

Zaklan, Aleksandar; Ellerman, Denny; Valero, Vanessa

Conference Paper

An Analysis of Allowance Banking in the EU ETS

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Environmental Economics III, No. C13-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Zaklan, Aleksandar; Ellerman, Denny; Valero, Vanessa (2015) : An Analysis of Allowance Banking in the EU ETS, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Environmental Economics III, No. C13-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/113034>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

An Analysis of Allowance Banking in the EU ETS

January 14, 2015

Abstract

The existence of some 2 billion unused EU Allowances (EUAs) at the end of Phase II of the EU's Emissions Trading System (EU ETS) has sparked considerable debate about structural shortcomings of the EU ETS. However, there has been a surprising lack of interest in considering the accumulation of EUAs in light of the theory of intertemporal permit trading, i.e. allowance banking. In this paper we adapt basic banking theory to the case of a linearly declining cap, as is common in greenhouse gas control systems. We show that it is perfectly rational for agents to decrease emissions beyond the constraint imposed by the cap initially, accumulating an allowance bank and then drawing it down in the interest of minimizing abatement cost over time. Having laid out the theory, we carry out a set of simulations for a reasonable range of key parameters, geared to the EU ETS, to illustrate the effects of intertemporal optimization of abatement decisions on optimal time paths of emissions and allowance prices. We conclude that bank accumulation as the result of intertemporal abatement cost optimization should be considered at least a partial explanation when evaluating the current discrepancy between the cap and observed emissions.

JEL classification: D92, F18, Q54

Keywords: Cap and Trade System, EU ETS, Intertemporal Trading.

1 Introduction

The existence of 2 billion unused allowances at the end of Phase II of the European Union's Emissions Trading System (EU ETS), roughly 20% of the five-year cap, is often cited as the cause of the current low price of European Union Allowances (EUAs), which is seen as problematic by some contributors to the policy debate. The existence of this permit bank is variously attributed to effects of the financial crisis of 2008 and the subsequent euro crisis on GDP growth, to the use of nearly 1.1 billion offsets, and to the promotion of energy from renewable sources that has displaced generation by CO₂ emitting coal and natural gas. The observed discrepancy between allowances distributed and allowances used has been the motivation for the debate about "back-loading" that has dominated discussion concerning the EU ETS in 2013, as well as for the proposal made in January 2014 to establish a Market Stability Reserve. Both of these measures would reduce the number of allowances available in some near term while putting the withdrawn allowances back into circulation at a later time.

The rules governing unused allowances at the end of Phase II (2008-2012) are different from what they were at the end of Phase I (2005-2007). The ETS Directive established from the beginning that any unused allowances in Phase II could be "banked," that is, carried over for use in subsequent years. In contrast, the rules developed for Phase I did not allow unused allowances to be carried over for use in Phase II. The bank at the end of Phase I was very small (approximately 100 million allowances, less than 2% of the three-year total) and the expectation of this surplus drove the price to zero well before the end of 2007. The price at the end of Phase II was much lower (around five euros) than expected at the beginning of Phase II (perhaps thirty euros), but the price was never driven to zero, presumably because holders of these allowances believed they had greater unit value than the lowest prices offered (around three euros).

Given the ability to bank allowances in Phase II for later use, it is surprising how little attention has been given to the application of allowance banking theory to the EU ETS. This neglect is the more surprising for the extent to which banking was recognized as a major factor in explaining agent behavior in the US SO₂ Emissions Trading Program (Schennach, 2000; Ellerman et al., 2000; Ellerman and Montero, 2007). That program was, however, structured very differently from the EU ETS, the main difference being a pre-determined sharp discontinuity in the cap between Phases I and II. In contrast, there is no similar discontinuity either in cap level or coverage in the EU ETS.¹ While the five-year cap was flat during Phase II, a linear reduction factor of 1.74% of the average total quantity of allowances issued annually during Phase II become effective at the start

¹Although the relative expansion in terms of affected units has not been as pronounced as between Phases I and II of Title IV, the EU ETS has been expanded considerably over time through the addition of both new countries and sectors.

of Phase III. This smoothly declining cap did not raise the specter of a sharp increase in the marginal cost of abatement that would create an obvious motivation to bank, as in the US SO₂ Program. Still, a steadily declining cap could be expected to lead to rising marginal abatement cost over time and this was the general expectation as evidenced by no small number of modelling studies that sought to predict EUA prices beyond the next few years. The question of whether the expected increase in marginal cost would be enough to warrant banking (and if so, at what levels) seems simply never to have been asked.

This paper seeks to remedy that failure of analysis. We do not propose to attempt a definitive answer that settles the question once and for all time, but only to raise the question of whether allowance banking could be expected to occur in a cap and trade system such as the EU ETS. We characterize the main determinants of permit banking before exploring its implications, especially concerning the levels of banked EUAs that might be expected at the end of Phase II under some specific assumptions. Doing so requires first of all that the theory of allowance banking (Rubin, 1996; Cronshaw and Kruse, 1996; Kling and Rubin, 1997; Leiby and Rubin, 2001) be fitted to the peculiarities of the EU ETS, in particular, the smoothly and modestly declining cap. The next section presents and explains the theory of allowance banking as it would be applied to the EU ETS. Section 3 discusses reasonable values for critical parameters. Section 4 presents simulations of plausible banking outcomes for the EU ETS. Section 6 discusses these results, while section 7 concludes.

2 Intertemporal Trade with a Linearly Declining Cap

2.1 An Intuitive Explanation of Allowance Banking

Allowance banking is a manifestation of trading through time instead of the more familiar trading across space in the same time period. Intertemporal trading could involve borrowing, but in the EU ETS, as in most allowance trading programs, borrowing is not allowed.² Thus, agents cannot use tomorrow's allowances to cover today's emissions, but they can use any allowances not used to cover today's emissions to cover tomorrow's emissions.

The basic intuition of allowance banking is similar to that for spatial trading. If the marginal cost of abating a ton of emissions is higher at one place or time than at another, cost savings can be gained by reducing more than required where or when it is cheaper and using these allowances to cover emissions where and when it is more costly. Trading

²It has been possible (and continues to be for some agents) to borrow one year ahead because of the schedule by which freely allocated allowances are distributed in relation to when allowances must be surrendered against emissions. Intertemporal trading usually implies multi-year horizons so that this exception does not invalidate the basic rule against borrowing. Besides, with free allocation being phased out in Phase III, the ability to use this near-term flexibility is being progressively reduced.

across time invokes two further considerations, namely, whether borrowing is permitted and the discount rate for comparing today's and tomorrow's costs. When borrowing is not allowed, the calculation is only whether tomorrow's expected cost is higher than today's given the discount rate. If tomorrow's discounted expected cost is higher than today's cost, it is worth holding allowances, whether obtained by abating more today or by purchase, and using them to either cover some of tomorrow's emissions or sell them later on. If that expectation is not true, then it is not worth doing so. It would be more profitable to use or sell any allowances held, however obtained, at today's price and earn the return represented by the discount rate than to continue holding them.

