Hedging and the Temporal Permit Issuance in Cap-and-Trade Programs: The Market Stability Reserve under Risk Aversion

Oliver Tietjen[a,*], Kai Lessmann[b], Michael Pahle[b]

[a]*Potsdam Institute for Climate Impact Research (PIK), Technical University Berlin*
[b]*Potsdam Institute for Climate Impact Research (PIK)*

Abstract

Cap-and-trade programs expose firms to considerable risks due to the often politically induced permit price variability. We develop an intertemporal equilibrium model to analyze the implications of risk aversion by regulated firms. We show that the resulting risk premium significantly shapes the permit price path. The size of the premium depends on the endogenously time-varying hedging demand for and availability of permits. We apply the model to the European Union's Emission Trading System and offer an explanation for the price hike after the recent reform. Shifting permits to the future, as with the introduced Market Stability Reserve, increases the hedging value of permits. Yet, this comes with a lower growth rate and thus the price increase may not be sustainable. The hedging demand also implies that firms want to bank more permits. Therefore, we find a stronger impact of the Market Stability Reserve, e.g. in terms of permit cancellations, than previous analyses suggest.

Keywords: cap-and-trade, emission trading system, risk aversion, hedging, EU ETS, Market Stability Reserve

JEL codes: D25, H23, Q02, Q54, Q58.

---

[*]Corresponding author. E-mail: oliver.tietjen@pik-potsdam.de

July 11, 2019

1. Introduction

According to standard theory, cap-and-trade programs lead to a cost-effective achievement of a given emission cap. Permit trading between regulated firms implies static efficiency (Montgomery 1972) and the possibility to bank permits implies dynamic efficiency (Cronshaw and Kruse 1996; Rubin 1996). The latter further entails that the permit price should rise at the discount rate due to intertemporal arbitrage. Yet, real world programs such as the European Union Emissions Trading System (EU ETS) substantially deviate from the price paths suggested by theory: the EU ETS price fell from 30 EUR/t in mid 2008 to about 5 EUR/t until mid 2017 and has risen since then to about 25 EUR/t until mid 2019.

Several reasons have been put forth for the price drop after 2008 (Koch et al. 2014; Ellerman et al. 2016; Fuss et al. 2018): the financial crisis and corresponding lower economic growth rates (reduced baseline emissions), companion policies (e.g. faster expansion of emission-free renewable energies) and the use of international offset credits as part of the Kyoto Protocol led to lower abatement costs within the EU ETS. Empirical papers only explain a relative small part of the price movements indicating that other reasons affect prices as well (see Hintermann et al. 2016 and Friedrich et al. 2018 for literature reviews). A related view is thus that market and regulatory failures distort the EU ETS, leading to inefficiently low prices. The literature points to myopia, regulatory uncertainty and excessive discounting as potential distortions for intertemporal efficiency (see Fuss et al. (2018) for an overview).

In this paper, we take a similar view and focus on the potential distortions from hedging by risk averse firms when markets for risk are incomplete. Specifically, we consider the case in which firms bank permits to reduce their profit risk exposure (i.e. hedging). We show that such hedging with permits can have strong effects on the permit price path since prices incorporate an additional hedging value. Based on this we analyze the recent reform of the EU ETS, namely the introduction of the Market Stability Reserve (MSR). The MSR was implemented because according to the EU, the high "supply-demand imbalance" leads to a surplus (i.e. large permit bank) that destabilizes the market (European

Parliament and Council 2015). Hence if the bank is not within the "hedging corridor" (Neuhoff et al. 2012), the argument goes, price signals are distorted. As response the MSR reduces the permit bank by issuing less permits, that go into the reserve instead, if the bank exceeds the hedging corridor. If the bank is below the corridor, permits are released from the reserve again. While the original MSR was designed to be cap-neutral, the reformed MSR may reduce the cap through cancellation of permits (European Parliament and Council 2018). With these measures the EU aims for higher permit prices, lower price variability and more low carbon investments (European Commission 2014; European Parliament and Council 2018). We analyze how the shifting of permits to the future (original MSR) and the additional cancellation of permits (new MSR) contribute to these targets when firms want to hedge profits.

But why do firms hedge their profits to begin with? While there is no need for hedging when markets are complete[1], the corporate hedging literature provides a list of exceptions. For example, market imperfections arising from costs associated with financial distress, principal-agent incentive problems or progressive tax rates.[2] Here, the Modigliani-Miller theorem does not hold, and thus firm value may increase due to hedging (Bessembinder 1991; Froot et al. 1993). For the EU ETS, a survey among market participants suggests that hedging is indeed the most important motive for trading (KfW and ZEW 2016) and interviews conducted by Schopp and Neuhoff (2013) find that electricity producers hold permits for hedging profits several years ahead. Hintermann (2012) derives and estimates an options pricing formula for EU ETS permits and finds empirically that permits have a hedging value. The inability of the regulator to commit to a long-term cap, frequent policy interventions and the resulting large uncertainty about the future price development (Koch et al. 2016) may give even stronger hedging incentives compared to other commodity markets.

While the hedging demand for permits may be innocuous if markets for risk are

[1]With complete markets firms should behave risk neutral since risk is borne by shareholders who diversify their income streams to reduce risks (Diamond 1967).

[2]For example, Cronqvist et al. (2012) and Cain and McKeon (2016) show that the risk-taking decisions of firms are affected by managers risk preferences.

complete, we argue that they are not complete in realty. According to Staum (2007) incompleteness arises because of (1) missing markets due to a lack of securities that span all risks that firms want to hedge, (2) market frictions (e.g. capital constraints) and (3) ambiguity (unknown stochastic model). Our work is motivated by the first two reasons, see Quemin (2017) for an analysis of ambiguity aversion in a cap-and-trade program. While permits in the EU ETS are certainly a frequently traded asset, a liquid derivative market is available at most for the next six years. Furthermore, capital market frictions may restrict the risk-taking capacity of financial traders, who typically are the counterparts of producers, leading to "limits to arbitrage" (Shleifer and Vishny 1997). Consequently, producers also face limits to hedging (Acharya et al. 2013). Put differently, since speculators cannot absorb the whole risk, a risk premium remains which has to be paid by producers who want to offload their risk. In cap-and-trade markets as the EU ETS this problem is amplified because of the large regulatory risks. Such risks are by their very nature difficult to assess, which arguably increases the required compensations (risk premiums) by financial counter parties. Several papers find empirical evidence for such risk premiums in different commodity markets (e.g. Acharya et al. 2013; Bessembinder and Lemmon 2002; Hamilton and Wu 2014) and in particular also in the EU ETS (Chevallier 2010; Pinho and Madaleno 2011; Chevallier 2013; Kamga and Schlepper 2015; Trück and Weron 2016).

Against that background, we develop a stochastic intertemporal model where heterogeneous firms, regulated by a cap-and-trade program, produce a homogeneous good (electricity) using either a relative clean (gas) or dirty (coal) technology. They have to invest in capacity in order to produce the good, which happens with a time lag. We show that such capacity constraints amplify the impact of hedging since they increase the permit price variability. Our setup is motivated electricity markets that are typically covered by existing cap-and-trade programs as in the EU ETS. We model uncertainty as the regulator's inability to commit to the cap. Specifically, the regulator may adapt its future supply of permits in each period.

The permit price in our model has an endogenous risk premium component, which is a function of the hedging demand for permits, the permit price variability and the permit bank. This gives rise to a distinct temporal permit price profile: Initially, the dominating hedging demand of dirty coal firms creates a negative risk premium such that the expected price may not grow or even fall. Over time, the market becomes cleaner, implying a declining hedging demand of dirty firms and, in addition, firms build up a permit bank which reduces risks. In consequence, the (negative) risk premium declines and may turn positive. Yet, the price path strongly depends on the permits available for hedging purpose, which in turn depends on the time plan (schedule) of issuing permits by the regulator. In our simulation of the EU ETS we find a declining price in early years especially if the impact of the MSR is considered, resulting in a U-shaped price path. Therefore, the recently observed price hike in the EU ETS may not be sustainable.

The permit schedule has also been analyzed in the context of the original MSR without cancellation of permits. A main result of these studies, which stands in contrast to our findings, is that the temporal issuance is irrelevant as long as the overall cap remains unchanged and banking and borrowing constraints do not bind (e.g. Salant 2016). Perino and Willner (2016) accordingly find that a cap-neutral MSR only lifts the (short-term) permit price if the borrowing constraint binds earlier due to the MSR. Since long-term prices are lower they also conclude that low carbon investments with long lead time may decline (see also Perino and Willner 2017b). We find that investments into relative clean gas capacities are hardly affected and investments in coal capacity significantly decline in the short-term and are higher in the long-term because of the MSR. This can be traced back to worse hedging conditions for dirty capacities in early years and price level effects related to the risk premium.

Moreover, Perino and Willner (2016), Kollenberg and Taschini (2019) and Richstein et al. (2015) find that the MSR increase price variability, though Fell (2016) and Quemin and Trotignon (2019) find the opposite. Kollenberg and Taschini (2019) go a step further and relate the price variability positively to the risk premium for banking permits. In consequence, the MSR may even lead to lower prices in the short-term since firms want to use more permits early on due to the higher discount rate. Our approach differs from this work by deriving an endogenous (time-dependent) risk premium rather than assuming a positive relationship between price variability and risk premium. In doing so, we find different results: even the cap-neutral MSR significantly rises short-term prices because the hedging value of permits increases. This is because the risk premium becomes smaller (or more negative) reflecting that firms require a lower return for holding permits due to the hedging value. Hedging in the context of the EU ETS and MSR is also analyzed by Schopp and Neuhoff (2013) and Schopp et al. (2015). Their approach has been criticized for not explicitly accounting for risk and inconsistent price jumps (Salant 2016). We overcome these drawbacks by explicitly including regulatory risk and risk aversion.

Several papers[3] analyze the new MSR including the cancellation of permits. However, all except Quemin and Trotignon (2019) assume given discount rates and neglect the impact of uncertainty. Quemin and Trotignon (2019) analyze the impact of myopia and limited sophistication of firms in understanding the impacts of the MSR. They find that the MSR is never cap-neutral, even without cancellations. While myopia in a sense increases the applied discount rate, the hedging demand in our model may also reduce the discount rate. Indeed, we find that risk premiums are always negative in our simulation of the EU ETS. Therefore, we also find a relative higher number of MSR cancellations compared to other papers because with lower discount rates the permit bank and this influx into the MSR is larger in early years.

The remainder of this paper is structured as follows: After presenting the general model setup in section 2.1 we derive formal results in a simplified two-period version of the model in section 2.2, while we focus on a cap-neutral stylized shift of permits to the future. In section 3, the model is numerically applied to the EU ETS for multiple periods, while we explicitly take the original and new MSR mechanic into account. Finally, we discuss the results and conclude in section 4.

[3] Beck and Kruse-Andersen (2018), Bocklet et al. (2019), Bruninx et al. (2018), Carlén et al. (2019), Quemin and Trotignon (2019), Perino and Willner (2017a), Rosendahl (2019) and Silbye and Sørensen (2018)

2. The Model

We assume heterogeneous (dirty and clean) firms competing in a goods market, which is regulated by a cap-and-trade program. Since the regulator cannot fully commit to future targets, the amount of permits issued in the future is uncertain and therefore also the firms' profits. This creates a demand for hedging profits when firms are risk averse. The model is described in detail in the following.

2.1. General Model Setup

We consider I competitive firms, indexed i, that produce a homogeneous and non-storable good (best thought of as electricity) x_{it} at T dates, indexed by t. Demand is given by $D\left(w_t\right)$ with price w_t for which holds $D^{'} < 0$. The equilibrium condition

$$\sum_i^I x_{it} = D\left(w_t\right). \tag{2.1}$$

is always fulfilled. Firms are heterogeneous with respect to the production technology which implies a different hedging demand for permits as we see below. The production costs, given by the function $C_{Xi}\left(x_{it}\right)$ with $C_{Xi}^{'} > 0$, therefore varies between the firms. In order to produce x_{it} units, firms also need at least k_{it} units of capacity, for which capacity costs are given by $C_{Ki}\left(k_{it}\right)$ with $C_{Ki}^{'} > 0$. Defining $\zeta_{it} \equiv \frac{x_{it}}{k_{it}}$ as the utilization rate of the capacity, production is thus constrained by

$$1 \geq \zeta_{it} \geq 0. \tag{2.2}$$

While the utilization rates can be immediately adjusted within a date, investment decisions in capacity (e.g. coal or gas plants), $I_{Kit} \geq 0$, are adding to the existing capacity stock with a lag of one period,

$$k_{it} = \left(1-\delta\right) k_{it-1} + I_{Kit-1} \tag{2.3}$$

where δ is a depreciation rate.

Moreover, the production of each unit of x_{it} causes ϕ_i units of emission. Due to the use of different technologies, emission factors ϕ_i also vary between firms. This is important for our analysis, because permit supply uncertainty affects dirty (coal) firms (high ϕ_i) differently than (relative) clean (gas) firms (low ϕ_i), which has significant implications for the hedging behavior and its consequences as shown below. Overall emissions are capped since the goods market is regulated by a cap-and-trade program at all dates. Compliance requires that at the end of each date, firms need as least as many permits y_{it} as emissions $x_{it}\phi_i$. At the beginning of each period t , the number of S_t permits are auctioned[4] by a regulator at price p_t such that in equilibrium it holds that

$$S_t = \sum_i^I y_{it}. \tag{2.4}$$

The regulator also announces the permit supply for all future periods τ, $S_\tau \; \forall \; \tau > t$ at the beginning of each period. However, as mentioned we assume that the regulator is unable to commit to her announcement such that the actual supply may deviate while uncertainty resolves at the beginning of each date. At the beginning of any period t, total permits (in addition to already issued permits) expected to be auctioned in this and future periods are:

$$E_t\left[\bar{S}\right] = S_t + \sum_{\tau>t}^T E\left[S_\tau\right] \tag{2.5}$$

Since in this and the following section we are interested in the effects of the (expected) temporal permit issuance for a given (expected) amount of permits, we define a *cap-neutral temporal reallocation* as

$$\Delta E_1\left[\bar{S}\right] = \Delta S_1 + \sum_{\tau>1}^T \Delta E\left[S_\tau\right] = 0. \tag{2.6}$$

That is, any change in permit supply in any period is fully compensated by the (announced) supply in other periods such that the total expected cap is always the

[4]Throughout the paper we assume that the initial allocation of permits is through auctioning, i.e. there is no free allocation. While the allocation method affects results if firms are risk averse (Baldursson and von der Fehr (2004; 2012)), we leave this out for future research and concentrate on auctioning. In this paper we focus on the EU ETS and in particular on the electricity sector therein, where in principle all permits are auctioned (European Parliament and Council 2018).

same from the perspective of date 1. This corresponds to the original cap-neutral MSR (cancellation is introduced in section 3).

