Cap-and-Trade Programs under Delayed Compliance: Consequences of Interim Injections of Permits

Makoto Hasegawa*
Department of Economics, University of Michigan

Stephen Salant†
Department of Economics, University of Michigan, and Resources for the Future

December 18, 2012

Abstract

Previous analyses of cap-and-trade programs regulating carbon emissions assumed that firms must surrender permits as they pollute. If so, then the price of permits may remain constant over measurable intervals if the government injects additional permits at a ceiling price or may even collapse if more permits are injected through an auction. However, no cap-and-trade program actually requires continual compliance. The three federal bills and California's AB-32, for example, instead require that firms surrender permits only periodically to cover their cumulative emissions since the last compliance period. Anticipated injections of additional permits during the compliance period should have different effects than under continual compliance. We develop a methodology for analyzing the effects of such permit injections. Using it, we explain why sales provisions of one federal bill might generate a speculative attack in the permit market and while provisions of AB-32 may undermine the very existence of an equilibrium.

Keywords: emissions trading, marketable permits, price collar, safety valve, price ceiling, price floor

---

*E-mail: makotoh@umich.edu.

†Corresponding author. E-mail: ssalant@umich.edu. Address: Department of Economics, University of Michigan, 611 Tappan St., Ann Arbor, MI 48109-1220. Phone: (734) 223-7439. Fax: (734) 764-2769.

1 Introduction

Cap-and-trade programs are being utilized as the main vehicle to combat global warming by national and state governments of advanced countries. Such regulations are sometimes exceedingly complex. Nonetheless they share some common features. First, although firms subject to the regulations are required to surrender permits to cover their carbon emissions, they are not required to surrender permits on a continuing basis ("continual compliance") but only periodically. As a result, a firm may emit carbon without possessing the permits to cover its emissions as long as it acquires sufficient permits by the compliance date. We refer to this aspect of the regulations as "delayed compliance." In the case of the California law (AB-32), for example, the compliance period is initially two years and subsequently three years (although a fraction of the permits must be surrendered earlier as a down-payment). In the case of the three federal bills that failed to become law, the compliance period was one year.[1] Second, while some permits are issued at the outset of a compliance period, provision is made in most recent proposals for the government to inject *additional* permits into the market later in the compliance period.[2] Third, while permits may be stored ("banked") over time for later use, these programs prohibit or severely restrict the opportunity to borrow from future allocations. For example, California AB-32 allows unlimited banking and prohibits borrowing from the future compliance periods.

These common features have consequences that have escaped notice. Virtually all previous analyses have assumed that firms must be in continual compliance.[3] Under this assumption, a sizable literature has developed to assess the welfare benefits of holding back some of the permits that could have been allocated at the outset and using them subsequently to hold down the price through auctions or sales at fixed prices. Such policies are classified as "price collars" or "safety valves."[4] Burtraw et al. (2010) find that a price collar (also called

---

[1] Waxman-Markey's "American Clean Energy and Security Act of 2009,", Kerry-Boxer's "Clean Energy Jobs and American Power Act of 2009," and Kerry-Lieberman's "American Power Act of 2010."

[2] We do not discuss the European trading program since it has no interim injections during its annual compliance period. However, such injections have been proposed. Stavins (2012) regards the absence of a safety valve or price collar in the European system as a "design flaw" and Hone (2012) notes that "One approach is...to [require] a sufficient proportion of allowances to be auctioned, instead of being allocated free of charge and auctions to be held periodically throughout the commitment period. It is too late to do this for Phase III of the EU ETS (2013-2020) but it could be introduced as part of the expected legislative process to set the parameters for Phase IV (2021 and beyond, probably extending to 2030)."

[3] The single exception is the contemporaneous working paper of Holland and Moore (2012), which complements our work. They consider a wide variety of cap-and-trade programs, including the carbon-trading programs of primary concern to us. After classifying these programs in terms of their compliance timing, banking, and borrowing provisions, Holland and Moore provide a sufficient condition for delayed and continual compliance to yield the *same* equilibrium price path. Although their sufficient condition holds in programs to limit SO_x and NO_x (e.g. RECLAIM and the Acid Rain Program), their condition is violated in all four of the carbon-trading programs we analyze as they acknowledge when discussing our work.

[4] For a valuable explanation of the origins of the safety valve concept and its evolution in the climate

a "symmetric safety valve" in the paper) outperforms a safety valve in a static setting. Fell and Morgenstern (2010) and Fell et al. (2012) simulate a dynamic stochastic model of a cap-and-trade program with a price collar or a safety valve.[5] Fell and Morgenstern (2010) find that price collar mechanisms are more cost-effective than both purely quantity-based mechanisms and safety valve mechanisms for a given level of expected cumulative emissions. They also find that the combination of a price collar with banking and borrowing systems can achieve expected cost as low as a tax with lower emissions variance. Fell et al. (2012) find that *hard* collars, which ensure unlimited supply of reserve allowances to defend a ceiling price yield lower net present value of expected abatement costs than *soft* collars, price collars with limited supply of reserve allowances, for the same level of the expected cumulative emissions net of offsets. Most recently, Hasegawa and Salant (2012) have shown that if firms must cover their emissions on a continuing basis, then in the competitive equilibrium, the price path of permits may remain constant over periods while the government is selling additional permits at a ceiling price or may even collapse in response to a government auction. Clearly, no rational private agent would hold permits in the face of such capital losses. But the government sales of additional permits would enable firms to acquire the necessary permits to remain continually in compliance.[6]

Despite this sizable literature analyzing the effects of permit auctions and sales under a regime of continual compliance, such policies remain to be investigated under the actual regime of delayed compliance. With delayed compliance, firms purchase the permits they will ultimately need only at those instants within the compliance period when the permit price has the lowest capitalized value at the compliance date.