While the theory and basic intuition of allowance banking is clear, any particular application requires fitting the theory to the structure of the particular allowance trading program. For the EU ETS, the essential features of the cap are that it starts at a level at or only slightly below counterfactual emissions and declines continuously at a pre-specified linear reduction rate (LRF). The LRF implies, under the usual assumption of rising costs with increasing amount of abatement, and with all other things equal, that the marginal cost of abatement will rise over time as the cap declines. How much marginal cost (or the price) will rise depends not only on how much abatement is implied by the cap, but also on the marginal abatement cost function, which determines the rate at which marginal cost rises as more abatement is required. The issue for the individual agent is whether the expected increase in marginal cost over time is greater than the discount rate. If it is, banking is justified and vice versa.

Whether the expected rising cost would justify banking (and if so, how much) depends a great deal of the relationship between the cap and counterfactual emissions. Consider, for instance, an initial cap that is set at a level 1% below counterfactual emissions and that then declines at 1% per year while counterfactual emissions are constant. In the first year, the required abatement is 1% of counterfactual emissions, 2% in the second year, 3% in the third year and so on. Assuming a simple linear marginal cost function, marginal cost will double in the second year, increase by 50% in the next year, 33% in the following year and so forth in a continually declining series as the additional abatement required in each succeeding year becomes smaller in proportion to abatement in the past year. The growth rate of the marginal cost should be compared at the discount rate. When the marginal abatement cost increases at a rate greater than the discount rate, the agent banks allowances. In other words, she emits less than required and saves allowances for future use (or sale). We call this the accumulation phase. However, at some point in the future the rise in the marginal cost of abatement will be less than the discount rate. Agents will emit more than imposed by the cap and use their banked allowances to cover the difference, until the bank is exhausted. This is called the the draw-down phase. After the point of exhaustion agents would like to continue emitting above the cap and borrowing from their future allocation to cover the shortfall in available permits.

However, borrowing is not allowed, so that their emissions will be equal to their allocation.

The period of time during which firms bank allowances and then draw down the accumulated bank is called the banking period. How much banking will occur and for how long, and with what effect on the allowance price, depends on the values embedded in the agents' expectations. To answer such questions and to put more structure on the manifestations and effects of allowance banking, we must turn to the theory as applied to the specifics of the EU ETS.

2.2 The Model of Intertemporal Trading

We develop a model representing a cap and trade system for greenhouse gas emissions. Contrary to the SO₂ program analysed by Rubin (1996) and Schennach (2000), the amount of allowances is limited by a cap which is reduced over time.

As stated in previous literature, the firms' level of emissions can be characterized as if a single central planner were making the decision.³ The latter wants to minimize the presented discounted abatement cost with respect to the aggregate level of emissions when banking is allowed. He thus faces the following infinite-horizon dynamic abatement cost-minimization problem (P):

$$\min_{\{e_t\}} \left\{ \int_0^{\infty} e^{-rt} C(u_t - e_t) dt \right\}$$

$$\text{s.t. } \dot{B}_t = Y_t - e_t \quad (1)$$

$$B_t \geq 0 \quad (2)$$

where C is the abatement cost, i.e. the cost of reducing emissions where $C'(\cdot) > 0$ and $C''(\cdot) > 0$.⁴ The quantities abated correspond to the difference between u_t , the aggregate counterfactual emissions (i.e. the emissions emitted without any restriction on emissions by all firms participating in the system) and e_t , the aggregate level of emissions at time t . The variable r corresponds to the discount rate, assumed to be constant over time. The variable B_t refers to the number of allowances in the bank, assumed non-negative, which evolves according to the state equation (1). It states that the annual change in the bank is the difference between each period's cap Y_t and emissions e_t .⁵

As is the case in the EU ETS, we assume that the cap Y_t decreases each period at a constant rate a (known as the LRF) such that:

$$Y_t = e^{-at} Y_0$$

³See Rubin (1996) and Schennach (2000).

⁴We assume no technological progress so the cost function does not change with time.

⁵Note that \dot{B}_t can be either positive or negative with positive values signaling a build-up of the bank and negative values indicating a draw-down of the bank.

where Y_0 is the initial cap.

We also assume that the aggregate counterfactual emissions u_t increase each period at a constant rate b such that:

$$u_t = e^{bt}u_0$$

where u_0 is the initial level of counterfactual emissions.

We then define the Lagrangian expression of the problem (P):⁶

$$\mathcal{L} \equiv e^{-rt}C(u_t - e_t) + \Lambda_t(Y_t - e_t) - \Phi_t B_t$$

where Λ_t and Φ_t are multipliers associated with the constraints on the change in size of the bank and on borrowing, respectively. As shown by Rubin (1996) and Schennach (2000), differentiation of the Lagrangian to obtain the first-order conditions and further rearrangement of terms leads to the following equation for the change in the marginal cost (mc) of abatement through time:

$$\dot{mc}(u_t - e_t) = r mc(u_t - e_t) - \Phi_t \quad (3)$$

It states that the change in marginal cost over time along the optimal path will be equal to the discount rate minus the shadow price of the constraint on borrowing. The conditions for the solution of the optimal path are such that, after the banking period (i.e. for $t \geq \tau$) when the bank is exhausted ($B_t = 0$) and the borrowing constraint ($B_t \geq 0$) becomes operative, the shadow price is strictly positive $\Phi_\tau > 0$. Consequently, marginal cost mc increases at a rate less than the discount rate. Since reducing emissions is costly for firms, the aggregate level of emissions e_t will be equal to the cap such that:

$$e_t = Y_t \quad (4)$$

These same conditions imply that when agents are banking a positive quantity of allowances $B_t > 0$ during the banking period (i.e. for $0 \leq t < \tau$), the multiplier Φ_t is zero so that the marginal cost mc increases at the discount rate. During this time interval and given initial marginal cost, the evolution of marginal cost can be expressed as:

$$mc(u_t - e_t) = e^{rt}mc(u_0 - e_0) \quad (5)$$

From equations (4) and (5), we derive the path of the allowance price P_t over the entire time horizon from 0 to ∞ such that:⁷

$$P_t = \begin{cases} e^{rt}mc(u_0 - e_0) & \text{for } t < \tau \\ mc(u_t - Y_t) & \text{for } t \geq \tau \end{cases}$$

⁶To solve this continuous time minimisation problem, we use optimal control theory.