If firms hold more permits than needed, additional permits can be transferred to the next date (banking), while borrowing from the future is not allowed,

$$b_{it} = b_{it-1} + y_{it} - k_{it}\zeta_{it}\phi_i \tag{2.7}$$

$$b_{it} \geq 0. \tag{2.8}$$

with b_{it} as the banked permits at the end of date t. Besides investing in capacity, k_{it}, and permit, b_{it}, stocks, firms also hold a risk-free asset stock l_{it}, providing a save return r. The latter serves as alternative investment opportunity, allowing for a risk-free allocation of wealth over time. Denoting investments in the risk free asset as I_{Lit}, the risk free asset stock is

$$l_{it} = (1 + r)\, l_{it-1} + I_{Lit}. \tag{2.9}$$

Given this setup the firms' profits at date t are

$$\pi_{it} = w_t k_{it}\zeta_{it} - C_{Xt}\left(k_{it}\zeta_{it}\right) - C_{Kt}\left(k_{it}\right) - p_t y_{it} - I_{Lit}. \tag{2.10}$$

where the first term is the revenues of selling the good, the second and the third terms are costs for producing and for plant capacities[5] and the forth and fifth terms are costs (> 0) or revenues (< 0) for trading permits and the risk-free asset.

We assume that firms have concave preferences about their profits that can be described by a von Neumann–Morgenstern utility function $U_{it}\left(\pi_{it}\right)$ with $U_{it}' > 0$ and $U_{it}'' < 0$. This reflects market imperfections as for example, costs associated with financial distress or principal agent issues that result in higher utility from a more stable profit, which let

[5]Note that we assume for simplicity that there are no costs for investing in plant capacity I_{Kit} but only capacity costs. That is investment costs are implicitly allocated to the capacity costs.

firms behave risk averse. The problem of the firms is

$$\max_{\zeta_{it},y_{it},I_{Kit},I_{Lit}} \sum_{t=1}^{T} \frac{1}{(1+r)^{t-1}} E\left[U_{it}\left(\pi_{it}\right)\right] \tag{2.11}$$

subject to

$$1 \geq \zeta_{it} \geq 0 \quad I_{Kit} \geq 0 \quad b_{it} \geq 0$$

$$k_{it} = (1-\delta)\, k_{it-1} + I_{Kit-1} \tag{2.12}$$

$$b_{it} = b_{it-1} + y_{it} - k_{it}\zeta_{it}\phi_i$$

$$l_{it} = (1+r)\, l_{it-1} + I_{Lit}$$

For the analysis below it is convenient to rewrite the profit by using the intertemporal banking condition (2.7),

$$\pi_{it} = \pi_{it}^{plant} + p_t\left(b_{it-1} - b_{it}\right) - I_{Lit} \tag{2.13}$$

with $\pi_{it}^{plant} = w_t k_{it}\zeta_{it} - C_{Xt}\left(k_{it}\zeta_{it}\right) - C_{Kt}\left(k_{it}\right) - p_t k_{it}\zeta_{it}\phi_i$. Hence we decompose the profit (or losses) into plant profits, π_{it}^{plant}, profits made from holding excess permits, $p_t\left(b_{it-1} - b_{it}\right)$ and profits due to the risk-free asset, I_{Lit}.

2.2. Two-Period Model

In order to derive analytical results we solve the model for two periods, $t = 1, 2$ in this section. In addition, we make the following assumptions: the goods demand is linear, $D\left(w_t\right) = A - aw_t$. There are only two firms, $i = c, d$ a clean and a dirty firm with $\phi_d > \phi_c$ and the production and capacity cost functions are given by $C_{Xt}\left(x_{it}\right) = \frac{\beta_i}{2}\left(\zeta_{it}k_{i,t}\right)^2$ and $C_{Kt}\left(k_{it}\right) = c_i k_{i,t}$, respectively. The firms' costs functions are again motivated by the electricity sector in which capacity costs typically exhibit constant marginal costs per unit of capacity and marginal costs increase with production. Moreover, to arrive at closed-form results we assume a quadratic utility function in some cases, $U_{it}\left(\pi_{it}\right) = \pi_{it} - \pi_{it}^2$. For the numerical application to the EU ETS in section 3, we extend the model to multiple periods and show that the results also hold for utility exhibiting constant relative risk aversion.

2.2.1. Date 2 Equilibrium

We solve the model backwards and start at date 2. Note that all derivations can be found in Appendix A.

At date 2, there are no further investments in capacity k_{it} and all available permits are used or sold (assuming a strictly positive permit price p_2) and, similar, the stock of the risk-free asset is depleted, implying $b_{i,2} = l_{i,2} = 0$. Uncertainty has resolved and the firms' problem is thus to maximize $U_{i,2}(\pi_{i,2})$ over $\zeta_{i,2}$ and $y_{i,2}$ subject to the constraints in (2.12). Taking the first order conditions (see Appendix A.1), the utilization rate and the permits purchases or sales can be written as

$$\zeta_{i,2} = \frac{w_2 - p_2\phi_i}{\beta_i k_{i,2}} - \frac{\mu_{i,2}}{U'_{i,2}\beta_i k_{i,2}^2} \tag{2.14}$$

$$y_{i,2} = \phi_i\left(\frac{w_2 - p_2\phi_i}{\beta_i} - \frac{\mu_{i,2}}{U'_{i,2}\beta_i k_{i,2}}\right) - b_{i,1} \tag{2.15}$$

where $\mu_{i,2}$ is the shadow value of the capacity constraint which is positive if the capacity is fully utilized, $\zeta_{i,2} = 1$, and zero otherwise:

$$\mu_{i,2} = \begin{cases} U'_{i,2}\left(k_{i,2}w_2 - \beta_i k_{i,2}^2 - p_2\phi_i k_{i,2}\right) & if\ \zeta_{i,2} = 1 \\ 0 & if\ 1 \geq \zeta_{i,2} \geq 0 \end{cases}. \tag{2.16}$$

It indicates scarcity of capacity k_{it}, which cannot be increased within a date due to the time lag for investments. For (2.15) we assume that the cap is always binding and therefore there is always a positive permit price p_2. Note that risk aversion, reflected by the marginal utility $U'_{i,2}$, has no effect at date 2 (in (2.14) either $\mu_{i,2} = 0$ or $U'_{i,2}$ cancels out due to (2.16)). It only adjusts the shadow value of the capacity which, however, triggers no changes in the firm behavior because they cannot change their capacity level.

By making use of the equilibrium condition of the goods market, $\sum_i^I k_{i,2}\zeta_{i,2} = D_2 = A - aw_2$, the goods price reads:

$$w_2 = \frac{1}{(\beta_d + \beta_c + \beta_c\beta_d a)}\left(A\beta_c\beta_d + p_2\left(\beta_d\phi_c + \beta_c\phi_d\right) + \beta_d\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}} + \beta_c\frac{\mu_{d,2}}{U'_{d,2}k_{d,2}}\right). \quad (2.17)$$

Similarly, the permit price can be derived from using (2.15) in the permit equilibrium condition, $S_t = \sum_i^I y_{it}$, and by additionally considering (2.17) we get:

$$\begin{aligned} p_2 \;&=\; \frac{A\left(\beta_d\phi_c + \beta_c\phi_d\right) - \left(\beta_d + \beta_c + \beta_c\beta_d a\right)\left(b_{c,1} + b_{d,1} + S_2\right)}{\left(\phi_c - \phi_d\right)^2 + a\left(\beta_c\phi_d^2 + \beta_d\phi_c^2\right)} \\ &+\; \frac{\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}}\left(\phi_d - \phi_c\left(1 + \beta_d a\right)\right) + \frac{\mu_{d,2}}{U'_{d,2}k_{d,2}}\left(\phi_c - \phi_d\left(1 + \beta_c a\right)\right)}{\left(\phi_c - \phi_d\right)^2 + a\left(\beta_c\phi_d^2 + \beta_d\phi_c^2\right)}. \end{aligned} \quad (2.18)$$

Intuitively, the goods price is a positive function of demand, reflected by A (the intercept of the demand function), and the permit price p_2. Shocks on the permit price are thus transferred to the consumers via the goods price. The only source of uncertainty (form the perspective of date 1) is the supply of permits at date 2 S_2. Ignoring the capacity constraints (reflected by the last two terms in (2.17) and (2.18)), a positive shock on S_2 (less ambitious policy) leads to a lower permit price and vice versa as can directly be seen from (2.18). Concerning the utilization rates, permit price shocks have the following effects.

Lemma 1. *If the capacity constraints do not bind, $\mu_{i,2} = 0$, a positive permit price shock leads to (1) a higher capacity utilization by the clean firm, $\frac{d\zeta_{c,2}}{dS_2} > 0$, if $\phi_d > \phi_c\left(1 + \beta_d a\right)$ holds and (2) a lower capacity utilization by the dirty firm, $\frac{d\zeta_{d,2}}{dS_2} < 0$. For a negative permit price shock, the contrary holds.*

While the dirty firm always produces more if there are more permits and vice versa, for the clean firm it depends on parameters. Specifically, the condition $\phi_d > \phi_c\left(1 + \beta_d a\right)$ implies that if the demand reaction to price changes in the goods market is strong enough, reflected by a high a, or the clean firm is not clean enough (i.e. ϕ_c is too large), it produces more if the permit price is low like the dirty firm. However, we consider the case in which $\phi_d > \phi_c\left(1 + \beta_d a\right)$ holds and thus the clean firm increases production as soon as the permit price increases. In electricity markets, for example, there is a fuel switch: a higher permit price leads to less coal (dirty) and more gas (clean) production. Note that for lemma 1 we also assume that the capacity constraints do not bind. Of course, if the capacity constraints bind, firms cannot increase their production after the shock. However, typically capacity constraints do not bind in expectation. Power plants, for example, are not always fully utilized because demand varies on a short time scale such that most of the time large parts of the capacity is not (fully) utilized. Hence the utilization rate in our model should be interpreted as a long-term (e.g. annual) utilization rate.

For the analysis of hedging with permits the relationship between plant profits and permit price is important.

Lemma 2. *If the capacity constraints do not bind, $\mu_{i,2} = 0$, a positive permit price shock leads to (1) higher plant profits for the clean firm and thus $Cov\left[\pi_{c,2}^{plant}, p_2\right] > 0$, if condition $\phi_d > \phi_c(1 + \beta_d a)$ holds and (2) lower plant profits for the dirty firm and thus $Cov\left[\pi_{d,2}^{plant}, p_2\right] < 0$. For a negative shock, the contrary holds.*

If condition $\phi_d > \phi_c(1 + \beta_d a)$ is fulfilled, and thus the clean firm increases its production level after a positive permit price shock, it also gains higher plant profits. The dirty firm produces less (lemma 1) and has higher costs and therefore, it always loses from higher ETS prices. While we ignore capacity constraints for lemma 1 and lemma 2 because they do not change the nature of the results, they have an important impact on the price variability.

Lemma 3. *If production is constrained by capacity, price variability is higher.*

Intuitively, a higher price variability also increases the profit variability and therefore capacity constraints amplify the effect of hedging. The intuition of lemma 3 is that capacities partly lock-in the production levels. This implies that firms have less flexibility to react to shocks. For instance, after a negative permit supply shock, the production of the clean firm increases less if capacity constraints bind. In order to comply with the cap, the permit price thus must rise to a higher level than without capacity constraints, because abatement is achieved with more expansive technologies (i.e. via a lower goods demand in our model).

2.2.2. Date 1 Equilibrium

At date 1, firms have to make decisions under uncertainty of the permit supply. Uncertainty in the permit supply also creates uncertainty of the permit and goods prices and therefore utilization rates are also uncertain. Firms maximize utility in (2.11) for $T = 2$ and subject to (2.12). Due to the time lag for investments in capacity, we further assume for simplicity that there are sufficient initial capacities $k_{i,1}$ such that the capacity constraints do not bind at date 1.

Note that for a risk neutral firm that maximizes expected profits the optimality conditions are the same as in the risk averse case but with constant marginal utility, i.e. $U'_{i,1} = E\left[U'_{i,2}\right] = 1$. The risk neutral case only serves as a benchmark to analyze the effects of market imperfections due to risk aversion and incomplete markets for risks.

While the capacity utilization rate $\zeta_{i,1}$ and the permits trades $y_{i,1}$ must fulfill the same condition as at date 2, firms additionally decide about the optimal permit bank level $b_{i,1}$, capacity level for date 2 $k_{i,2}$ and the amount invested in the risk-free asset $l_{i,1}$ (see Appendix A.1 for first order conditions). First, we analyze how firms hedge with banking permits while we ignore capacity effects. Second, we show the impact of hedging on the (expected) price levels and dynamics again without capacity effects. Thereafter, we discuss the capacity effects.

