On the way to climate neutrality?

Intertemporal market and regulatory failures in the European Union's Emission Trading System and beyond

vorgelegt von
M. Sc.
Oliver Tietjen

an der Fakultät VI – Planen Bauen Umwelt
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften
– Dr. rer. oec. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Johann Köppel
Gutachter: Prof. Dr. Ottmar Edenhofer
Gutachter: Prof. Dr. Grischa Perino

Tag der wissenschaftlichen Aussprache: 22. Februar 2022

Berlin 2022

Summary

Since the implementation of the European Union's (EU) Emission Trading System (ETS) in 2005, the number of such cap-and-trade programs is on the rise globally. However, experience indicates that these programs often do not work as economic textbook theory suggests. In this thesis, underlying failures, their implications and potential solutions are analyzed.

First, an extensive literature review is carried out that alludes to regulatory risk and financial market failures as important price drivers in the EU ETS. As both have a bearing on the discount rate applied by market agents, the discount rate is at the core of this thesis. Using a simple equilibrium model, it is shown that a too-high discount rate applied to allowance banking leads to inefficient high carbon damage during the transition to emission neutrality. In this case, the waterbed effect of overlapping policies is reconsidered: The widespread and often criticized policy mix of ETS and subsidies for clean energy can be welfare-enhancing compared to an ETS-only policy.

As it reflects the cost of capital, the discount rate also affects the technology costs. Therefore, this thesis highlights that if the monetary policy becomes stricter and the discount rate rises accordingly, the trend of falling renewable costs may be reverted. As a result, investments in clean technologies can be significantly postponed. In turn, this puts the durability of the ETS cap at risk, because emission-intensive firms can lobby more effectively for a weaker climate policy if the transformation of the economy is concentrated towards the end of an ETS program.

This work also sheds light on financial market failures affecting the discount rate. Relying on a theoretical equilibrium model, it is shown that such frictions can have a significant impact on the ETS price path via risk premiums incorporated in the discount rate. In this context, this thesis examines the Market Stability Reserve (MSR), an instrument aiming to stabilize the EU ETS. Although the MSR has a more positive effect on the short-term price when the frictions are considered, the MSR does not overcome the more fundamental problems of the EU ETS in the long term. An analysis with the detailed model LIMES-EU further reveals which MSR parameters are of greater importance for the achievement of long-term emission targets.

Overall, a central result of this thesis is that free intertemporal allowance banking causes inefficiencies and puts the transition to a low-carbon economy at risk. The MSR is not an appropriate solution; instead, the EU ETS and other cap-and-trade programs should be complemented by a price-responsive allowance supply, such as a price floor.

Zusammenfassung

Seit der Einführung des Emissionshandelssystems (ETS) in der Europäischen Union (EU) im Jahr 2005, ist die Zahl solcher Cap-and-Trade-Programme weltweit gestiegen. Die Erfahrungen zeigen jedoch, dass diese Programme oft nicht der ökonomischen Theorie entsprechend funktionieren. In der vorliegenden Arbeit werden mögliche Versagen, deren Auswirkungen und Lösungen analysiert.

In einem ersten Schritt wird ein Literaturüberblick erstellt, welcher auf regulatorisches Risiko und Finanzmarktversagen als wichtige Preistreiber im EU-ETS hindeutet. Beides beeinflusst die von Marktteilnehmenden verwendete Diskontrate, weshalb diese Rate im Zentrum dieser Dissertation steht. Mittels eines einfachen Gleichgewichtsmodells wird gezeigt, dass eine zu hohe Diskontrate zu einem ineffizient hohen Klimaschaden während des Übergangs zur Emissionsneutralität führt. Für diesen Fall wird der Wasserbetteffekt sich überlappender Politiken neu bewertet: Die weit verbreitete und oft kritisierte Kombination aus ETS und Subventionen für emissionsfreie Energie kann im Vergleich zu einer reinen ETS-Politik wohlfahrtssteigernd sein.

Über die Kapitalkosten wirkt sich die Diskontrate auch auf die Technologiekosten aus. Diese Arbeit weist darauf hin, dass bei einer Erhöhung des Zinssatzes der Zentralbanken, und des damit verbundenen Anstiegs der Diskontrate, der Trend sinkender Kosten für erneuerbare Energien umgekehrt werden kann. Infolgedessen können Investitionen in klimaneutrale Technologien aufgeschoben werden. Dies gefährdet wiederum die Stabilität des ETS-Deckels, da emissionsintensive Unternehmen wirksamer für eine schwächere Klimapolitik lobbyieren können, wenn sich die Dekarbonisierung der Wirtschaft am Ende des ETS-Programms konzentriert.

Weiterhin werden in dieser Dissertation Finanzmarktversagen betrachtet, die sich auf die Diskontrate auswirken. Mit Hilfe eines theoretischen Gleichgewichtsmodells wird gezeigt, dass solche Versagen erhebliche Auswirkungen auf den ETS-Preispfad haben können, weil sie durch Risikoprämien die Diskontrate verändern. In diesem Zusammenhang wird auch die Marktstabilitätsreserve (MSR) untersucht – ein Instrument zur Stabilisierung des EU-ETS. Obwohl die MSR unter Berücksichtigung der Finanzmarktversagen kurzfristig einen positiveren Preiseffekt aufweist, überwindet die MSR langfristig nicht die grundlegenden Probleme des EU-ETS. Eine Analyse mit dem detaillierten Modell LIMES-EU zeigt darüber hinaus, welche MSR-Parameter für die Erreichung langfristiger Emissionsziele von groSSer Bedeutung sind.

Zusammenfassend ist ein zentrales Ergebnis dieser Dissertation, dass der unbeschränkte intertemporale Zertifikatehandel Ineffizienzen verursacht und den Übergang zu einer klimaneutralen Wirtschaft gefährdet. Die MSR ist keine geeignete Lösung für die genannten Probleme. Stattdessen sollte das EU-ETS wie auch andere Cap-and-Trade-Programme durch ein preisabhängiges Angebot von Zertifikaten, etwa durch einen Minimalpreis, ergänzt werden.

List of papers

This thesis includes nine chapters, of which Chapters 2 to 8 are based on individual research papers. Chapter 2 and Chapters 4 to 8 are the result of collaborations between the author of this thesis and colleagues.

Chapter 2 is based on Friedrich, M., Mauer, E.-M., Pahle, M., Tietjen, O. (2020). From fundamentals to financial assets: the evolution of understanding price formation in the EU ETS. Econstor working paper, `http://hdl.handle.net/10419/225210`.

Chapter 3 is based on Tietjen, O. (2020). Reducing the cost of delay: on the interaction of cap-and-trade and subsidies for clean energy. SSRN, `http://dx.doi.org/10.2139/ssrn.3580673`.

Chapter 4 is based on Schmidt, T. S., Steffen, B., Egli, F., Pahle, M., Tietjen, O., Edenhofer, O. (2019). Adverse effects of rising interest rates on sustainable energy transitions. Nature Sustainability, Vol. 2, pages 879-885, `https://doi.org/10.1038/s41893-019-0375-2`.

Chapter 5 is based on Pahle, M., Tietjen, O., Osorio, S., Egli, F., Steffen, B., Schmidt, T. S., Edenhofer, O. (2020). The risk of softening the cap in emissions trading systems. Mimeo.

Chapter 6 (preprint) is based on Tietjen, O., Lessmann, K., Pahle, M. (2021). Hedging and temporal permit issuances in cap-and-trade programs: The Market Stability Reserve under risk aversion. Resource and Energy Economics, Vol. 63, `https://doi.org/10.1016/j.reseneeco.2020.101214`.

Chapter 7 (preprint) is based on Osorio, S., Tietjen, O., Pahle, M., Pietzcker, R., Edenhofer, O. (2021). Reviewing the Market Stability Reserve in light of more ambitious EU ETS emission targets. Energy Policy, Vol. 158, `https://doi.org/10.1016/j.enpol.2021.112530`.

Chapter 8 (preprint) is based on Flachsland, C., Pahle, M., Burtraw, D., Edenhofer, O., Elkerbout, M., Fischer, C., Tietjen, O., Zetterberg, L. (2020). How to avoid history repeating itself: the case for an EU Emissions Trading System (EU ETS) price floor revisited. Climate Policy, Vol. 20 (1), pages 133-142, `https://doi.org/10.1080/14693062.2019.1682494`.

Contents

Chapter 1

Introduction

Over the past 15 years, the number of carbon pricing policies to address climate change has risen considerably (World Bank 2020). Among the two prototypes of carbon pricing, emission trading systems (ETS) and carbon taxes, the former alone will cover about 14% of global greenhouse gas (GHG) emissions when the Chinese national program begins its operation in 2020-21 (ICAP 2020).

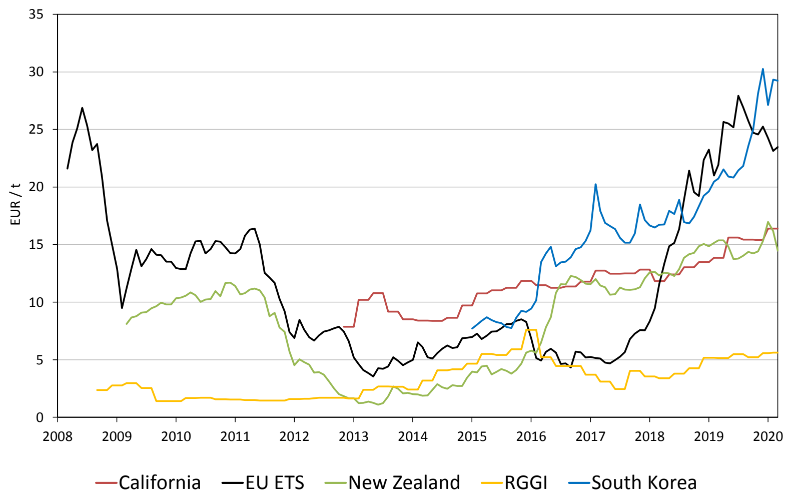
The European Union's (EU) ETS was implemented in 2005 and is the longest-lived program to regulate GHG emissions. The EU ETS is a classical cap-and-trade program: policymakers issue tradable certificates[1] that allow emission of one ton of carbon dioxide equivalent. Each year, the number of issued allowances declines such that emissions eventually should go down as well. This principle was copied in similar forms in several jurisdictions around the globe, including New Zealand, South Korea and the Regional Greenhouse Gas Initiative (RGGI), which consists of several US states (ICAP 2020).

The EU ETS covers more than 40% of the EU's emissions and is the cornerstone of EU climate policy (European Parliament and Council of the European Union 2018). Currently, the possibility of raising the 2030 emission reduction target to 55% from 40% relative to 1990 is being discussed. In this context, the scope of the ETS will possibly be extended to more sectors (European Commission 2019; European Parliament 2020). Therefore, the ETS can be considered the major instrument for the EU's aspiration to reach climate neutrality by 2050 (European Commission 2018; European Parliament 2019; European Council 2019).

In this thesis, I consider transitional challenges of the EU ETS – and cap-and-trade programs in general – on the path to long-term decarbonization. In particular, I investigate the intertemporal dimension of emission trading from an economic perspective. The EU ETS, such as many other cap-and-trade programs, allows firms to store unused certificates for future compliance. Such

[1] The terms "certificates," "allowances" and "permits" are used synonymously in this thesis.

Figure 1.1: Allowance prices in selected emission trading systems



Note: RGGI stands for Regional Greenhouse Gas Initiative, which includes the following US states: Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island and Vermont. Source: ICAP, https://icapcarbonaction.com/en/ets-prices (accessed on 30/07/2020)

intertemporal banking implies that the price between periods is connected via the discount rate according to Hotelling's rule. As a result, ambitious emission targets in the long term should already be reflected in today's prices.

However, the starting point of this thesis is the constantly low price level of the EU ETS from 2011 to 2018 (see Figure 1.1) despite ambitious long-term goals, which can also be observed in other cap-and-trade programs. The low price level created a debate about the functioning of the EU ETS (cf. Ellerman et al. 2016): Some argue that the EU ETS works as intended and the price merely reflects low abatement costs, for example, due to a faster expansion of renewable energies and lower emissions caused by the economic crisis. Others argue that the price is too low to incentivize investments for a more profound decarbonization in the long term (Fuss et al. 2018).

In this thesis, I explore regulatory and market failures that may explain the low price level. I consider the implications of such failures concerning the efficiency and effectiveness of ETS, and I examine instruments that intend to overcome such problems. My central thesis is that free intertemporal allowance banking in cap-and-trade programs, such as the EU ETS, is a problematic feature causing inefficiencies and putting the transition to a low carbon economy at risk. The recently implemented Market Stability Reserve (MSR) does not heal the EU ETS; instead, the MSR should be substituted or complemented by a price support mechanism (e.g., a price floor).

A first step of this thesis is to explore the actual price drivers of the EU ETS based on an extensive review of the econometric literature (Chapter 2). The literature suggests that abatement costs alone can hardly explain the price movements and alludes to regulatory and market failures. In particular, financial market imperfections and regulatory risk seem to be important price drivers, both of which are related to the discount rate applied to intertemporal allowance banking. The discount rate is a crucial variable for the functioning of ETS from an intertemporal perspective as it determines the growth rate of the allowance price, and thus, the discount rate is at the center of this thesis.

In the second step, I therefore examine regulatory and market failures that may explain why the discount rate applied to allowance banking is not optimal. In this regard, I consider a first failure of ETS with free intertemporal allowance banking in Chapter 3: Since the allowance price grows at the discount rate, but the optimal carbon price grows at a lower rate, unconstrained banking is generally not efficient (Pizer and Prest 2020). I show that the widespread and often criticized policy mix of ETS and subsidies for clean energy (Böhringer and Rosendahl 2010; Fankhauser et al. 2010) can indeed be welfare-superior compared to an ETS-only policy.

Furthermore, as the discount rate also reflects the cost of capital, it has a significant effect on the costs of renewable energy (Hirth and Steckel 2016; Schmidt 2014). The monetary policy of central banks thus can have a significant effect on renewable costs because it determines the general interest rate level, which is part of the discount rate applied by firms. Therefore, we find in Chapter 4 that a potentially tighter monetary policy may outweigh the cost reductions for renewable energies through technological advancements. Based on the large-scale numerical model LIMES-EU, we show in Chapter 5 that a higher discount rate significantly postpones abatement efforts in the EU ETS. We argue that the resulting concentration of abatement in the mid and end of the transition to climate neutrality may lead to a softening of the ETS cap, as emission-intensive firms may lobby for a looser cap.

Another failure considered in Chapter 6 deals with financial market imperfections as a source for an inefficient discount rate applied to allowance banking. More specifically, frictions in futures markets lead to a risk premium becoming part of the discount rate. We show theoretically how the risk premium is affected by the size of the allowance bank and hedging demand of the firms and thus may cause a time-varying growth rate of the ETS price.

In the third step of this thesis, instruments to improve the functioning of the EU ETS are analyzed. The EU considers the prevailing large allowance bank as the reason that prevents "the EU ETS from delivering the necessary investment signal [...] in a cost-efficient manner" (European Parliament and Council of the European Union 2015). As as result, the EU implemented the MSR, which reduces the allowance bank. The MSR caused a price surge after it was tightened and amended by an allowance cancellation mechanism by the end of 2017 (see Figure 1.1). We show in Chapter 6 that the price increase can indeed be explained to a certain degree by the reduction of the allowance bank through the MSR if hedging value of allowance for firms is considered. However, because the hedging demand reduces the growth rate of the allowance price, the price may not rise or even decline for a considerable time.

The MSR is examined further in Chapter 7, where we focus on the allowance cancellation mechanism. We show that certain MSR parameters and the discount rate have a strong effect on cancellations and therefore achieving specific emission targets becomes more complicated due to the MSR. Overall, we argue in Chapter 8 that the MSR is not a suitable instrument to cure the EU ETS. Instead, we suggest implementing a price floor, because it improves the long-term credibility of the EU ETS, reduces price uncertainty and, in doing so, creates a market environment that fosters clean investments.

The remainder of this chapter is organized as follows. In Section 1.1, I lay out the theoretical foundation of emission trading and discuss regulatory and market failures and their implications, maintaining a focus on the intertemporal dimension. Section 1.2 deals with ETS design features that may overcome intertemporal failures, namely a price-responsive allowance supply and the MSR. Section 1.3 provides an outline and the objectives of this thesis, where I introduce and connect Chapters 2 to 8 in more detail and briefly touch upon the methods used.

1.1 Emission trading as an intertemporal market: foundations, failures and their implications

From an economic perspective climate change is an externality issue. Ignoring other market failures, the problem can be solved by a carbon price set equal to the marginal social damage of emissions. An ETS and a Pigouvian tax are the two idealized instruments favored by economists in order to establish a carbon price. In this thesis, the focus is on ETS but only those ETS featuring intertemporal trading of allowances. Such allowance banking provides firms temporal flexibility for their abatement efforts and is a standard attribute of actually established programs such as the EU ETS (ICAP 2020).

In the following section, I outline the fundamentals of emission trading and discuss why free intertemporal trading is cost-effective (ignoring market failures), but generally not welfare-optimal. In this sense, intertemporally unconstrained allowance trading can be considered a regulatory failure. In Section 1.1.2, I elaborate on further regulatory and market failures that have a bearing on the price path, and Section 1.1.3 deals with the implications of such failures.

1.1.1 Cost-effective vs. efficient emission trading

The foundations of emission trading can be traced back to Coase (1960), who argues that bargaining in markets is an efficient way to deal with externalities if transactions costs are sufficiently low and property rights are well defined. This concept was later transferred to markets in rights to pollute the environment (Crocker 1966; Dales 1968; Montgomery 1972), which is the basis of modern emission trading. The core idea is that the state issues a certain amount of allowances that permit the pollution of the environment. For example, in the EU ETS, one such certificate (European Union Allowances, EUA) allows emission of one ton of carbon. Therefore, the total number of issued allowances effectively is a cap on emissions.