⁷In this dynamic equilibrium model, we assume perfect competition so that the allowance price equals the marginal abatement cost.

From this, the path of emissions can also be deduced:

$$e_t = \begin{cases} u_t - e^{rt}(u_0 - e_0) & \text{for } t < \tau \\ Y_t & \text{for } t \geq \tau \end{cases} \quad (6)$$

During the banking period, emissions will be equal to counterfactual emissions less the amount of abatement forthcoming at each point in time. Thereafter, the cap will determine emissions. The remaining task is to determine the initial level of emissions e_0 , which depends on the discount rate and the end of the banking period τ .

By prior reasoning and the conditions for defining an optimal banking program, the allowance price increases at the discount rate during the banking period. We thus can rewrite the path of price P_t depending on the end of the banking period τ such that:

$$P_t = mc(u_t - e_t) = e^{-r(\tau-t)}mc(u_\tau - Y_\tau) \quad (7)$$

which leads to $P_0 = e^{-r\tau}mc(u_\tau - Y_\tau)$ and $e_0 = u_0 - e^{-r\tau}mc(u_\tau - Y_\tau)$.

Also, since all the allowances issued over the banking period must be equal to the cumulative emissions at τ , we have:

$$\begin{aligned} \int_0^\tau e_t dt &= \int_0^\tau Y_t dt \\ \Leftrightarrow \int_0^\tau e_t dt &= \frac{Y_0}{a}(1 - e^{-a\tau}) \end{aligned} \quad (8)$$

Finding τ is then an iterative process of finding the point in time when the condition expressed by equation (8) is met. Any point in time before τ implies a lower price during the banking period, less abatement, more cumulative emissions than allowances issued until then, and therefore a violation of this condition. Conversely, any point in time after τ implies a higher price path, more abatement, and fewer cumulative emissions than allowances, also a violation of the condition. Once τ is found, everything else follows, as illustrated in the simulations. The analytical solution when assuming a linear marginal abatement cost function is presented in the Appendix.

3 Intertemporal Trading in the EU ETS

Applying the allowance banking theory explained in Section 2 to the EU ETS and simulating how agents might be expected to behave requires assumptions about several parameters. The goal of the simulations presented in the next section should be considered exploratory and expository rather than an attempt to model accurately the past and future development of the EUA bank.

We remind the reader that the simulations correspond to the model as presented in the previous section of this paper, i.e. the world is perfectly deterministic and there are

no shocks or changes in expectations concerning the relevant variables. Of course, in reality shocks do occur and expectations may and typically will change over time. Still, the fundamental insights and motivations for banking remain and the basic structure will remain unchanged. For these reasons, no attempt is made to calibrate parameter values to match the paths observed to date. However, the effect of reasonable variations in parameter values are presented and discussed to provide the reader with a sense of how results are affected by changes in these values.

3.1 Discount Rate

The futures market for EUAs provides data concerning the discount rate that agents apply in valuing present and future values at least for the term of the various EUA futures contracts. Unfortunately, these contracts extend only three to four years into the future, less than what the banking period would likely be, but they are a starting point. Moreover, the yield curves in these contracts are remarkably stable, much as would be expected given the Hotelling-like price path that is predicted by equation (3), and very unlike the variations from backwardation to contango that can be observed in the futures markets for commodities (Ellerman, Marcantonini and Zaklan, 2014).⁸ The implied discount rates are always positive and since the beginning of 2008 they have varied between a low of 1.4% and a high of 9.1% with most observations falling between 2.5% and 5.5%.⁹ The relationship to the euro interest rate set by the European Central Bank, which started the period at 4% and had been reduced to 0.75% in 2012 and more recently to 0.25%, is decidedly tenuous. For this reason it is not clear how to approximate the discount rate. We thus chose a range of discount rates for the simulations that follow, applying a central rate of 4%, with variations of 3% and 5%.

3.2 Growth in Counterfactual Emissions

Counterfactual emissions, the evolution of emissions in the absence of the EU ETS, are not known with any certainty, but some estimate of these emissions and therefore of the abatement occasioned by the EU ETS, is implicit in today's price, not to mention future prices. As discussed in Ellerman et al. (2010), reconstructions of pre-2005 emissions for the ETS sectors indicated a rate of increase of about 1% per annum since 2000 at a time when EU15 GDP was increasing at a rate of about 2%. These data imply a 1% rate of

⁸A potential explanation is that an inventory stock-out is virtually impossible in allowance markets since allowances are not required inputs at the time of emissions (although the liability is then incurred) and agents typically have a grace period of several months after the close of each compliance period before allowances equivalent to the compliance-period emissions must be surrendered.

⁹The average discount rate for the prompt and next contracts (the most thickly traded) for weekly observations since the beginning of 2008 is 3.98% with a median of 3.92% and a standard deviation of 1.53%.

decline in CO₂ emissions per unit of GDP and are consistent with the regularly observed secular improvement in the carbon efficiency of economies which is often used modelling the relation between GDP growth and CO₂ emissions. For the purpose of the simulations presented here, high and low growth assumptions are presented, as well as an intermediate value. These assumptions can be taken to reflect the relatively optimistic expectations concerning economic growth that prevailed until the financial crisis of late 2008 and the considerably dampened expectations that have prevailed since. Counterfactual emissions are assumed to increase at a constant rate b , for which the upper and lower values are 1.0% and 0.1% annual growth with 0.5% as the intermediate value. Assuming an underlying trend of 1.0% improvement in the carbon efficiency of the economy, these values imply growth rates for output in the ETS sectors between 1% and 2% per annum. Again, no claim is made that these values are accurate reflections of the expectations that have informed market and banking behavior. They are plausible, but their value for our present purpose is illustrative.

3.3 Slope Coefficient of the Declining Cap

In contrast to counterfactual emissions and discount rate, the cap is always known, embedded in legislation with as much certainty as implemented policy provides. In the case of the EU ETS, the rate of decline is specified in the amended ETS Directive, 1.74% annually.^{10,11} The relation of this precisely defined cap to initial counterfactual relations is not known, but it is widely believed that the initial EU ETS cap was at or very little below business-as-usual emissions. For simplicity, we assume that the initial level of both the cap and counterfactual emissions is the same and that the cap declines indefinitely at a constant rate a of 1.74% annually.¹²

3.4 Specification of the Marginal Abatement Cost Function

The shape of the aggregate marginal abatement cost function is perhaps the least researched aspect of the EU ETS. Every model representing the EU ETS implicitly has

¹⁰In its recent communication concerning the 2030 framework for climate and energy policy, the European Commission has suggested that a decline rate of 2.2% starting in 2021 would be consistent with the proposed legally binding target of reducing EU greenhouse gas emissions to 40% below 1990 levels by 2030. However, a specific proposal to do so was not put forward at this time. That will be one of the many measures to be decided following the Parliamentary elections in May 2014 and the installation of a new Commission thereafter.