Banking and Hedging . In contrast to date 2, firms also decide about their bank level with their permit purchases. The number of permits they buy is equal to their date 1 emissions plus the desired bank at the end of date 1:

$$y_{i,1} = \phi_i \zeta_{i,1} k_{i,1} + b_{i,1} \tag{2.19}$$

where the banking demand can be written as follows:

$$b_{i,1} = \frac{E\left[p_2\right] - p_1\left(1+r\right)}{\lambda_i Var\left[p_2\right]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var\left[p_2\right]} - \frac{\left(1+r\right)\varphi_{i,1}}{U'_{i,1} Var\left[p_2\right]} \tag{2.20}$$

for which we assume quadratic utility and with $\lambda_i = -\frac{U''_{i,1}}{U'_{i,1}}$ as the coefficient for absolute risk aversion. The third term on the right-hand side in (2.20) includes the shadow price of the borrowing constraint $\varphi_{i,1}$ (due to inequality (2.8)), which is positive if firms want to borrow but cannot and zero otherwise. The first term reflects the intertemporal arbitrage or speculation motive. If the expected discounted price exceeds today's price $E[p_2] - p_1(1+r) > 0$, firms want hold a positive bank for pure speculation reasons and vice versa. The second term is the hedging demand, determined by the covariance of plant profits with the date 2 permit price. It reflects the (positive or negative) amount of permits that firms want to bank in order to reduce their risk exposure. For this hedging demand we have the following proposition.

Proposition 1. *For pure hedging purpose, the dirty firm wants to hold a positive amount of permits $b_{d,1} > 0$ (banking) and the clean firm a negative amount, $b_{c,1} < 0$ (borrowing). Since borrowing is not allowed, the clean firm holds no permits, $b_{c,1} = 0$, or a positive amount of permits if $E[p_2] - p_1(1+r) > \lambda_i Cov\left[\pi_{c,2}^{plant}, p_2\right]$.*

Intuitively, dirty firms want to hold a long position in the permit market (i.e. banking) because they are short with respect to the permit price in the goods market and for clean firms the contrary holds (see lemma 2). This is reflected by the hedging demand, the second term in equation (2.20), which is positive for dirty firms because $Cov\left[\pi_{d,2}^{plant}, p_2\right] < 0$ and negative for clean firms because $Cov\left[\pi_{c,2}^{plant}, p_2\right] > 0$. However, since we assume that borrowing is not allowed, clean firms cannot hedge their goods market profits by trading permits. Only when the speculative demand exceeds the hedging demand, i.e. if $E[p_2] - p_1(1+r) > \lambda_i Cov\left[\pi_{i,2}^{plant}, p_2\right]$, clean firms also bank because the expected profit for banking compensates for the higher risk exposure due to banking.

Note that in the risk neutral model the individual bank volume is undetermined since only the total bank of all firms is relevant. In this case the incentives to bank are for all firms identical since they have no specific hedging demand and the speculation demand is the same for all firms due to constant marginal utility.

Price Effects. In this section, we analyze how hedging affects permit prices. The price dynamics can be decomposed into three parts,

$$\frac{E\left[p_2\right]-p_1}{p_1}=r+\frac{(1+r)\,\varphi_{i,1}}{p_1 E\left[U_{i,2}^{\prime}\right]}+q_1. \tag{2.21}$$

The first term is the risk-free rate r, which reflects the opportunity to invest in the alternative asset $l_{i,1}$. The second term is only present if the borrowing constraint binds. In this case, the shadow price is positive $\varphi_{i,1}>0$, and therefore (while ignoring q_1) the growth rate is lower than the interest rate r. This is a standard result in the deterministic or risk neutral case (Rubin 1996; Schennach 2000; Fell 2016). The third term q_1 is the risk premium at date $t=1$ which emerges endogenously due to the hedging demand of the firms. With a general utility function it is $q_1=-\frac{Cov\left[U_{i,2}^{\prime},p_2\right]}{E\left[U_{i,2}^{\prime}\right]p_1}$ and thus it depends on the risk preferences of the firms, reflected by the marginal utility $U_{i,t}^{\prime}$ and the relationship of the firm's marginal utility to the permit price, reflected by the covariance term. Assuming quadratic utility and considering the permit market clearing in equation (2.4), the equilibrium risk premium can be expressed as follows,

$$q_1=\frac{\Lambda}{p_1}\left(Cov\left[\pi_{d,2}^{plant},p_2\right]+Cov\left[\pi_{c,2}^{plant},p_2\right]+Var\left[p_2\right]B_1\right) \tag{2.22}$$

where $B_1=b_{d,1}+b_{c,1}$ is the total bank and Λ is a parameter that reflects the risk-taking capacity of the market. The risk-taking capacity if both firms bank is $\Lambda=\left(\lambda_d^{-1}+\lambda_c^{-1}\right)^{-1}$ and if only one firm banks it is $\Lambda=\lambda_i$ (recall that λ_i is the coefficient of absolute risk aversion). In case of risk neutrality the risk premium disappears, $\Lambda=0$. Besides the risk-taking capacity of the market, equation (2.22) further shows that the risk premium is a function of the total bank B_1, permit price variability, $Var\left[p_2\right]$, and the hedging demand of the firms, given by the covariance terms. The price variability has a positive effect on the risk premium because it increases the risk of permit banking and thus firms require a higher return for banking. Similarly, a higher overall bank in isolation increases the volume of risky permits for which firms require a larger risk premium. The hedging demand, in contrast, may have a positive or negative effect on the risk premium.

Clean firm's hedging demand increases and dirty firm's hedging demand decreases the risk premium, since $Cov\left[\pi_{c,2}^{plant}, p_2\right] > 0$ and $Cov\left[\pi_{d,2}^{plant}, p_2\right] < 0$ (see lemma 2).

However, recall that the clean firm only banks if the risk premium is positive (cp. proposition 1) and thus the sign of the risk premium only depends on the strength of dirty firm's hedging demand and the risk of banking permits $\left|Cov\left[\pi_{d,2}^{plant}, p_2\right]\right| \lesseqgtr Var\left[p_2\right] b_{d,1}$. If the former exceeds the latter, the risk premium is negative. This is because banking has the additional benefit of a lower risk exposure for dirty firms in this case. Therefore, they are willing to accept a lower return for banking permits (potentially even a negative one). In turn, if the permit price variability and the banked volume is too high such that $\left|Cov\left[\pi_{d,2}^{plant}, p_2\right]\right| < Var\left[p_2\right] b_{d,1}$ holds, the risk premium is positive and also the dirty firm requires a risk premium for holding permits.

Proposition 2. *The risk premium increases with the permit price variability* $Var\left[p_2\right]$*, and the hedging demand of the clean firm* $Cov\left[\pi_{c,2}^{plant}, p_2\right]$*. It is decreasing in the absolute value of the (generally negative) hedging demand of the dirty firm* $\left|Cov\left[\pi_{d,2}^{plant}, p_2\right]\right|$*. The sign of the risk premium is positive if* $\left|Cov\left[\pi_{d,2}^{plant}, p_2\right]\right| < Var\left[p_2\right] b_{d,1}$ *and vice versa.*

Note that in the absence of capacity constraints a positive risk premium always leads to a lower price and higher emissions at date 1, and a higher (expected) price and lower emissions at date 2. By rewriting (2.21) to $p_1 = \frac{E[p_2]}{(1+r+q_1)}$, it becomes obvious that the risk premium has the same effect as the risk-free rate and thus a positive risk premium increases the applied discount rate and leads to a steeper price path and vice versa.

Next we are interested in the effect of a cap-neutral date 1 permit reallocation in the sense of equation (2.6) on the permit price. By issuing more permits at date 1, rather than date 2, the regulator increases the permit bank at the end of date 1 and vice versa. By using the first order conditions we get the following relation between the permit prices, $p_1 = \frac{E\left[U'_{i,2} p_2\right]}{(1+r)U'_{i,1}}$. Taking the partial derivative with respect to the bank gives

$$\frac{\partial p_1}{\partial b_{i,1}} = \frac{E\left[U''_{i,2} p_2^2\right] U'_{i,1} + E\left[U'_{i,2} p_2\right] U''_{i,1} p_1}{U'^2_{i,1}} < 0 \tag{2.23}$$

because $U'_{i,1} > 0$ and $U''_{i,2} < 0$ due to the concavity of the utility function. Hence, if the bank volume increases, p_1 decreases. This is because firms require a larger return for holding more permits (higher risk premium, see (2.22)) which is achieved with a lower price at date 1. Intuitively, a lower permit price at date 1 leads to more emissions at date 1. If the total (expected) amount of permits is given, this implies that the expected emissions at date 2 must decrease and in turn the expected date 2 permit price must rise. We summarize this in the following proposition.

Proposition 3. *A temporal reallocation of permits by the regulator to date 1 in the sense of equation* (2.6) *such that the bank B_1 increases leads to a lower permit price and higher emissions at date 1 and a higher expected permit price and lower expected emissions at date 2 and vice versa.*

Hence the regulator has real production effects by deciding about the temporal schedule of permit issuance even if the borrowing constraint is not affected. The earlier the issuance of permits, the more risk of the permit budget is privatized. At a low level of available permits, dirty firms are willing to pay for holding a bank (negative risk premium) in order to reduce their risk exposure and if many permits are available they require a positive risk premium for holding permits. These hedging or risk costs are incorporated into the permit price, such that they emit less if less than desired permits for hedging are available and emit more if many permits are available.

These results hinge on the risk tolerance of the market as a higher tolerance decreases the risk premium (Colla et al. 2012). The risk tolerance can be increased by futures markets and speculators that trade permits or derivatives as we show in Appendix A.4. Yet, as we argue in the introduction it is implausible that speculators hold open positions in a risky asset as permits without compensation, especially if the risk is largely politically driven. Hence futures markets and speculators may reduce but not eliminate risk premiums.

Capacity Effects. In this section we look at the effect of capacities which we have ignored so far. Optimal capacity investments can be decomposed into three parts:

$$k_i = \frac{1}{c_i}\left(E\left[\zeta_{i,2}\right]E\left[\mu_{i,2}^{RN}\right] + Cov\left[\zeta_{i,2}, \mu_{i,2}^{RN}\right] + \frac{1}{U_{i,1}'}Cov\left[U_{i,2}', \mu_{i,2}^{RN}\right]\right) \qquad (2.24)$$

where $\mu_{i,2}^{RN}$ is the marginal capacity value in the risk neutral case (i.e. $\mu_{i,2}$ if $E\left[U_{i,2}'\right] = 1$, see equation (2.16)). The first two terms on the right hand side in (2.24) reflect optimal capacities when firms are risk neutral. Specifically, the effect of uncertainty in the risk neutral case compared to the deterministic case is given by $Cov\left[\zeta_{i,2}, \mu_{i,2}^{RN}\right]$. Since $Cov\left[\zeta_{i,2}, \mu_{i,2}^{RN}\right]$ is strictly positive, uncertainty has a positive impact on capacity investments, ceteris paribus. The intuition for this is that $\mu_{i,2}^{RN}$ reflects the scarcity of capacity. As such $\mu_{i,2}^{RN}$ is bounded at zero but has no upper bound. Hence capacity constraints imply an asymmetric impact of symmetric shocks if the shocks are large enough leading to higher expected profits reflected by a higher capacity value. All else the same, this leads to more investments in capacity.

The third term represents the effect of risk aversion. If firms do not bank permits $Cov\left[U_{i,2}', \mu_{i,2}^{RN}\right] \leq 0$ holds and thus risk aversion has a negative impact on investments, ceteris paribus. This is intuitive since capacity investments are risky and firms are risk averse. We summarize these results as follows.

Proposition 4. *When confronted with uncertainty about the permit price a risk neutral firms invest more in capacity compared to the deterministic case, ceteris paribus. Risk aversion has a negative impact on capacities if firms do not bank permits, ceteris paribus.*

In the next step, we combine the effects of hedging and capacity investments. It can be shown that the effect of banking permits on $Cov\left[U_{i,2}', \mu_{i,2}^{RN}\right]$ is positive for the dirty and negative for the clean firm and therefore capacity investments by the former increase and by the latter decrease, ceteris paribus. Intuitively, banking hedges dirty plant profits, but increases risk for clean firms and investment incentives change accordingly which we summarize as follows.

Proposition 5. *Banking has a positive effect on dirty capacities and a negative effect on clean capacities, ceteris paribus.*

Based on this, we can discuss the long-run (capacity) effects of a cap-neutral temporal reallocation of permits by the regulator in the sense of equation (2.6). Proposition 3 shows that if more permits are issued at date 2 rather than date 1, the short-term price (date 1) is higher and the expected long-term price is lower. Intuitively, a lower expected long-term price implies more investments in dirty capacities and less in clean capacities. However, due to proposition 5, we have an additional effect: since a reallocation of permits to date 2 decreases the bank volume of the firms, and this effect in isolation decreases dirty and increases clean capacities, the combined effect on dirty and clean capacities may be positive or negative, depending on the respective effects strengths.

3. Numerical Application to the EU ETS

In this section we apply the model of the previous sections to the EU ETS. In particular we focus on the effects of the Market Stability Reserve (MSR) which shifts permits into the future and cancels permits if too many permits are in the MSR. The MSR was introduced to increase (short-term) permit prices, decrease price variability and spur cleaner investments. Before we assess the impact of the MSR regarding these targets, we explain the model implementation and important assumptions in the following section.

3.1. Model Implementation

We solve the firms' problems given by equations (2.11) and (2.12) for $i = c, d$ a representative gas and coal firm. We focus on the time period between 2018 and 2057 (last time permits are issued) but solve the model until 2102 to reduce side effects due to our terminal horizon approach. The model explicitly considers only every fifth year such that we have $T = 17$ model periods, while we write $t = 2020, 2025, ..., 2100$ for every fifth period and $y = 2018, 2019, ...2102$ for annual periods. We believe modeling five year steps is only a mild limitation since we focus on mid- to long-term regulatory risks, while we are still able to solve a long time horizon with uncertainty and risk aversion.