In an idealized market, the tradability of allowances implies that the marginal abatement costs of regulated entities are equalized. Market agents buy allowances as long as the allowance price is below their marginal abatement costs and vice versa such that in equilibrium the allowance price equates the marginal abatement costs of all regulated entities. As a result, the emission cap is achieved at lowest possible cost (Montgomery 1972). That is, in the absence of further market failures, emission trading is a cost-effective tool to achieve a given emission target.

Cronshaw and Kruse (1996) and Rubin (1996) extend emission trading to a dynamic setting by allowing market agents to store allowances for future use. If a firm holds more allowances than it requires to comply with the periodic (e.g., annual) cap, the surplus can be banked.[2] This gives firms greater flexibility because they can spread their abatement efforts over time. Specifically, firms exploit intertemporal arbitrage and thus equalize the discounted marginal abatement costs across periods. The allowance price therefore grows at the rate of interest, which is called a Hotelling price path.

An implication is that emission trading is also cost-effective in an intertemporal sense: the overall certificate budget is efficiently allocated over time such that the budget is achieved at the lowest possible costs. This requires that the "availability condition" is satisfied (Perino and Willner 2016; Salant 2016). This is the case if allowances are temporally issued in such a way that the allowance bank is always positive until the total allowance budget is exhausted. Otherwise the binding borrowing constraint would disrupt the optimal allocation of the emission budget over time.

[2] Rubin (1996) also considers allowance borrowing from the future, which, however, is usually prohibited in real ETS.

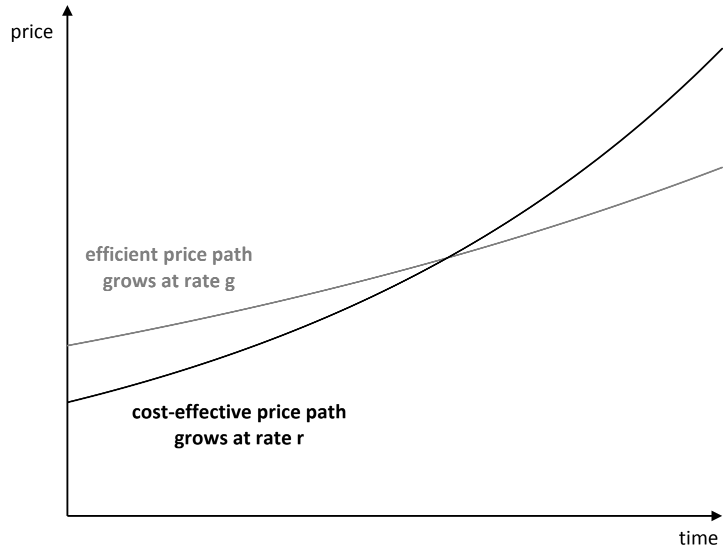
However, cost-effectiveness does not imply an efficient outcome if carbon damage is considered. A robust result from the social cost of carbon (SCC) literature is that the growth rate of the optimal carbon price (i.e., the SCC) is lower than the interest rate (e.g., Rezai and Van der Ploeg 2016; van den Bijgaart et al. 2016; Dietz and Venmans 2019). Golosov et al. (2014) are the first to show that the growth rate of the optimal carbon price is equal to the growth rate of the economy g. At the same time, the interest rate r is determined by the Ramsey-rule in neoclassical growth models: $r = \rho + \eta g$ with ρ for the pure rate of time preferences and η for the elasticity of marginal utility. In a standard Ramsey-type model, it must hold $r > g$; otherwise, the transversality condition would not be fulfilled (Acemoglu 2009). In the estimates of the Interagency Working Group on the Social Cost of Greenhouse Gases (2016), for instance, the interest rate is about twice as high as the growth rate of the SCC.

As a result, the price in ETS programs with free intertemporal trading, although cost-effective, is inefficient because it is too steep as it grows at the rate of interest r. For a given emission budget, the cost-effective allowance price is lower than the efficient carbon price at the beginning of the ETS program, as depicted in Figure 1.2. Therefore, emissions accumulate faster than optimal, which causes inefficient high damage until the allowance budget is exhausted.

Theoretically, the efficient price path including carbon damage can be achieved with emission trading. A straightforward solution would be to prohibit banking and instead implement the optimal periodical caps. However, banking has welfare advantages under uncertainty as shocks spread over to more periods, which stabilizes the price (Fell et al. 2012; Weitzman 2020). Alternatively, allowance banking can be enhanced by intertemporal trading ratios (Kling and Rubin 1997; Leiby and Rubin 2001; Yates and Cronshaw 2001). These ratios adjust the amount of banked allowances by a certain factor and thereby correct the growth rate of the allowance price. For example, if firms get 1.05 allowances in $t+1$ for each allowance banked in t, they require a 5% lower interest rate for allowances and the growth rate of the price declines accordingly. However, intertemporal trading ratios are complex because they alter the size of the cap and thus need to be aligned with the number of issued allowances. Moreover, in the presence of uncertainty the intertemporal trading ratio and the number of issued allowances need to be adapted after a shock occurs in order to achieve the optimal solution (Pizer and Prest 2020). Probably because of the high complexity, but maybe also because countries often consider climate change a cost-effectiveness problem, trading ratios hardly play a role in actual policy.

Instead, ETS programs are often combined with other policies, such as clean energy subsidies. Although additional measures can in principle be justified by further market failures, for example innovation and network effects (Fischer

Figure 1.2: Cost-effective and efficient carbon price paths



and Newell 2008; Jaffe et al. 2005), they are often considered to be "overlapping" (Böhringer et al. 2008), implying that they distort the allowance price. This is called the waterbed effect because for given emissions (cap), overlapping policies only change the temporal distribution of emissions (Perino 2018). In contrast to this view, I show in Chapter 3 that such a temporal shift of emissions due to overlapping policies can in fact be welfare-enhancing. By postponing emissions, additional measures bring the abatement path closer to the optimum and reduce the cost of excess carbon damage in the case of free intertemporal allowance trading.

1.1.2 Regulatory and market failures affecting the price path

In the context of the EU ETS, several regulatory and market failures are discussed to be responsible for the low price level until 2018. Although the price has recovered since then, the failures are unlikely to be settled and may still cause inefficiencies. In the following, I outline the main failures – again with a focus on intertemporal aspects.[3]

A widespread concern is that market agents do not take the long run properly into account because they are myopic or have a truncated planning horizon

[3]The presented list is non-exhaustive. For example, market power (Hintermann 2017) and the allowance allocation method (Baldursson and von der Fehr 2004; Böhringer and Lange 2005) are other often discussed issues.

(Fuss et al. 2018; Quemin and Trotignon 2019; Perino and Willner 2019; Willner 2018). Because in the far future the allowance supply is typically lower, but the marginal abatement costs are higher, shortsightedness implies that firms want to bank fewer allowances. That is, they use too large a share of the overall allowance budget in the near term because they do not take high costs in the future into account. As a result, the allowance price is too low early on and eventually needs to rise to high levels later to comply with the cap. A similar price path may be caused by regulatory risk, as shown by Salant (2016): If there is a risk of a downward price jump at an unknown time in the future due to a regulatory intervention, the price rises too fast. That is, in the case of myopia and regulatory risk, the price path can be bent more than the cost-effective path in Figure 1.2, inducing a too-low price level initially and a too-high price level later.

Financial market failures, and in particular, incomplete markets for risk are another reason why the allowance price may not rise at the rate of interest. If firms are risk averse, they can bank allowances to hedge their profits. Schopp and Neuhoff (2013) argue that when the overall allowance bank in the EU ETS exceeds the hedging demand of firms, speculators hold the excess allowances, for which they require a high return. In consequence, the allowance price must rise at a high rate, which again leads to low prices early on. In Chapter 6, we also relate the size of the bank to the allowance price, but explicitly build on the financial economics literature that analyzes the interaction of producers and speculators in commodity markets (e.g., Anderson and Danthine 1979; Bessembinder and Lemmon 2002; Ekeland et al. 2019). In contrast to Schopp and Neuhoff (2013), this hedging pressure theory explicitly considers uncertainty and the desire of producers to hedge risky profits by trading with speculators. In this case, producers have to pay a risk premium to speculators that has a significant effect on the allowance price path. The root cause of risk premiums due to hedging are financial market frictions that prevent an ideal allocation of risk (Hirshleifer 1990), for which we consider capital constraints for speculators similar to Acharya et al. (2013). Our econometric literature review in Chapter 2 alludes to further financial market failures: behavioral finance aspects such as herding behavior or under- and overreaction to new information may cause inefficiencies.

Another driver of the allowance price path is the monetary policy of central banks. While monetary policy itself is certainly not a failure, it is triggered by failures like rigid prices, as in Keynesian models (Walsh 2010). Beginning with the financial crisis in 2007-08, the European Central Bank, similar to other central banks, such as the Federal Reserve in the US, introduced a very loose monetary policy in order to stimulate the economy. As a result, the interest rate level dropped significantly, which caused two effects considered in this thesis: first, a lower interest rate level affects abatement costs because it reduces financing costs. This is of importance as renewable energies exhibit high capital costs, and accordingly financing costs play a significant role. Egli et al. (2018) show that the declining interest rate level was indeed an important reason for the rapidly falling renewable energy costs in the electricity sector in the past. Yet, the interest level may rise again in the coming years with a tighter monetary policy. In Section 4 of this thesis, we show that this can reverse the trend of falling renewable energy costs and thus slow down the transition. The second effect of monetary policy considered in this thesis bears on the growth rate of the allowance price. A potentially tighter monetary policy in the coming years may increase the interest rate level. In turn, the allowance price path would become steeper.

1.1.3 Implications of intertemporal failures

The failures described in the two previous sections have in common that they have a bearing on the discount rate applied to allowance banking. The discount rate, in turn, affects the level and growth rate of the allowance price, leading to the following implications.

First, a too-high discount rate causes too low allowance prices early on and thus postpones abatement. As a result, emissions accumulate too fast in the atmosphere, implying excessive carbon damage during the transition to climate neutrality (Section 1.1.1). This also holds absent market failures and is a direct consequence of free intertemporal trading, because the allowance price grows at a higher rate than the external costs of carbon. Regulatory and market failures may exacerbate this problem by further increasing the discount rate due to regulatory risks or incomplete risk markets.

Second, by distorting the allowance price, such failures also impair the cost-effectiveness of reaching a given carbon budget (Section 1.1.2). More specifically, they violate the condition required to minimize abatement costs: If the allowance price does not rise at the discount rate, (discounted) marginal abatement costs are not equalized across periods. Furthermore, too low prices early on, as observed in the EU ETS, may incentivize investments in emission-intensive capital with a long lifetime. This may lock in emission-intensive infrastructures such that switching to cleaner technologies becomes more costly in the future (Fuss et al. 2018). Postponed action also raises adjustment costs because it leads to a concentration of abatement in a relatively short time at the end of the cap-and-trade program (Vogt-Schilb et al. 2018).

A third implication of insufficient action in the beginning is an increased risk that the ETS policy will be dismantled. Based on political science, we argue in Chapter 5 that too steep an allowance price path may cause negative policy feedback on the persistence of the cap of the emission trading program.

When the allowance price is too low early on, the incumbent brown (emission-intensive) coalition remains powerful for a relatively long time, because they still own large market shares. At the same time, the new green (emission-free) coalition hardly grows in the beginning. Under such conditions, a fast rising price in the course of the emission trading program may not be viable as the still powerful brown coalition lobbies for a softer cap. Also, the incentives to lobby for a softer cap are large if abatement is concentrated at the end due to high adjustment costs and stranded assets (Rozenberg et al. 2020).

Lastly, in the face of low ETS prices, policy makers may believe the ETS is ineffective and does not work properly. This may encourage them to implement further overlapping policies. While this can in principle be a reasonable second-best option (see Section 3), it contains the risk that the ETS price is further depressed due to the waterbed effect. A fourth possible implication is therefore that the credibility of the ETS is undermined due to continuous policy changes (Fankhauser et al. 2010). Specifically, low prices motivate the implementation of additional policies, which decrease prices even further. Such a feedback effect endangers the credibility of the cap (Pahle et al. 2018).

1.2 Enhancing emission trading systems

In this section, I introduce two additions to ETS intended to tackle the problems described in the previous section. Both have in common that they adjust the supply of allowances, but they differ in how the adjustment is triggered: The supply is either price- or quantity-responsive. While the first transforms emission trading schemes into a classical hybrid instrument and is implemented in ETS in the US, the latter is implemented as MSR in the EU ETS.

1.2.1 Price-responsive allowance supply

While a pure ETS fixes the quantity (emissions) and leaves the price to the market, a tax fixes the price and leaves the quantity to the market. A fundamental insight of Weitzman (1974) is that both lead to the same outcomes under certainty, but differ under uncertainty. The tax is dominant when the slope of the marginal benefit curve of abatement is flat, and the quantity regulation is superior when the slope is steep. Both can be viewed as two extremes of a market-based regulation, where a pure ETS has a completely inelastic allowance supply and a tax has (implicitly) a completely elastic allowance supply. Weitzman (1974) realized that a combination of price (tax) and quantity (ETS) regulation should be superior compared to either of them in isolation, which was later shown by Roberts and Spence (1976) and Weitzman (1978). In the context of climate change, such a hybrid policy is typically considered to be a price-responsive allowance supply in an ETS. In the following, I present different forms of a price-responsive allowance supply and discuss why it can help to overcome the problems described in the previous section.

Pizer (2002) was among the first to analyze a hybrid policy to address climate change. He shows that adding a price ceiling as a safety valve to an ETS strongly increases welfare because it avoids high costs if abatement turns out to be more costly than expected (see Webster et al. 2010 for a similar result). The price ceiling works by an unlimited allowance supply if the ETS price reaches the ceiling, and thus, the allowance supply becomes price responsive. Burtraw et al. (2010) show that a one-sided price ceiling reduces the expected price if the ceiling binds with a positive probability and since it also does not insure against low prices, implementing only a price ceiling has adverse effects on abatement. They argue that a hybrid policy should consist of a symmetric safety valve with a price ceiling and floor and find that this significantly increases welfare (see also Fell and Morgenstern 2010).[4] A price floor introduces another step in the allowance supply function, as the allowance supply is reduced when the price falls below the floor.

In contrast to hard price ceilings and floors that enforce that the price is never higher or lower, Murray et al. (2009) propose to limit the supply of additional allowances when the price ceiling is met. In addition to such a soft price ceiling, Fell et al. (2012) analyze a soft price collar in which price support is limited from both sides. They find that a soft collar allows compromise between a hard collar (lower abatement costs) and no collar (emission target is exactly met).[5] Such price collars are a design feature, for example, in California and RGGI in the US (ICAP 2020). Price floors are usually implemented as auction reserve prices, which limit the supply reduction to the number of auctioned allowances. Price ceilings are typically soft because the additional supply is limited by allowances available from a reserve. Burtraw et al. (2018) go a step further and suggest making the allowance supply more responsive by a step-wise or even continuous supply curve. A similar approach is proposed by Traeger et al. (2020) who, however, want to ban allowance banking at the same time.

Overall, the literature suggests that hybrid instruments exhibit significant efficiency advantages compared to a pure quantity-based ETS. If the regulatory and market failures described in Section 1.1 are considered as well, a price collar has further advantages. For one, inefficiently low prices can be avoided or at least increased and thus, the problem of delayed abatement is mitigated: A price floor reduces excess damages early on and enhances investment incentives

[4] Abrell and Rausch (2017) and Abrell et al. (2019) additionally show that a price floor can increase welfare if there are two sectors, where one is regulated by an ETS (with price floor) and the other one is regulated separately.

[5] Relatedly, Grüll and Taschini (2011) suggest implementing a price collar via financial options as these allow keeping the environmental target.

in clean technologies. Clean investments are further facilitated by reduced investment risks, as the ETS price cannot drop below a hard price floor or is less sensitive to shocks when the allowance supply is price responsive. An increasing price floor or collar over time also signals the commitment of the regulator to the long-term emission targets, and therefore, may reduce the regulatory risk and, in doing so, improves the viability of the cap. Moreover, the waterbed effect is reduced as well with a price-responsive allowance supply, since the cap shrinks when the price drops as a result of overlapping policies. In Section 8, we elaborate in more detail on the advantages of a price floor concerning these problems.

1.2.2 Quantity-responsive allowance supply: the Market Stability Reserve

As a quantity-responsive supply is not a classical hybrid instrument, the literature on this instrument has only started to evolve with the implementation of the MSR in the EU ETS (European Parliament and Council of the European Union 2015). In this section, I outline how the MSR works and discuss whether it can solve (some of) the previously mentioned problems of the EU ETS.

The MSR makes the allowance supply quantity-responsive because the number of auctioned allowances depends on the size of the allowance bank held by market agents. The justification for such a measure is the EU's interpretation of the large allowance bank as a "structural supply-demand imbalance," which prevents "the EU ETS from delivering the necessary investment signal to reduce CO2 emissions in a cost-efficient manner" (European Parliament and Council of the European Union 2015). That is, the large bank has adverse effects and therefore is reduced by the MSR through the reduction of the allowance supply. Specifically, the originally planned version of the MSR works by reducing the allowance supply if the bank exceeds 833 Mt allowances. Instead of being auctioned, allowances equal to 12% of the bank of the previous year are put in the MSR, where they are not available for compliance. If the bank is lower than 400 Mt, 100 Mt allowances per year are released from the MSR and additionally auctioned. Only if the bank is between 400 and 833 Mt does the MSR not absorb or release allowances.