¹¹This linear reduction factor has been interpreted as a constant decrement equal to 1.74% of the average annual cap in Phase II. In the interest of simplicity, we ignore what will be an accelerating decrement when expressed as a proportion of each year's cap. This detail will modify specific results but not change the basic nature of banking or the conclusions presented here.

¹²While the cap was flat during Phase II and the linear reduction factor of 1.74% was applied starting with the 2013 compliance year, the theoretical analysis and simulations in this paper assume a linearly declining cap from the beginning to simplify the exposition.

such a function, which is inevitably the result of values assigned to various elasticities or to cost engineering data. The common feature of all is the economic intuition that marginal costs rise as the amount of abatement increases. As others have done (cf. Schennach, 2000, and Ellerman and Montero, 2007), we assume a time-invariant, linear marginal cost function with parameters chosen for the underlying total cost function to yield prices roughly in line with observations.

4 Simulations

4.1 Emissions and Prices in the Baseline Case

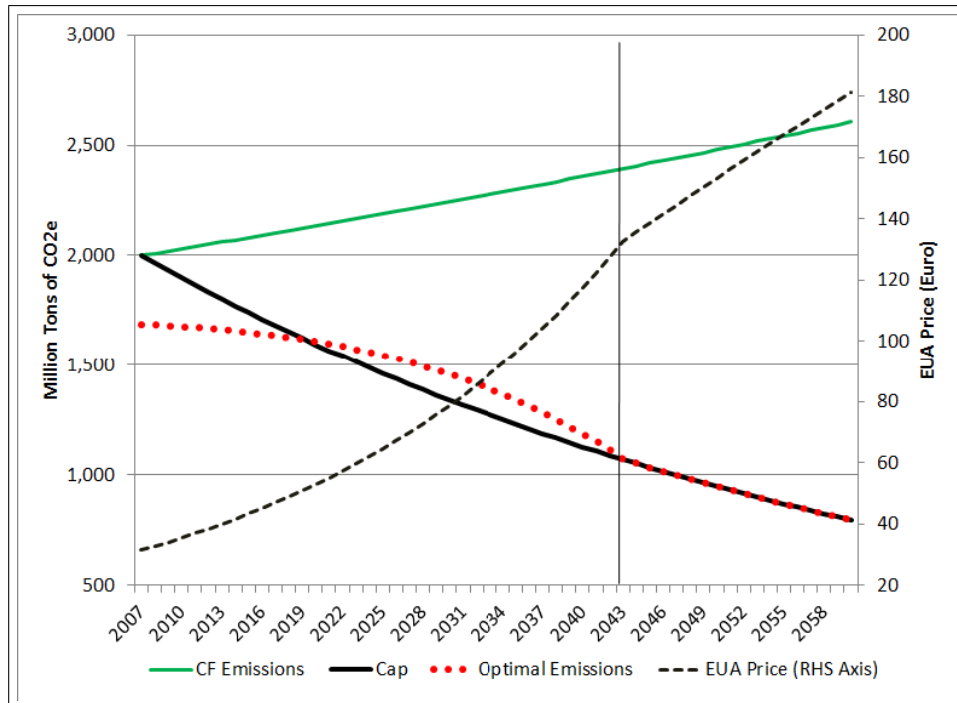


Figure 1: Cap, Emissions, Price Path, Baseline Case

Figure 1 illustrates the simulation results for the baseline case, where counterfactual emissions grow at an annual rate of 0.5% and the representative agent discounts the future at a rate of 4%. In this figure and subsequent ones, results are measured both by quantities and prices. In Figure 1, the cap, counterfactual emissions, and the optimal emission path are to be read against the millions of tons of CO2 on the left-side scale. The price path should be read against the right-side scale which is in euros per ton of CO2. The evolution of the bank itself in this baseline case is the middle line in Figure 2.

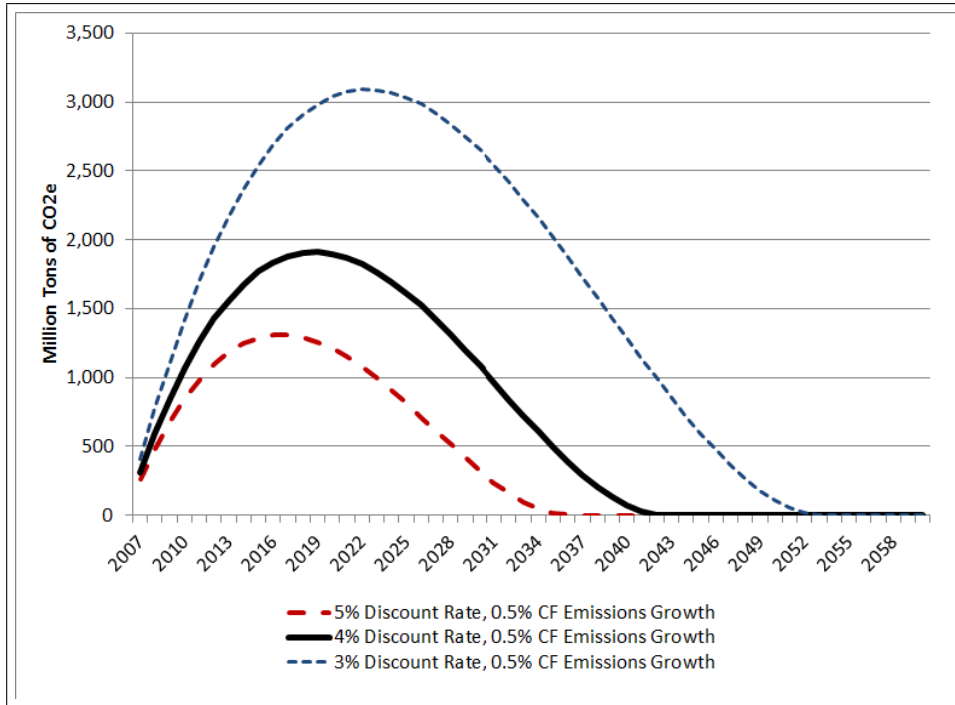


Figure 2: Optimal Cumulative Bank, Varying Discount Rates

The vertical line in Figure 1 marks the end of the banking period (2043). This is when the emissions path becomes permanently coincident with the cap because of the inability to borrow and a kink is observed in the price path when the rate of increase in allowance prices becomes less than the discount rate. Comparing the optimal emissions path with the cap shows that emissions will be below the cap for an initial sub-period of accumulation, ending in 2019 for this simulation. Thereafter, the accumulated bank is drawn down over a period of 24 years and emissions exceed the cap in these years. Of course, for the banking period as a whole, cumulative emissions equal cumulative allowed emissions. The cross-over year (when emissions momentarily equal the cap and the draw-down of the bank starts) is also when the bank reaches its peak value. Note also that the accumulation phase is considerably shorter (11 years) than the draw-down phase (24 years), which implies that average amount by which emissions are below the cap in the accumulation phase is roughly twice the average by which emissions will be above the cap in the subsequent draw-down phase. In the following we explore the effects of varying the discount rate, rate of growth in counterfactual and slope of the cap on the optimal banking and price paths.