At the beginning of the first year the regulator announces to issue $\hat{S}_y$ permits each year (for $\hat{S}_t$ we take the average of the corresponding five years). Specifically, $\hat{S}_y$ is in line with the (announced) permits to be auctioned in the EU ETS between 2018 and 2057, with a linear reduction factor of 1.74% until 2020 and 2.2% thereafter, which implies the last time permits are issued is in 2057. Note that we only consider the power sector for which permits are auctioned[6]. According to European Parliament and Council (2018) the auction share of all issued permits shall be 57% which we assume to be constant over the whole time horizon.

Regulatory uncertainty is reflected by an adjustment of the permit auction schedule every five years. While the amount of auctioned permits are known with certainty in the first period S_{2020}, it is uncertain thereafter, $S_t = \hat{S}_t + \theta_t$ where θ_t is the following shock process,

$$\theta_t = \theta_{t-1} + \epsilon_t \quad \forall\, 2045 \geq t \geq 2025 \tag{3.1}$$

with $\epsilon_t \in \{-0.35\hat{S}_t, 0.35\hat{S}_t\}$ which is either a positive or negative shock, while both having equal probability. The process implies that realized shocks fully persist (the autoregressive parameter is unity) and the magnitude of new shocks declines as they are a proportion (35%) of the initially announced permits ($\hat{S}_t$). We assume that the last shock emerges in period $t = 2045$ (2043-2047). While the choice of 35% is somewhat arbitrary, it leads to permit price range of about 18 EUR/t in 2025 and 42 EUR/t in 2030 which, we think, can be justified given the high uncertainty surrounding the EU ETS.

[7]

[6] Other sectors as the energy-intensive industry receive permits for free mainly based on grandfathering. The assessment of the hedging demand in these sector is difficult because firms operate in different markets and produce heterogenous goods. For simplictly, we exclude these sectors by assuming that the grandfarthered permits perfectly meet their permit demand which is so far roughly in line with reality (EEA 2018).

[7] In order to avoid a negative auction supply, we set potential negative auction values due to the shocks to zero. This implies that the in the first period overall cap of 24.3 Gt is higher than the initially announced cap of 22.1 Gt. Yet, this is in line with the often reported low credibility of the EU ETS since it can be interpreted as higher expected permit supply by firms than the announced by the regulator.

Moreover, the actual auctioned permits may further be adjusted by the MSR which we implement according to European Parliament and Council (2015; 2018). If the aggregate firm bank in the previous year B_{y-1} is larger than 0.833 Gt, a share γ_y of that bank is deducted from the auctioned permits in year y (if there are enough permits to be auctioned). The share is $\gamma_y = 0.24$ until $y = 2023$ and $\gamma_y = 0.12$ thereafter. Permits that are not auctioned due to this mechanism are denoted by M_y^{in} and go into the MSR bank denoted byM_y. If the banked permits in the previous year are lower than 0.4 Gt an amountM_y^{out} is released from the MSR bank and added to the auctioned permits. This amount is equal to 0.1 Gt (if there are enough permits in the MSR). If the bank in the previous year lies within the hedging corridor, $0.4 < B_{y-1} < 0.833$, the permit supply is not adjusted. The actual issued permits after the impact of the MSR S_y^M therefore reads

$$S_y^M = \begin{cases} S_y - \min\left(\gamma_y B_{y-1}; S_y\right) & if\ B_{y-1} > 0.833\, Gt \\ S_y + \min\left(0.1; M_{y-1}\right) & if\ B_{y-1} < 0.400\, Gt \\ S_y & otherwise, \end{cases} \tag{3.2}$$

and the MSR bank is given by

$$M_y = M_{y-1} + M_y^{in} - M_y^{out} - \max\left(M_y - S_{y-1}^M; 0\right) \tag{3.3}$$

The last term in (3.3) reflects the cancellation of permits which was introduced by the latest MSR reform. From 2023 onward, if there are more permits in the MSR than were auctioned in the previous year, these permits are invalidated, implying that the overall cap of the ETS is tightened. The MSR starts to operate in 2019 and is in the first year approximately filled with $M_{2019} = 1.525$ Gt permits.[8]. In order to avoid a positive MSR bank in the terminal model period, we assume that from 2058 onward the outtake of the MSR increases to 1 Gt.[9]

[8]Initially 0.9 Gt backloaded permits between 2014 and 2016 plus other unallocated permits which are estimated to be between 0.55 to 0.7 Gt (European Commission 2015) are placed in the MSR.

[9]Since the MSR rules are likely to be adjusted anyway and we focus on the time until 2055 we think this a mild assumption. Note that if cancelation is active, the MSR bank is zero at that time anyway because no further permits are auctioned after 2057 implying all permits in the MSR are cancelled.

The dirty and relative clean firms are assumed to generate electricity either with coal or gas, respectively. Based on European Commission (2019) we set the initial bank volume to $B_{2018} = 1.655$ Gt while we assume that initially all permits are held by the dirty coal firm which is consistent with our model results because in early years the gas firm is clean and thus does not bank (cp. proposition 1). The risk-free rate is assumed to be $r = 3\%$. For more details on assumed parameters we refer to Appendix B. To solve the model with the MSR we first run the model with the pre-MSR auction schedule S_t.[10] The resulting bank volumes B_t are then used to compute the MSR adjustments according to (3.2) and (3.3). The model is solved again with the adjusted permit issuance S_t^M. This procedure is iterated until it converges (similar to e.g. Beck and Kruse-Andersen 2018).

3.2. Results

In order to disentangle the effects of permit shifting over time and cancellation both due to the MSR, we run run a scenario without MSR, with MSR but without cancellation and a scenario with MSR and cancellation. We further differentiate between scenarios with risk aversion (RA) and risk neutrality (RN). Hence we have six scenarios in total denoted by RN, RA (both without MSR), RN MSR, RA MSR (MSR without cancellation), RN MSR + cancel and RA MSR + cancel (MSR with cancellation). In scenarios without MSR the initial MSR bank is added to the initial bank level of the coal firm. First we focus on price effects due to the hedging demand in general and permit shifting (MSR without cancellation), and thereafter on the cancellation effects. Finally, we analyze capacity and variability effects of the MSR. Since we are interested in the mid-term effects and for clarity we show results until 2055, for the full time horizon see Appendix C.

3.2.1. MSR without cancellation

First we concentrate on scenarios without MSR cancellation. Figure 3.1 (a) shows both scenarios with risk aversion (RA, RA MSR) have a similar price pattern compared

[10] The model is implemented with the software GAMS as Extended Mathematical Programming (EMP) model with the solver JAMS. The code is available...

to their risk neutral counterparts (RN, RN MSR). Initially prices are higher with risk aversion, then decline or do not rise and therefore are lower compared to risk neutrality from 2030 to 2035 onward. Deviations are driven by the firms' hedging demand, as reflected by risk premiums, shown in Figure 3.1 (b). In early years, risk premiums are highly negative and thus permit prices rise at a rate below the risk neutral (risk-free) rate. Risk premiums are as low as -7.4% because in early periods the electricity market is relatively emission intensive implying a high coal production level and thus high hedging demand of the coal firm. The available permits for banking do not suffice to cover the high hedging demand of the coal firm. As a result, the firm accepts a reduced return for holding permits reflected by the negative risk premiums.

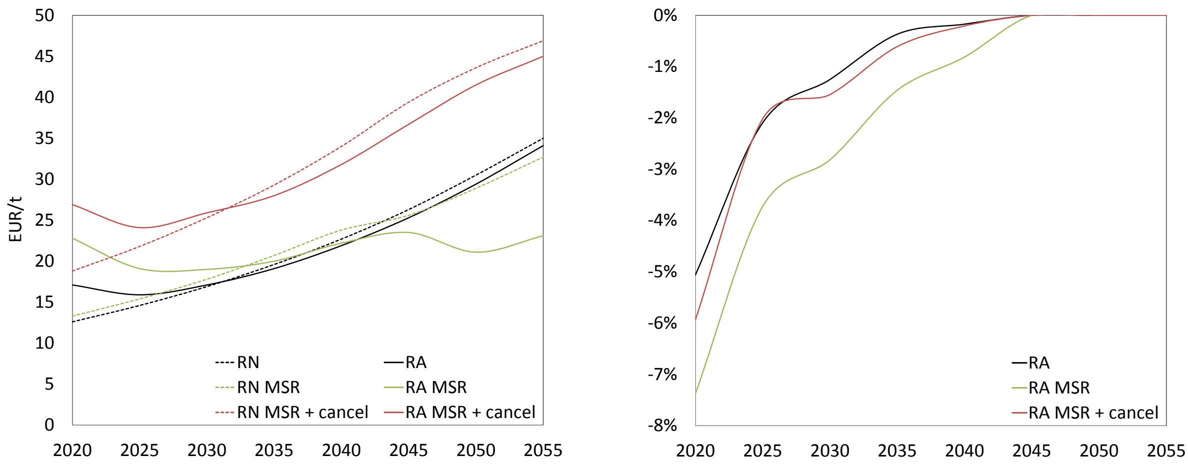
Over time, the hedging demand declines along the falling emission intensity of the market and, additionally, the bank volume rises (see Figure 3.2 (a)). Consequently, the (absolute) risk premiums decline, however, they never turn positive. The negative risk premiums indicate that the permit issuance schedule is in all cases too flat, meaning that in early years not enough permits are issued to cover the hedging demand. With a steeper permit issuance (for a given cap) risk premiums could also be positive.

Figure 3.1 (a) further shows the positive impact of the MSR on near-term prices, but negative effects on long-term prices if cancellation is not active (green vs. black lines). For risk neutrality this can be explained by an earlier binding borrowing constraint for permits (see Perino and Willner 2016). With risk aversion the effect of the MSR is significantly amplified: instead of a price increase of only 0.70 EUR/t in 2020 (RN MSR vs. RN), the price increases by 5.70 EUR/t (RA MSR vs. RA) if the hedging demand of firms is considered. The larger price increase in the near-term is outweighed by a larger price decline or lower growth rate thereafter such that the price level in RA MSR is in 2040 approximately as high as in 2020 and as high as in case without MSR.

The reason for the strong effects of the MSR even without cancellation is the reduced firm bank level as shown in Figure 3.2 (a). Instead of firms holding permits, a large amount of permits is transferred into the MSR bank (see Figure 3.2 (b)) where they cannot cover the firms hedging demand. This implies higher (negative) risk premiums, leading to higher short-term and lower long-term prices as explained above.

Figure 3.1: Expected permit price and risk premium

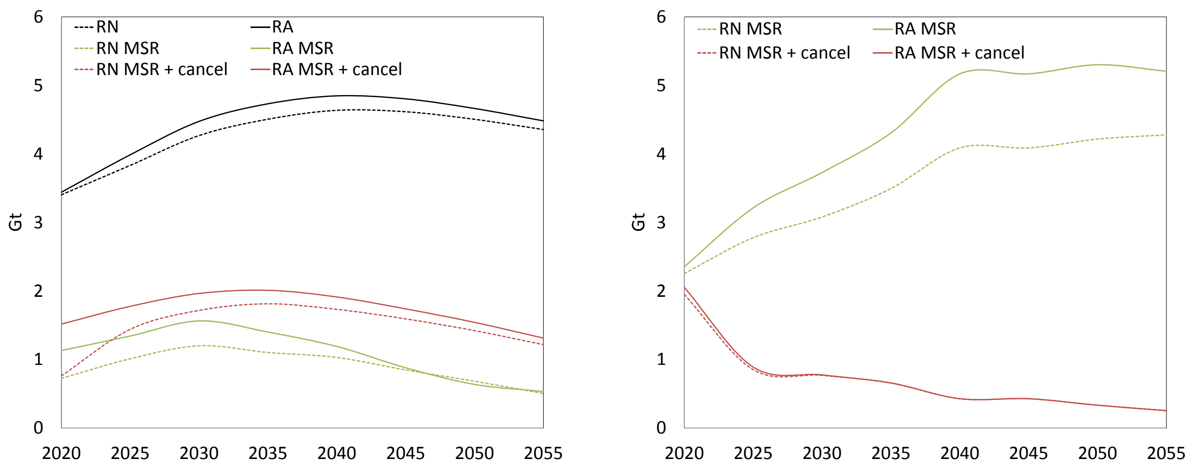
(a) price (b) annual risk premium



note: the lower growth rate or price deline after 2045 is due to binding borrowing constraints.

Figure 3.2: Expected firm and MSR bank

(a) aggregate firm bank (b) MSR bank



note: at the beginning of the first period firm banks in RA and RN, as well as in the four MSR scenarios are the same, respectively. The figure shows bank levels at the end of each period and thus the lines in the figure do not start from the same point. The same holds for the MSR bank. In line with proposition 1 the gas firm does not bank in RA scenarios before 2025 or 2030, depending on the scenario.

3.2.2. MSR with cancellation

If cancellation is active, a similar price pattern as without cancellation can be observed but on a higher level (see Figure 3.1 (a)). However, cancellation mitigates the price drop after 2020 and thus the price level of 2020 is already reached in 2035 instead of 2040 as without cancellation. This can be traced back to the higher price level induced by cancellations: higher prices imply less coal production and thus a lower hedging demand. Moreover, less coal production also implies a higher firm bank level (see Figure 3.2 (a)) and thus the mismatch between hedging demand and permit availability is lower compared to RA MSR. In turn, absolute risk premiums are lower as shown in Figure 3.1 (b). Overall, the price starts at a higher level and declines less and therefore prices are strictly higher than without MSR.