Much of the literature assumes discrete time and defines the period length in a way that obscures the distinction between delayed compliance and "contemporaneous" (the discrete-time analog of "continual") compliance. To understand this distinction, consider a discrete-time model where one period represents one day. If the government will inject permits on some of the days within the next year, the policy of requiring that permits be surrendered every day to match that day's emissions ("contemporaneous compliance") *differs* from the policy of requiring that permits be surrendered once every 365 days to cover cumulative

---

context, see Jacoby and Ellerman (2004).

[5] In a dynamic context, intertemporal trading of emissions permits matters for economic efficiency. Cronshaw and Kruse (1996) and Rubin (1996) show that emissions trading allowing banking and borrowing of emission permits achieves the least-cost outcome.

[6] If unlimited "borrowing" were permitted, such price drops would not occur since permits could be borrowed from a future low-price period and sold earlier at a higher price. We assume here that a firm cannot borrow permits it expects to acquire in the future to cover current emissions under continual compliance. A distinction between delayed and continual compliance would remain even if borrowing was allowed provided it was constrained enough that the constraint on it was binding.

emissions during the entire year. If, as in most of the discrete-time literature, the length of each period is *defined* to be the same as the length of the compliance period, nothing by definition can happen between periods and, by this modeling choice, one prevents oneself from investigating the consequences of a government injection of permits *between* one compliance period and the next. To consider the effects of government policies conducted *within* a compliance period, we adopt a continuous-time formulation as less cumbersome than its discrete-time counterpart.

In contrast to the case of continual compliance, under delayed compliance prices can never rise *slower* than the rate of interest. For suppose the contrary. Then the highest capitalized price would *strictly exceed* the lowest capitalized price. But then everyone with an initial allocation of permits would want to sell them at the highest capitalized price and there would be no one on the other side of the market willing to buy permits at that price; as a result, there would be massive excess supply.[7] Under both compliance regimes, prices can never rise faster than the rate of interest in the absence of uncertainty; otherwise traders would attempt to buy low and sell high on an infinite scale. It follows that in any equilibrium under delayed compliance, prices must rise throughout the compliance period at the rate of interest. Anticipated government auctions or sales from a finite reserve at a fixed price will not slow this rate of price appreciation although, as we will show, they will determine the position of the price path or, equivalently, the permit price prevailing at the compliance date.

The equilibrium permit price at the compliance date equates the demand for permits required to cover the cumulative emissions which have occurred since the last compliance date with the cumulative supply of permits provided by the government over that period.[8] The following algorithm can then be used to determine the permit price at the compliance date.

For each possible terminal price, determine the (unique) associated price path over the compliance period. To determine the cumulative demand for permits along that path, note that at every instant firms will abate up to the point where their marginal cost of abatement capitalized to the compliance date equals the permit price anticipated to prevail at that date. Compute the aggregate cumulative emissions of the regulated entities over time. Firms will need a matching number of permits at the compliance date. This procedure provides one

---

[7] This argument implicitly assumes that private agents are the only purchasers of permits. We assume that the government never purchases permits since none of the delayed compliance programs we consider envision that. If the government did purchase permits, the equilibrium price path under delayed compliance could rise slower than the rate of interest.

[8] As an analytical simplification, we assume that it is illegal or unprofitable to carry permits from one compliance period to the next. The following algorithm can then be used to determine the permit price at the compliance date. However, the algorithm in the text is easily modified if carryovers between compliance periods is permitted.

price-quantity pair on the cumulative demand curve for permits. Repeat the procedure to generate the other points on the demand curve.[9]

Deriving the cumulative supply of permits as a function of the terminal price is somewhat trickier. Since all prices on each associated price path will have the same capitalized value, private agents will not care when they sell as long as they hold zero permits after the compliance period ends. Hence *cumulative* supply of permits over the period will consist of the initial allocations *plus* the subsequent injections of additional permits. These injections depend on the fine details of particular regulations as we will illustrate using provisions from California's cap-and-trade program AB-32 which begins later this year and from the three Congressional bills which died in Congress. All four programs envision an initial allocation of permits supplemented by subsequent injections of additional permits during the compliance period. The programs differ, however, in the rules governing these injections. For example, while all these programs prescribe a periodic sequence of auctions with pre-announced reserve prices, the programs differ in whether permits unsold in one auction can be re-offered in subsequent auctions. As we show, California AB-32 has a troublesome provision for offering unsold permits in subsequent auctions that induces a jump in the supply of permits; under this rule, there may be no price path that will clear the permit market.