4.2 Changing the Discount Rate

Varying the discount rate generates substantial changes both in the size of the bank and in the length of the banking period (Figure 2).

For instance, applying a discount rate of 5% shortens the banking period by 7 years compared to the baseline case using a 4% discount rate and decreases the maximal bank by about one third. Applying a 3% discount rate lengthens the banking period by 10 years and increases bank's maximum size by more than 50%.

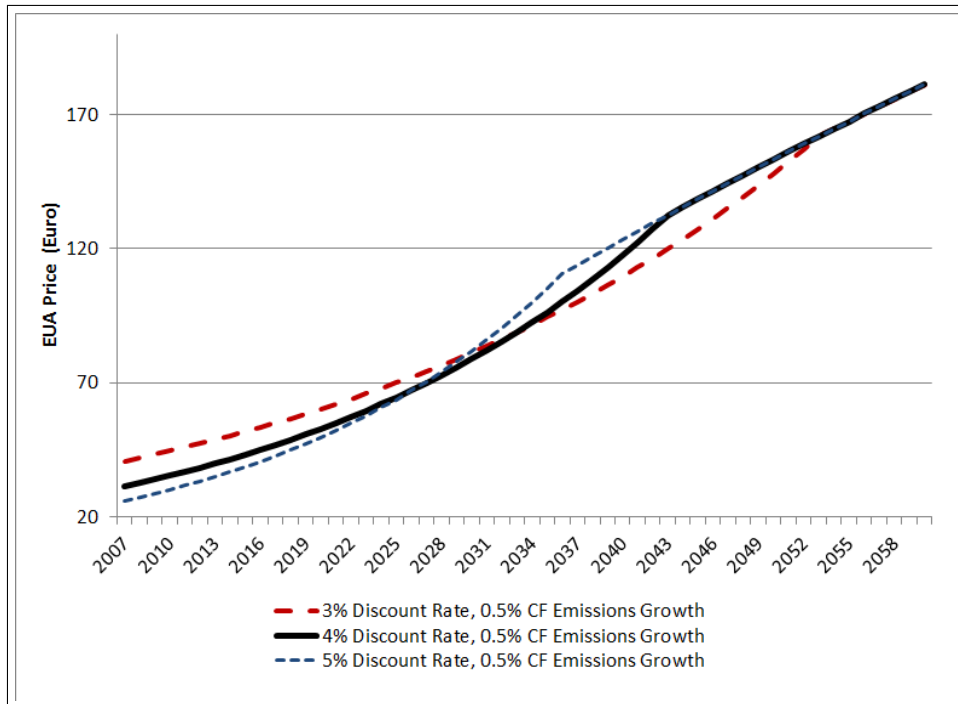


Figure 3: Optimal Price Path, Varying Discount Rates

Variations in discount rates also have an impact on the optimal permit price path (Figure 3). Changing the discount rate not only alters the rate of price increase during the banking period but also changes the length of the banking period determined by and thereby the associated reference price that anchors the price path during the banking period. Mechanically, a lower discount rate means that the point when marginal abatement cost rises less than the discount rate is farther in the future. The lower discount rate makes future abatement more costly relative to the present thereby justifying more early abatement to mitigate those costs, a higher initial price, a longer banking period and a larger bank. A higher discount rate has the opposite effect.

4.3 Varying the Growth Rate in Counterfactual Emissions

Varying the growth rate of counterfactual emissions mainly affects the maximal size of the bank (Figure 4) and has, in comparison to changes in the discount rate, relatively little impact on the length of the banking period. Whether the growth rate is 0.1% or 1.0% changes the length of the banking period by only 4 years. However, the effect on the maximal bank size is more substantial. Changing the counterfactual emissions growth from the middle value of 0.5% to 1% increases the maximal size of the bank by almost 50%, and a change of the growth rate in the opposite direction, from 0.5% to 0.1%, decreases the maximal bank size by about 25% with respect to the baseline case. Changing the growth rate in counterfactual emissions also strongly affects the price path (Figure 5), since greater growth in counterfactual emissions requires more abatement and shifts up the price path without changing the rate of increase in marginal cost and price.