Cancellation amounts to 7.60 Gt and 8.59 Gt in case of RN MSR + cancel and RA MSR + cancel, respectively. Hence if the hedging demand is considered, cancellation is about 1 Gt higher. This can be explained by the higher value of permits in early years in case of risk aversion due to firms' hedging demand. Specifically, due the hedging value the price in 2020 is significantly higher (red lines in Figure 3.1 (a)) leading to less emissions and a higher private bank (red lines Figure 3.2 (a)). In turn, the influx into the MSR is higher and thus more permits are canceled. The lower prices after 2030 cannot outweigh this effect since cancellation mainly takes place before 2030.

Compared to the results from the literature, in which cancellations range from 1.7 Gt to 10 Gt, our results are at the higher end.[11]

While differences can be explained by several model assumptions (e.g. baseline emission, see discussions in Bocklet et al. (2019) and Silbye and Sørensen (2018)), a particular important one is the choice of the discount rate. For example, discount rates of 10% and 8% in Perino and Willner (2017a) and Bocklet et al. (2019), respectively, are one reason why cancellations are relatively low. Since higher discount rates imply that firms put less weight on the future, they use more permits initially. Consequently, prices and firm bank levels are lower in early years, and thus less permits flow into MSR and accordingly cancellations are lower as well. In our model the discount rate reflects the risk-free rate for an alternative asset which we explicitly model. Due to low interest rates in Europe, we feel a low rate of 3% is adequate. Moreover, the premium (implicitly) put on top of

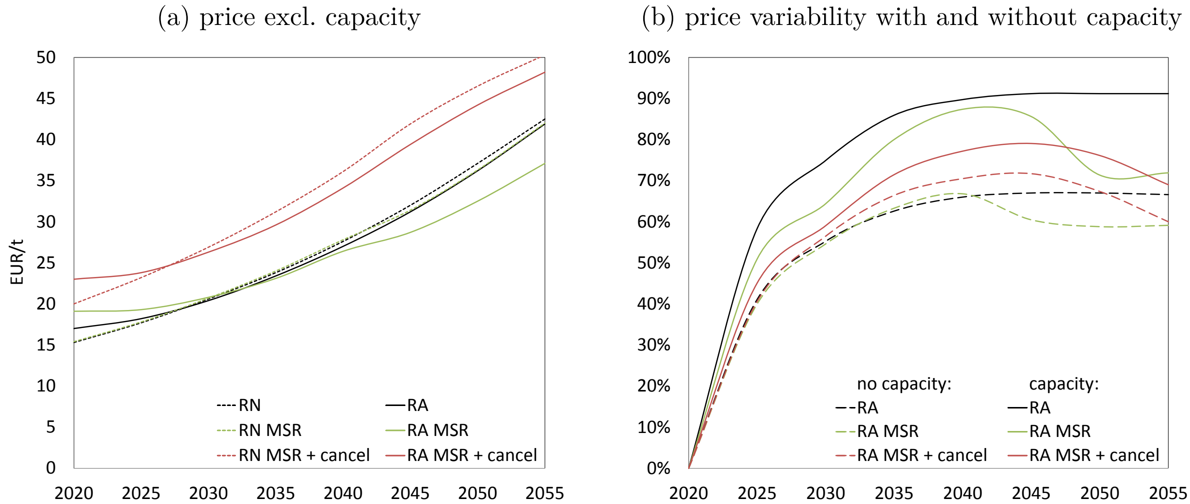
[11]Results from the literature (considering only reference or baseline scenarios) in ascending order are 1.7 Gt (Perino and Willner 2017a), 1.9Gt (Bruninx et al. 2018), 2 Gt (Bocklet et al. 2019), 2.9 Gt (Carlén et al. 2019), 4.5 Gt (Rosendahl 2019), 5 Gt (Silbye and Sørensen 2018), 6 Gt (Beck and Kruse-Andersen 2018) and 5 Gt or 10 Gt (if MSR impacts are anticipated and depending on the type of myopia) in Quemin and Trotignon (2019).

the risk-free rate in other papers, is to a certain degree endogenous in our approach since we explicitly account for risk and risk aversion. This, however, even leads to a lower applied discount rate because of the negative risk premiums.

We also run a scenario with 5% interest rate in which cancellations are with 6.23 Gt (RN MSR + cancel) and 6.81 Gt (RA MSR + cancel) significantly lower. Furthermore, if futures contracts are added as described in Appendix A.4 the risk premium is smaller because the risk-taking capacity of the market increases. Cancellations are with 8.27 Gt in RA MSR + cancel somewhat lower compared to 8.59 Gt without futures market (in case of risk neutrality futures markets have no effect), because the effect of risk aversion becomes weaker. However, both modifications do not change the nature of our results since the price pattern is similar (see Appendix C).

3.2.3. Capacity and price variability effects

Figure 3.3: Expected permit price (excl. capacity) and permit price variability



note: the price variability is measured as relative standard deviation from the perspective of the first period.

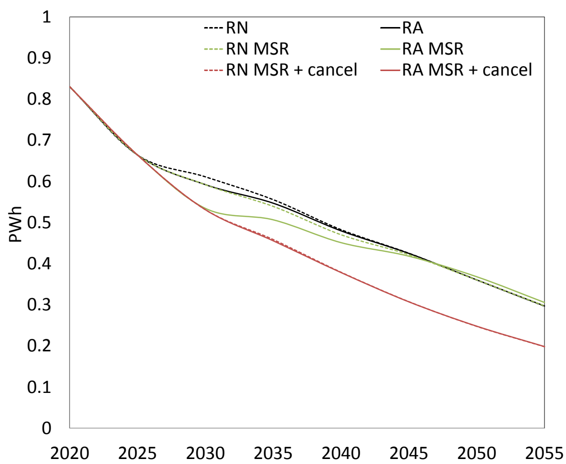
In order to highlight the effect of capacity constraints, we run an additional scenario without capacity constraints. As Figure 3.3 (a) shows the effect of the hedging demand on the price paths are weaker without capacity constraints (compared to results in Figure 3.1 (a)). This can be traced back to the price variability increasing effect of capacity constraints (lemma 3) as shown in Figure 3.3 (b) because higher risks increase the hedging demand.

In this regard, a target of the MSR is to lower the price variability. While findings in the literature are mixed as Fell (2016) and Quemin and Trotignon (2019) find lower and Kollenberg and Taschini (2019), Perino and Willner (2016) and Richstein et al. (2015) a higher variability due to the MSR, our findings suggest that the MSR is more likely to reduce price variability if capacity constraints are considered, see Figure 3.3 (b).

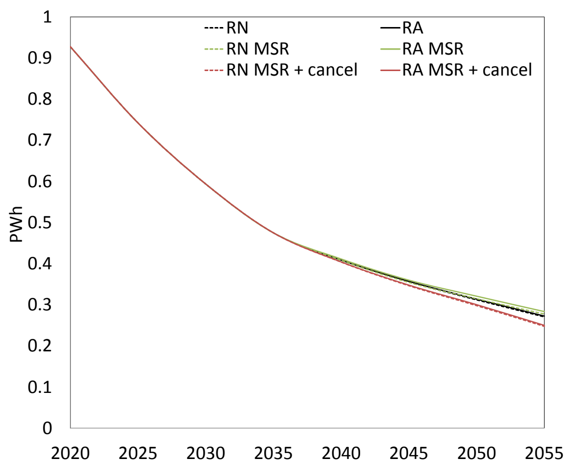
Another aim of the MSR is to spur low carbon investments. While we do not explicitly consider low carbon investments, we can analyze the effect on the capacity paths of coal and gas, as shown in Figure 3.4. They reflect how investments in high carbon (coal) and relatively low carbon (gas) are affected by the MSR. First observe that there are no new investments in gas capacity before 2035 and therefore, overall effects on gas capacities are weak. For coal the MSR effects are much stronger. Intuitively, if cancellation is active, there is significantly less coal capacity because of the higher permit price. If cancellation is not active, the MSR has a significant effect in case of risk aversion (RA MSR): due to the higher permit price until about 2040 and the worse hedging opportunities (see section 2.2.2), there are less coal capacities compared to RA. From 2045 onward, however, there is more coal capacity because the MSR leads to lower prices in the long-term. In case of risk neutrality the effects are similar but much weaker, because the MSR has only a small effect on prices.

Figure 3.4: Expected capacity

(a) coal capacity



(b) gas capacity



4. Conclusion

We analyze the impact of firm level hedging on the permit price path of a cap-and-trade program in an intertemporal stochastic equilibrium model. Hedging demand arises from uncertain profits due to regulatory risk regarding the future permit supply and has different implications for relative clean (gas) and dirty (coal) firms. Hedging by dirty firms via permit banking has a negative effect on the risk premium of the permit price – the sign is opposite for the clean firm's hedging via borrowing. If, as usual, permit borrowing is not allowed, the dirty firm's hedging demand becomes decisive for the permit price path: When the hedging demand exceeds the available permits, the resulting permit price is higher than in the risk neutral case but rises at a lower rate. When the dirty firm's hedging demand falls short of the permit supply, the opposite holds. Since the hedging demand of dirty firms is typically high in early years (implying price growth at low rate) of a cap-and-trade program and low in latter years (implying high growth rate), the expected growth rate of the permit price likely has a hockey stick or even U-shape.

We apply the model numerically to the European Union's Emission Trading System (EU ETS) and in particular to the recently introduced Market Stability Reserve (MSR). The core components of the MSR are a shift of permits to the future and cancellation of permits if the aggregate permit bank exceeds certain thresholds. In our stylized model, the hedging demand of the dirty coal firm always exceeds the available permits and thus risk premiums are always negative. The MSR rises the (negative) risk premiums further because of the permit transfer to the future and the associated reduction of available permits for banking in the near-term. The results offer an explanation for the recent permit price hike in the EU ETS following the EU ETS reform[12] since higher (negative) risk premiums lead to higher short-term prices. While higher short-term prices are intended by policy makers, we show that the price level might not be sustainable. Specifically, the negative risk premiums may even lead to a price decline in the mid-term. This also questions whether the MSR is an appropriate instrument to incentivize cleaner

[12] The permit price in the EU ETS has risen since mid 2017, reaching about 25 EUR/t in mid 2019.

investments. We find that a pure shifting of permits to the future (without cancellation of permits) increase coal capacities in the long-term.

Besides the higher hedging value of permits due to the MSR another important reason for the recent price increase is certainly the cancellation of permits from 2023 onward. Here we find that cancellations may be higher than previous analyses suggest: due to the hedging demand, firms apply a lower discount rate for banking permits. This implies that they want to build up a larger bank, which, in turn, increases the MSR cancellations. We also stress the role of capacity constraints which prevail in electricity market. Specifically, we show that they increase the permit price variability and thereby amplify the effects of risk aversion.

Shifting permits as a policy instrument, as with the MSR, is further complicated by informational challenges. Determining the specific volumes, timing, and threshold levels of the MSR becomes a difficult task since permits need to be allocated in accordance to parameters which are partly unknown to the regulator. The hedging demand in general is private information and additionally depends on several factors (e.g. the availability of other hedging opportunities, also in related markets), making its estimation difficult. Moreover, our simulation suggest that the hedging corridor for permit banking is set too low because firms want to hold more permits for hedging purpose. However, increasing the hedging corridor implies a general weaker impact of the MSR, which questions whether it can be designed to reach its targets. While our numerical simulation is based on a simplistic model ignoring several factors (e.g. freely allocated permits, alternative hedging opportunities), future research could refine the estimation of the hedging demand and optimal issuance of permits. This is especially important to provide a basis for the upcoming review of the MSR in the early 2020s.

Overall, the MSR is a rather complex mechanism and works indirect via the bank volume and thus is at most a second best solution. If the aim of the ETS reform is to increase the permit price level, reduce its variability and incentivize more clean investments, it seems advisable to approach these issues more directly by tackling the flawed price. In this regard, price collars are promising since they can directly help to reach these targets and are in addition less complex than the MSR.

Acknowledgments

The authors would like to thank Dallas Burtraw, Christian Gambardella, Simon Quemin and participants of the WCERE conference 2018 in Gothenburg for valuable comments. Oliver Tietjen's contribution was funded from the European Union's Horizon 2020 research and innovation program under grant agreement No 730403 (INNOPATHS). Michael Pahle was funded by Stiftung Mercator Foundation under the research project AHEAD.

Kai's funding!?

Thanks to Ottmar!?

Appendix A. Derivations

Appendix A.1. First Order Conditions

Date 2. The first order conditions of the firms' problem at date 2 are:

$$U_{i,2}^{'}\left(k_{i,2}w_2-\beta_i\zeta_{i,2}k_{i,2}^2\right)-\rho_{i,2}\phi_i k_{i,2}-\mu_{i,2}=0 \quad (\zeta_{i,2}) \tag{A.1}$$

$$U_{i,2}^{'}p_2-\rho_{i,2}=0 \quad (y_{i,2}) \tag{A.2}$$

Date 1. The first order conditions of the firm's problem at date 1 are:

$$U_{i,1}^{'}\left(k_{i,1}w_1-\beta_i\zeta_{i,1}k_{i,1}^2\right)-\rho_{i,1}\phi_i k_{i,1}-\mu_{i,1}=0 \quad (\zeta_{i,1}) \tag{A.3}$$

$$U_{i,1}^{'}p_1-\rho_{i,1}=0 \quad (y_{i,1}) \tag{A.4}$$

$$\frac{E\left[U_{i,2}^{'}p_2\right]}{1+r}-\rho_{i,1}+\varphi_{i,1}=0 \quad (b_{i,1}) \tag{A.5}$$

$$E\left[U_{i,2}^{'}\left(\zeta_{i,2}w_2-\beta_i\zeta_{i,2}^2k_i-p_2\phi_i\zeta_{i,2}-c_i\right)\right]=0 \quad (k_{i,2}) \tag{A.6}$$

$$U_{i,1}^{'}-E\left[U_{i,2}^{'}\right]=0 \quad (l_{i,1}) \tag{A.7}$$

Appendix A.2. Date 2 Equilibrium

Lemma 1. Inserting the goods market price (2.17) in the utilization rate (2.14), if the capacity constraints do not bind, $\mu_{c,2} = \mu_{d,2} = 0$, yields

$$\zeta_{i,2} = \frac{A\beta_c\beta_d + p_2(\beta_d\phi_c + \beta_c\phi_d)}{(\beta_d + \beta_c + \beta_c\beta_d a)\beta_i k_i} - \frac{p_2\phi_i}{\beta_i k_i} \tag{A.8}$$

Considering the case for $i = c$, and taking the derivative with respect to the permit price yields:

$$\frac{d\zeta_{c,2}}{dp_2} = \frac{\phi_d - \phi_c(1 + \beta_d a)}{(\beta_d + \beta_c + \beta_c\beta_d a)k_c} > 0 \tag{A.9}$$

if $\phi_d - \phi_c(1 + \beta_d a) > 0$. Similar, for the dirty firm we get

$$\frac{d\zeta_{c,2}}{dp_2} = \frac{\phi_c - \phi_d(1 + \beta_c a)}{(\beta_d + \beta_c + \beta_c\beta_d a)k_d} < 0 \tag{A.10}$$

since by definition $\phi_d > \phi_c$.