In addition to auctions, the programs envision sales at pre-determined prices; but here too the terms of these sales differ. The California plan contemplates sales of specified amounts at specified prices from an "Allowance Price Containment Reserve" shortly after each quarterly auction whereas the Kerry-Lieberman (2010) bill proposed sales of permits at a fixed price over a designated time interval or until the "Cost Containment Reserve" was depleted.

In the next section, we discuss the determination of cumulative demand for permits as a function of the last price on a price path rising at the rate of interest. We then discuss the cumulative supply of permits as a function of the last price on that path. We show how the supply curve depends on the particular provisions of the emissions trading program. The last price on the equilibrium price path is determined by the intersection of the cumulative demand and supply curves. We will also note when the equilibrium price path under continual compliance differs from the path under delayed compliance. Such differences occur when

---

[9] To simplify the exposition, we assume that firms do not abate by investing in new or altered technology. Otherwise a firm's current abatement decision would affect its future cost of emissions, and each firm would have a dynamic investment problem to solve. Accounting for this would complicate the derivation of cumulative emissions if the permit price rises at the rate of interest and hence the aggregate cumulative demand for permits, but it would not alter any of our points. The equilibrium price path under delayed compliance would still be determined at a terminal price that equates cumulative supply to the altered cumulative demand, and this path would fail to equilibrate markets under continual compliance for the reasons we discuss. We have chosen, therefore, to abstract from this real world complication. As noted in the concluding section, one cannot abstract from this complication when conducting welfare analysis.

firms would not receive injections of permits soon enough to surrender them under continual compliance. In such cases, excess demand occurs and permit prices must initially be higher (and emissions per unit time initially lower) under continual compliance.

2 Preliminaries

Under delayed compliance, firms will purchase permits at the lowest price, capitalized to the date of compliance. Since, as explained previously, the equilibrium price path under delayed compliance must rise at the rate of interest, every price is lowest and we may index such paths by the price expected to prevail at the date of compliance. Denote that price as $\mathbf{P}$.

We assume that firm i ($i \in \{1, \ldots, n\}$) can reduce its emissions per unit time to $e_i(t)$ at time t by abating at cost $c_i(e_i(t))$, where firm i's cost is a strictly decreasing, strictly convex, differentiable function of emissions. To avoid corners, we assume the Inada condition holds: $-c_i'(e) \to \infty$ as $e \to 0$. Moreover, at a sufficiently high level of emissions ("baseline emissions," $\bar{e}_i$), the firm's cost declines to zero and approaches that level at a zero slope: $c_i(\bar{e}_i) = c_i'(\bar{e}_i) = 0$. Then firm i chooses its emissions path $e_i(t)$ to minimize its total cost of complying with the cap-and-trade regulation. It minimizes $c_i(e_i(t))e^{r(T-t)} + e_i(t)\mathbf{P}$, where T is the compliance date. Its optimal emissions path therefore solves:

$$-c_i'(e_i(t)) = \mathbf{P}e^{r(t-T)}, \text{ for } t \in [0, T] \text{ and } i = 1, \ldots n. \tag{1}$$

Given the properties of the n cost functions, the emissions of each firm at any instant are a continuous, strictly decreasing function of $\mathbf{P}$. From the emissions paths of the firms, we can determine the cumulative aggregate demand for permits through time τ as a function of $\mathbf{P}$:

$$D(\mathbf{P}, \tau) = \int_{t=0}^{\tau} \sum_{i}^{n} e_i(t)dt \text{ for } \tau \in [0, T]. \tag{2}$$

$D(\mathbf{P}, \tau)$ is continuous, strictly decreasing in its first argument and strictly increasing and strictly concave in its second argument. The intercepts are $D(\mathbf{P}, 0) = 0$ and $D(0, \tau) = \tau \sum_{i=1}^{n} \bar{e}_i$. We will make extensive use of this function in the subsequent analysis.

For any particular government method of injecting permits, we can define $S(\mathbf{P}, \tau)$ as the government's cumulative supply of permits until time τ on a price path rising at the rate of interest and ending at $\mathbf{P}$. Under delayed compliance, any price path such that $D(\mathbf{P}, T) = S(\mathbf{P}, T)$ equilibrates the market.

To determine when the equilibrium price path under delayed compliances generates a disequilibrium under continual compliance, we will have to compute the cumulative supply and demand for permits at any time τ under continual compliance when the price rises throughout at the rate of interest, reaching $\mathbf{P}$ at T. A given method of injecting permits will generate the same cumulative supply $S(\mathbf{P}, \tau)$ under the two regimes. Moreover, under continual compliance firm i's demand for permits at τ is also given by equations (1) and (2). Provided the price path rises at the rate of interest, each firm's cumulative demand from the outset to time T will be the same under the two regimes.[10]

However, equilibrium under continual compliance requires that $D(\mathbf{P}, \tau) \leq S(\mathbf{P}, \tau)$ for *all* $\tau \in [0, T)$ in addition to $D(\mathbf{P}, T) = S(\mathbf{P}, T)$. That is, in the continual compliance regime, agents must be provided enough permits to be able to cover their emissions at *every* instant and not merely the last one. Since the requirement of equilibrium is more restrictive under continual compliance, price paths that equilibrate the market under delayed compliance but where $D(\mathbf{P}, \tau) > S(\mathbf{P}, \tau)$ for some $\tau \in [0, T)$, fail to equilibrate it under continual compliance.