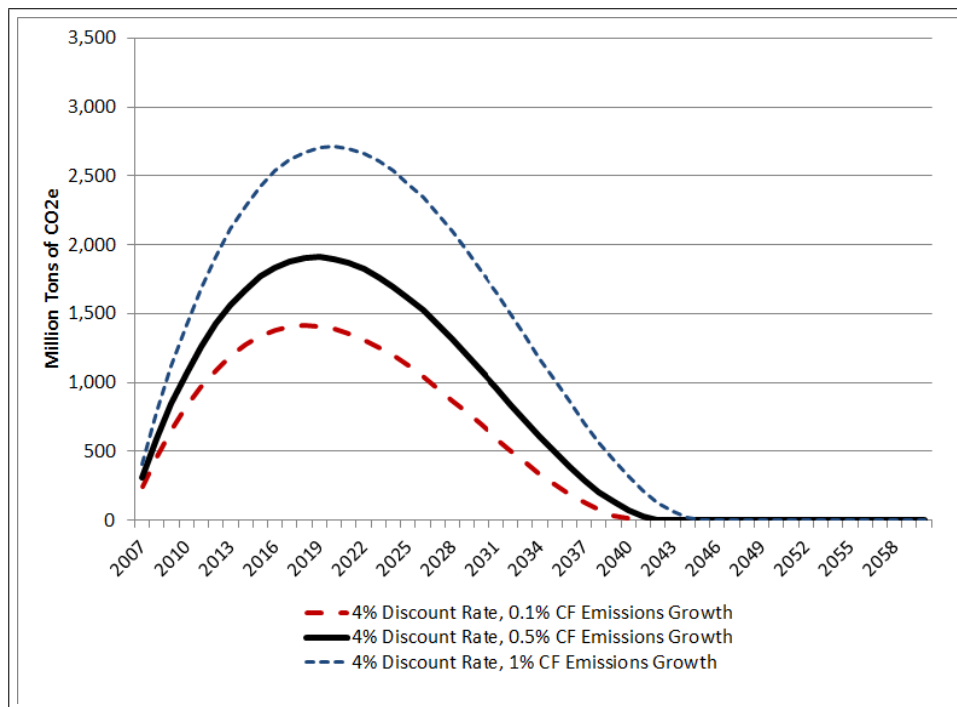


Figure 4: Optimal Cumulative Bank, Varying Growth in Counterfactual Emissions

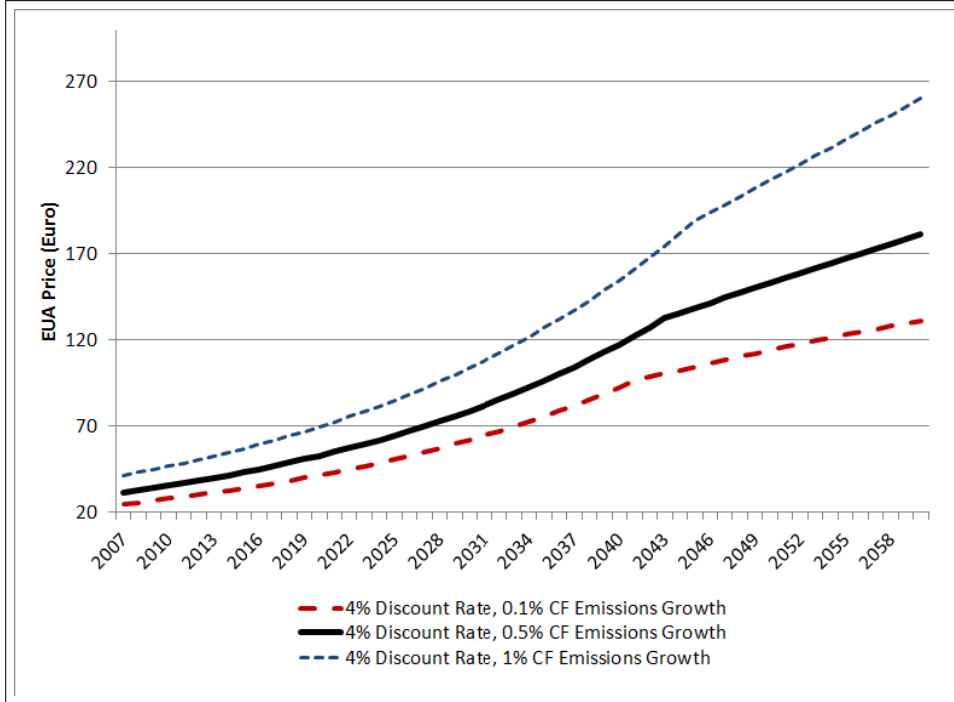


Figure 5: Optimal Price Path, Varying Growth in Counterfactual Emissions

5 Extension

5.1 Change in the Growth of Counterfactual Emissions

In this part we examine the impact of a one-time potential change in the growth rate of counterfactual emissions on the optimal level of emissions over time. Specifically, we consider a change in the growth rate b at time t' . Counterfactual emissions can either change to b' , with known probability μ , or remain at b , with probability $1 - \mu$. We denote the difference between the two growth rates $\Delta b \equiv b' - b$. Thus, counterfactual emissions are such that:

$$u_t = \begin{cases} e^{bt} u_0 & \text{for } 0 < t \leq t' \\ (1 - \mu)(e^{b(t-t')} u_{t'}) + \mu(e^{b'(t-t')} u_{t'}) & \text{for } t > t' \end{cases}$$

where $u_{t'}$ is the level of counterfactual emissions at time t' .

In our model the values of all parameters are known ex ante and since the optimization occurs in one shot the shape of the optimal path of permit banking will not change qualitatively compared to a scenario in which no change occurs. Only the amounts banked and the length of the banking period will change. Recall that when determining the

optimal level of emissions the firm minimizes the compliance cost over an infinite time horizon. She therefore considers the level of counterfactual emissions over the entire interval of time, implying that t' does not play any role in the determination of u_t because it will always be closer to zero than to infinity. Quantitatively this means that the level of counterfactual emissions will be a weighted average between the growth rates prior to and after the change respectively, the weights being the values of μ and $1 - \mu$.

The size of the bank and the length of the banking period will depend on the difference between both growth rates, Δb , and the probability that the change occurs, μ . When $\Delta b < 0$ counterfactual emissions decrease and require less abatement with respect to the no-change scenario. The maximal bank size is reduced and the length of banking period shortened slightly. Overall the size of the bank is decreased. The reverse will be true for $\Delta b > 0$. When μ tends to zero we get the bank size and length of banking period corresponding to the original scenario with counterfactual emissions growing at the rate b . For μ approaching one bank size and length of banking period will be according to the lower-growth scenario of counterfactual emissions, b' .

Figure 6 illustrates this situation for an anticipated drop in the growth rate of counterfactual emissions from 0.5% to 0.2% with probability $\mu = 0.5$. This corresponds to a case in which counterfactual emissions grow at the rate of 0.35% throughout. For comparison Figure 6 also depicts the two optimal banking paths for the cases of 0.5% and 0.2%.

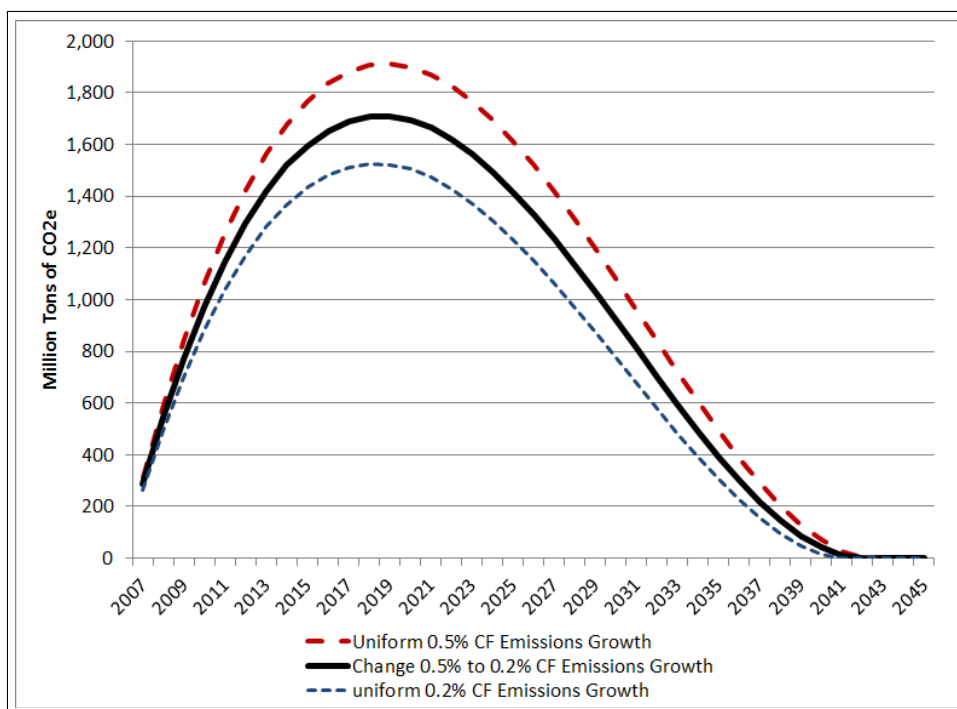


Figure 6. Anticipated Change in the Growth of Counterfactual Emissions

We observe that the new optimal banking path lies precisely between the paths for the two original cases. As the probability the change occurs equal to 0.5, the new optimal path is at equidistance between both lines representing the no-change and the change scenario.¹³