Lemma 2. Using the utilization rate (2.14) and the goods price (2.17), the plant profit can be written as

$$\pi_{i,2}^{plant} = \frac{(A\beta_c\beta_d + p_2(\beta_d\phi_c + \beta_c\phi_d))^2}{2\beta_i(\beta_d + \beta_c + \beta_c\beta_d a)^2} - \frac{(A\beta_c\beta_d + p_2(\beta_d\phi_c + \beta_c\phi_d))p_2\phi_i}{\beta_i(\beta_d + \beta_c + \beta_c\beta_d a)} + \frac{p_2^2\phi_i^2}{2\beta_i}. \tag{A.11}$$

for which have assumed that the capacity constraints do not bind. For the clean firm it can be shown that the profit increases with the ETS price, $\frac{d\pi_{c,2}^{plant}}{dp_2} > 0$, if

$$A\beta_d + p_2(\phi_d - \phi_c(1 + \beta_d a)) > 0 \tag{A.12}$$

which is always the case if $\phi_d - \phi_c(1 + \beta_d a) > 0$ holds. From $\frac{d\pi_{c,2}^{plant}}{dp_2} > 0$ directly follows that $Cov\left[\pi_{c,2}^{plant}, p_2\right] > 0$. For the dirty firm the profit decreases with the ETS price, $\frac{d\pi_{d,2}^{plant}}{dp_2} < 0$, if

$$A\beta_c + p_2\left(\phi_c - \phi_d\left(1+\beta_c a\right)\right) > 0 \tag{A.13}$$

holds. If the price is

$$p_2 = \frac{A\beta_c}{\left(\phi_d\left(1+\beta_c a\right) - \phi_c\right)}, \tag{A.14}$$

profits are not affected, i.e. $A\beta_c + p_2\left(\phi_c - \phi_d\left(1+\beta_c a\right)\right) = 0$. If the price is larger than in (A.14), the dirty firm does not produce and thus profits are not affected as well. This can be seen by inserting (A.14) in the utilization rate (A.8) which yields $\zeta_{d,2} = 0$. The same is true for higher prices because of lemma 1. For lower prices than in (A.14), condition (A.13) is fulfilled and thus $\frac{d\pi^{plant}_{d,2}}{dp_2} < 0$ and $Cov\left[\pi^{plant}_{d,2}, p_2\right] < 0$ hold.

Lemma 3. The effect of capacity constraints on the ETS price is given by the second line in (2.18), which we replicate here for convenience

$$\frac{\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}}\left(\phi_d - \phi_c\left(1+\beta_d a\right)\right) + \frac{\mu_{d,2}}{U'_{d,2}k_{d,2}}\left(\phi_c - \phi_d\left(1+\beta_c a\right)\right)}{\left(\phi_c - \phi_d\right)^2 + a\left(\beta_c\phi_d^2 + \beta_d\phi_c^2\right)}. \tag{A.15}$$

The first line in (2.18) is the same as without capacities. There are four cases:

Case 1: Before the shock on S_2 is realized the capacity constraint of the clean firm binds, $\mu_{c,2} > 0$, and the capacity constraint of the dirty firm does not bind, $\mu_{d,2} = 0$. A negative shock implies that $\mu_{d,2} = 0$ still holds after the shock because of lemma 1. For the effect on the constraint of the clean firm, we make use of (2.18) and (2.17) in (2.16) such that we get

$$\begin{aligned}\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}} &= \frac{A\phi_d\left(\phi_d - \phi_c\right) - k_{c,2}\left(\left(\phi_d - \phi_c\right)^2 + a\left(\beta_d\phi_c^2 + \beta_c\phi_d^2\right)\right)}{\phi_d^2 a} \\ &\quad + \frac{\left(b_{c,1} + b_{d,1} + S_2\right)\left(\phi_c\left(1+\beta_d a\right) - \phi_d\right)}{\phi_d^2 a}.\end{aligned}$$

The effect of a change in the permit supply is given by

$$\frac{d\left(\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}}\right)}{dS_2} = \frac{\phi_c\left(1+\beta_d a\right)-\phi_d}{\phi_d^2 a} < 0 \tag{A.16}$$

because $\phi_d - \phi_c\left(1+\beta_d a\right) > 0$ and thus capacity constraints lead to a larger price increase due to a negative shock on S_2 compared to the model without capacity constraints.

Case 2: Before the shock is realized it holds $\mu_{c,2} = 0$ and $\mu_{d,2} > 0$. A negative shock implies that $\mu_{c,2}$ rises or may still be zero, $\mu_{c,2} \geq 0$, and that the dirty capacity constraint does not bind anymore, $\mu_{d,2} = 0$ (lemma 1). To see that a declining $\mu_{d,2}$ and a rising $\mu_{c,2}$ leads to a stronger ETS price increase in (2.18) consider that $\phi_c - \phi_d\left(1+\beta_c a\right) < 0$ and $\phi_d - \phi_c\left(1+\beta_d a\right) > 0$ hold.

Case 3: Before the shock is realized it holds $\mu_{c,2} > 0$ and $\mu_{d,2} > 0$. As in case 2 the dirty constraint cannot bind anymore after a negative shock has emerged, which again has a positive effect on the price. As in case 1, $\mu_{c,2}$ rises which also has a positive effect on the price.

Case 4: Before the shock is realized it holds $\mu_{c,2} = \mu_{d,2} = 0$. As in case 1 a negative shock implies that $\mu_{d,2} = 0$ still holds and $\mu_{c,2} \geq 0$. Hence if $\mu_{c,2} > 0$ after the shock, capacity constraints have a positive effect on the price and no effect if $\mu_{c,2} = 0$.

In sum, a negative shock on S_2 leads to a stronger or the same price effect than in the case without capacity constraints. A positive shock on S_2 has opposite effects and thus leads to a stronger or equal price decline. Therefore, the price variability is amplified due to capacity constraints.

Appendix A.3. Date 1 Equilibrium

Appendix A.3.1. Banking and Hedging

Combining first order conditions (A.7), (A.4) and (A.5) yields

$$\frac{E\left[p_2\right]-p_1}{p_1} = r + \frac{\left(1+r\right)\varphi_{i,1}}{p_1 E\left[U'_{i,2}\right]} - \frac{Cov\left[U'_{i,2},p_2\right]}{E\left[U'_{i,2}\right]p_1}. \tag{A.17}$$

Assuming quadratic utility, $U_t\left(\pi_{it}\right) = \pi_{it} - \pi_{it}^2$, we can write the covariance as $Cov\left[U'_{i,2},p_2\right] = -2\left(Cov\left[\pi_{i,2}^{plant},p_2\right] + Var\left[p_2\right]b_{i,1}\right)$. Inserting the covariance in (A.17) and rearranging yields:

$$b_{i,1} = \frac{E\left[p_2\right] - p_1\left(1+r\right)}{\lambda_i Var\left[p_2\right]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var\left[p_2\right]} - \frac{\left(1+r\right)\varphi_{i,1}}{U'_{i,1} Var\left[p_2\right]} \tag{A.18}$$

Assuming a zero risk premium $E\left[p_2\right] - p_1\left(1+r\right) = 0$, the pure banking or borrowing demand is only due to the second term. Because of lemma 2 we have $Cov\left[\pi_{c,2}^{plant}, p_2\right] > 0$ and $Cov\left[\pi_{d,2}^{plant}, p_2\right] < 0$. Obviously, if firms bank, the borrowing constraint does not bind and thus $\varphi_{i,1} = 0$. It follows that dirty firms want to bank $b_{i,1} > 0$ and clean firms want to borrow $b_{i,1} < 0$ permits for hedging reasons. However, clean firms cannot borrow by assumption. Instead, clean firms bank permits only if the expected profit is at least as high as the costs of risks of this action, $E\left[p_2\right] - p_1\left(1+r\right) > \lambda_i Cov\left[\pi_{c,2}^{plant}, p_2\right]$.

Appendix A.3.2. Price Effects

Consider that the permit demand can be written as

$$\begin{aligned} y_{i,1} &= \frac{\phi_i A \beta_c \beta_d}{\left(\beta_d + \beta_c + \beta_c\beta_d a\right)\beta_i} + p_1\left(\frac{\left(\beta_d\phi_c + \beta_c\phi_d\right)\phi_i}{\left(\beta_d + \beta_c + \beta_c\beta_d a\right)\beta_i} - \frac{\phi_i^2}{\beta_i} - \frac{\left(1+r\right)}{\lambda_i Var\left[p_2\right]}\right) \\ &+ \frac{E\left[p_2\right]}{\lambda_i Var\left[p_2\right]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var\left[p_2\right]} \end{aligned} \tag{A.19}$$

for which we have used (2.20) and (A.8) (but for date 1) in (2.19) and we assumed non-binding capacity constraints. Inserting this permit demand for the clean and dirty firm in the permit equilibrium condition (2.4) yields after same algebra the permit price,

$$p_1 = \frac{E\left[p_2\right]}{\left(1+r\right)} - \frac{\Lambda}{\left(1+r\right)}\left(Cov\left[\pi_{d,2}^{plant}, p_2\right] + Cov\left[\pi_{c,2}^{plant}, p_2\right] + Var\left[p_2\right] B_1\right) \tag{A.20}$$

The whole term in brackets is the absolute risk premium and divided by p_1 the relative risk premium as shown in (2.22).

Appendix A.3.3. Capacity Effects

The first order condition for $k_{i,2}$ can be reformulated as

$$
\begin{aligned}
E\left[\zeta_{i,2}\right]\left(E\left[w_2\right]-E\left[\zeta_{i,2}\right]\beta_i k_i-E\left[\zeta_{i,2}\right]\phi_i\right)-c_i & \\
+Cov\left[\zeta_{i,2},w_2-\zeta_{i,2}\beta_i k_i-p_2\phi_i\right] & \\
+\frac{1}{U'_{i,1}}Cov\left[U'_{i,2},\zeta_{i,2}w_2-\zeta_{i,2}^2\beta_i k_i-p_2\zeta_{i,2}\phi_i\right] &= 0
\end{aligned}
\tag{A.21}
$$

for which we have used covariance properties and the first order condition for the risk-free asset (A.7). By further noting that the marginal capacity value in the risk neutral case is $\mu_{i,2}^{RN}=k_{i,2}w_2-x_{i,2}k_{i,2}\beta_i-p_2k_{i,2}\phi_i$ (see equation (2.16) and consider that in case of risk neutrality $U'_{i,2}=1$.), we can rewrite this further and finally get (2.24).

Effect of Uncertainty in Risk Neutral Case. Risk neutrality implies $U'_{i,2}=1$ and thus $Cov\left[U'_{i,2},\mu_{i,2}^{RN}\right]=0$. Therefore, only the first two terms in (2.24) matter in the risk neutral case. Moreover, in the deterministic case $Cov\left[\zeta_{i,2},\mu_{i,2}^{RN}\right]=0$ holds and investments are only determined by the first term in (2.24) (ignoring the expectation operator). Hence given risk neutrality the effect of uncertainty compared to the deterministic case is given by the second term, $Cov\left[\zeta_{i,2},\mu_{i,2}^{RN}\right]$. We can rewrite this term as $Cov\left[\zeta_{i,2},\mu_{i,2}^{RN}\right]=E\left[\mu_{i,2}^{RN}\right]\left(1-E\left[\zeta_{i,2}\right]\right)$ for which we have used $E\left[\mu_{i,2}^{RN}\right]=E\left[\zeta_{i,2}\mu_{i,2}^{RN}\right]$ because $\mu_{i,2}^{RN}$ is only positive if $\zeta_{i,2}=1$ and zero otherwise. Since $0<E\left[\zeta_{i,2}\right]<1$ and $E\left[\mu_{i,2}^{RN}\right]>0$ it holds $Cov\left[\zeta_{i,2},\mu_{i,2}^{RN}\right]>0$.