3 Auctions with Reserve Prices

Throughout we will assume that g permits are "grandfathered" at the outset and that the number grandfathered is smaller than the cumulative emissions that would have occurred without a cap-and-trade program ($g < T \sum_{i=1}^{n} \bar{e}_i$).

In this section, we assume that the government commits at the outset to conduct a sequence of auctions. The date, amount, and reserve price of each auction is announced at the outset. Let t_i denote the date of the i^{th} auction, a_i its amount, and $\underline{p}_i$ its reserve price (assumed strictly positive) for $i = 1, \ldots A$, where A is the total number of auctions to be held during the compliance period, $[0, T]$.

To determine the equilibrium price path under delayed compliance, we construct the cumulative demand and cumulative supply curves and determine their unique point of intersection. The cumulative demand curve is simply $D(\mathbf{P}, T)$, which is downward-sloping with respect to $\mathbf{P}$. The cumulative supply curve $S(\mathbf{P}, T)$ is a step-function. For the price path with the terminal price of zero, aggregate supply consists of the g grandfathered permits. As the terminal price is increased, it eventually equals lowest capitalized reserve price. At that terminal price, the cumulative supply is indeterminate—as small as g and as large as

---

[10] The cumulative demands no longer coincide on price paths that rise somewhere more slowly than the rate of interest. Suppose, for example, that the price is constant at $\mathbf{P}$ until T. Then at every instant under continual compliance emissions would solve $-c_i'(e_i(t)) = \mathbf{P}$, for $t \in [0, T]$ and $i = 1, \ldots n$ which is strictly smaller than the solution to (1); hence, cumulative demand until T would be strictly smaller on such a price path under continual compliance. However, this observation is unimportant since the equilibrium price path under delayed compliance must always rise at the rate of interest and we will be checking whether such a price path equilibrates the permit market under continual compliance.

g plus the amount offered at the auction with the lowest capitalized reserve price. If the terminal price is slightly higher, the cumulative supply equals the upper end of this interval. Cumulative supply would remain at that level until the terminal price reached the next-to-the-lowest capitalized reserve price. A sufficiently high terminal price will equal the highest of the capitalized reserve prices of the A auctions. Any higher terminal price will elicit the maximal supply of $g + \sum_{i=1}^{A} a_i$ permits.

There exists a unique equilibrium price path and terminal price, $\mathbf{P}$. Existence follows since a zero terminal price would generate excess cumulative demand (by assumption, $T \sum_{i=1}^{n} \bar{e}_i > g$) while a sufficiently high terminal price would generate excess cumulative supply (cumulative supply $g + \sum_{i=1}^{A} a_i$ is bounded away from zero and cumulative demand approaches zero for sufficiently high $\mathbf{P}$). Moreover the intersection point must be unique since, at any higher price, demand is strictly smaller and supply weakly larger while, at any lower price, demand is strictly larger and supply weakly smaller.

To construct the supply curve geometrically, proceed as follows: (1) on a diagram with time on the horizontal axis and price per permit on the vertical axis (see Figure 1), record the date and reserve-price pair $(t_i, \underline{\mathrm{p}}_i)$ of each of the A auctions; (2) determine the capitalized value ($\mathbf{P_i}$) of each reserve price by drawing through each of these A points a price path rising at the rate of interest and noting its height at T ($\mathbf{P_i} = \underline{\mathrm{p}}_i e^{r(T-t_i)}$). For terminal prices smaller than the smallest capitalized reserve price, only the g grandfathered permits are supplied to the market. For higher prices, the cumulative supply function $S(\mathbf{P}, T)$ will have a horizontal step of length a_i at height $\mathbf{P_i}$ for $i = 1, \ldots A$.

Our methodology can be used to predict the consequences of *any* exogenous path of auction reserve prices. For example, suppose the auction reserve price rises exactly at the rate of interest. Then, if bids at one auction strictly exceed its reserve price, bids at the other auctions will strictly exceed their reserve prices. Conversely, if no bids meet the reserve price in one auction, none will meet it in any other auction.

If instead the exogenous auction reserve price rises faster than the rate of interest, then if bids fail to meet the reserve price in one auction, no bids will be acceptable in subsequent auctions while if bids *do* meet the reserve price in one auction, bids in prior auctions will also be acceptable. Consequently, when the exogenous reserve price rises faster than the rate of interest, auctions where bids fail to meet the reserve price cluster at the end of the compliance period.

If instead the exogenous auction reserve prices rise slower than the rate of interest, then if bids fail to meet the reserve price in one auction, no bids will be acceptable in earlier auctions while if bids do meet the reserve price in one auction, bids in subsequent auctions will also be acceptable. Consequently, when the exogenous reserve price rises slower than the rate of interest, auctions where bids fail to meet the reserve price cluster at the beginning of the compliance period.