Perino and Willner (2016) find that this original cap-preserving MSR has only a limited price-increasing effect in the short term and a price-decreasing effect in the long term. Therefore, long-term investments in clean technologies might even be deterred (see also Perino and Willner 2019). Perino and Willner (2016) also show that the price volatility increases, similar to Mauer et al. (2019), Kollenberg and Taschini (2019) and Richstein et al. (2015), though Fell (2016) finds the opposite. Accounting for the firms' risk aversion, Kollenberg and

Taschini (2019) find that the increased volatility leads to higher risk premiums, implying lower short-term and higher long-term prices. However, in Chapter 6 we find the opposite: The short-term price increases and the long-term price decreases due to the MSR because of the hedging demand for allowances, which determines the risk premium in our approach.

Even before the MSR came into effect in 2019, it was reformed in early 2018 (European Parliament and Council of the European Union 2018). The major innovation in this new version of the MSR is the permanent cancellation of allowances from 2023, if more allowances are in the MSR than were auctioned in the previous year. This cancellation mechanism makes the cap endogenous and dependent on the size of the bank. While the cancellation mechanism clearly raises the price level because it reduces the cap, it is difficult to determine by how much. For one, the exact number of cancellations depends on the allowance bank level and the complex MSR mechanism. In Section 7, we provide an overview of cancellation estimates from the literature ranging from 1.7 Gt to 13 Gt, while we find 5.1 Gt in our own analysis. Given that the total cap (without MSR cancellations) for the period 2018 to 2057[6] is 40.1 Gt, the large range of cancellation results implies that the impact of the MSR on the ETS price is hard to foresee.

Moreover, several papers examine the question of how cancellations are affected by overlapping policies, including renewable energy support (e.g., Beck and Kruse-Andersen 2018; Carlén et al. 2019; Burtraw et al. 2018; Perino 2018). A main result is that the new MSR "punctures the waterbed" (Perino 2018), meaning that additional policies can reduce the cap now because they increase the allowance bank level, which in turn triggers more cancellations by the MSR. However, Gerlagh et al. (2019) show that policies announced today but effective in the future may even increase the cap because they reduce the current bank and therefore fewer allowances are canceled. According to the analysis by Pahle et al. (2019), the German coal phase-out could be an example of such a paradoxical effect. Similarly, Perino et al. (2019) find that overlapping policies can raise the cap due to internal carbon leakage: Emission reductions of an unilateral policy in a country can be overcompensated if the policy triggers emissions in another (neighboring) EU ETS country due to more exports. As a result, the unilateral policy reduces the total allowance bank, implying fewer cancellations. Moreover, Bruninx et al. (2019) show that the MSR cancels more allowances when abatement costs are high and vice versa, which is at odds with economic theory that suggests the opposite to enhance welfare. Specifically, an efficient hybrid instrument (price collars, see above) reduces the cap (more cancellations) when abatement costs are low and vice versa.

[6] 2057 is the year when allowances are issued for the last time, according to current regulation.

In total, the literature and our own analyses cast doubt on whether the MSR can tackle the ETS problems described above. Although the MSR has a positive price effect, implying less climate damage and more clean investments, our analysis in Chapter 6 shows that the expected price can also drop again if the hedging demand of firms is considered. Moreover, we show in Chapter 7 that the potential cancellation of allowances has a huge range, which creates additional risks for market agents, but also for policy makers as it becomes more difficult to achieve desired emission targets. As the MSR also does not adequately address the waterbed effect of overlapping policies, it is probably not a suitable instrument to solve the ETS problems. We discuss the MSR in more detail in Chapter 8.

1.3 Outline and objectives of this thesis

The outcomes of emission trading systems such as the EU ETS are often not in line with the standard cap-and-trade theory. The main objective of this thesis is to improve the understanding of the underlying inefficiencies, analyze implications and examine potential solutions, where I focus on the intertemporal dimension of ETS. The analysis is divided into seven research papers, which provide answers to the following main research questions:

Chapter 2: What drives the EU ETS price?

Chapter 3: What are the implications of a discount rate that is too high for the instrument choice to regulate the climate externality? Can overlapping policies be justified in this case, and how should they be designed?

Chapter 4: How does the discount rate affect the abatement costs, and what is the role of monetary policy?

Chapter 5: What determines the regulatory risk that the ETS cap is softened, and why is the discount rate important in this regard?

Chapter 6: How do incomplete risk markets and firm hedging affect the discount rate? What is the impact of the MSR given this failure?

Chapter 7: What MSR parameters are of importance for the number of canceled allowances? How should the linear reduction factor be adjusted to achieve more ambitious 2030 targets under consideration of the MSR cancellation?

Chapter 8: Overall, is the MSR a recommendable instrument to overcome the problems of the EU ETS, or should it be complemented or substituted by a price floor?

The aim of Chapter 2 is to find actual price drivers of the EU ETS based on a review of the empirical literature. Summarizing the knowledge of this field indicates gaps in the understanding of the price determination serving as motivation and basis for further analyses carried out in this thesis. Specifically, the review shows that the price is affected by regulatory risk and financial market frictions. Since both affect the discount rate applied to allowance banking, the discount rate is a core element in the following chapters: In Chapters 3 to 6, regulatory and market failures affecting the discount rate and their implications are at the center, and Chapters 7 and 6 focus on the MSR and price floors as potential solutions.

In particular, I analyze the effect of an inefficiently high discount rate applied to allowance banking in Chapter 3, and show that the waterbed effect of overlapping policies can enhance welfare in this setting. For this purpose, I construct a simple theoretical and numerical model, which includes a representative dirty (emission-intensive) and clean (emission-free) firm competing in an energy market. A welfare-maximizing regulator wants to internalize the damage of a stock pollutant (carbon emissions) by choosing the optimal levels of different policy instruments. While a carbon tax implements the first-best solution, the ETS with free intertemporal trading fails to do so: The discount rate determines the growth rate of the allowance price, but the optimal carbon price equals the marginal social damage of carbon (the social cost of carbon), which grows at a lower rate than the discount rate. For a given cap, the too high discount rate implies that the allowance price is too low in the early phase of the ETS and too high at the end.

In this setting, I reconsider the waterbed effect of overlapping policies by showing that complementing the ETS with subsidies for clean energy is welfare-enhancing. The reason is that, for a given ETS cap, subsidies shift emissions to the future and thus, reduce the inefficient high carbon damage. Optimally set subsidies do not even have an effect on the ETS price if the regulator is able to commit. However, the time-consistent subsidy path lowers the ETS price, but still increases welfare compared to the ETS in isolation. I also find that subsidies mitigate the ETS price volatility as subsidies prolong the banking phase of the ETS such that shocks spread over more periods.

A further effect of a high discount rate is considered in Chapter 4: It raises the relative costs of renewable energy technologies like wind mills and solar photovoltaics compared to fossil fuel plants. The discount rate reflects the cost of capital for investments and because renewable energies exhibit a large capital share, their overall costs are strongly affected by the cost of capital. As central banks lowered the interest rate level in the aftermath of the financial crisis in 2007-08, renewable energy currently benefits from the low cost of capital. However, based on a simple spreadsheet model, we show that an increase of the interest rate back to the pre-crisis level could raise renewable energy costs in the coming years despite technological progress. As a result, the competitiveness of renewable energies would deteriorate and the transition to emission neutrality would slow down.

In Chapter 5, we go a step further and argue that high discount rates increase the risk that the ETS cap will be softened. Relying on economics and political science, we derive a political feedback effect the allowance price path could have on the risk that the cap is softened. For this purpose, we consider two scenarios, which we also quantify with the large-scale electricity and industry sector model LIMES-EU. In the first scenario, the discount rates applied to allowance banking and investments in renewable energy are low. In the second scenario, the discount rates are increased, which triggers the two above mentioned effects: a larger growth rate of the allowance price and higher renewable energy costs. The second scenario leads to a lower allowance price and less investments in renewable energies early on and a faster increasing allowance price over time. This more hockey stick-shaped price path induces a higher risk that the cap is softened, because it postpones the development of a green (emission-free) coalition that wants to keep the cap. In addition, the adjustment costs for the brown (emission-intensive) coalition are higher and concentrated in the late ETS phase. Thus, the brown coalition has more relative power and stronger incentives to lobby for a softer cap if the price is hockey stick-shaped.

Chapter 6 continues the analysis of discount rates in ETS by considering financial market frictions and risk-averse firms that want to hedge their profits. If speculators – as firms' trading counterparties – face liquidity constraints, hedging is costly, which is reflected by a risk premium in the allowance price. We construct a simple theoretical model with dirty coal and relatively clean gas firms whose hedging demands have opposing effects on the risk premium. In our numerical application to the electricity sector of the EU ETS, the risk premium is highly negative in the early phase but increases over time. In consequence, the price is higher but does not rise, or it even declines in the early ETS phase compared to the case without hedging demand. Because the risk premium becomes less negative over time, the price path may be U-shaped.

We also show that the size of the allowance bank affects the risk premium, because the dirty firm uses the bank to hedge its profits. This is of particular relevance in the EU ETS due to the MSR, which shifts the issuance of allowances to the future and thus, reduces the bank. In doing so, the MSR increases the hedging value of allowances reflected by a more negative risk premium. As a result, hedging implies that the short-term effect of the MSR on the price is more positive, but the long-term effect is more negative (compared to risk neutrality). In addition, hedging also leads to more cancellations by the MSR compared to the risk-neutral case, since hedging implies that firms want to hold a larger bank, which, in turn, leads to more influx into the MSR.

Chapter 7 also provides an analysis of the MSR, but with a focus on its envisaged review in 2021 and in view of more ambitious emission targets in the EU. Using the LIMES-EU model, we numerically examine the impact of several MSR parameters on the number of canceled allowances. We find that especially the allowance bank thresholds beyond which the influx in the MSR is triggered have a considerable impact on cancellations. In contrast, higher intake rates only weakly affect cancellations, but may increase uncertainty because high intake rates induce oscillatory behavior to intake volumes. In the second step, we show that reducing the cap by a higher linear reduction factor also raises MSR cancellations. However, this relationship, as well as the number of cancellations in general, are subject to large uncertainties, which implies that achieving specific emission targets becomes more complicated due to the MSR. In particular, the discount rate plays an important role because a high rate reduces cancellations significantly.

In Chapter 8 we argue based on the literature – including chapters of this thesis and other studies by the author of this thesis – and discussions in several workshops with stakeholders from academia, policy, industry and NGOs, that the MSR does not solve the problems of the EU ETS. The MSR does not reduce regulatory risk and may even increase price variability. It also does not solve the waterbed problem but, on the contrary, the MSR can even exacerbate it. As an alternative or in addition to the MSR, we propose to implement a price floor. We argue that the main advantage of a price floor is that it would improve the long-term investment environment because it enhances the regulators' commitment signal and reduces regulatory risk and price uncertainty in general. It also mitigates inefficiencies due to too low prices in the early phase of the ETS due to excessive discounting. Similarly, the price declining effect of overlapping policies can also be reduced by such a price-responsive allowance supply.

In Chapter 9, I summarize the main results of this thesis. In addition, I provide a discussion on the economic models used, as well as the novelty of the results and their policy relevance. Finally, I propose avenues for future research and provide an outlook on ETS as a policy instrument, in particular in the EU.

References

Abrell, J. and S. Rausch (2017). Combining price and quantity controls under partitioned environmental regulation. *Journal of Public Economics 145*, 226–242.

Abrell, J., S. Rausch, and H. Yonezawa (2019, apr). Higher Price, Lower Costs? Minimum Prices in the EU Emissions Trading Scheme. *The Scandinavian Journal of Economics 121*(2), 446–481.

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton, NJ, USA: Princeton University Press.

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.

Anderson, R. W. and J.-P. Danthine (1979). Hedging and Joint Production: Theory and Illustrations. *The Journal of Finance 35*(2), 487–498.

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.

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. *Working Paper 2018:2*.

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

Böhringer, C., H. Koschel, and U. Moslener (2008). Efficiency losses from overlapping regulation of EU carbon emissions. *Journal of Regulatory Economics 33*(3), 299–317.

Böhringer, C. and A. Lange (2005). On the design of optimal grandfathering schemes for emission allowances. *European Economic Review 49*(8), 2041–2055.

Böhringer, C. and K. E. Rosendahl (2010). Green promotes the dirtiest: On the interaction between black and green quotas in energy markets. *Journal of Regulatory Economics 37*(3), 316–325.

Bruninx, K., M. Ovaere, K. Gillingham, and E. Delarue (2019). The unintended consequences of the EU ETS cancellation policy. *MPRA Paper No. 96437*.

Burtraw, D., C. Holt, K. Palmer, A. Paul, and W. Shobe (2018). Quantities with Prices. *RFF working paper*.

Burtraw, D., A. Keyes, and L. Zetterberg (2018). Companion Policies under Capped Systems and Implications for Efficiency - The North American Experience and Lessons in the EU Context. *RFF Report*.

Burtraw, D., K. Palmer, and D. Kahn (2010, sep). A symmetric safety valve. *Energy Policy 38*(9), 4921–4932.

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

Coase, R. H. (1960, oct). The Problem of Social Cost. *The Journal of Law and Economics 3*, 1–44.

Crocker, T. (1966). Structuring of Atmospheric Pollution Control Systems. In H. Wolozin (Ed.), *The Economics of Air Pollution*, pp. 61–86. New York: W.W. Norton & Co.

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

Dales, J. (1968). *Pollution, Property and Prices: An Essay in Policy-making and Economics*. Toronto: University of Toronto Press.

Dietz, S. and F. Venmans (2019). Cumulative carbon emissions and economic policy: In search of general principles. *Journal of Environmental Economics and Management 96*, 108–129.

Egli, F., B. Steffen, and T. S. Schmidt (2018). A dynamic analysis of financing conditions for renewable energy technologies. *Nature Energy 3*(12), 1084–1092.

Ekeland, I., D. Lautier, and B. Villeneuve (2019). Hedging pressure and speculation in commodity markets. *Economic Theory 68*(1), 83–123.

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 (2018). A Clean Planet for all. A European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Soc. *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0773*.

European Commission (2019). The European Green Deal. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN*.

European Council (2019). European Council conclusions, 12 December 2019. *https://www.consilium.europa.eu/media/41768/12-euco-final-conclusions-en.pdf*.

European Parliament (2019). Climate change. European Parliament resolution of 14 March 2019 on climate change âĂŞ a European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy in accordance with the Paris Agreement. *https://www.europarl.europa.eu/doceo/document/TA-8-2019-0217_EN.html*.

European Parliament (2020). The European Green Deal. European Parliament resolution of 15 January 2020 on the European Green Deal. *https://www.europarl.europa.eu/doceo/document/TA-9-2020-0005_EN.html*.

European Parliament and Council of the European Union (2015). Decision (EU) 2015/1814. *Official Journal of the European Union 58*.

European Parliament and Council of the European Union (2018). Directive (EU) 2018/410. *Official Journal of the European Union 61*.

Fankhauser, S., C. Hepburn, and J. Park (2010). Combining Multiple Climate Policy Instruments: How not to do it. *Climate Change Economics 1*(3), 209–225.

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

Fell, H., D. Burtraw, R. D. Morgenstern, and K. L. Palmer (2012). Soft and hard price collars in a cap-and-trade system: A comparative analysis. *Journal of Environmental Economics and Management 64*(2), 183–198.

Fell, H., I. A. MacKenzie, and W. A. Pizer (2012). Prices versus quantities versus bankable quantities. *Resource and Energy Economics 34*(4), 607–623.

Fell, H. and R. D. Morgenstern (2010). Alternative Approaches to Cost Containment in a Cap-and-Trade System. *Environmental and Resource Economics 47*(2), 275–297.

Fischer, C. and R. G. Newell (2008). Environmental and technology policies for climate mitigation. *Journal of Environmental Economics and Management 55*(2), 142–162.

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.

Gerlagh, R., R. Heijmans, and K. E. Rosendahl (2019). Endogenous Emission Caps Always Induce a Green Paradox. *CESifo Working Paper No. 7862*.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica 82*(1), 41–88.

Grüll, G. and L. Taschini (2011). Cap-and-trade properties under different hybrid scheme designs. *Journal of Environmental Economics and Management 61*(1), 107–118.

Hintermann, B. (2017). Market Power in Emission Permit Markets: Theory and Evidence from the EU ETS. *Environmental and Resource Economics 66*(1), 89–112.

Hirshleifer, D. (1990). Hedging Pressure and Futures Price Movements in a General Equilibrium Model. *Econometrica 58*(2), 411–428.

Hirth, L. and J. C. Steckel (2016). The role of capital costs in decarbonizing the electricity sector. *Environmental Research Letters 11*(11), 114010.

ICAP (2020). *Emissions Trading Worldwide: Status Report 2020.* Berlin, Germany: International Carbon Action Partnership.

Interagency Working Group on the Social Cost of Greenhouse Gases (2016). *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866*. Washington, DC: Interagency Working Group on the Social Cost of Carbon.

Jaffe, A. B., R. G. Newell, and R. N. Stavins (2005). A tale of two market failures: Technology and environmental policy. *Ecological Economics 54*(2-3), 164–174.

Kling, C. and J. Rubin (1997). Bankable permits for the control of environmental pollution. *Journal of Public Economics 64*(1), 101–115.

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.

Leiby, P. and J. Rubin (2001). Intertemporal permit trading for the control of greenhouse gas emissions. *Environmental and Resource Economics 19*(3), 229–256.

Mauer, E.-M., S. J. Okullo, and M. Pahle (2019). Evaluating the performance of the EU ETS MSR. *Working paper*.

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

Murray, B. C., R. G. Newell, and W. A. Pizer (2009, sep). Balancing Cost and Emissions Certainty: An Allowance Reserve for Cap-and-Trade. *Review of Environmental Economics and Policy 3*(1), 84–103.