Effect of Risk Aversion without Banking. Next, we consider the effect of risk aversion given by the third term in (2.24). Since $\frac{1}{U'_{i,1}}>0$, the sign of the effect of risk aversion depends on $Cov\left[U'_{i,2},\mu_{i,2}^{RN}\right]$. Assuming that the permit bank is zero, marginal utility $U'_{i,2}$ depends only on the plant profit $\pi_{i,2}^{plant}$ and the risk-free asset returns. The latter do not affect the covariance since they are certain. Due to the concavity of $U_{i,2}$, it follows $\frac{dU_{i,2}}{d\pi_{i,2}^{plant}}<0$ and thus the sign of $Cov\left[U'_{i,2},\mu_{i,2}^{RN}\right]$ is inversely related to $Cov\left[\pi_{i,2}^{plant},\mu_{i,2}^{RN}\right]$ which is positive,

$$Cov\left[\pi_{i,2}^{plant},\mu_{i,2}^{RN}\right]=Cov\left[\zeta_{i,2}\left(w_2-\frac{\beta_i}{2}x_{i,2}-p_2\phi_i\right),\zeta_{i,2}\left(w_2-\beta_i x_{i,2}-p_2\phi_i\right)\right]k_{i,2}^2\geq 0 \tag{A.22}$$

since firms only increase their utilization rate if this covers at least their marginal cost (A.1). Therefore also $Cov\left[U'_{i,2},\mu_{i,2}^{RN}\right]\leq 0$ holds.

Effect of Risk Aversion with Banking. We consider again $Cov\left[\pi_{i,2},\mu_{i,2}^{RN}\right]$ which becomes with banking,

$$Cov\left[\pi_{i,2},\mu_{i,2}^{RN}\right]=Cov\left[\pi_{i,2}^{plant},\mu_{i,2}^{RN}\right]+Cov\left[p_2b_{i,1},\mu_{i,2}^{RN}\right]. \tag{A.23}$$

Compared to case without banking, there is an additional effect $Cov\left[p_2b_{i,1},\mu_{i,2}^{RN}\right]$. Firms invest, ceteris paribus, more if $Cov\left[p_2b_{i,1},\mu_{i,2}^{RN}\right]<0$ and less if $Cov\left[p_2b_{i,1},\mu_{i,2}^{RN}\right]>0$ because a lower $Cov\left[\pi_{i,2},\mu_{i,2}^{RN}\right]$ implies a higher $Cov\left[U'_{i,2},\mu_{i,2}^{RN}\right]$ due to the concavity of the utility function. Due to lemma 1, dirty firms always produce less if there is a positive permit price shock and therefore $Cov\left[p_2b_{i,1},\mu_{i,2}^{RN}\right]<0$. For clean firms the opposite holds.

If capacity constraints are strictly binding such that firms cannot produce more in case of positive (clean firm) or negative (dirty firm) price shocks or stick to the fully capacity utilization in the opposite case we get $Cov\left[p_2b_{i,1},\mu_{i,2}^{RN}\right]=\left(Cov\left[p_2,w_2\right]-Var\left[p_2\right]\phi_i\right)b_{i,1}k_{i,2}$. Hence if permit price shocks are disproportionately transferred to the goods market price ($Cov\left[p_2,w_2\right]-Var\left[p_2\right]<0$), firms want to invest even more given that they are dirty enough (large ϕ_i). Very clean firms, in contrast, with $\phi_i\approx 0$, always want to invest less in plant capacity if they bank.

Appendix A.4. Extension with Futures Market

In addition to the dirty and clean firms there are also speculators active in the futures market who only trade the risk-free assets and permit futures. In the futures market, firms (generators and speculators) trade f_{it} units at date t, where $f_{it}>0$ means firms buy and $f_{it}<0$ means they sell futures. We only consider futures contracts with an expiry in the next period for which the underlying is next period's permit price, p_{t+1}. Hence the expected profit of trading permit futures is $E_t\left[\pi^f_{it+1}\right] = \left(E\left[p_{t+1}\right] - p^f_t\right) f_{it}$, where p^f_t is the price of one contract at date 1.

The generating firms' problem is the same as before but futures expenditures, $-f_{i,1}p^f_1$, and expected payoffs, $f_{i,1}E\left[p_2\right]$, are added to $\pi_{i,1}$ and $\pi_{i,2}$, respectively. While the date 2 problem of generating firms remains unchanged, firms additionally maximize utility via $f_{i,1}$ at date 1 and thus we have in addition to the previous first order conditions (see Appendix A.1),

$$U'_{i,1}p^f_1 - \frac{E\left[U'_{i,2}p_2\right]}{1+r} = 0 \quad \left(f_{i,1}\right). \tag{A.24}$$

Assuming quadratic utility, the futures demand of generating firms is

$$f_{i,1} = \frac{E\left[p_2\right] - p^f_1\left(1+r\right)}{\lambda_i Var\left[p_2\right]} - \frac{Cov\left[\pi^{plant}_{i,2}, p_2\right]}{Var\left[p_2\right]} - b_{i,1}. \tag{A.25}$$

Since the payoff of the futures contract and banking one permit is identical, they are perfect substitutes in terms of hedging and thus $-b_{i,1}$ appears in (A.25). The speculator's, $i = s$, problem is $\max_{f_{s,1},l_{s,1}} U_{s,1}\left(\pi_{i,1}\right) + \frac{1}{(1+r)}E\left[U_{s,2}\left(\pi_{s,2}\right)\right]$ with $\pi_{s,1} = -p_1 f_{s,1} - l_{s,1} + \left(1+r\right)l_{s,0}$ and $\pi_{s,2} = p_2 f_{s,1} + \left(1+r\right)l_{s,1}$. The first order conditions for the speculator are (A.7) (see Appendix A.1) and (A.24). In case of quadratic utility its futures trades are

$$f_{s,1} = \frac{E\left[p_2\right] - p^f_1\left(1+r\right)}{\lambda_s Var\left[p_2\right]}. \tag{A.26}$$

The futures market extension introduces two differences to our model. First, the clean firm can also reduce its risk exposure by taking short positions, $f_{it} < 0$, in the futures market. Second, the speculator increases the risk tolerance of the market. By making use of (A.25) and (A.26) in the equilibrium condition of the futures market, $\sum_i^I f_i = 0$, it can be shown that the equilibrium futures price is:

$$p_1^f = \frac{E\left[p_2\right]}{(1+r)} - \frac{\Lambda^f}{(1+r)}\left(Cov\left[\pi_{d,2}^{plant}, p_2\right] + Cov\left[\pi_{c,2}^{plant}, p_2\right] + Var\left[p_2\right] B_1\right) \quad (A.27)$$

The difference to the permit price without futures market (see equation (A.20) in Appendix A.3.2) is due to Λ^f. Essentially, speculators increase the risk tolerance since $\Lambda^f = \left(\lambda_d^{-1} + \lambda_c^{-1} + \lambda_s^{-1}\right)^{-1} < \Lambda = \left(\lambda_d^{-1} + \lambda_c^{-1}\right)^{-1}$ and thus they reduce the risk premium (cp. Colla et al. 2012).

Appendix B. Parameter assumptions for EU ETS simulation

Cost function parameters are chosen in line with coal and gas power plants for the representative dirty and clean firm, respectively (for parameters see Table (B.1)). For the electricity demand function, $D(w_t) = A - a_t w_t$, we assume the intercept to be $A = 3,462$ TWh, which is total electricity generation in the EU28, Iceland and Norway (the EU ETS countries except Lichtenstein) in 2017 according to Eurostat. Deviations from A due to aw_t are interpreted as production from other plant types (mostly nuclear and renewable energy), which we do not model explicitly. Therefore, a higher a_t means that other technologies gain a higher market share. This parameter leaves a degree of freedom to calibrate the model to recent EU ETS outcomes. Specifically, we calibrate the model such that the outcomes of the first period (2018 - 2022) of the scenario RA MSR + cancel (the actual EU ETS) are in line with recent EU ETS values. For this purpose we set the initial value to $a_{2020} = 60$ and assume that it raises at a 9% rate every five years. The increase in a_t mainly reflects market entry of renewable energies (due to declining costs or support schemes).

These parameter assumptions lead to an ETS price of 26.9 EUR/t and 0.78 Gt emissions in the first model period of scenario RA MSR + cancel. The price is in line with actual (futures) prices between 2018 and 2022 (26.15 - 27.44 EUR/t)[13]. Our emission level is somewhat lower than the emissions due to combustion in the EU ETS in 2018,

[13]https://www.barchart.com/futures/quotes/CK*0/futures-prices (05-07-2019)

which are 1.1 Gt[14]. However, emissions are likely to fall due to recently rising ETS prices compared to 15.92 EUR/t on average in 2018. The production shares of gas (16.7%) and coal (17.8%) in the model are close to actual values in 2017 with 19.2% for gas and 19.1% for coal (Eurostat), which again are likely to be lower in 2018 - 2022 due to higher ETS prices and growing renewable energy output.

Table B.1: Firm data

	clean	dirty
production costs (EUR/GWh):β_i	0.050	0.020
capacity costs (EUR/GWh): c_i	0.0049	0.0084
emission factor (t/GWh): ϕ_i	333	950
initial capacities (TWh): $k_{i,2020}$	830.2	927.5
capacity depreciation: δ	0.2	0.2

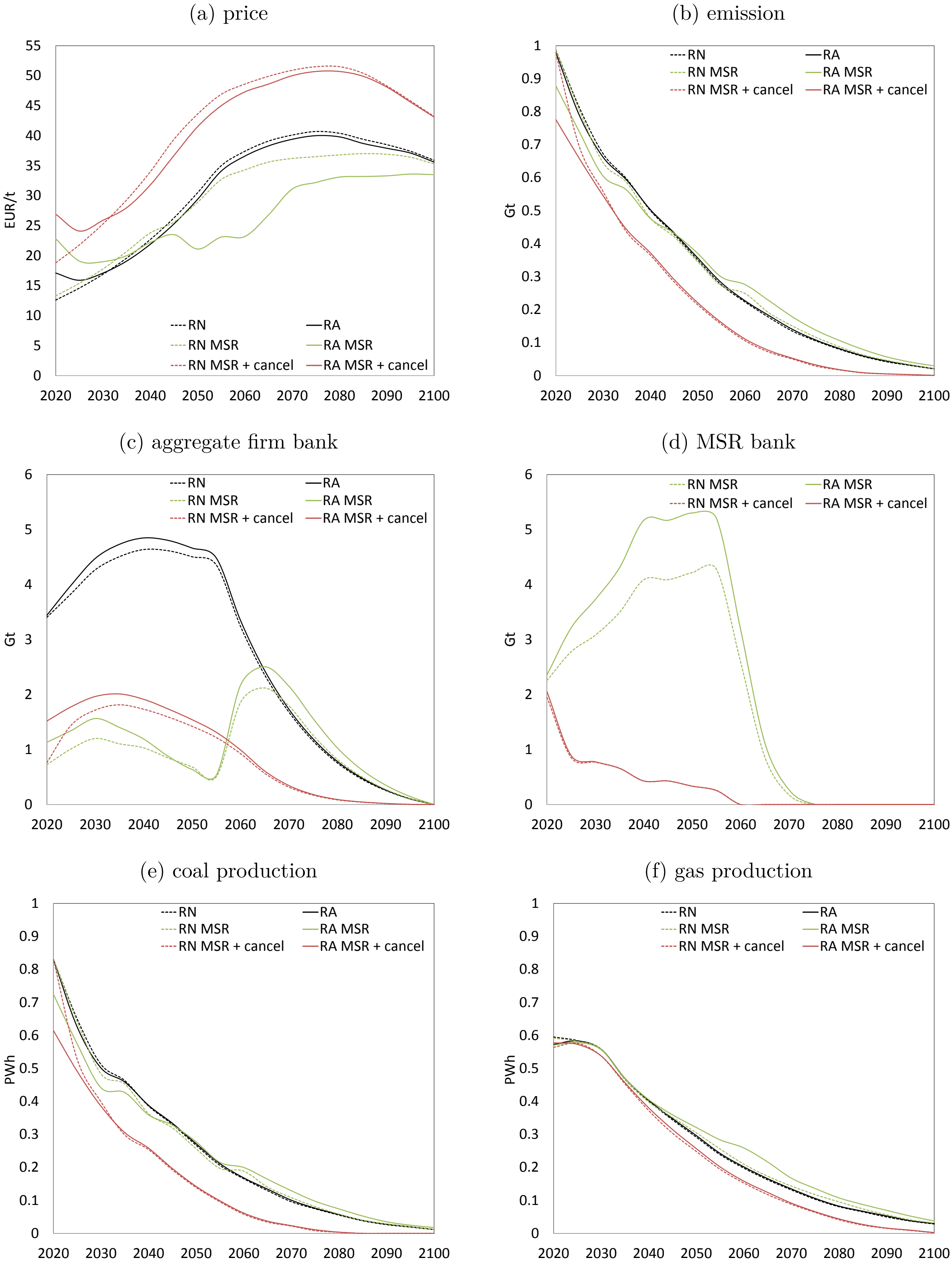
note: emission factors are based on UBA (2014) and divided by conversion efficencies (fuel to electricity) of 60% for gas and 40% for coal. Capacity costs are based on IEA (2016) but converted to annuities by considering plant lifetimes of 40 years and a 3% discount rate. Capacity costs are further converted from TW to TWh by assuming that plants are used 70% of the time. The production cost parameters β_i are roughly in line with gas and coal production costs (excl. emission costs). Initial capacities are from Eurostat for 2017: values for 'steam' (coal) and 'Gas turbine' and 'Combined cycle' (gas) are converted from W to Wh by multiplying the respective value with (8760*0.7), i.e. hours per year times the assumed utilization of 70%.

Regarding risk aversion, we assume in contrast to the analytical part a more common functional form. Specifically, we assume $U_{it} = \frac{\pi_{it}^{1-\eta}-1}{1-\eta}$ with constant relative risk aversion η. In line with empirical estimates we set relative risk aversion to $\eta = 1.5$ (cp. Gandelman and Hernández-Murillo 2015). We further assume an initial endowment of $l_{i,2020} = 40$ billion EUR. This value roughly corresponds to the profit made with the plant and permit trades in the the first period which is between 23 and 38 billion EUR for the coal and 41 and 42 billion EUR for the gas firm, depending on the scenario. That is, we assume the firms made a comparable profit in the previous (not modeled) period which is at their disposal in the first model period.