Although our methodology can be applied to any exogenous path of reserve prices, we illustrate it below in the simplest manner. Hence, we assume that all the reserve prices are the same. In Figure 1, all auctions have the same reserve price ($\underline{\mathrm{p}}_i = \underline{\mathrm{p}}_j$). Since these reserve prices rise by less than the rate of interest, $\mathbf{P_1} > \mathbf{P_2} > \mathbf{P_3} > \mathbf{P_4}$. In the example portrayed, the equilibrium terminal price $\mathbf{P}^*$ is contained in the open interval $(\mathbf{P_4}, \mathbf{P_3})$. Hence, no bids are accepted at the first three auctions but all of a_4 permits are sold at the fourth auction at the price $\mathbf{P}^* e^{r(t_4 - T)} > \underline{\mathrm{p}}$. Therefore, in equilibrium emissions equal $g + a_4$.

[Figure 1 Goes Here]

If the government had auctioned no permits at t_4 but had instead added these a_4 permits to the number grandfathered, then the cumulative supply curve would become $g + a_4$ for terminal prices below $\mathbf{P_4}$ but the modified cumulative supply curve would still intersect the unchanged cumulative demand curve at the same point. Hence, the equilibrium price path would not change under delayed compliance nor would the cumulative emissions it induces. Alternatively, if the government had grandfathered no permits but had instead auctioned these g permits along with the a_4 permits at t_4, then the cumulative supply at prices below $\mathbf{P_4}$ would be zero and the cumulative supply at $\mathbf{P_4}$ would be as large as $g + a_4$. Nonetheless this modified cumulative supply curve would still intersect the cumulative demand curve at the same point. Neither change would affect the equilibrium under delayed compliance.

If the entire sum of permits $(g + a_4)$ was grandfathered at the outset, then this same price path would also equilibrate the market under *continual* compliance. But if all of these permits were made available instead at t_4, then a disequilibrium would inevitably occur since, for some $\tau < T, D(\mathbf{P}, \tau) > S(\mathbf{P}, \tau)$. In Figure 2, we plot cumulative supply and demand until τ along the equilibrium price path under delayed compliance (the price path ending at $\mathbf{P}^*$). Since the path generates an equilibrium under delayed compliance, the two curves intersect at T and $D(\mathbf{P}^*, T) = g + a_4$. An equilibrium under continual compliance requires in addition that the cumulative supply curve lies nowhere strictly below the cumulative demand curve for $\tau \in [0, T)$. Let $\tilde{g}$ be the number of permits the government chooses to grandfather initially. We have drawn the boundary case where $\tilde{g} = D(\mathbf{P}^*, t_4)$ permits are grandfathered and $(g + a_4) - \tilde{g}$ are auctioned at t_4. If the government grandfathered strictly less than $D(\mathbf{P}^*, t_4)$ permits and added them instead to the amount auctioned at t_4, then the equilibrium under delayed compliance could no longer be supported as an equilibrium under continual compliance. Instead, the price path would consist of segments rising at the rate of interest, separated by a downward jump at t_4.[11]

[Figure 2 Goes Here]

3.1 Non-existence of Equilibrium Induced by the Rules of California's AB-32

Returning to the case of delayed compliance, suppose that cumulative demand was so large that $\mathbf{P}^* \in (\mathbf{P_2}, \mathbf{P_1})$. That is, every auction after t_1 sells out, but bids in this first auction are below its reserve price ($\underline{\mathrm{p}}_1$). Under the rules of California's AB-32, the a_1 permits which failed to sell in the first auction would be returned to the "Auction Holding Account."[12] Some of these permits would be available for sale in the fourth auction since it would have occurred "after two consecutive auctions have resulted in an auction settlement price greater than the applicable Auction Reserve Price."[13] However, not all of the a_1 permits could be made available. At most the number of permits which can be added to the auction at t_4 is 25% of a_4.[14] It is not clear to us how many of these permits would be offered and who decides, but at the old equilibrium price excess supply would occur because these unsold additional permits would be offered in an auction where the market price exceeds the reserve price. As a result, the equilibrium price path under delayed compliance would rise to a lower terminal price.

Offering unsold permits for sale if and only if permits are sold at two preceding auctions in a row can create a situation where no competitive equilibrium exists. Suppose, for example, that every bid is strictly below the reserve price in the first auction but the next two auctions sell out. Suppose cumulative demand is sufficiently high that in the absence of the rule regarding unsold permits that $\mathbf{P}^* \in (\mathbf{P_2}, \mathbf{P_1})$. Under this rule $\min(a_1, 0.25a_4)$ of the permits from the first auction can be offered in the fourth auction. If $\min(a_1, 0.25a_4) \geq D(\mathbf{P_2}) - g - a_2 - a_3 - a_4 = D(\mathbf{P_2}) - D(\mathbf{P}^*)$ then there will be excess supply at any terminal price

---

[11] We note that, although we have assumed throughout that g permits are grandfathered at the outset, distributing some of these g permits later would not affect the equilibrium price path under delayed compliance. No matter when the g permits are distributed the cumulative supply curve $S(\mathbf{P}, T)$ will remain unchanged. On the other hand, under continual compliance, distributing some of the g permits at a subsequent date may induce a higher price initially and a drop of the permit price when the subsequent permit injection occurs (Hasegawa and Salant, 2012).