Pahle, M., D. Burtraw, O. Tietjen, C. Flachsland, and O. Edenhofer (2018). Preserving the integrity of the EU ETS under unilateral action. *Paper presented at the WCERE 2018 in Gothenburg.*.

Pahle, M., O. Edenhofer, R. Pietzcker, O. Tietjen, S. Osorio, and C. Flachsland (2019). Die unterschätzten Risiken des Kohleausstiegs. *Energiewirtschaftliche Tagesfragen 69*(6), 31–34.

Perino, G. (2018). New EU ETS Phase 4 rules temporarily puncture waterbed. *Nature Climate Change 8*(4), 260–261.

Perino, G., R. A. Ritz, and A. van Benthem (2019). Understanding Overlapping Policies: Internal Carbon Leakage and the Punctured Waterbed. *NBER Working Paper No. 25643*, 26.

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 (2019). Rushing the Impatient: Allowance Reserves and the Time Profile of Low-Carbon Investments. *Environmental and Resource Economics 74*(2), 845–863.

Pizer, W. A. (2002). Combining price and quantity controls to mitigate global climate change. *Journal of Public Economics 85*(3), 409–434.

Pizer, W. A. and B. C. Prest (2020). Prices versus Quantities with Policy Updating. *Journal of the Association of Environmental and Resource Economists 7*(3), 483–518.

Quemin, S. and R. Trotignon (2019). Emissions Trading with Rolling Horizons. *Working paper*.

Rezai, A. and F. Van der Ploeg (2016). Intergenerational Inequality Aversion, Growth, and the Role of Damages: Occam's Rule for the Global Carbon Tax. *Journal of the Association of Environmental and Resource Economists 3*(2), 493–522.

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.

Roberts, M. J. and M. Spence (1976). Effluent charges and licenses under uncertainty. *Journal of Public Economics 5*(3-4), 193–208.

Rozenberg, J., A. Vogt-Schilb, and S. Hallegatte (2020, mar). Instrument choice and stranded assets in the transition to clean capital. *Journal of Environmental Economics and Management 100*.

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.

Schmidt, T. S. (2014). Low-carbon investment risks and de-risking. *Nature Climate Change 4*(4), 237–239.

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

Traeger, C., G. Perino, K. Pittel, T. Requate, and A. Schmitt (2020). The Flexcap. An Innovative CO2 Pricing Method for Germany. *ifo DICE Report 18*(1), 42–48.

van den Bijgaart, I., R. Gerlagh, and M. Liski (2016). A simple formula for the social cost of carbon. *Journal of Environmental Economics and Management 77*, 75–94.

Vogt-Schilb, A., G. Meunier, and S. Hallegatte (2018). When starting with the most expensive option makes sense: Optimal timing, cost and sectoral allocation of abatement investment. *Journal of Environmental Economics and Management 88*, 210–233.

Walsh, C. E. (2010). *Monetary Theory and Policy*. Cambridge MA, USA: MIT Press.

Webster, M., I. Sue Wing, and L. Jakobovits (2010). Second-best instruments for near-term climate policy: Intensity targets vs. the safety valve. *Journal of Environmental Economics and Management 59*(3), 250–259.

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

Weitzman, M. L. (1978). Optimal rewards for economic regulation. *American Economic Review 68*(4), 683–691.

Weitzman, M. L. (2020). Prices or Quantities Can Dominate Banking and Borrowing. *The Scandinavian Journal of Economics 122*(2), 437–463.

Willner, M. (2018). Consulting the chrystal ball: Firm's foresight and a cap-and-trade scheme with endogenous supply adjustments. *WiSo-HH Working Paper Series 46*.

World Bank (2020). *State and Trends of Carbon Pricing 2020*.

Yates, A. J. and M. B. Cronshaw (2001). Pollution permit markets with intertemporal trading and asymmetric information. *Journal of Environmental Economics and Management 42*(1), 104–118.

Chapter 2

From fundamentals to financial assets: the evolution of understanding price formation in the EU ETS[1]

Marina Friedrich
Eva-Maria Mauer
Michael Pahle
Oliver Tietjen

[1]Submitted to *The Energy Journal*.

From fundamentals to financial assets: the evolution of understanding price formation in the EU ETS*

Marina Friedrich†[a], Eva-Maria Mauer[b,c], Michael Pahle[b], and Oliver Tietjen[b,c]

[a]Vrije Universiteit Amsterdam
[b]Potsdam Institute for Climate Impact Research – Member of the Leibniz Association
[c]Technical University of Berlin

October 15, 2020

Abstract

Price formation in the EU Emission Trading System (EU ETS) has persistently puzzled economists and policy makers. In recent years, the empirical literature investigating this topic has expanded considerably, but a synthesis of what could be learned about price formation as a whole including the last wave of research is still missing. To fill this gap, we review the empirical literature structured along three categories of price drivers and related econometric methods. For better guidance of the reader, we draw on a simple theoretical model of price formation that we subsequently extend to connect the three different strands of literature: demand-side fundamentals, regulatory intervention and finance. In particular the insights from the second and third strand challenge the widespread view that allowance markets primarily reflect marginal abatement costs. Accordingly, the next wave of research should focus on shedding light on the complex interplay of compliance, regulatory uncertainty and financial trading motives.

JEL classifications: Q48, Q50, Q56, Q58
Keywords: emission trading, EU ETS, price formation, literature review

*The authors would like to thank Nicolas Koch, Fabian Hein and Christian Gambardella for valuable comments. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 642147 (CD-LINKS). Michael Pahle's contribution was supported by Stiftung Mercator Foundation under the research project AHEAD. Oliver Tietjen's contribution was additionally funded from the European Union's Horizon 2020 research and innovation program under grant agreement No 730403 (INNOPATHS).

†Corresponding author. Address: Virje Universiteit Amsterdam, Department of Econometrics and Data Science, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. E-mail: m.friedrich@vu.nl.

1 Introduction

The European Union's (EU) Emissions Trading System (ETS) is the world's largest and longest-lived cap-and-trade program to regulate greenhouse gas (GHG) emissions (ICAP, 2019). It is the flagship instrument of the climate policy of the EU. Now close to its forth compliance period, we can look back at more than 14 years of existence of this market. Since the onset of the second trading period in 2008, the price of allowances (European Emission Allowances, EUAs) as shown in Figure 1 has experienced a downward trend until the end of 2017, which some consider as an indicator for inefficiencies (Fuss et al., 2018), though others think the EU ETS works efficiently and they explain the low prices by the large cap or the low allowance demand (Hintermann et al., 2016). In this paper, we review the empirical literature on the EU ETS to examine which factors have actually determined the price.

In particular, the considerable price decline that started in 2011 came as a surprise. In 2008, when the EU's 2020 climate target was adopted, 2013 EUA futures prices were at a level of around 30€/t (Ellerman et al., 2016). Furthermore, the accompanying regulatory impact assessment pointed to a price of around 40€/t in 2020 (Delbeke et al., 2009; Capros et al., 2011). This suggested that prices would rise rather than decline. The following period of low prices until the end of 2017 gave rise to concern that the EU ETS does not work as intended and it is in need of reform (Edenhofer, 2014). After some smaller reforms[1], the EU ETS for Phase IV (2021-2030) was enacted in early 2018, entailing a tightening of the cap and the strengthening of the Market Stability Reserve (MSR).[2] Presumably as a consequence, prices started to rise by the end of 2017 when the political decision was taken.

Nevertheless, it remains controversial whether the ETS is functioning well, i.e. if price formation is efficient. For instance, the COVID-19 shock led to a massive price decline in March

[1]European Commission (2019): Report on the functioning of the European carbon market https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52019DC0557R(01)).

[2]The MSR is a mechanism within the EU ETS that absorbs allowances if the number of allowances held by market agents (i.e. the allowance bank) exceeds a certain threshold. Based on certain rules, the allowances in the MSR either come back to the market later on or are ultimately cancelled.

2020. A first (theoretical) paper indicates that the recent ETS reform indeed improved the performance of the ETS during such economic crises because the MSR cancels (some of the) allowances that are additionally on the market due the crisis (Gerlagh et al., 2020) and thus, the price quickly recovered until June 2020. However, concerns remain that fundamental flaws still prevail that distort price formation (Flachsland et al., 2020). In fact, our review sheds light on this issue, and we will come back to it in the conclusions.

Figure 1: EUA price development from 2008 to September 2020 (Weekly closing prices based on spot-month continuous contract #1 from ICE via Quandl)

Against this background, this paper aims to answer the question of what actually drives prices in the EU ETS. For this purpose, we conduct an extensive literature review and link empirical results to theory. Christiansen et al. (2005) offers an early analysis of the main price drivers in Phase I (2005-2007) of the EU ETS, structured along policy and regulatory issues, fundamentals, and technical indicators. The long phase of persistently low prices until the end of 2017 motivated new empirical research to explain this development, in particular with regard to the role of regulation. Moreover, the EU ETS also received more attention of financial economists that analyze EUAs through an asset pricing lens. In parallel, the number of studies that examine fundamental price drivers grew as well. We systematically review this empirical literature with a view on how it can help to explain the price.

Our review is structured according to the three mentioned explanatory factors for the ETS price: In Section 2, we look at studies investigating demand-side fundamental price drivers such as coal and gas prices. In Section 3, we review the literature focusing on the supply of allowances by the regulator and, more broadly, on the impact of political and regulatory events (e.g., announcements about planned changes of the cap). In Section 4, we consider the empirical finance literature devoted to this market. In this vast literature we concentrate on hedging, speculation and behavioral aspects. In each section, we select a representative group of main papers (see Table 1), for which we summarize the methodology before presenting the results. Based on this, we briefly compare the results of other papers. At the end of each of the three sections, we synthesize insights and discuss implications for the EUA price, while also paying attention to methodological limitations of the studies.

We focus on the empirical literature that helps to explain the ETS price level and development and thus, we leave out many related and in other regards highly relevant papers. For instance, we exclude papers that mainly focus on the price behavior itself (e.g., price volatility) for which the implications for the price level or development are hard to grasp (Benz and Trück, 2009; Chevallier, 2011; Dutta, 2018). We also do not consider literature that analyzes the effect of the ETS on other variables as the economic performance of firms (Commins et al., 2011; Mo et al., 2012; Marin et al., 2018), technological innovation (Calel and Dechezleprêtre, 2016; Rogge et al., 2011), emissions abatement (Ellerman and Buchner, 2008; Petrick and Wagner, 2014; Guo et al., 2020) or, more recently, carbon leakage (Borghesi et al., 2020; Koch and Basse Mama, 2019). Furthermore, we ignore the non-empirical literature and papers on other cap-and-trade markets. For one, other ETS markets are much smaller or less mature so that only a very limited number of studies exist, though the Chinese ETS pilots are an exemption (Ji et al., 2021; Chang et al., 2018; Cong and Lo, 2017; Wen et al., 2020; Zhao et al., 2017). However, because the characteristics of the ETS programs, their maturity and the market environments are fairly different from each other, a comparison is beyond the scope of this paper.

We complement existing reviews not only by considering more recent work and by having a broader topical coverage, but also by linking the different strands through a simple theoretical model of price formation that we extend step by step. The first reviews of the topic conducted by Zhang and Wei (2010) and Bertrand (2014) only covered Phase I. More recent reviews by Zhang (2016) and Hintermann et al. (2016) cover Phase II, but pay limited attention to the finance literature. In this strand particularly, a substantial number of papers has come out over the last years. Furthermore, by extending a theoretical model alongside the empirical literature we develop an incremental understanding of price formation. This gradual approach enhances understanding of the complex interplay of different price drivers in the EU ETS.

2 Demand-side fundamental price drivers

The starting point for this review is a simple theoretical model on intertemporal price formation in emission trading systems based on the classical paper by Rubin (1996). The purpose of referring to this model is to explain how different price drivers influence the price path in theory. It also severs as the backdrop for reviewing the empirical literature investigating the specific price drivers. Initially, the model merely covers demand-side fundamentals, which have been the traditional focus of empirical analyses. In the next section, we extend the model to incorporate further price drivers.

Emission trading programs work by constraining emissions x to a regulatory defined cap $\overline{G}$

$$\int_0^T x(p)dt \leq \overline{G}, \tag{2.1}$$

where T is the lifetime of the ETS program.[3] The cap is translated into tradable allowances

[3] We disregard expectation operators for simplicity.

Section	Paper	Method	Data
2.1	Koch et al. (2014)	*Linear regression (with dummy variables)*	Jan 2008 - Oct 2013
	Rickels et al. (2014)	*Linear regression (with variable pre-selection)*	Dec 2008 - Jul 2012
	Aatola et al. (2013)	*Linear regression (with instrumental variables)*	Jan 2005 - Dec 2010
2.2	Lutz et al. (2013)	*Markov regime-switching model*	Jan 2008 - Dec 2012
	Creti et al. (2012)	*Cointegration analysis*	Jun 2005 - Dec 2010
	Keppler and Mansanet-Bataller (2010)	*Granger causality analysis*	Mar 2005 - Dec 2009
3	Hitzemann et al. (2015)	*Event study*	Jan 2007 - Dec 2012
	Koch et al. (2016)	*Event study with Dynamic Model Selection*	Mar 2008 - Apr 2014
	Deeney et al. (2016)	*Event study*	Oct 2007 - Feb 2014
4.1	Rittler (2012)	*Cointegration and Granger causality analysis*	May 2008 - Dec 2009
	Trück and Weron (2016)	*Linear regression*	Apr 2008 - Dec 2012
4.2	Kalaitzoglou and Ibrahim (2013)	*Autocorrelated duration model*	Jan 2005 - Dec 2008
	Balietti (2016)	*Trading activity-volatility relation*	Sep 2005 - May 2007
	Palao and Pardo (2017)	*Herding intensity measure*	Dec 2012 - Dec 2015

Table 1: Paper overview per section, with methods and time span of empirical analysis

or certificates which allow firms to emit. For example, in the EU ETS one European Emission Allowance entitles to emit one tonne of CO_2. If allowances are scarce, i.e. equation (2.1) holds with equality, the market establishes the allowance price p. The price reflects marginal abatement costs c' since profit maximization of firms implies, in equilibrium, that $p = c'(\overline{x} - x)$, where $\overline{x}$ are baseline emissions (uncapped emissions) and $\overline{x} - x$ is abatement.

Moreover, in an intertemporal market such as the EU ETS, firms can bank certificates for future use if they hold more allowances than needed for compliance in any period. As long as the market-wide allowance bank is positive, market agents exploit intertemporal arbitrage and therefore the allowance price at time t is given by

$$p_t = p_0 e^{rt}. \tag{2.2}$$

In equilibrium, the allowance price grows at the rate of interest r as it reflects the opportunity costs of banking. For any initial price level p_0, the entire (expected) price path is established.

According to the model, the price is determined by marginal abatement costs and the cap which can be called demand- and supply-side market fundamentals, respectively. In addition, the interest rate determines how allowances are used over time and therefore, it has a important impact on the growth rate and price level. In this section, we present empirical results for several demand-side fundamentals as part of c' while ignoring effects of $\overline{G}$ and r. In Section 3, the focus is on the supply side. Finally, Section 4 concentrates on factors that determine the applied discount rate and other price drivers.

A first challenge for empirical studies is that many price drivers are not directly observable because the current allowance price depends on future abatement costs. For instance, the expected development of low-carbon technologies affects marginal abatement costs (and thus, the price), but neither technological development, nor expectations about it, are observable. Hence innovations in abatement technology can hardly be considered in empirical studies although they clearly influence the EUA price. Empiricists therefore need to rely on observable information variables to analyze the impact on the allowance price. On the demand side, these variables include past coal and gas prices as main factors. They play a major role for electricity generation which so far was the most important sector covered by the EU ETS. Since coal is emission intensive, the coal price has, in theory, a negative impact on the allowance price. A higher coal price reduces marginal abatement costs. Since gas is a cleaner alternative to coal, the gas price should have a positive impact on the allowance price. Another frequently considered demand-side price driver is economic activity. A higher economic activity has a positive impact on production and thus emissions, leading to increasing marginal abatement cost and, in turn, higher prices.

In addition, factors such as weather conditions including hot or cold periods, wind speed or precipitation can also have an impact, for example via electricity generation from renewables. While extreme temperatures should have a positive impact on the allowance price due to an increase in electricity demand, higher wind speed or more sunny days should have a negative impact due to increased electricity generation from renewable sources. These opposing effects make it hard to empirically investigate. Some papers which directly include data on electricity generation from renewables find a negligible or statistically insignificant effect which might be due to the lack of reliable (Europe-wide) data.

The remainder of this section elaborates on the corresponding empirical findings based on six selected papers. Each paper differs in the applied methods and/or the set of variables it considers. First, we focus on linear regression approaches. Second, we present results of papers using alternative techniques.

2.1 Empirical evidence and challenges using linear models

To empirically investigate the relations identified by economic theory, linear regression analysis is a natural starting point. However, the empirical literature shows that it is a non-trivial task to find a good model for allowance prices. Early evidence is provided by a group of papers investigating the relationship between allowance prices and abatement related fundamentals using Phase I data (e.g., Alberola et al., 2008, Mansanet-Bataller et al., 2007, Alberola and Chevallier, 2009, Chevallier, 2009 and Hintermann, 2010). As a common result, all papers find the gas price to have a positive and significant influence on the allowance price; other considered variables differ among studies.

While some might argue that the market was not mature enough in Phase I to establish a strong relationship between allowance prices and their fundamentals, Koch et al. (2014) face similar challenges using Phase II data. The authors initially find very limited explanatory power of fundamentals, although the set of included variables is extensive. It consists of coal and gas futures, a stock index as measure of economic activity, renewables generation from two production types (hydro, wind and solar) and the number of issued Certified Emission Reductions (CERs).[4]

The regression exercises reveal that the gas price and economic activity can be identified as clear price drivers. Surprisingly, the coal price does not significantly affect allowance prices. Although the significant factors show the anticipated effect, the overall explanatory power of the models is low. They increase the model fit by accounting for the effect of major policy events using dummy variables. This implies that certain observations disturb or change the estimated relationship such that taking them out improves the findings of a linear model.