The period during which such a change would happen, t' , will play a role only if we consider an unexpected change. However, the model in its current form is not suited to considering this case. As it rests on the assumption of perfect foresight there is nothing unexpected that the agent might have to face. To consider the effects of unexpected changes in the growth of counterfactual emissions, or in any of the other parameters, we would first have to relax the assumption of perfect foresight. This is left to future research.

5.2 Change in the Linear Reduction Factor

We now examine the impact of a change in the slope of decline in the cap on banking. The analysis in this sub-section reflects the proposal by the European Commission of tightening the LRF to 2.2% at the start of Phase IV. We denote the probability p that the LRF changes at some point t'' during the banking period. In this case, the number of allowances allocated to participating firms decreases each year by a constant rate a from period 0 to t'' and then to a different rate a' starting in t'' :

$$Y_t = \begin{cases} e^{-at}Y_0 & \text{for } 0 \leq t \leq t'' \\ e^{-a'(t-t'')}Y_{t''} & \text{for } t > t'' \end{cases}$$

In the case of the EU ETS a' would be -2.2% per year, compared with -1.74% currently such that $a > a'$. If we continue the example, t'' would be 2021. As in the previous extension the size of the potential change and the probabilities are known.

From equations (7) and (8), we obtain:

$$\tilde{u}\tau - \frac{(1 - e^{-r\tau})}{r}(\tilde{u} - Y_\tau) = (1-p) \left[\frac{Y_0}{a}(1 - e^{-a\tau}) \right] + p \left[\frac{Y_0}{a}(1 - e^{-at''}) + \frac{Y_0}{a'}e^{-at''}(e^{-a't''} - e^{-a'\tau}) \right]$$

When the cap is tightened the quantity of abatement increases. Therefore, the firms will bank more allowances initially to decrease their discounted future marginal cost, increasing the total size of the bank. Tightening the cap also lengthens the banking period. The discussion from the previous section applies here, too. The effects of the change are stronger the closer t'' is to the beginning of the banking period and as p increases.¹⁴ When the change in the slope of decline in the cap is certain, i.e. $p = 1$, we

¹³When one or the other scenario is more likely the new path will be closer to the case with the higher probability of occurrence.

¹⁴We omit a figure illustrating this situation as the case is similar to that presented in the previous sub-section.

obtain:

$$\tilde{u}\tau - \frac{(1 - e^{-r\tau})}{r}(\tilde{u} - Y_\tau) = \frac{Y_0}{a}(1 - e^{-ac}) + \frac{Y_0}{a'}e^{-ac}(e^{-a'c} - e^{-a'\tau})$$

6 Discussion

An examination of Figures 2 and 4 shows that, under the assumptions used, a stock of allowances would have been built up during the first five years of the ETS and that the optimal level at the end of 2012 would have been somewhere between 1.3 billion and 3 billion depending on parameter values. Moreover, all of these simulations suggest that the banking build-up has several years to go. On this evidence alone, allowance banking would seem to offer at least a partial explanation for the allowance stock accumulated during Phase II. Perhaps, the current level of the bank is larger (or smaller) than what our analysis suggests, but the fact that allowance banking would lead to comparable numbers should at the very least give pause to those who would suggest that the existence of such a surplus indicates some deep flaw in the design or functioning of the EU ETS. Instead, our results suggest that a considerable bank may accumulate as a rational response by agents facing future scarcity who anticipate rising costs at a rate higher than the discount they apply in evaluating present and future costs.

Observed allowance prices offer another point of comparison between theory and observed practice and serve to highlight the limitations of our model. Our model is focused on the long run and assumes that demand increases and supply decreases at constant rates, respectively. We focus on understanding the basic mechanics of banking assuming parameters that are either constant or only subject to one-time anticipated changes. The model abstracts from shocks to either the supply or the demand sides of the permit market. Shocks to the EU ETS have certainly occurred in recent years, for instance the expansion of supply through the rapid influx of Kyoto offsets, as well as a decrease in counterfactual emissions on account of the recent economic crisis, which led to a decrease in demand. Either of these shocks would have had a negative effect on the permit price, so that there is no doubt about the direction of the resulting price or quantity effects, namely a fall in the equilibrium price and an increase in the equilibrium quantity.

As our model does not take into account such shocks it cannot capture their effects on either prices or banking. However, we know that a fall in prices would further increase the incentive to bank permits, since the difference between present and future equilibrium marginal abatement cost will increase, given the increasing amount of required abatement due to the decline in the cap. Thus, given the direction of the recent shocks our deterministic model tends to under-predict the build-up of the bank quantitatively, since it assumes higher prices than we observe. However, qualitatively the result is unambiguous. The model therefore provides a lower bound for explaining the amount of permits

accumulated, with no downward shocks to prices. Adjusting the model to account for the effects of shocks quantitatively is left for future research.¹⁵

Some further qualifications are in order, highlighting the need for future research in key areas. The first is about the effective time horizon over which agents optimize. The theory as we have applied it assumes that agents are so far-seeing as to be thinking about 2043 for instance. Agents may be easily assumed to be forward-looking, but do their horizons extend thirty years and even more into the future? And, if the reality is truncated horizons that are updated and moved forward as time progresses, what is the relevant time span? More importantly for our purposes, how would such behavior change the equilibrium paths that have been presented in the body of this paper?

