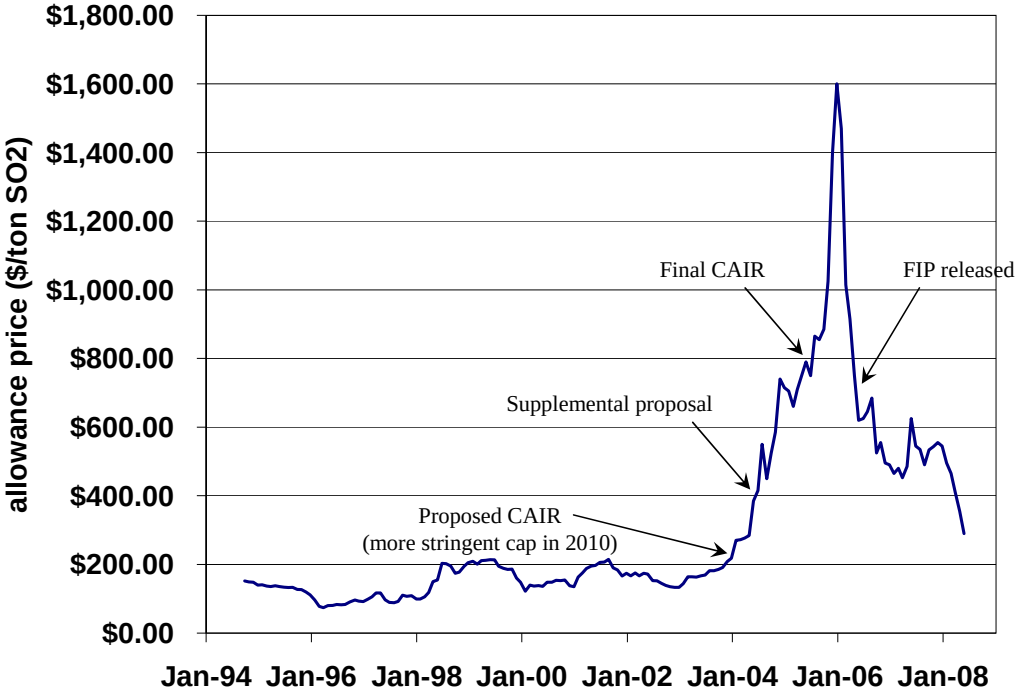
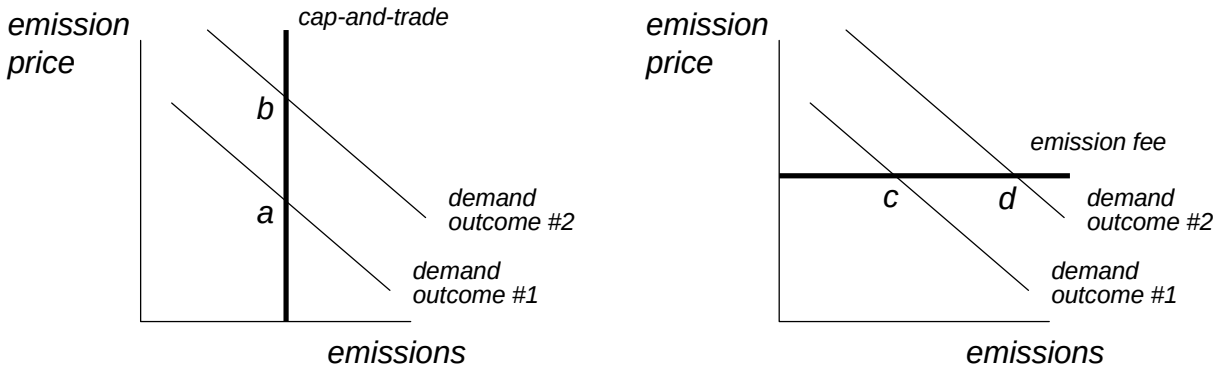


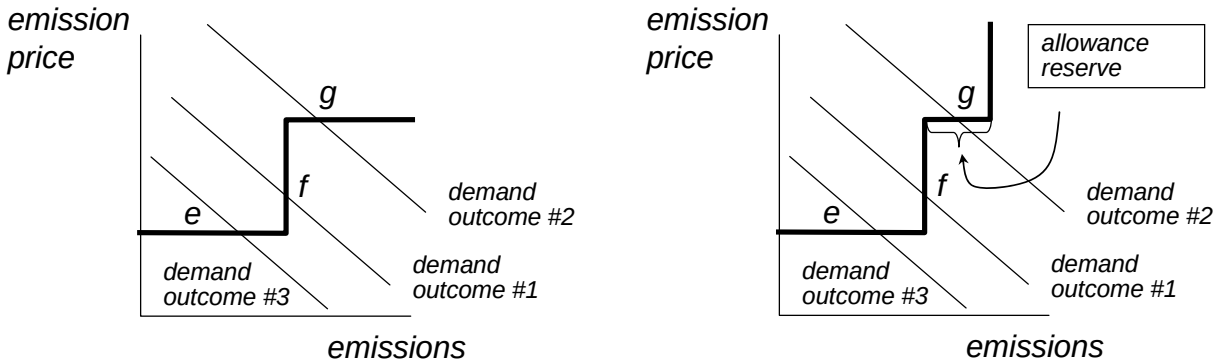
Figure 1: Historic Prices in the Sulfur Dioxide Allowance Market

Note: CAIR: Clean Air Interstate Rule; FIP: Federal Implementation Plan.



Allowance supply under cap-and-trade *Allowance supply under emissions fee*

Appendix Figure 1. Emissions and Price Outcomes under Cap-and-Trade versus Emissions Fee/Tax



Allowance supply based on Roberts and Spence (1976) *Allowance supply with allowance reserve*

Appendix Figure 2. Emissions and Price Outcomes under Hybrid Price–Quantity Policies