A second paper we consider is Rickels et al. (2014), because contrary to most other studies it pays special attention to the multitude of data series that exist. The authors show that empirical papers, whose conclusions regarding the role of fundamentals differ, often use different price series. They point out that, particularly, the coal price can differ quite substantially and it is not obvious which series to choose, as the market lacks transparency. In their empirical study, the authors carefully select each price series by running auxiliary regressions of each candidate series on the allowance price. In the final model specification, they find a significant positive effect of the fuel switching price[5], a significant but negligible effect of renewables and a positive

[4]CERs are so-called carbon credits which can be used instead of emission allowances by covered firms for compliance. They can be earned by companies that engage in certain GHG mitigation projects in developing countries.

[5]The fuel switching price is a linear combination of the coal and the gas price with constants depending on the coal and gas plant's efficiency and emission factor. Hence, using this price can be seen as a restriction which is imposed on the respective coefficients of the coal and gas prices.

effect of economic activity as measured by the oil price and a stock index.

The third paper we would like to review is Aatola et al. (2013). The authors additionally use electricity prices as potential price drivers. The previous two papers do not consider this variable in their empirical analysis due to endogeneity concerns caused by the fact that the relationship between electricity prices and allowance prices might run two-ways: electricity prices drive allowance prices, and allowance prices are likely to have an impact on electricity prices. To address this issue, Aatola et al. (2013) apply a two-stage least squares procedure using appropriate instrumental variables for electricity prices. They find a positive and significant effect of electricity prices. Additionally, they identify significant effects of both coal and gas prices in the direction predicted by economic theory. To the best of our knowledge, this is the only paper which establishes significant effects of both coal and gas prices using Phase II data which is robust throughout all considered specifications.

In Koch et al. (2014), we can see the importance of controlling for political uncertainty, which will be discussed in Section 3. Moreover, all three papers find the residuals of their models to have a non-constant variance. This is addressed by the authors in different ways. Aatola et al. (2013) use a GARCH specification, while Koch et al. (2014) and Rickels et al. (2014) rely on Newey-West robust standard errors. Heteroskedasticity is frequently encountered in empirical studies involving financially traded assets. The allowance price series show several characteristics of financial data which is discussed in more detail in Section 4.

2.2 Alternative approaches

In the previous section, we saw that linear regression approaches need to be adapted to account for time-varying volatility, outliers related to news events as well as possible endogeneity. Here, we present papers that focus on other aspects: time-variation, nonlinearity and instability of the relation between allowance prices and fundamentals.

The first paper is Lutz et al. (2013). It investigates possible nonlinearities in the relationship between the EUA price and its fundamentals during Phase II. They distinguish two different pricing regimes - one applies during periods of high volatility and the other during periods of low volatility. The model allows for two distinct sets of coefficients. The set of explanatory variables is composed of coal and gas futures, oil prices, a stock index, a commodity price index and deviation from average temperature.

In both regimes, the authors find the same set of relevant price drivers. Coal and gas prices, oil prices and the stock index are statistically significant determinants of the EUA price. In Regime 2, which is characterized by low and constant volatility, all significant price drivers show the anticipated sign. Regime 1, however, shows high uncertainty and time-varying volatility. The results on price drivers are similar in this regime, except for the effect of the coal price, which is now positive. This goes against economic theory which predicts the effect to be negative.[6]

Creti et al. (2012) investigate the question whether the relation between EUA prices and its fundamentals has been stable over the course of Phases I and II and might be evolving towards a long-term equilibrium relationship. Rather than working with stationary data by transforming integrated price series into returns, they analyze the non-stationary price data using cointegration techniques. Previous work finds evidence of a cointegration relationship in Phase II, while evidence for Phase I is mixed. This is confirmed in Bredin and Muckley (2011) who find a cointegration relationship in Phase II but not in Phase I or the whole sample. Hintermann (2010) and Rickels et al. (2007) also find no evidence of cointegration in Phase I.

Creti et al. (2012) consider fuel switching prices, oil prices as well as a stock index. They look at their whole sample (2005-2010) as well as two sub-periods corresponding to the different compliance phases. They find a clear cointegration relationship in Phase II with positive and significant coefficients for all fundamentals. For Phase I, they can only find a relationship if they allow for a break in 2006. The nature of the relationships differ between Phase I and II. They find a negative effect of the stock index in Phase I and an insignificant effect of fuel switching. Overall, these findings indicate an increasing role of fundamentals over time while there is no

[6] A recent paper by Jiao et al. (2018) follows the idea of different regimes by looking at EUA return distributions in two regimes defined by economic states. However, they do not investigate the impact of fundamentals, but use predictions of future economic states together with the past return behavior for Value at Risk forecasting.

clear evidence for a stable long-term relationship.

There is a small body of more recent papers which confirm this conclusion. Rickels et al. (2014) find cointegration relationships to be dependent on the choice of data series. Fell et al. (2015) find cointegration relationships among electricity, EUA, coal and gas prices. The resulting vector error-correction model (VECM) shows, however, insignificant responses of EUA prices to shocks in coal and gas prices. Carnero et al. (2018), who estimate a VECM on Phase III data, find a negative relationship between the allowance prices and gas prices. This is to our knowledge the first paper to find a negative relationship between EUA and gas prices, it is also the first paper which only focuses on Phase III data. Overall, caution should be applied when considering results from cointegration analysis in this market due to the short duration and a potential seasonal pattern caused by European rules.[7]

Keppler and Mansanet-Bataller (2010) analyze the interplay between EUA, electricity, gas and coal prices based on a Granger causality analysis. A time series $\{x_t\}_{t=1}^n$ Granger causes another time series $\{y_t\}_{t=1}^n$, if the past of x has an effect on the present of y. Keppler and Mansanet-Bataller (2010) include both EUA spot and futures prices from 2005 to 2008, as well as gas and coal futures, peak and base load electricity prices, the clean dark and spark spread (CDS, CSS) as well as a stock index and several temperature variables.[8]

For Phase I, Keppler and Mansanet-Bataller (2010) find that returns on EUA futures Granger cause spot returns, while spot returns do not Granger cause futures. In addition, EUA futures are Granger caused by CSS and CDS for peak-load electricity as well as by the temperature index, but not by gas, coal or electricity returns. Conversely, EUA futures Granger cause electricity futures for peak and base load, which, in turn, Granger cause the stock index. Based on the previous causality analysis, the authors decide to run a regression with electricity futures as dependent variable. The allowance price, which is usually the dependent variable, enters as regressor together with the CSS and the gas price. They find positive and significant effects of

[7] We thank an anonymous referee for pointing this out.

[8] CDS refers to the revenue a coal-fired generator makes by selling power after having bought coal and the required number of EUAs. CSS represents the same quantity for a gas-fired generator.

all regressors. Results are different for Phase II, but since the analysis is restricted to one year we do not include the detailed results in our review.

Because econometric analysis is still unavailable, we have to leave out the most recent episode of the EU ETS that has puzzled many market observers: The price plunge and shot-term rebound following the onset of the COVID crisis in March 2020. One can look to work on previous shocks of similar scale for potential explanations. Work by Zhu et al. (2015) finds that the financial crises lead to a structural break point. However, this only makes clear that the relative role of different price drivers changed, but it does not explain why and what the mechanisms behind it could be.

2.3 Insights and implications

The following two main insights can be drawn from the reviewed studies. (1) In general, fundamental price drivers have relatively little explanatory power. Specifically for coal, some studies even find a positive correlation, contradicting the prediction of economic theory. This might be due to the fact that the coal price is location-specific and not uniform across Europe as pointed out by Rickels et al. (2014), or that the relationship between coal and allowance prices might change over time as indicated by the results of Lutz et al. (2013). (2) Methodologically, linear regression models come with the limitation that they can only indirectly account for time variation or important political events by using dummy variables. The reviewed papers indicate that the inclusion of such dummy variables can improve the fit of such models, while the overall explanatory power of abatement-related fundamentals remains low.

These insights have two implications. (1) Price formation in the EU ETS is driven by other drivers than purely by fundamentals. This begs the question what these drivers could be. As mentioned above, some papers point to the role of political decisions and changes in the regulatory framework. In fact, in recent years a literature emerged on this topic, which we will review in Section 3. (2) Alternatively, the fact that the explanatory power of abatement-related fundamentals remains low might be due to data and methodological limitations. This underlines the need for better data and more flexible approaches.

3 Political and regulatory changes

The previous section alluded to regulatory uncertainty as another factor influencing price formation, which has gained more attention in face of the low EUA prices and the difficulty of explaining the price development with the help of demand-side fundamentals.

We follow Salant (2016) to introduce regulatory uncertainty in our theoretical model. In this setting, a regulatory intervention implies that the price will either jump up with probability φ or jump down with probability $(1-\varphi)$. The new expected price p^A is then given by

$$p^A = \varphi p^H + (1-\varphi) p^L, \tag{3.1}$$

where $p^L \geq 0$ and $p^H \geq p^L$ denote the lower and higher price after a jump, respectively. The risk for a regulatory intervention which causes such a jump is ongoing and it is determined by the hazard rate $\alpha > 0$. Salant (2016) shows that no arbitrage considerations cause the price change in anticipation of the jump to be

$$\frac{\dot{p}_t}{p_t} = r + \alpha \left(1 - \frac{p^A}{p_t}\right). \tag{3.2}$$

That is, the growth rate of the price changes from r, as in the previous section, to equation (3.2). Assuming that the downward price jump is more severe than the upward price jump, the new expected price is lower: $p^A < p_t$. Consequently, the growth rate of the allowance price is larger than r because the term in the brackets in equation (3.2) is positive. Given that the cap remains unchanged, the price path can be written as

$$p_t = p_0 e^{(r+\alpha)t} - \frac{\alpha p^A}{r+\alpha}\left(e^{(r+\alpha)t} - 1\right), \tag{3.3}$$

which adds the two new factors, α and p^A, to the price equation (2.2) from the previous section. The higher α and the lower p^A, the faster the price increases. Moreover, for a given cap, a higher growth rate implies that the current price is lower and the price in the far future is larger compared to the case without regulatory risks. Therefore, even if the regulatory intervention never occurs, current prices are depressed.

There are different plausible causes of the price jumps. For instance, a new policy that reduces the demand for allowances such as coal phase-outs or support schemes for renewable energies. Such policies have already been implemented in the EU, but similar might follow in the future and thus, cause a further reduced allowance demand and associated price drop. A very important cause is certainly also a change in the cap where a higher cap reduces the price and vice versa. Hence, regulatory uncertainty can be included in the model by assuming that the cap is uncertain. The new market balance equation reads

$$\int_0^T x(p)dt \leq \widetilde{G}, \tag{3.4}$$

where the tilde reflects uncertainty compared to (2.1). For example, a news announcement gives market participants new information from which they infer a change of $\widetilde{G}$, leading to price jumps. Such price jumps are in practice reflected by abnormal returns which is the basis for the empirical studies we discuss in the remainder of this section.

The empirical papers focus on the impact of different types of regulatory announcements on allowance prices. They look at supply-side fundamentals as regulatory decisions that may affect the cap or are viewed as signal for the long-term cap setting. Demand-side fundamentals as announcements of realized emissions are considered as well. While many papers discussed in the previous section already incorporated some aspects related to political decisions, the papers discussed here take a more direct approach. Two of them perform an event study and one uses a dummy variable approach.

In general, event studies can uncover price changes caused by a specific event. The main idea relies on a comparison of the price change that would be expected in the absence of the event - the normal return - to the actual change in prices. If the difference aggregated over a pre-specified event window is large enough, there has been a significant price effect caused by the studied event. This difference between actual and normal returns is called the abnormal return. Formally, the abnormal return at time t can be defined as $AR_t = r_t - \mathbb{E}_t(r_t)$, where r_t denotes the actual return and $\mathbb{E}_t(r_t)$ is the expected normal return at time t. If the event window ranges from t_1 to t_2, cumulative abnormal returns are obtained as $CAR_{t_1,t_2} = \sum_{t=t_1}^{t_2} AR_t$. Those returns are the main quantity of interest and they need to be carefully estimated. Since actual returns are an observed quantity, it is the estimation of normal returns that is crucial. We do not know what the price would have been if the event had not taken place. To obtain estimates of normal returns, the following papers use different approaches.

3.1 Realized emissions

The first paper we present is Hitzemann et al. (2015). The authors look at the effect of emission announcements on EUA returns. Once a year, in April, the quantity of realized emissions of the previous year are publicly announced. This information affects prices if the number of realized emissions differs from expectations of market agents since more emissions imply a higher allowance demand and therefore higher prices and vice versa. In addition, the market may adapt its expectation about future emissions as well if past emissions are considered as indicator for the future which would affect prices accordingly.

In order to capture the news related effect, Hitzemann et al. (2015) define five dummy variables for each announcement. They are designed to capture the effect on the day of the announcement as well as the period before and after the event. The dummy variables are used as explanatory variables in a regression on absolute abnormal returns of EUA futures. This procedure provides an estimate of the immediate effect as well as the effect directly prior to and after the announcement. They are calculated as the difference between actual returns and the overall average return. Using absolute abnormal returns, Hitzemann et al. (2015) run the dummy variable regression which is also carried out on trading volumes and implied as well as realized volatility. The latter two are a measure of intra-day volatility. As a robustness check, Hitzemann et al. (2015) also calculate abnormal returns using average returns over a rolling window, which yields similar results as using the total average.

Overall, the authors find significant abnormal returns on the event day. They also find increased trading volumes and intra-day volatility on the same day. The return response is particularly high in 2008 and 2012. There is no significant effect on abnormal returns before or after the announcement day. This finding shows an immediate market response and thus the market seems to incorporate new information efficiently into prices. In addition, they observe low trading volumes and low intra-day volatility prior to the announcement. Hitzemann et al. (2015) interpret this as a "calm-before-the-storm" effect. The results are in line with findings from other energy and commodity markets.

3.2 Backloading and cap-updating

A second paper we highlight is Koch et al. (2016). The paper investigates two types of policy events: backloading and updating of the cap. Backloading refers to the decision to postpone the auctioning of allowances. While updating of the long-term cap should lead to a price reaction, backloading is cap-neutral and should not affect allowance prices according to the theoretical model introduced in Section 2. However, if market participants have a short foresight horizon or use allowances for hedging purpose (see Section 4), then backloading can have an effect. In addition, incorporating regulatory risk into the theoretical framework, as in Section 3, says that also backloading announcements may have an effect if they are perceived as indicator for the credibility of the cap.

The analysis in Koch et al. (2016) is a classical event study approach, but instead of simply using average returns as an estimate of normal return, they rely on model predictions. As we saw in the previous section, there is not one preferred model for EUA returns. This complicates the question of model choice. To solve this problem, Koch et al. (2016) rely on a flexible procedure called Dynamic Model Selection. From a vast amount of different models, including potentially different sets of regressors at different time periods, the procedure selects the one with the best fit. For each event, this model is used to predict normal returns for a 7-day event window. Although the general idea is similar, this approach is substantially more involved than the dummy variables in the previous paper. The set of possible regressors consists of oil, coal, electricity, commodity and stock prices, interest rates, corporate bond spread, CER prices and a volatility index.

They find that events related to backloading explain many jumps in the data. Four backloading events cause a significant price drop and two a significant price increase, while only two events related to long-term cap changes trigger a statistically significant price effect. The latter are the agreement on 2020 targets and the Green Paper on 2030 targets - both having a positive effect. In summary, they conclude that policy events can explain the existence and timing of jumps in EUA prices. However, many events do not cause an effect in the anticipated direction. The goal to increase the price by backloading has not been achieved, because there is an overall negative effect on prices. Koch et al. (2016) argue that expectations about the degree of commitment plays an important role in allowance pricing.

3.3 Decisions by the European Parliament

The third paper in this section is Deeney et al. (2016). The authors look at the effect of announcements of the European Parliament (EP) on EUA returns. They categorize events according to three main criteria. First, they distinguish between ”party-political” and ”non-party-political”. Party-political decisions concern resolutions put forward by the seven political groups of the EP. Non-party-political decisions come from the European Commission or the European Council. Second, they construct a measure of EUA market sentiment and label events as high or low sentiment according to the resulting index. Third, they measure market attention, or news exposure, which leads to the third and final category. The events are divided between high and low news exposure.

The event window is chosen to consist of 11 days, the day of the Decision by Parliament as well as five days before and after this day. To calculate abnormal returns, they use a zero log return model as well as a constant log return model. In the first model, normal returns are assumed to be zero during the event window, $\mathbb{E}_t(r_t) = 0$. In this case, $AR_t = r_t$. In the second model, normal returns are constant and equal to the mean return during the estimation window, which consists of the 20 days before the start of the event window. Both models are easy to implement and yield very similar results in this application. In addition, the authors investigate volatility effects using a GARCH model equipped with dummy variables for the period before the event, the event day itself and the period after the event.

Deeney et al. (2016) find significant negative abnormal returns as well as an increase in volatility due to the announcements related to EP decisions. Looking at the different categories, these findings seem driven by non-party-political events. Most party-political events have no significant effect. A possible explanation according to Deeney et al. (2016) is that party-political decisions get more media coverage and attract more attention in advance than non-party-political resolutions. This suggests that the party-political decision does not come as a surprise and that prices already reflect this information. Additionally, they find the same effects after events in times of low market sentiment and when market attention is low. When market attention is high, there is no significant abnormal return, but a decrease in volatility after the announcement. Both findings are relevant for the timing and extent to which political decisions are revealed to market participants.