Another issue is the effect of agent heterogeneity on outcomes in the permit market, one example being the rate at which different firms discount the future. If we assume that each firm's discount rate is related to the relevant central bank interest rate then the question is how this will affect a permit market in a multi-national setting with agents using different currencies in countries with differing monetary regimes. Our model assumes a single uniformly applicable discount rate for all agents as if all operated within a single state. EUAs are denominated and traded in euros and most of the agents in the EU ETS operate in countries where the currency is the euro and where the discount rates used by agents might be assumed to reflect with varying premia the discount rate established by the European Central Bank. However, a sizeable minority of installations operate in countries where the euro is not the currency and under monetary regimes with different discount rates. While agents in these countries may buy and sell EUAs denominated in euros, their accounting is ultimately in their own currency mediated by an exchange rate and perhaps applying discount rates different from those of their counterparts in member states using the euro. In our simulations, a range of discount rates emerging from the relatively short-term pricing of EUA futures is used and this might be viewed as the expression of the market aggregation of these potentially heterogeneous discount rates, but our choice has no rationale other than that it can be directly connected to market pricing of EUAs in different time periods. As noted, this discount rate bears little relation to the interest rate set by the European Central Bank during Phase II. This discussion shows the need to carefully think about how agents value the present relative to the future and how this may affect the permit market. Heterogeneity regarding other attributes of firms may exist, e.g. in terms of risk aversion.

¹⁵Of course, this discussion assumes that the forward-looking model presented in this paper is a fair representation of actual agent behavior.

7 Conclusion

Allowance banking has been a neglected subject on the research agenda concerning the EU ETS. Our hope is that our analysis will convince researchers interested in the EU ETS, or more generally in carbon markets, that the topic belongs on that agenda. As the preceding comments indicate, there are plenty of open questions. What should not be open, however, is whether allowance banking must be considered in explaining observed phenomena in the EU ETS and in particular the stock of unused allowances that has been accumulated in the course of Phase II and into Phase III. For too long, the facile explanation of « over-allocation » has been used when the reality is more complicated and involves economic choices by optimizing agents, rather than purely a failure of administrative systems, as the term over-allocation may suggest. One key lesson of this analysis of allowance banking in the EU ETS is that it is rational to decrease emissions below the cap at the start of the banking period to minimize abatement costs over time. The observed EUA bank at the end of Phase II falls within the range of values indicated by the illustrative simulations presented in this paper suggesting behavior by agents consistent with intertemporal cost minimization in a perfect-foresight model.

This is good news for it reveals a form of voluntary early action triggered by the particular structure of the cap in the EU ETS, namely, one that starts out near or at the level of initial business-as-usual emissions and declines steadily thereafter. This structure can be found in one form or another in other proposed and implemented CO₂ emissions trading systems and it seems likely to characterize future greenhouse gas trading systems, given the nature of the problem being addressed and the available technology. The logic of allowance banking would suggest that when banking is allowed and agents are faced with a credible prospect of future scarcity, they will reduce emissions initially more than required in order to capture the gains that come from intertemporal cost minimization.

As discussed previously, the impact of bringing the model closer to reality, e.g. by considering uncertainty, shorter planning horizons or allowing for firm heterogeneity, on the optimal paths of emissions and price need to be explored. It is our hope that this paper will stimulate further work, both theoretical and empirical, on banking in the context of permit markets featuring a smoothly declining cap.

References

- [1] Cronshaw, M. B. and, J. B. Kruse (1996), "Regulated Firms in Pollution Permit Markets with Banking," *Journal of Regulatory Economics*, 9, 179-189.
- [2] Ellerman, A. D. and J. P. Montero (2007), "The Efficiency and Robustness of Allowance Banking in the US Acid Rain Program," *The Energy Journal*, 47-71.
- [3] Ellerman, A. D., F. J. Convery and C. de Perthuis (2010), "Pricing Carbon: The European Union Emissions Trading Scheme", Cambridge University Press.
- [4] Ellerman, A. D., C. Marcantonini and A. Zaklan (2014), "The EU ETS: Eight Years and Counting," Robert Schuman Centre for Advanced Studies Working Paper 2014-04, European University Institute.
- [5] Kling, C. and J. Rubin (1997), "Bankable Permits for the Control of Environmental Pollution," *Journal of Public Economics*, 64, 101-115.
- [6] Leiby, P. and J. Rubin (2001), "Intertemporal Permit Trading for the Control of Greenhouse Gas Emissions," *Environmental and Resource Economics*, 19, 229-256.
- [7] Rubin, J. (1996), "A Model of Intertemporal Emission Trading, Banking, and Borrowing," *Journal of Environmental Economics and Management*, 31, 269-286.
- [8] Rubin, J. and C. Kling (1993), "An Emission Saved is an Emission Earned: An Empirical Analysis of Emission Banking for Light-Duty Engine Manufacturers," *Journal of Environmental Economics and Management*, 25, 257-274.
- [9] Schennach, S.M. (2000), "The Economics of Pollution Permit Banking in the Context of Title IV of the 1990 Clean Air Act Amendments," *Journal of Environmental Economics and Management*, 40, 189-210.
- [10] Stevens, B. and A. Rose (2002), "A Dynamic Analysis of the Marketable Permits Approach to Global Warming Policy: A Comparison of Spatial and Temporal Flexibility," *Journal of Environmental Economics and Management*, 44, 45-69.
- [11] Yates, A. J. and M. B. Cronshaw (2001), "Pollution Permit Markets with Intertemporal Trading and Asymmetric Information," *Journal of Environmental Economics and Management*, 42, 104-118.

Appendix: Analytical Solution

In this appendix we provide an analytical solution for a simple case, to facilitate the analysis and to obtain tractable results. The marginal cost mc is assumed to be a linear function. As Schennach (2000), we rewrite mc as:

$$mc(u_t - e_t) = (\tilde{u}_t - e_t)B$$

where the variable $\tilde{u}_t = A_t/B + u_t$. The paths of emissions and price are such that:

$$P_t = \begin{cases} (\tilde{u}_0 - e_0) B e^{rt} & \text{for } t < \tau \\ (\tilde{u}_t - Y_t) B & \text{for } t \geq \tau \end{cases}$$

$$e_t = \begin{cases} \tilde{u}_t - e^{rt}(\tilde{u}_0 - e_0) & \text{for } t < \tau \\ Y_t & \text{for } t \geq \tau \end{cases}$$

The equation (7) can be rewritten as:

$$\tilde{u}_\tau - e^{r\tau}(\tilde{u}_0 - e_0) = Y_\tau$$

and the equation (8) as:

$$\tau\tilde{u} + (\tilde{u}_0 - e_0)\frac{(1 - e^{r\tau})}{r} = \frac{Y_0}{a}(1 - e^{-a\tau})$$

Putting equation (7) into (8), we obtain:

$$\tilde{u}_\tau - \frac{(1 - e^{-r\tau})}{r}(\tilde{u} - Y_\tau) = \frac{Y_0}{a}(1 - e^{-a\tau})$$

We can then substitute for e_0 and τ in the path of price and emissions.