[12] Final Regulation Order, §95911. Format for Auction of California GHG Allowances. (b) (4) (A)

[13] Quoted from Final Regulation Order, §95911. Format for Auction of California GHG Allowances. (b) (4) (B)

[14] Final Regulation Order, §95911. Format for Auction of California GHG Allowances. (b) (4) (C)

strictly exceeding $\mathbf{P_2}$.[15] But at any terminal price equal to or strictly below $\mathbf{P_2}$, there will be excess demand since, in the absence of two consecutive auctions where permits are sold at prices strictly higher than the reserve price, *none* of the unsold permits from auction 1 can be offered for sale in the fourth auction, and then $D(\mathbf{P}) > g + a_2 + a_3 + a_4$ holds for all $\mathbf{P} \leq \mathbf{P_2}$. We illustrate a situation with no equilibrium in Figure 3.[16]

[Figure 3 Goes Here]

4 Sales at Specified Prices

Permits can also be injected during the compliance period by sales at a specified price, which we denote $\bar{p}$. To simplify, we assume in this section that all permits not grandfathered at the outset are injected by such sales. Such sales can occur over a specified time interval which commences at t_c and finishes at t_f or until all of the R permits in the "Cost Containment Reserve" have been sold. The Kerry-Lieberman bill envisioned such sales over a finite interval. They resemble a continuum of auctions with reserve price $\underline{\text{p}}$ over the time interval $[t_c, t_f]$ but with R available in the initial auction, and *everything* unsold in one auction immediately available for sale in subsequent auctions.

Since the sales price over the interval is constant, the price at t_f has the smallest capitalized value ($\mathbf{P_f} = \bar{p}e^{r(T-t_f)}$). In Figure 4, we depict the interval of offers and the sales price. As in Figure 1, we depict $\mathbf{P_f}$ by noting the terminal price on the path through the point $(t_f, \bar{p})$ rising at the rate of interest. To derive the cumulative supply curve, note that if the terminal price is strictly smaller than $\mathbf{P_f}$, then nothing would sell during this time interval and the cumulative supply would just be g. If the terminal price is strictly larger, then the cumulative supply would be $g + R$. If the terminal price is exactly $\mathbf{P_f}$ then the cumulative supply is any number of permits in the closed interval $[g, g + R]$.

[Figure 4 Goes Here]

Suppose cumulative demand is sufficiently large that under delayed compliance the terminal price strictly exceeds $\mathbf{P_f}$. Then R permits sell instantaneously, either at some interior

---

[15] We use the excess supply condition at any terminal price higher than $\mathbf{P_2}$: $D(\mathbf{P_2}) \leq g + \min(a_1, 0.25a_4) + a_2 + a_3 + a_4$ and the definition of $\mathbf{P}^*$: $D(\mathbf{P}^*) = g + a_2 + a_3 + a_4$.

[16] Our analysis clarifies that existing regulations should be changed. Indeed, it suggests a remedy: re-word the regulation so permits may be sold as long as the settlement price in each of the two preceding auctions equals or exceeds the auction's reserve price.

date $\tau^* \in (t_c, t_f)$ or at the first moment of the sale (t_c), where τ^* is determined so that $\mathbf{P}^* e^{-r(T-\tau^*)} = \bar{p}$ for $\mathbf{P_f} < \mathbf{P}^* < \mathbf{P_c}$. In either case, such purchases at an infinite rate are just like "first-generation" speculative attacks which have been widely discussed in the literatures on foreign exchange markets and on commodity agreements.[17]

In commodity markets or foreign exchange markets, defending the ceiling price typically requires the government to sell at a slow rate over a finite interval prior to the attack. Then, at some endogenously determined date, further defense of the ceiling requires the instantaneous sale of the remaining stock to buyers who store it for later re-sale to private agents at higher prices. If government stocks are sufficiently small, however, the attack occurs as soon as the ceiling price is reached.

In the market for emissions trading under delayed compliance, defense of the ceiling never involves selling permits at a slow rate over a finite interval prior to the attack. For no one would buy those permits. No one needs to surrender permits until the compliance date and, if the government did sell permits over an interval at the same ceiling price, it would always be cheaper to postpone their purchase until the end of the interval. With emissions permits under delayed compliance, therefore, the speculative attack occurs the moment the ceiling price is reached. The permits are then held by private agents until the compliance date, when they are returned to the government to cover emissions since the end of the previous compliance period.