3.4 Other related papers

Other studies in this direction are e.g., Mansanet-Bataller and Pardo (2009), Mansanet-Bataller and Sanin (2014) and Fan et al. (2017). The first paper shows that news announcements during Phase I had an influence on allowance prices on both the announcement day and on previous days, while they find no effects on the volatility of returns. Mansanet-Bataller and Sanin (2014) find a strong impact of announcements by the European Commission, in particular, regarding the Phase II announcements of National Allocation Plans and the global cap for Phase III. Fan et al. (2017) look at a wide range of announcements regarding regulatory updates in an event study using adjusted mean returns as a measure of normal returns. They find 24 out of the 50 events they consider to have caused significant abnormal returns. Moreover, according to this study, impacts of events having negative impacts are higher than those having positive impacts. Another recent contribution is Creti and Joëts (2017) who also use an event study. However, before the event study, they test for periods of exuberance in the allowance price data and find evidence for several short periods of explosive behavior. Events that offer possible explanations for these episodes are then used in an event study in which no abnormal returns are found.

In addition, Conrad et al. (2012) find that decisions regarding the allocation of allowances has a strong and immediate impact on EUA prices. This finding is based on high-frequency data and the use of surprise variables which are constructed with the help of market expectations obtained from surveys.

Sanin et al. (2015) apply a different approach which is, in essence, a combination of methods used by papers in the previous section on fundamentals. The authors use an ARMA model for allowance prices with fundamentals as exogenous regressors and a GARCH component. To the GARCH model they add a jump component that allows for sudden jumps in volatility which they relate to supply announcements by the European Commission. The fact that their focus lies on volatility dynamics rather than prices is due their focus on financial market aspect of the EU ETS.

3.5 Insights and implications

The following insights can be drawn from the reviewed studies. (1) There is clear evidence that, in general, the market reacts to a variety of regulatory news with changes in returns and volatility. However, some events triggered a response in a direction contradicting theoretical predictions. Offered explanations are that the information has already been priced in (party-political EP decisions), or that it more profoundly signaled a lack of commitment (backloading). (2) Methodologically, core to all approaches is the notion of abnormal returns. There exists a plurality of ways how to estimate this unobserved quantity, e.g., zero or constant returns and Dynamic Model Selection. Respective results can differ substantially and some methods may find significant effects where others do not. Moreover, all studies employ time windows spanning over several days. Accordingly, it remains unanswered how persistent the effects are.

These insights have two implications. (1) The way in which the market responds to news alludes to the potential role of information processing and belief formation. Prices may not respond to an event because it has been anticipated, or they respond indirectly by means of adapted beliefs. This implies that identifying the true underlying cause of the market response remains challenging. (2) Information processing and belief formation also imply that market reactions depend on capabilities and access to information of different trader types. This is reflected in their trading behavior, which is covered by the finance literature that we review in the next section.

4 Emission allowances as a financial asset

Besides its role as a compliance market, the EU ETS is also a financial market. The main purpose of such a financial commodity market is risk reduction (hedging), speculation and price discovery. Moreover, a financial market introduces new agents to the ETS: financial traders or speculators who aim to make profit from trading allowances or derivatives such as options and futures. This raises the question in how far speculation, hedging and (in)efficient price discovery affect the price formation in the EU ETS. Analyzing the EU ETS from a financial market perspective thus offers useful insights into the functioning of the market.

We focus on two different aspects. The first strand of literature (Section 4.1) considers market frictions and the price discovery processes. Financial economic theory suggests that markets are efficient if prices reflect all available information such that there are no arbitrage opportunities. Typical inefficiencies which impair information transmission are transaction costs. Other market frictions are convenience yields and risk premiums that may lead to under- or overvalued prices, compared to the idealized first-best market solution. In this case, the temporal availability of allowances and the hedging demand of firms can affect prices.

The second strand of literature (Section 4.2) that we discuss is on behavioral aspects of the

EU ETS. By now, non-regulated actors make up a large share of the overall trading volume in the market. Given that trading accounts held by financial actors tend to be more active than those of compliance traders (Berta et al., 2017; Betz and Schmidt, 2016), their behavior is potentially an important factor for price formation. Consequently, during the last years, the behavior of these actors and possible differences in their trading strategies have become a major interest in the literature.

Although the relationship of interest is different, it is worth noting that many of the methods which we saw in Section 2 reappear in this section. Previously, they were applied to analyze the allowance price and its relationship to fundamentals. In particular, cointegration analysis, Granger causality tests and (variations of) GARCH models are popular also in this strand of the literature. The main difference is that the methods are applied to different data series. Studies in the section analyze, for example, the price volatility, bank volume or the duration of trades.

4.1 Financial market frictions

Financial market frictions as considered in this section do not directly affect the demand- and supply-side fundamentals of the previous sections. Instead, they affect how allowances are evaluated over time. Accordingly, empirical papers in this field typically analyze the relationship between ETS spot and futures prices. There are two (non-exclusive) theoretical views on this relationship (Fama and French, 1987). First, according to the theory of storage, price differences should reflect the forgone interest due to investing in a commodity, its storage costs and a convenience yield (Kaldor, 1939; Working, 1949; Brennan, 1958). Since storage costs for EUAs are virtually zero, price differences should be only due to the interest rate and the convenience yield. The latter arises because of a potential benefit of holding EUAs rather than futures. This benefit exists due to potential stock-outs (i.e. a zero allowance bank) in the future which can lead to positive price shocks because firms cannot borrow from future compliance periods. The second view on the relationship between spot and futures prices is the hedging pressure theory (Keynes, 1930, Hicks, 1939, Hirshleifer, 1990). In this case, futures prices consist of the expected spot price and a risk premium. The latter has to be paid by producers to financial traders that take the contrary position in the market. The risk premium reflects the producers' demand for risk reduction (due to risk aversion). They thus accept a lower return which is the profit of speculators. Both theories can be incorporated in a standard cost-of-carry model implying no-arbitrage between spot and futures prices. Specifically, the futures price is

$$p_t^{fut} = p_t e^{(r-\gamma+\lambda)(S-t)} \tag{4.1}$$

where p_t is the ETS spot price, S is the expiry date of the futures contract, γ is a convenience yield and λ is a risk premium. Due to arbitrage this relationship holds for allowances prices in general and thus we can write the allowance price path as

$$p_t = p_0 e^{(r-\gamma+\lambda)t}. \tag{4.2}$$

Hence, similar as the hazard rate in the previous section, the risk premium and the convenience yield change the price path and enter the original equation (2.2) through an additional term in the exponential function. A positive risk premium leads to a steeper price path and thus lower prices initially and higher prices in later periods. A negative risk premium and the convenience yield, in contrast, have opposing effects since they flatten the price path.

The empirical papers presented in this section often test whether the relationship between futures and spot prices hold as in equation (4.1) with $\gamma = \lambda = 0$. If this is not the case, it is interpreted as indicator for non-zero convenience yields γ or risk premiums λ that prevent a perfect arbitrage. An alternative interpretation is that the information transmission between spot and futures markets is inefficient due to transaction costs, implying also that the price discovery process is distorted.

The first paper we consider is Rittler (2012). In a first step, he derives the theoretical futures prices from observed spot prices based on the cost-of-carry model given by equation (4.1) assuming no convenience yields and risk premiums, $\gamma = \lambda = 0$. Theoretical and observed futures prices are then used to estimate a vector error correction model (VECM) to analyze cointegration of long-run prices. Subsequently, Rittler (2012) computes common factor weights as price discovery measures for the markets and conducts Granger causality tests for the short-term relationship.

Using daily data, the author finds no cointegration between prices, indicating the absence of a stable long-run relationship. This confirms the result by Chevallier (2010a) who also finds no cointegration for similar data. It is also consistent with Joyeux and Milunovich (2010), who provide results for Phase I. In contrast, Uhrig-Homburg and Wagner (2009) find evidence of a long-run relationship with daily data for Phase I. More recent studies by Charles et al. (2013) and Bredin and Parsons (2016) also conduct conintegration tests with daily data. The former find a significant relationship between spot and futures prices using data from March 2009 to January 2012. Bredin and Parsons (2016) use data from 2005 to 2014, and find only cointegration between observed and theoretical cost-of-carry futures in Phase I, while for Phases II and III, there is no cointegration. Overall, the results for the relationship between daily spot and futures prices are mixed and suggest some frictions preventing a perfect arbitrage. However, when using 10 or 30 minutes intra-day data, Rittler (2012) finds strong support for cointegrated prices. He suggests that markets are indeed closely linked but this can only be observed when exploiting information in high frequency rather than daily data. Furthermore, Rittler (2012) finds common factor weights of about 70% for the futures market, which means that it contributes more to the price discovery process than the spot market. Regarding short-term causality, he finds a bidirectional impact. Rittler (2012) concludes that the price discovery process is similar to other mature markets. This result is confirmed by Schultz and Swieringa (2014) who also use high frequency data. In addition, Schultz and Swieringa (2014) find that transaction costs are an important market friction that prevents faster price adjustments for some EU ETS securities. The study by Mizrach and Otsubo (2014) confirms the cointegration between EUA futures and spot prices, where the more liquid futures market leads the price discovery.

While Charles et al. (2013) find cointegration between spot and futures prices (see above), they reject the cost-of-carry model with zero convenience yield. They interpret this as market inefficiency because it implies arbitrage opportunities. A related strand of literature considers the presence of profitable trading strategies which should not exist in an efficient market when arbitrage is exploited. Daskalakis (2013) examines the EUA futures market and analyzes the relative performance of different trading strategies that aim at identifying price trends by looking at past prices. The results hint both at the failure of the efficient market hypothesis in the period from 2008 to 2009 as well as at an increase in efficiency from 2010 onwards. However, even for 2011, the trading strategies produced positive returns, although these were lower than those of the reference sell and hold strategy. This implies that the market became more mature over the years and thus closer to being in line with weak market efficiency. Related to this, Crossland et al. (2013) consider the daily EUA spot prices in Phase II and analyze the presence of profitable trading strategies based on momentum (price trend continuing) and overreaction (price trend reversing). They find the occurrence of momentum in the short-term and overreaction in the medium-term, both phenomena that contradict the efficient market hypothesis. Mizrach and Otsubo (2014), Narayan and Sharma (2015), Niblock and Harrison (2013) and Aatola et al. (2014).

Trück and Weron (2016) explicitly take non-zero convenience yields and risk premiums into account that may explain profitable trading strategies of other studies. They calculate the implied convenience yield using observed spot and futures prices based on equation (4.1), yielding

$$\gamma = r - \frac{\ln(p_t^{fut}) - \ln(p_t)}{(S-t)} \tag{4.3}$$

where γ may also include a risk premium λ. They find that after a short positive period in 2008, the convenience yield turns highly negative between -2% and -7%. In a second step, Trück and Weron (2016) regress the implied convenience yields on several factors by applying a pooled OLS regression. They use the allowance surplus in the market and risk measures as independent variables. They find that a higher allowance surplus decreases the convenience yield and interpret this result as consistent with the theory of storage, since generally more allowances should lead to lower risk of a stock-out. They also find a negative effect of the EUA price variance on the convenience yield. This is seen as evidence for the impact of the hedging demand: firms are willing to pay higher prices for futures to reduce their risk exposure.

Other papers that empirically consider risk premiums are Chevallier (2010b, 2013), Kamga and Schlepper (2015) and Pinho and Madaleno (2011). They find on average positive risk premiums and suggest that this indicates that investors want to hedge against rising prices. The role of hedging is also confirmed by Hintermann (2012). He derives an option pricing formula in which the price depends on the penalty for non-compliance and the probability of a non-binding cap. He applies it to Phase I data and finds that it can explain large parts of the price development. Therefore, he concludes that hedging against paying the penalty was an important price driver in Phase I.

4.2 Behavioral aspects

This section sheds light on how the behavior of different market actors affects ETS prices. An important question in this context is how market participants form their expectations and beliefs about how the price will evolve. Behavioral aspects covered here comprise the existence of different trading types, non-rational behavior such as herding, and the use of trading strategies aiming at exploiting price patterns. Two important theoretical papers in this context are Barberis et al. (1998) and De Long et al. (1990). Both derive price formulas for assets in which they distinguish between the fundamental value of the asset and the actual price distorted by the behavior of a part of the market participants.

Barberis et al. (1998) ask how market participants form beliefs about the probability of future changes – or rather how these beliefs are updated in response to new information. Standard models implicitly assume that updating happens instantaneously and with full confidence about the effect on prices in equilibrium. Yet it is known that there is both overreaction and underreaction of stock prices to new information. Barberis et al. (1998) propose a model of investor sentiment based on psychological evidence to explain this behavior. While earnings from the asset actually follow a random walk, investors believe that earnings are either following a trend or reverting to the mean. With every new information, investors are updating their beliefs. Barberis et al. (1998) show that their framework can explain both under- and overreaction. They link both phenomena to concepts from psychology: conservatism (hesitance to update model in view of new information) and representativeness (a small part of a process is interpreted as being representative for the overall process).

De Long et al. (1990) analyze the effect of noise traders in financial markets. They assert two types of traders in the market: sophisticated investors and noise traders, whereby the latter falsely believe they have special information about the future price of the risky asset and misperceive the true expected price. This misperception leads to persistent irrational trading behavior that distorts prices. This in turn creates a noise-trading induced risk for sophisticated traders, which even further distorts prices.

Applying such behavioral aspects to allowance price models is an interesting avenue for future research. For this paper we simply denote by p_t^F the original price path determined by fundamentals, as in the previous equations (2.2), (3.3) and (4.2) and add an additional behavioral term to our model

$$p_t = p_t^F + \mathcal{B}_t. \tag{4.4}$$

That is, $\mathcal{B}_t$ represents any changes from the fundamental allowance price p_t^F due to behavioral aspects such as herding behavior, different trading strategies or even speculative bubbles. From an empirical perspective it is challenging to distinguish the two parts in equation (4.4). This is because, as we have seen in previous sections, p_t^F is difficult to determine since demand- and supply-side fundamentals as well as financial frictions affect the price but are hardly observable. Nonetheless a number of empirical papers analyze behavioral aspects in the EU ETS which we present in the following.

Kalaitzoglou and Ibrahim (2013) identify different types of agents active in the EU ETS futures market that can be clearly distinguished by their trading behavior. The authors analyze the duration of trades examining in how far clustering of duration characteristics correlates with the trading behavior of market participants.

The duration between single transactions is modeled with a smooth transition mixture autocorrelated duration (STM-ACD) model. By incorporating smooth transitions into the model, the dynamics between two regimes (where a regime is dominated by a certain type of trader) can be captured. The presence of three different trader types in the market is examined: The informed traders receive private information to which they react by trading in large volumes. The uninformed have no access to this information and hence initiate their trades randomly. Lastly, while the fundamental traders are also uninformed, they are able to extract information from the market by examining past trades.

The trader types associated with the three regimes are identified by analyzing the shape of the hazard rate, which measures the probability of a trade being initiated after the arrival of exogenous information as a function of time. In the case of the informed traders, the hazard rate is decreasing, for the uninformed it is flat. For the fundamental traders, who extract information with a delay by analyzing informed trades, the hazard rate is increasing. Regarding the smooth transition mechanism included in the model, the findings suggest smoother transitions between the informed and the fundamental regime in Phase II compared to Phase I. This implies that learning by the uninformed happened faster in later stages of the EU ETS and, as a result, greater market depth.

Balietti (2016) also considers the presence of different trading behaviors. Specifically, the author estimates in how far the relation between trading activity and volatility varies with different trader types. In contrast to Kalaitzoglou and Ibrahim (2013), the author differentiates between different trader types according to the specific design of the EU ETS as a compliance market. This market is characterized by actors who are regulated by the EU ETS, and hence obliged to participate, and financial actors, who participate either as intermediaries or to make profit from speculation. Moreover, actors are exposed to different (product) markets depending on whether they are active in the energy, industry, or financial sector. The initial endowment of certificates relative to their baseline emissions is also taken into account.

In order to examine how the specific characteristics of the market actors translate into differences in their trading behavior, Balietti (2016) regresses the volatility on the trading activity of the different participants. The trading activity-volatility relation is estimated by simultaneously estimating returns and volatility. Therefore, two equations are iterated: Equation 1 estimates the price changes conditional on autoregressive terms and lagged volatilities, while equation 2 estimates the conditional standard deviation based on lagged volatilities, lagged price changes, and trading activity. As a proxy for trading activity, Balietti (2016) uses both the daily transferred EUA volumes (fitted by an ARIMA process) and the number of daily permit transfers.

The regression on the daily spot price differences hints at a lack of market efficiency in Phase I of the EU ETS: The coefficients of the lagged price differences are significant and negative, i.e. large price differences in the past come along with smaller price changes in the present. The regression on price volatility and permit trading shows that when distinguishing between the three sectors (energy, industry, finance), the trading activity-volatility relation differs with trader type. While the energy sector trades more when volatility is high, the industry sector tends to be more active when volatility levels are low. The financial sector seems to act as a flexible counterpart, trading more with the energy sector when volatility is high and more with the industry sector when volatility is low. However, all in all, many actors seem to have remained inactive during Phase I especially when volatility was high, suggesting that a large share of actors was unwilling to trade when a lot of information arrived in the market.

Also concerned with the microstructure of the carbon futures market is Ibrahim and Kalaitzoglou (2016). In the light of certain findings in the literature, such as autocorrelation in the price level and order flow (Benz and Hengelbrock, 2008) and the presence of intra-day price patterns (Ibikunle et al., 2016), they propose an asymmetric information microstructural model of intra-day price changes in order to analyze the effect of expected trading intensity on intraday price changes. In their model, the price responds dynamically to information and liquidity with every transaction, as traders form their expectations about subsequent trades based on trading activity and characteristics of previous trades. Specifically, when formulating price quotes, traders take into account trading intensity, information content, and volatility of previous trades. The authors find that the autocorrelation of returns and of the volatility of returns can be explained to a large extent by the predictability in the persistence of trading intensity. A similar positive autocorrelation in the trade sign has been found by Benz and Hengelbrock (2008), Mizrach and Otsubo (2014), Medina et al. (2014).