[14]https://www.eea.europa.eu/data-and-maps/dashboards/emissions-trading-viewer-1 (05-07-2019)

Appendix C. Additional simulation results

Figure C.1: Results for the full time horizon



note: the volatile permit price after 2045 in scenario RA MSR is due to the binding borrowing constraint (declining price) and the assumed higher output parameter for the MSR (rising price), which increases from 0.1 Gt to 1 Gt (see section 3.1). This also explains the quickly declining MSR bank and fast rising firm bank from 2055 onward.

Figure C.2: Expected permit price and risk premium with risk-free rate of 5%

(a) price (b) annual risk premium

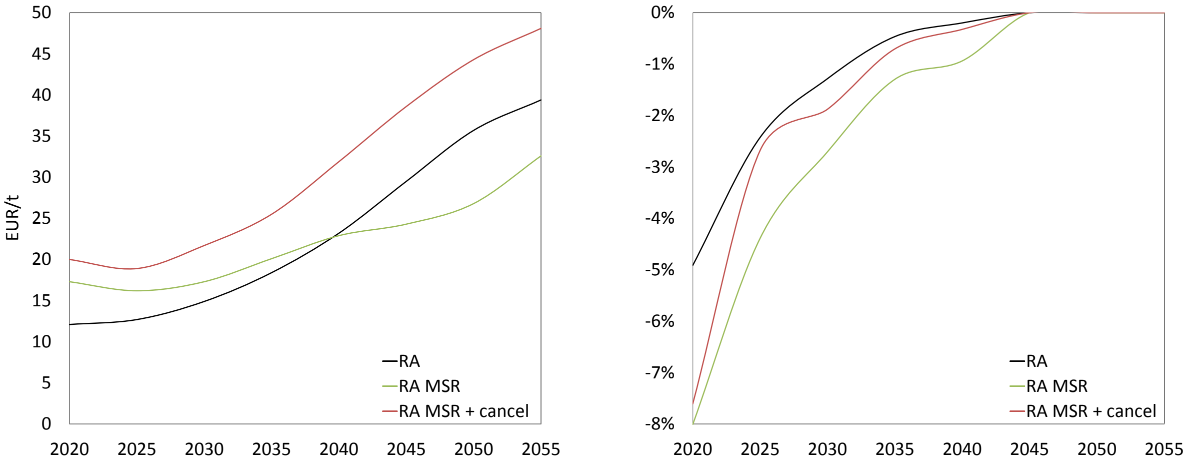
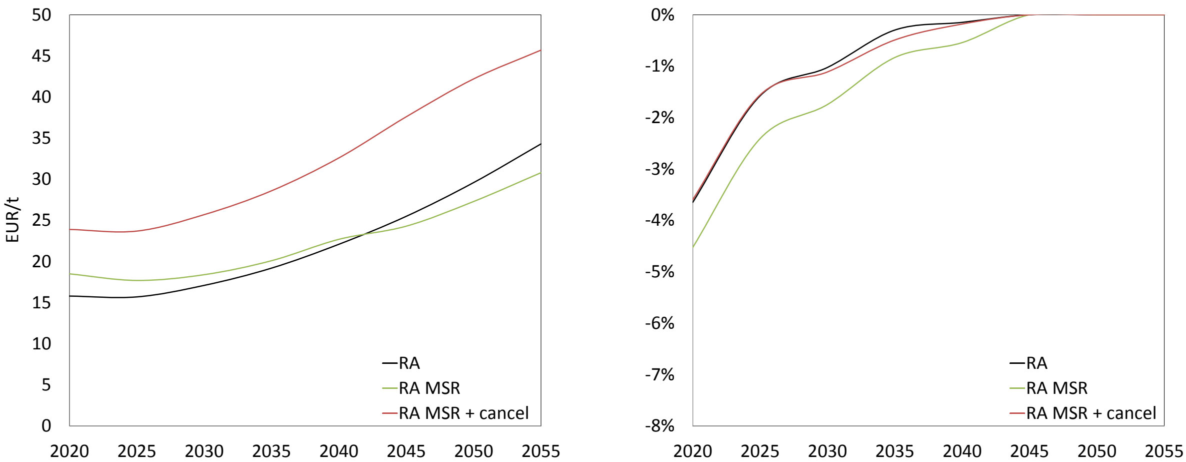


Figure C.3: Expected permit price and risk premium with futures contracts

(a) price (b) annual risk premium



note: speculators and futures contracts as described in Appendix A.4 are added to the standard model (3% interest rate). For simplicity we assume that speculators have the same initial endowment and level of risk aversion as the generating firms.

References

Acharya, V. V., L. A. Lochstoer, and T. Ramadorai (2013). Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics 109*(2), 441–465.

Baldursson, F. M. and N.-H. M. von der Fehr (2004). Price volatility and risk exposure: on market-based environmental policy instruments. *Journal of Environmental Economics and Management 48*(1), 682–704.

Baldursson, F. M. and N. H. M. von der Fehr (2012). Price Volatility and Risk Exposure: On the Interaction of Quota and Product Markets. *Environmental and Resource Economics 52*(2), 213–233.

Beck, U. R. and P. K. Kruse-Andersen (2018). Endogenizing the cap in a cap-and-trade system: Assessing the agreement on EU ETS phase 4.

Bessembinder, H. (1991). Forward Contracts and Firm Value: Investment Incentive and Contracting Effects. *The Journal of Financial and Quantitative Analysis 26*(4), 519.

Bessembinder, H. and M. L. Lemmon (2002). Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets. *The Journal of Finance 57*(3), 1347–1382.

Bocklet, J., M. Hintermayer, L. Schmidt, and T. Wildgrube (2019). The Reformed EU ETS - Intertemporal Emission Trading with Restricted Banking. *EWI Working Paper* (19).

Bruninx, K., M. Ovaere, and E. Delarue (2018). A First Analysis Of the Market Stability Reserve in the European Emission Trading System.

Cain, M. D. and S. B. McKeon (2016). CEO Personal Risk-Taking and Corporate Policies. *Journal of Financial and Quantitative Analysis 51*(1), 139–164.

Carlén, B., A. Dahlqvist, S. Mandell, and P. Marklund (2019). EU ETS emissions under the cancellation mechanism - Effects of national measures. *Energy Policy 129*(February), 816–825.

Chevallier, J. (2010). Modelling risk premia in CO2 allowances spot and futures prices. *Economic Modelling 27*(3), 717–729.

Chevallier, J. (2013). Variance risk-premia in CO2 markets. *Economic Modelling 31*(1), 598–605.

Colla, P., M. Germain, and V. Van Steenberghe (2012). Environmental Policy and Speculation on Markets for Emission Permits. *Economica 79*(313), 152–182.

Cronqvist, H., A. K. Makhija, and S. E. Yonker (2012). Behavioral consistency in corporate finance: CEO personal and corporate leverage. *Journal of Financial Economics 103*(1), 20–40.

Cronshaw, M. and J. Kruse (1996). Regulated firms in pollution permit markets with banking. *Journal of Regulatory Economics 9*(2), 179–189.

Diamond, P. A. (1967). The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty. *American Economic Review 57*(4), 759–776.

EEA (2018). *Trends and projections in the EU ETS in 2018.* Number 14.

Ellerman, A. D., C. Marcantonini, and A. Zaklan (2016). The European Union Emissions Trading System: Ten Years and Counting. *Review of Environmental Economics and Policy 10*(1), 89–107.

European Commission (2014). Impact Assessment Accompanying the document concerning the establishment and operation of a market stability reserve for the Union greenhouse gas emission trading scheme and amending Directive 2003/87/EC.

European Commission (2015). Proposal for a Directive of the European Parliament and of the Council amending Directive 2003/87/EC to enhance cost-effective emission reductions and low- carbon investments.

European Commission (2019). Publication of the total number of allowances in circulation in 2018 for the purposes of the Market Stability Reserve under the EU Emissions Trading System established by Directive 2003/87/EC.

European Parliament and Council (2015). Decision (EU) 2015/1814 of the European Parliament and of the Council of 6 October 2015 concerning the establishment and operation of a market stability reserve for the Union greenhouse gas emission trading scheme and amending Directive 2003/87/EC. *Official Journal of the European Union 264/1-264/.*

European Parliament and Council (2018). Directive (EU) 2018/410 of the European Parliament and of the Council of 14 March 2018 amending Directive 2003/87/EC to enhance cost-effective emission reductions and low-carbon investments, and Decision (EU) 2015/1814. *Official Journal of the European Union L 76/3 - L.*

Fell, H. (2016). Comparing policies to confront permit over-allocation. *Journal of Environmental Economics and Management 80*, 53–68.

Friedrich, M., E. M. Mauer, M. Pahle, and O. Tietjen (2018). Price formation in the EU ETS: a literature review. *Working paper*.

Froot, K., D. Scharfstein, and J. C. Stein (1993). Risk Management: Coordinating Corporate Investment and Financing Policies. *The Journal of Finance 48*(5), 1629–1658.

Fuss, S., C. Flachsland, N. Koch, U. Kornek, B. Knopf, and O. Edenhofer (2018). A Framework for Assessing the Performance of Cap-and-Trade Systems: Insights from the European Union Emissions Trading System. *Review of Environmental Economics and Policy 12*(2), 220–241.

Gandelman, N. and R. Hernández-Murillo (2015). Risk Aversion at the Country Level. *Federal Reserve Bank of St. Louis Review 91*(1), 53–66.

Hamilton, J. D. and J. C. Wu (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance 42*, 9–37.

Hintermann, B. (2012). Pricing emission permits in the absence of abatement. *Energy Economics 34*(5), 1329–1340.

Hintermann, B., S. Peterson, and W. Rickels (2016). Price and market behavior in phase II of the EU ETS: A review of the literature. *Review of Environmental Economics and Policy 10*(1), 108–128.

IEA (2016). World Energy Outlook 2016.

Kamga, M. K. and K. Schlepper (2015). Risk Premia in Carbon and Energy Futures Markets. *Working paper*.

KfW and ZEW (2016). KfW/ZEW CO2 Barometer 2016 - Carbon Edition. How the EU ETS can contribute to meeting the ambitious targets of the Paris Agreement.

Koch, N., S. Fuss, G. Grosjean, and O. Edenhofer (2014). Causes of the EU ETS price drop: Recession, CDM, renewable policies or a bit of everything?-New evidence. *Energy Policy 73*, 676–685.

Koch, N., G. Grosjean, S. Fuss, and O. Edenhofer (2016). Politics matters: Regulatory events as catalysts for price formation under cap-and-trade. *Journal of Environmental Economics and Management 78*, 121–139.

Kollenberg, S. and L. Taschini (2019). Dynamic supply adjustment and banking under uncertainty in an emission trading scheme: The market stability reserve. *European Economic Review 118*, 213–226.

Montgomery, W. (1972, dec). Markets in licenses and efficient pollution control programs. *Journal of Economic Theory 5*(3), 395–418.

Neuhoff, K., A. Schopp, R. Boyd, K. Stelmakh, and A. Vasa (2012). Banking of Surplus Emissions Allowances. *DIW Discussion Papers*.

Perino, G. and M. Willner (2016). Procrastinating reform: The impact of the market stability reserve on the EU ETS. *Journal of Environmental Economics and Management 80*, 37–52.

Perino, G. and M. Willner (2017a). EU-ETS Phase IV: allowance prices, design choices and the market stability reserve. *Climate Policy 17*(7), 936–946.

Perino, G. and M. Willner (2017b). Why the EU Market Stability Reserve deters long-term low-carbon investments. *WiSo-HH Working Paper Series 44*.

Pinho, C. and M. Madaleno (2011). Links between spot and futures allowances: ECX and EEX markets comparison. *International Journal of Global Energy Issues 35*(2/3/4), 101.

Quemin, S. (2017). Intertemporal Abatement Decisions underAmbiguity Aversion in a Cap and Trade. *Working Paper*.

Quemin, S. and R. Trotignon (2019). Intertemporal emissions trading and market design: an application to the EU ETS. *Working paper 348*.

Richstein, J. C., É. J. Chappin, and L. J. de Vries (2015). The market (in-)stability reserve for EU carbon emission trading: Why it might fail and how to improve it. *Utilities Policy 35*, 1–18.

Rosendahl, K. E. (2019). EU ETS and the new green paradox. *Working Papers No. 2/ 2019*.

Rubin, J. D. (1996). A Model of Intertemporal Emission Trading, Banking, and Borrowing. *Journal of Environmental Economics and Management 31*(3), 269–286.

Salant, S. W. (2016). What ails the European Union's emissions trading system? *Journal of Environmental Economics and Management 80*, 6–19.

Schennach, S. M. (2000). The Economics of Pollution Permit Banking in the Context of Title IV of the 1990 Clean Air Act Amendments. *Journal of Environmental Economics and Management 40*(3), 189–210.

Schopp, A., W. Acworth, D. Huppmann, and K. Neuhoff (2015). Modelling a market stability reserve in carbon markets.

Schopp, A. and K. Neuhoff (2013). The Role of Hedging in Carbon Markets. *DIW Discussion Papers 1271*, 24.

Shleifer, A. and R. W. Vishny (1997). The Limits of Arbitrage. *The Journal of Finance 52*(1), 35–55.

Silbye, F. and P. B. Sørensen (2018). Towards a more efficient European carbon market.

Staum, J. (2007). Chapter 12 Incomplete Markets. In *Handbooks in Operations Research and Management Science*, Volume 15, pp. 511–563.

Trück, S. and R. Weron (2016). Convenience Yields and Risk Premiums in the EU-ETS-Evidence from the Kyoto Commitment Period. *Journal of Futures Markets 36*(6), 587–611.

UBA (2014). Entwicklung der spezifischen Kohlendioxid-Emissionen des deutschen Strommix in den Jahren 1990 bis 2013. Technical report.