Suppose in the equilibrium under delayed compliance that the terminal price is $\mathbf{P}^*$ and cumulative emissions are $g + R$. Hence, $D(\mathbf{P}^*, T) = g + R$. Suppose the speculative attack occurs at the interior point $\tau^* > t_c$. Reallocating the $g + R$ permits between the initial allocation and the Cost Containment Reserve will not alter the equilibrium price path or the date of the attack under delayed compliance. Such reallocation may however affect the price path under continual compliance. In Figure 5, we depict the boundary case where the government chooses to grandfather $\tilde{g} = D(\mathbf{P}^*, \tau^*)$ and to stock the cost containment reserve with the remaining permits. Hence, $g + R - \tilde{g} = D(\mathbf{P}^*, T) - D(\mathbf{P}^*, \tau^*)$. If $\tilde{g}$ were reduced so that more permits were moved from the initial allocation to the Cost Containment Reserve, the equilibrium under continual compliance will differ from the equilibrium under delayed compliance. In that case, the equilibrium price path under continual compliance will have a segment rising at the rate of interest until t_c and then (weakly) dropping to $\bar{p}$ for an endogenous interval of time before rising continuously from $\bar{p}$ again at the rate of interest. A speculative attack must occur here too but it occurs later than under delayed compliance.

---

[17] For a discussion of speculative attacks on commodity ceilings defended by bufferstock sales, see Salant and Henderson (1978) and Salant (1983). For discussions of how their idea was developed further in the international finance literature, see Krugman (1999) and Flood et al. (2012).

We also note that when a speculative attack occurs ($\mathbf{P}^* > \mathbf{P_f}$), marginally changing the level of the price ceiling $\bar{p}$ does not affect the equilibrium price path and only alters the date of the attack under delayed compliance.

[Figure 5 Goes Here]

5 Conclusion

Cap-and-trade programs rather than emissions taxes are being utilized as the main vehicle to combat global warming by national (and state) governments of advanced countries. In the United States, some permits are withheld from the initial allocation and injected subsequently into the market by auctions or sales at fixed prices in an attempt to limit price increases (through so-called "price collars" or "safety valves"). The effect of these subsequent injections depends on whether the program requires regulated firms to be in compliance continually or merely periodically. Until now, virtually all analyses have assumed continual compliance even though actual programs always require delayed compliance.

The purpose of our paper has been to develop a methodology for analyzing the effects of such injections in a regime of delayed compliance. In the process of illustrating the use of this methodology, we identified two consequences of the provisions of cap-and-trade programs (potential speculative attacks and nonexistence of equilibrium) that have escaped notice.

We have also clarified when the equilibrium under continual compliance differs from the equilibrium under delayed compliance. We have not described in detail the effects of such injections under continual compliance since no actual programs require that.[18]

We have assumed away all forms of uncertainty in the current analysis but will address this issue in the future. Permit markets may be subject to three kinds of uncertainty: (1) uncertainty about the aggregate demand for permits that will be resolved by an information disclosure at a fixed date in the future; (2) aggregate demand shocks in each period; and (3) regulatory uncertainty.

The consequences of disclosing information at a known time about the demand for permits is illustrated by the collapse of the permit price in Europe following the disclosure of low demand for permits. In the case of demand shocks each period, the price path would become stochastic rather than deterministic. But if agents are risk neutral, little would change. If one works backward from the compliance period, assuming that on that date (1) all permits

---

[18] Interested readers are referred to our companion paper (Hasegawa and Salant, 2012) where such a detailed analysis is presented.

will be surrendered to the government and (2) agents with uncovered cumulative emissions must pay a well-specified penalty then, to equilibrate the market in the penultimate period under delayed compliance, the penultimate price must equal the discounted price expected in the next period. For, if that expected price were strictly higher, there would be excess demand for permits in the current period; and, if that expected price were strictly lower, there would be excess supply in the current period. But the same argument can be repeated as one works backward. In the stochastic equilibrium under delayed compliance, therefore, the price in every period must *equal* the discounted price then expected to prevail in any future period.

Regulatory uncertainty arises in part from the government's understandable goal of having the flexibility to cope with future circumstances. Regulators tend to avoid committing to future actions or policy rules. They prefer "discretion" to "precommitment." However, government flexibility, while understandable, distorts the intertemporal decision-making of private agents. This is true whether the private agents fully understand the regulator's objectives (Kydland and Prescott, 1977) or regard government actions as somewhat random (Salant and Henderson, 1978). McWilliams et al. (2011) have shown the importance of regulatory uncertainty in one permit market. Participants in the SO_2 market anticipated that at some unknown time in the future more permits would be required to cover each unit of SO_2 emissions and the price of permits would jump up. Anticipation of this uncertain event resulted in higher permit prices and more abatement; moreover, agents were willing to hold permits even though the permit price was rising by less than the rate of interest because of the capital gain they would receive when the uncertainty was resolved.

An important topic left for future work is welfare analysis. Under our assumptions, the cumulative emissions that arise in the equilibrium will be generated at least discounted costs since the marginal cost of abatement has the same present value in every period. However, under the plausible assumption that the stock of greenhouse gasses generates a flow of damages at each point in time, it has been shown (Kling and Rubin, 1997; Leiby and Rubin, 2001) that such an emissions path does not minimize the more relevant welfare functional—*the discounted sum of damages plus abatement costs.* In determining the socially optimal emissions path given such a damage function and in quantifying the welfare loss that would occur under a delayed compliance regime, we will have to take explicit account of when firms install their abatement technologies.[19]

Once we have calculated the welfare consequences of periods of delayed compliance of any given length, we can formally assess the optimal length of each compliance period. But some observations need not await formal analysis. Intuitively, the longer is each compliance

---

[19] See footnote 9.

period, the less likely is the government to wait for a period to end before intervening. In addition, the longer the compliance period, the greater the chance that (1) firms with uncovered pollution will go out of business before having to comply and (2) large utilities which have not complied before the end of the compliance period will evade regulation by threatening to shut down if forced to comply.