Herding behavior is one possible explanation for the autocorrelation in trade sign (Tóth et al., 2015). Palao and Pardo (2017) analyze the presence of herding behavior in the EU ETS in three parts: (i) detecting herding behavior in the futures market, (ii) identifying factors that influence herding behavior and (iii) analyzing the impact of herding behavior on the market.

Herding behavior can be detected by looking for persisting upward and downward runs in the price development: sequences of buy or sell trades. Palao and Pardo (2017) use the Herding Intensity Measure developed by Patterson and Sharma (2006) to identify the occurrence of such runs. They find that the herding effect decreases over time but is higher on days where price clustering is strong. Moreover, herding increases during speculative periods and when ETS-related news are published. Similarly, herding is positively correlated with trading frequency, uncertainty (as measured by intraday volatility) and the occurrence of extreme returns. The presence of herding, in turn, increases price volatility and entails overreaction.

A recent analysis of the price rally following the 2018 EU ETS reform also alludes to overreactions (Friedrich et al., 2020). The analysis suggests the shock (reform) has triggered the market into speculation about its price impacts, leading to an overreaction indicated by an episode of explosive price behavior. After the recent COVID shock in March 2020 the price plummeted, but quickly recovered until June 2020. This may also be explained by overreactions in response to the negative supply shock caused by COVID.

Further behavioral aspects examined in the EU ETS that are not in line with informational efficiency are the presence of price and size clustering as well as feedback trading (i.e. buying after a price rise and selling after a price fall). Palao and Pardo (2012) find that transactions are concentrated on prices ending on digits 0 and 5, tracing this back to the attraction theory (the preference for certain numbers without any rational explanation) and the negotiation hypothesis (where price clustering is used to limit transaction costs). Palao and Pardo (2014) complement this analysis by showing that not only prices, but also order sizes are rounded up in times of high uncertainty. While Chau et al. (2015) do not find significant feedback trading in emissions markets, Crossland et al. (2013), as mentioned beforehand, find short-term momentum and medium-term overreaction in the European carbon market.

4.3 Insights and implications

The following main insights can be drawn from the reviewed studies. (1) The idealized cost-of-carry model is violated by market frictions which affect the price level and its growth rate. Such frictions could be transaction costs, hedging demand of firms (and related risk premiums) as well as convenience yields. The latter are resulting from potential stock-outs because allowance borrowing from the future is not allowed in the EU ETS. (2) Methodologically, it is difficult to directly attribute these frictions to the violation of the cost-of-carry model. Frictions are typically not observable and therefore, deviations from the theoretically ideal predictions are open to several interpretations. However, the relatively large allowance bank over the past years suggests that the likelihood of a stock-out was low and thus convenience yields probably did not play an important role. This alludes to other frictions, such as hedging demand, as important price drivers. (3) Studies on behavioral aspects reveal the presence of different trader types with distinct trading strategies. Moreover, the results point to possibly non-rational behavior such as size and price clustering, herding as well as under- and overreaction to new information. Taking into account these trader types and behaviors sheds light on several observations about the EUA market: price jumps, persisting price trends including the occurrence of crashes and bubbles, excessive volatility and inadequate (or insufficient) intertemporal trading strategies by compliance actors.

These insights have the following implications. (1) In light of the relevance of frictions, the recently introduced Market Stability Reserve (MSR) may have a significant impact on the permit price beyond its effect through lowering the cap.[9] As the MSR reduces the allowance bank level in the coming years, it might affect the costs of hedging reflected by risk premiums. A smaller bank implies that less permits are available for hedging purposes. However, the lower bank should not significantly increase the relevance of convenience yields because allowances leave the MSR and enter the market again when the bank level becomes lower. (2) In view of the political nature of the EU ETS and the ongoing reform process, the question of how actors form expectations and beliefs about the future and how they respond to new information may be even more important than in other financial markets. For example, if investors tend to overreact to regulatory announcements, and shocks in general, then stabilizing expectations might be essential for the efficiency of the ETS.

5 Discussion and conclusion

Covering three different strands of empirical literature on the EU ETS, the structure of this review mirrors the different angles to analyze price formation in the EU ETS. Our findings challenge the widespread view that short-term price formation in the EU ETS is primarily driven by observable marginal abatement costs: Some important explanatory factors do not show a significant effect, or, in some cases, the effect is opposite to what is predicted by theory (Section 2). As shown in Section 3, price explanation in the EU ETS can be improved by considering regulatory or political events. Yet, the price response to such events is not always in line with theoretical considerations, which raises the question how market participants process information or form their beliefs. For instance, some may perceive a political decision as support

[9] The MSR has two main effects: (1) It lowers the allowance bank level by shifting the supply of EUAs to the future and (2) it reduces the overall cap ($\overline{G}$ in our model) by cancelling EUAs. While (2) clearly raises the price level, (1) has no effect on the price in an idealized market without frictions (Salant, 2016).

for the EU ETS (price), whereas others expected an even stronger supportive decision and thus, adjust their price expectations downwards. The finance literature, reviewed in Section 4, alludes to different trader types and different reactions to new information. Specifically, the market exhibits herding behavior and over- or underreaction to new information that may cause bubbles. To conclude, in order to understand price formation in the allowance market, it must (also) be viewed as a financial market – and analyzed and regulated as such.

This has important implications for policy design and the further evolution of the EU ETS. Speculation fueled by regulatory uncertainty and affected by financial market flaws could imply a substantial – and potentially persistent – deviation of prices from their fundamental value. Such deviations would impair the functioning of the system, in particular its dynamic cost-effectiveness. Moreover, theory suggests that such speculation could also destabilize the market by inducing excess volatility (De Long et al., 1990). According to our review, regulatory uncertainty could be a major source for this. Accordingly, avoiding speculation on changes of the design of the system including its targets – or creating uncertainty about future changes in the first place – seems key for policy design. Obviously, this can be accomplished through stable market rules, whose impact on prices is well predictable.

Specifically for the EU ETS, this suggests that the Market Stability Reserve (MSR) does not actually stabilize the market. First, it considerably increases the complexity of the market (Perino, 2018). Notably, cancellation depends on the expectation of future prices, leading to the counter-intuitive effect that future complementary policies may reduce cancellations (Rosendahl, 2019) – contrary to what policy maker and market participants had expected it to do. In addition, the MSR opens up the door for price manipulations and the magnitude of future cancellations is highly uncertain (Osorio et al., 2020; Quemin, 2020). In contrast, a price collar may be a more suitable option; see Flachsland et al. (2020) for a short review for the literature and design options. In particular, by clarifying ex-ante that the price will never surpass (ceiling) or fall below (floor) a certain level, price expectations can be substantially stabilized. In face of that, EU policy makers should seriously consider replacing the MSR with such a price control mechanism.

Finally, echoing the call by Hintermann et al. (2016), this review also identifies promising fields for future research in two additional directions. Firstly, in all strands of the empirical literature, it is evident that findings heavily depend on the method used and its specific restrictions. The fact that the market is maturing together with the existing empirical evidence calls for more flexible methods in the analysis of classical price drivers. Such methods should be able to identify and capture potential structural changes. So far, the analysis often keeps the relationship between the allowance price and its price drivers constant over time or it relies on restrictive assumptions which limit the form of potential transitions. Future approaches could, for example, include a thorough analysis of break points or smooth transitions in the relationship. Alternatively, it could be modeled nonparametrically or with a local trend model from the state-space literature.

Secondly, the theoretical finance literature emphasizes the importance of considering different trader types and points out the implications for price formation. While some work in this direction exists, e.g. on the role of banks (Cludius and Betz, 2020), the presence and impact of speculation is a particularly promising topic for future research. Analyzing other allowance market data beyond prices offers opportunity for such work. To begin with, Open Interest (OI) data together with volume data could be used to measure speculative and hedging activities in the futures market. This approach is quite common for other markets (Lucia and Pardo, 2010), but to the best of our knowledge there is just a single analysis of this type for the EU ETS (Lucia et al., 2015) so far. Moreover, pursuant to stronger financial market regulations such as MiFID II and MiFIR, Commitment of Traders (CoT) reports became mandatory for the EU ETS in 2018. They break down positions by types of traders and trading motives at the end of each business week for each market place (Bohl et al., 2019). Analyzing this new source of data promises even more accurate insights into hedging and speculation activities.

References

Aatola, P., Ollikainen, M., and Toppinen, A. (2013). Price determination in the EU ETS market: Theory and econometric analysis with market fundamentals. *Energy Economics*, 36:380–395.

Aatola, P., Ollikka, K., and Ollikainen, M. (2014). Informational efficiency of the EU ETS market - a study of price predictability and profitable trading. *Journal of Environmental Economics and Policy*, 3(1):92–123.

Alberola, E. and Chevallier, J. (2009). European Carbon Prices and Banking Restrictions: Evidence from Phase I (2005-2007). *The Energy Journal*, 30(3):51–79.

Alberola, E., Chevallier, J., and Chèze, B. (2008). Price drivers and structural breaks in European carbon prices 2005-2007. *Energy Policy*, 36(2):787–797.

Balietti, A. C. (2016). Trader types and volatility of emission allowance prices. Evidence from EU ETS Phase I. *Energy Policy*, 98:607–620.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49:307–343.

Benz, E. and Hengelbrock, J. (2008). Price Discovery and Liquidity in the European CO_2 Futures Market: An Intraday Analysis. *SSRN Working Paper*, AFFI/EUROF(Paris December 2008 Finance International Meeting).

Benz, E. and Trück, S. (2009). Modeling the price dynamics of CO_2emission allowances. *Energy Economics*, 31(1):4–15.

Berta, N., Gautherat, E., and Gun, O. (2017). Transactions in the European carbon market: a bubble of compliance in a whirlpool of speculation. *Cambridge Journal of Economics*, 41:575–593.

Bertrand, V. (2014). Modeling of emission trading: A literature review for dynamic partial equilibrium models with perfect competition. *Energy Studies Review*, 21(1).

Betz, R. A. and Schmidt, T. S. (2016). Transfer patterns in Phase I of the EU Emissions Trading System: a first reality check based on cluster analysis. *Climate Policy*, 16(4):474–495.

Bohl, M. T., Stefan, M., and Wellenreuther, C. (2019). An Introduction to ESMA's Commitments of Traders Reports: Do Hedgers Really Hedge?

Borghesi, S., Franco, C., and Marin, G. (2020). Outward Foreign Direct Investment Patterns of Italian Firms in the European Union's Emission Trading Scheme*. *The Scandinavian Journal of Economics*, 122(1):219–256.

Bredin, D. and Muckley, C. (2011). An emerging equilibrium in the EU emissions trading scheme. *Energy Economics*, 33(2):353–362.

Bredin, D. and Parsons, J. (2016). Why is Spot Carbon so Cheap and Future Carbon so Dear? The Term Structure of Carbon Prices. *The Energy Journal*, 37(3):83–108.

Brennan, M. J. (1958). The Supply of Storage. *The American Economic Review*, 48(1):50–72.

Calel, R. and Dechezleprêtre, A. (2016). Environmental Policy and Directed Technological Change: Evidence from the European Carbon Market. *The Review of Economics and Statistics*, 98(1):173–191.

Capros, P., Mantzos, L., Parousos, L., Tasios, N., Klaassen, G., and Van Ierland, T. (2011). Analysis of the EU policy package on climate change and renewables. *Energy Policy*, 39(3):1476–1485.

Carnero, M. A., Olmo, J., and Pascual, L. (2018). Modelling the dynamics of fuel and eu allowance prices during phase 3 of the eu ets. *Energies*, 11(11):1–23.

Chang, K., Lu, S., and Song, X. (2018). The impacts of liquidity dynamics on emissions allowances price: Different evidence from China's emissions trading pilots. *Journal of Cleaner Production*, 183:786–796.

Charles, A., Darné, O., and Fouilloux, J. (2013). Market efficiency in the European carbon markets. *Energy Policy*.

Chau, F., Kuo, J.-M., and Shi, Y. (2015). Arbitrage opportunities and feedback trading in emissions and energy markets. *Journal of International Financial Markets, Institutions and Money*, 36:130–147.

Chevallier, J. (2009). Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics*, 31(4):614–625.

Chevallier, J. (2010a). A Note on cointegrating and vector autoregressive relationships between CO_2 allowances spot and futures prices. *Economics Bulletin*, 30(2):1564–1584.

Chevallier, J. (2010b). Modelling risk premia in CO_2 allowances spot and futures prices. *Economic Modelling*, 27(3):717–729.

Chevallier, J. (2011). Detecting instability in the volatility of carbon prices. *Energy Economics*, 33(1):99–110.

Chevallier, J. (2013). Variance risk-premia in CO_2markets. *Economic Modelling*, 31(1):598–605.

Christiansen, A. C., Arvanitakis A., and Hasselknippe, H. (2005). Price Determinants in the EU Emissions Trading Scheme. *Climate Policy*, 5(1):15–30.

Cludius, J. and Betz, R. (2020). The Role of Banks in EU Emissions Trading. *The Energy Journal*, 41(2).

Commins, N., Lyons, S., Schiffbauer, M., and Tol, R. S. (2011). Climate Policy & Corporate Behavior. *The Energy Journal*, 32(4).

Cong, R. and Lo, A. Y. (2017). Emission trading and carbon market performance in Shenzhen, China. *Applied Energy*, 193:414–425.

Conrad, C., Rittler, D., and Rotfuß, W. (2012). Modeling and Explaining the Dynamics of European Union Allowance Prices at High-Frequency. *Energy Economics*, 34(1):316–326.

Creti, A. and Joëts, M. (2017). Multiple bubbles in the European Union Emission Trading Scheme. *Energy Policy*, 107:119–130.

Creti, A., Jouvet, P. A., and Mignon, V. (2012). Carbon price drivers: Phase I versus Phase II equilibrium? *Energy Economics*, 34(1):327–334.

Crossland, J., Li, B., and Roca, E. (2013). Is the European Union Emissions Trading Scheme (EU ETS) informationally efficient? Evidence from momentum-based trading strategies. *Applied Energy*, 109:10–23.

Daskalakis, G. (2013). On the efficiency of the European carbon market: New evidence from Phase II. *Energy Policy*, 54:369–375.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98(4):703–738.

Deeney, P., Cummins, M., Dowling, M., and Smeaton, A. F. (2016). Influences from the European parliament on EU emissions prices. *Energy Policy*, 88:561–572.

Delbeke, J., Klaassen, G., van Ierland, T., and Zapfel, P. (2009). The role of environmental economics in recent policy making at the European Commission. *Review of Environmental Economics and Policy*, 4(1):24–43.

Dutta, A. (2018). Modeling and forecasting the volatility of carbon emission market: The role of outliers, time-varying jumps and oil price risk. *Journal of Cleaner Production*, 172:2773–2781.

Edenhofer, O. (2014). Climate policy: Reforming emissions trading.

Ellerman, A. D. and Buchner, B. K. (2008). Over-Allocation or Abatement? A Preliminary Analysis of the EU ETS Based on the 2005-06 Emissions Data. *Environmental and Resource Economics*, 41(2):267–287.

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

Fama, E. F. and French, K. R. (1987). Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage. *The Journal of Business*, 60(1):55.

Fan, Y., Jia, J. J., Wang, X., and Xu, J. H. (2017). What policy adjustments in the EU ETS truly affected the carbon prices? *Energy Policy*, 103:145–164.

Fell, H., Hintermann, B., and Vollebergh, H. (2015). Carbon Content of Electricity Futures in Phase II of the EU ETS. *The Energy Journal*, 36(4):61–83.

Flachsland, C., Pahle, M., Burtraw, D., Edenhofer, O., Elkerbout, M., Fischer, C., Tietjen, O., and Zetterberg, L. (2020). How to avoid history repeating itself: the case for an EU Emissions Trading System (EU ETS) price floor revisited. *Climate Policy*, 20(1):133–142.

Friedrich, M., Fries, S., Pahle, M., and Edenhofer, O. (2020). Understanding the explosive trend in EU ETS prices - fundamentals or speculation? *arXiv working paper*, 1906.10572.

Fuss, S., Flachsland, C., Koch, N., Kornek, U., Knopf, B., and Edenhofer, O. (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.

Gerlagh, R., Heijmans, R. J. R. K., and Rosendahl, K. E. (2020). COVID-19 Tests the Market Stability Reserve. *Environmental and Resource Economics*, 76(4):855–865.

Guo, J., Gu, F., Liu, Y., Liang, X., Mo, J., and Fan, Y. (2020). Assessing the impact of ETS trading profit on emission abatements based on firm-level transactions. *Nature Communications*, 11(1):2078.

Hicks, J. (1939). *Value and Capital*. Oxford University Press, Cambridge.

Hintermann, B. (2010). Allowance price drivers in the first phase of the EU ETS. *Journal of Environmental Economics and Management*, 59(1):43–56.

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

Hintermann, B., Peterson, S., and Rickels, W. (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.

Hirshleifer, D. (1990). Hedging Pressure and Futures Price Movements in a General Equilibrium Model. *Econometrica*, 58(2):411–428.

Hitzemann, S., Uhrig-Homburg, M., and Ehrhart, K. M. (2015). Emission permits and the announcement of realized emissions: Price impact, trading volume, and volatilities. *Energy Economics*, 51:560–569.

Ibikunle, G., Gregoriou, A., Hoepner, A. G., and Rhodes, M. (2016). Liquidity and market efficiency in the world's largest carbon market. *British Accounting Review*, 48(4):431–447.

Ibrahim, B. M. and Kalaitzoglou, I. A. (2016). Why do carbon prices and price volatility change? *Journal of Banking and Finance*, 63:76–94.