6 Acknowledgements

We thank Andrew Stocking for alerting us to the delayed compliance provision and Soren Anderson for circulating an earlier draft of our paper to Stephen Holland and others concerned with AB-32. Severin Borenstein, Harrison Fell, Stephen Holland, Charles Kolstad, Gary Libecap, Carol McAusland, and Matthew Zaragoza provided helpful comments on an earlier draft as did seminar participants at University of California Energy Institute (Berkeley) and Santa Clara University. We benefited from discussions of our paper at the 13th Occasional Workshop and at the Design and Evaluation of Environmental Markets Workshop at Santa Barbara. Hasegawa acknowledges the financial support provided by the Nakajima Foundation and the Yamada Academic Research Fund. Salant completed this research while in residence at Santa Barbara's Bren School as a UCE3 Senior Fellow.

References

[1] Burtraw, Dallas, Karen Palmer, and Danny Kahn. 2010. "A Symmetric Safety Valve." *Energy Policy* 38(9), pp. 4921-4932.

[2] California Air Resources Board. 2011. "Final Regulation Order (corrected on December 21, 2011), Article 5: California Cap on Greenhouse Gas Emissions and Market-Based Compliance Mechanisms." available at http://www.arb.ca.gov/regact/2010/capandtrade10/finalrevfro.pdf.

[3] Cronshaw, Mark B. and Jamie B. Kruse. 1996. "Regulated Firms in Pollution Permit Markets with Banking." *Journal of Regulatory Economics* 9(2), pp. 179-189.

[4] Fell, Harrison, Dallas Burtraw, Richard D. Morgenstern, and Karen L. Palmer. 2012. "Soft and Hard Price Collars in a Cap-and-Trade System: A Comparative Analysis." *Journal of Environmental Economics and Management* 64(2), pp. 183-198.

[5] Fell, Harrison and Richard D. Morgenstern. 2010. "Alternative Approaches to Cost Containment in a Cap-and-Trade System." *Environmental and Resource Economics* 47(2), pp. 275-297.

[6] Flood, Robert, Nancy Marion, and Juan Yepez. 2012. "A Common Framework for Thinking about Currency Crises," In Lucio Sarno, Jessica James, and Ian Marsh (eds.), *Handbook of Exchange Rates*, John Wiley & Sons, Inc., Chapter 25.

[7] Hasegawa, Makoto and Stephen Salant. 2012. "Cap-and-Trade Programs under Continual Compliance." Discussion paper 12-33. Washington, DC: Resources for the Future.

[8] Holland, Stephen P. and Michael R. Moore. 2012. "Market Design in Cap and Trade Programs: Permit Validity and Compliance Timing." NBER Working Paper 18098.

[9] Hone, David. 2012. "The Case for an Auction Reserve Price." available at http://blogs.shell.com/climatechange/2012/06/auction/.

[10] Jacoby, Henry D. and A. Denny Ellerman. 2004. "The Safety Valve and Climate Policy." *Energy Policy* 32(4), pp. 481-491.

[11] Kling, Catherine and Jonathan Rubin. 1997. "Bankable Permits for the Control of Environmental Pollution." *Journal of Public Economics* 64(1), pp. 101-115.

[12] Krugman, Paul. 1999. "Currency Crises," available at: http://web.mit.edu/krugman/www/crises.html.

[13] Kydland, Finn E. and Edward C. Prescott. 1977. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85(3), pp. 473-492.

[14] Leiby, Paul and Jonathan Rubin. 2001. "Intertemporal Permit Trading for the Control of Greenhouse Gas Emissions." *Environmental and Resource Economics* 19(3), pp. 229-256.

[15] McWilliams, Michael, Michael R. Moore, Gloria Helfand, and Yimin Liu. 2011. "Explaining Price Movements of the SO_2 Market," work in progress.

[16] Rubin, Jonathan D. 1996. "A Model of Intertemporal Emission Trading, Banking, and Borrowing." *Journal of Environmental Economics and Management* 31(3), pp. 269-286.

[17] Salant, Stephen W. 1983. "The Vulnerability of Price Stabilization Schemes to Speculative." *Journal of Political Economy* 91(1), pp. 1-38.

[18] Salant, Stephen W. and Dale Henderson. 1978. "Market Anticipations of Government Policies and the Price of Gold." *Journal of Political Economy* 86(4), pp. 627-648.

[19] Stavins, Robert N. 2012. "Low Prices a Problem? Making Sense of Misleading Talk about Cap-and-Trade in Europe and the USA." available at http://www.robertstavinsblog.org/2012/.