ICAP (2019). Emissions Trading Worldwide: Status Report 2019. Technical report, Berlin: ICAP.

Ji, C.-J., Hu, Y.-J., Tang, B.-J., and Qu, S. (2021). Price drivers in the carbon emissions trading scheme: Evidence from Chinese emissions trading scheme pilots. *Journal of Cleaner Production*, 278:123469.

Jiao, L., Liao, Y., and Zhou, Q. (2018). Predicting carbon market risk using information from macroeconomic fundamentals. *Energy Economics*, 73:212–227.

Joyeux, R. and Milunovich, G. (2010). Testing market efficiency in the EU carbon futures market. *Applied Financial Economics*, 20(10):803–809.

Kalaitzoglou, I. and Ibrahim, B. M. (2013). Does order flow in the European carbon futures market reveal information? *Journal of Financial Markets*, 16(3):604–635.

Kaldor, N. (1939). Speculation and Economic Stability. *The Review of Economic Studies*, 7(1):1.

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

Keppler, J. H. and Mansanet-Bataller, M. (2010). Causalities between CO_2, electricity, and other energy variables during phase I and phase II of the EU ETS. *Energy Policy*, 38(7):3329–3341.

Keynes, J. (1930). *A treatise on money, Vol 2.* Macmillan, London.

Koch, N. and Basse Mama, H. (2019). Does the EU Emissions Trading System induce investment leakage? Evidence from German multinational firms. *Energy Economics*, 81:479–492.

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

Lucia, J. J., Mansanet-Bataller, M., and Pardo, Á. (2015). Speculative and hedging activities in the European carbon market. *Energy Policy*, 82:342–351.

Lucia, J. J. and Pardo, A. (2010). On measuring speculative and hedging activities in futures markets from volume and open interest data. *Applied Economics*, 42(12):1549–1557.

Lutz, B. J., Pigorsch, U., and Rotfuß, W. (2013). Nonlinearity in cap-and-trade systems: The EUA price and its fundamentals. *Energy Economics*, 40:222–232.

Mansanet-Bataller, M. and Pardo, A. (2009). Impacts of regulatory announcements on CO_2 prices. *The Journal of Energy Markets*, 2(2):75–107.

Mansanet-Bataller, M., Pardo, A., and Valor, E. (2007). CO_2 Prices , Energy and Weather. *The Energy Journal*, 28(3):73–92.

Mansanet-Bataller, M. and Sanin, M.-E. (2014). Regulation as determinant of EUA prices. *Energy Studies Review*, pages 66–89.

Marin, G., Marino, M., and Pellegrin, C. (2018). The Impact of the European Emission Trading Scheme on Multiple Measures of Economic Performance. *Environmental and Resource Economics*, 71(2):551–582.

Medina, V., Pardo, Á., and Pascual, R. (2014). The timeline of trading frictions in the European carbon market. *Energy Economics*, 42(2014):378–394.

Mizrach, B. and Otsubo, Y. (2014). The market microstructure of the European climate exchange. *Journal of Banking and Finance*, 39(1):107–116.

Mo, J.-L., Zhu, L., and Fan, Y. (2012). The impact of the EU ETS on the corporate value of European electricity corporations. *Energy*, 45(1):3–11.

Narayan, P. K. and Sharma, S. S. (2015). Is carbon emissions trading profitable? *Economic Modelling*, 47:84–92.

Niblock, S. J. and Harrison, J. L. (2013). Carbon markets in times of VUCA: a weak-form efficiency investigation of the phase II EU ETS. *Journal of Sustainable Finance & Investment*, 3(1):38–56.

Osorio, S., Tietjen, O., Pahle, M., Pietzcker, R. C., and Edenhofer, O. (2020). Reviewing the Market Stability Reserve in light of more ambitious EU ETS emission targets.

Palao, F. and Pardo, A. (2012). Assessing price clustering in European Carbon Markets. *Applied Energy*, 92:51–56.

Palao, F. and Pardo, A. (2014). What makes carbon traders cluster their orders? *Energy Economics*, 43:158–165.

Palao, F. and Pardo, A. (2017). Do carbon traders behave as a herd? *North American Journal of Economics and Finance*, 41:204–216.

Patterson, D. M. and Sharma, V. (2006). Do traders follow each other at the NYSE?

Perino, G. (2018). New EU ETS Phase 4 rules temporarily puncture waterbed. *Nature Climate Change*, 8(4):262–264.

Petrick, S. and Wagner, U. (2014). The impact of carbon trading on industry: Evidence from German manufacturing firms. Technical Report 1912.

Pinho, C. and Madaleno, M. (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. (2020). Using supply-side policies to raise ambition: the case of the EU ETS and the 2021 review. Centre for Climate Change Economics and Policy Working Paper 366/Grantham Research Institute on Climate Change and the Environment Working Paper 335.

Rickels, W., Duscha, V., Keller, A., and Peterson, S. (2007). The determinants of allowance prices in the European emissions trading scheme: Can we expect an efficient allowance market 2008? *Kiel Working Paper*, 1378.

Rickels, W., Görlich, D., and Peterson, S. (2014). Explaining european emission allowance price dynamics: Evidence from phase II. *German Economic Review*, 16(2):181–202.

Rittler, D. (2012). Price discovery and volatility spillovers in the European Union emissions trading scheme: A high-frequency analysis. *Journal of Banking and Finance*, 36(3):774–785.

Rogge, K. S., Schleich, J., Haussmann, P., Roser, A., and Reitze, F. (2011). The role of the regulatory framework for innovation activities: the EU ETS and the German paper industry. *International Journal of Technology, Policy and Management*, 11(3/4):250.

Rosendahl, K. E. (2019). EU ETS and the waterbed effect. *Nature Climate Change*, 9(10):734–735.

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.

Sanin, M.-E., Violante, F., and Mansanet-Bataller, M. (2015). Understanding volatility dynamics in the EU-ETS market. *Energy Policy*, 82(1):321–331.

Schultz, E. and Swieringa, J. (2014). Catalysts for price discovery in the European Union Emissions Trading System. *Journal of Banking and Finance*, 42(1):112–122.

Tóth, B., Palit, I., Lillo, F., and Farmer, J. D. (2015). Why is equity order flow so persistent? *Journal of Economic Dynamics and Control*, 51:218–239.

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

Uhrig-Homburg, M. and Wagner, M. (2009). Futures Price Dynamics of CO 2 Emission Allowances: An Empirical Analysis of the Trial Period. *The Journal of Derivatives*, 17(2):73–88.

Wen, F., Wu, N., and Gong, X. (2020). China's carbon emissions trading and stock returns. *Energy Economics*, 86:104627.

Working, H. (1949). The Theory of Price of Storage. *The American Economic Review*, 39(6):1254–1262.

Zhang, Y. J. (2016). Research on carbon emission trading mechanisms: current status and future possibilities. *International Journal of Global Energy Issues*, 39(1/2):89.

Zhang, Y. J. and Wei, Y. M. (2010). An overview of current research on EU ETS: Evidence from its operating mechanism and economic effect. *Applied Energy*, 87(6):1804–1814.

Zhao, X.-g., Wu, L., and Li, A. (2017). Research on the efficiency of carbon trading market in China. *Renewable and Sustainable Energy Reviews*, 79:1–8.

Zhu, B., Chevallier, J., Ma, S., and Wei, Y. (2015). Examining the structural changes of European carbon futures price 2005â“2012. *Applied Economics Letters*, 22(5):335–342.

Chapter 3

Reducing the cost of delay: on the interaction of cap-and-trade and subsidies for clean energy[1]

Oliver Tietjen

[1]Submitted to *Journal of Public Economics*.

Reducing the cost of delay: on the interaction of cap-and-trade and subsidies for clean energy

Oliver Tietjen[a,b]

[a] *Technical University of Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany*
[b] *Potsdam Institute for Climate Impact Research, Telegraphenberg A 31, 14473 Potsdam, Germany.*
E-mail: oliver.tietjen@pik-potsdam.de, phone: +49 331 288 20796

Abstract

Relying on analytical and numerical modeling, I show that subsidies for clean energy can be welfare enhancing when a cap-and-trade (CAT) program is in place and if there is only the carbon externality. The growth rate of the permit price in the CAT program is too high if intertemporal permit trading is unconstrained implying too low prices and too high emission levels early on. Subsidies shift emissions to the future and thus, reduce carbon damage during the transition to carbon neutrality. The optimal subsidy path does not directly affect the permit price, but it is not time-consistent. The time-consistent subsidy has a direct permit price-reducing effect but is still welfare enhancing compared to a CAT-only policy. Subsidies also affect the regulator's choice of the cap and reduce the permit price volatility. Overall, subsidies can be a reasonable second-best alternative if more efficient instruments are not available.

Keywords: cap-and-trade, renewable support, energy subsidy, carbon pricing, stock pollutant

JEL codes: D62, D92, H23, Q48, Q58

October 2020

1. Introduction

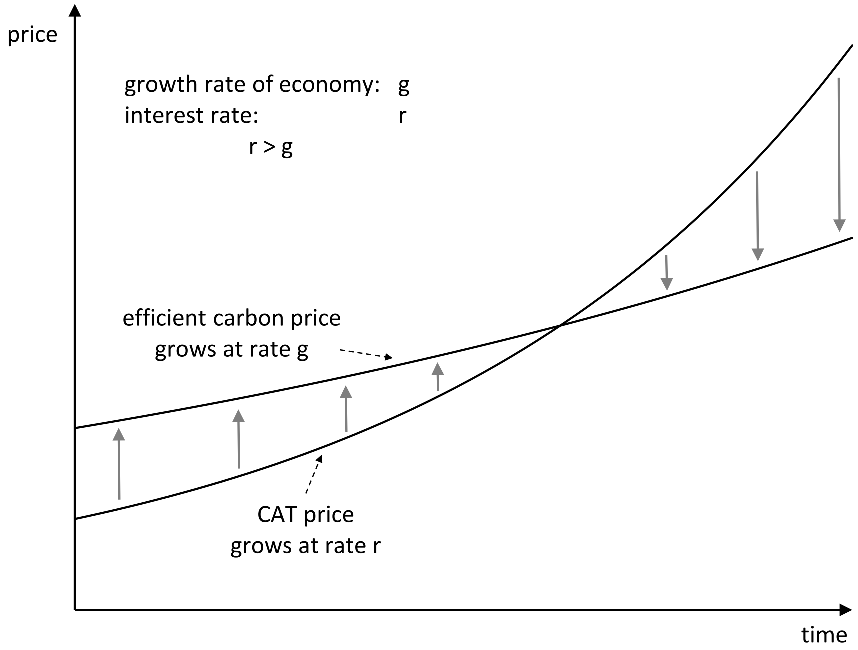
The economic textbook solution to address climate change is to price the carbon externality. The optimal price should reflect the present value of the marginal social damage of emissions[1]: the social cost of carbon (SCC). However, many countries rely on alternative policies such as subsidies for clean production to decarbonize (parts of) their economies. Virtually all jurisdictions that have implemented carbon pricing via cap-and-trade (CAT) programs, such as the European Union (EU) and California, subsidize clean energy in the same sector as well (ICAP 2019; REN21 2019). In the absence of other market failures, such overlapping policies are often considered inefficient and ineffective (Böhringer and Rosendahl 2010; Fankhauser et al. 2010): Production subsidies distort (reduce) the permit price of the CAT program, but they do not affect cumulative emissions if the cap is fixed.[2] This is also called the "waterbed effect" because the amount of emissions is fixed, and overlapping policies change only the distribution of the emissions (Perino 2018).

In this paper, I reconsider the waterbed effect based on a theoretical model and show that adding clean production subsidies to a sector regulated by a CAT program can be welfare-enhancing. This result also holds if the cap of the CAT program is chosen optimally and if there are no other market failures. The reason is that CAT with free intertemporal trading induces a permit price path that is steeper than the path of the SCC. Thus, abatement is delayed, and the carbon damage is inefficiently high in a CAT-only regime. I show that subsidies can complement such a CAT program because they correct the carbon price path as depicted in Figure 1.1. An initial high and declining subsidy over time raises the total carbon price (CAT price plus subsidy) early on, and reduces the total price later on. In doing so, subsidies shift emissions to the future and reduce the "cost of delay" (Goulder 2020).

[1] For simplicity, the terms "emission" and "carbon" are used interchangeably in this paper.

[2] But a subsidy can affect emissions in linked sectors that are not covered by the cap (Jarke and Perino 2017). Moreover, in some programs the cap is endogenous due to price collars or other instruments such as the Market Stability Reserve in the EU ETS. In this paper, I focus on CAT programs with fixed caps and ignore links to other industries.

Figure 1.1: Subsidies can correct the inefficiently steep price path of CAT programs with free banking



Why is the CAT-induced price path too steep? The permit price of a CAT program with unconstrained intertemporal trading rises at the rate of interest r (Cronshaw and Kruse 1996; Rubin 1996) because market agents exploit intertemporal arbitrage. Specifically, they equalize the discounted marginal value of permits over time, which implies that the permit value rises at the discount rate r.[3] The optimal carbon price (i.e., the SCC), in contrast, rises at the growth rate of the economy g (Golosov et al. 2014).[4] An important reason for this is that, in the SCC literature, the carbon damage is typically assumed to be proportional to the economy's output (Dietz and Venmans 2019; Nordhaus 2017). In numerical estimations of the SCC, the growth rate of the SCC is often significantly smaller than the interest rate, and consequently, the permit price path is too steep.[5] For example, in van der Ploeg's (2018) work, the annual growth rate of the

[3]This is also known as Hotelling's rule and can be traced back to Hotelling (1931), who analyzed exhaustible resources.

[4]Note that the SCC can also grow at other rates, for example, if uncertainty is considered (Dietz et al. 2018; Daniel et al. 2019). However, for the argument of this paper it suffices that the CAT price does not rise at the same rate as the SCC.

[5]Note that the SCC are usually estimated with neoclassical growth models in which the real interest rate must be larger than the growth rate of the economy, $r > g$; otherwise, the transversality condition would not hold (Acemoglu 2009).

optimal carbon price is 2%, while the interest rate is 4.4%, and in the estimates of the Interagency Working Group on the Social Cost of Greenhouse Gases (2016), the interest rate is also about twice as high as the growth rate of the SCC.

To correct the growth rate of the permit price, regulators could prohibit intertemporal permit trading, as this would avoid the permit price growing at the interest rate. However, under uncertainty banking has welfare advantages compared to fixed periodic caps because the impact of shocks is reduced (Fell et al. 2012; Weitzman 2020). In fact, banking is allowed in virtually all larger CAT programs (ICAP 2019). Although borrowing is usually prohibited, intertemporal trading is practically unconstrained in many CAT programs because of large permit banks, which implies that borrowing constraints do not play a role. An alternative to prohibiting banking is an intertemporal trading ratio that adjusts banked permits by a certain factor so that the growth rate of the permit price is corrected (Kling and Rubin 1997; Leiby and Rubin 2001; Yates and Cronshaw 2001). However, in the presence of uncertainty, intertemporal trading ratios lead to the optimum only if the cap and the trading ratio are constantly updated according to predefined policy rules, which seems challenging from a regulatory perspective (Pizer and Prest 2020). To the best of my knowledge, regular intertemporal trading ratios have never been implemented in the context of climate change.[6] Therefore, the case of free intertemporal trading combined with clean subsidies analyzed in this paper is of high practical relevance.

I use a theoretical model in which a regulator implements a policy or a policy mix to internalize a stock pollutant (carbon emissions) from an energy sector in an infinite horizon setting. Specifically, the regulator decides about the level of a subsidy for clean production, a cap of a CAT program and a carbon tax and, in doing so, acts as a Stackelberg leader for a dirty (carbon-emitting) and clean (carbon-free) firm. Based on analytical modeling, and in a numerical application to the EU electricity sector, I find

[6] Pizer and Prest (2020) outline where intertemporal trading ratios are or were discussed. In addition to the Waxman-Markey bill in the United States, which was never implemented, they mention China and New Zealand. In both countries, discount factors for permits are seen as a temporary measure to reduce overall banked permits. Interestingly, this implies a trading ratio of less than one, whereas I show that a ratio of greater than one in combination with a lower cap would be required for optimality.

that adding a subsidy when a CAT program is in place can significantly enhance welfare compared to the CAT-only case. The optimal subsidy path declines over time when added to a CAT program: In the beginning, a high subsidy complements the inefficient low permit price and thus postpones the carbon damage. As the permit price rises faster than the SCC, a lower subsidy is needed over time. Moreover, the optimal subsidy does not directly affect the permit price because the subsidy becomes negative at the end of the transition to carbon neutrality, which compensates for the permit price-decreasing effect of the positive subsidy at other times. However, the optimal subsidy path within the CAT program creates a time-consistency problem. The time-consistent (Markov) subsidy starts higher and never turns negative, and thus, the permit price is directly reduced by the waterbed effect. Nonetheless, the welfare effects are comparable to the commitment solution.

Under abatement cost uncertainty, subsidies have an additional advantage as they reduce the permit price volatility, because subsidies prolong the transition or permit banking phase such that shocks spread to more periods. Under subsidy updating (under time consistency), the abatement path is further stabilized because the subsidy and the permit price are negatively correlated. With that said, the commitment and the time-consistent subsidy also have adverse effects under uncertainty: As the former is not adjusted when new information arrives, and because the latter amplifies the permit price-reducing effect, both induce additional inefficiencies under uncertainty. However, in total, the welfare advantage of a subsidy and CAT mix is comparable to the case under perfect information in the numerical simulation.

This paper is related to different lines of research. First, I build on the findings from general equilibrium models that show that the optimal carbon price grows at the rate of the economy (Golosov et al. 2014), as explained above. Some papers also consider endogenous technological change, which justifies a policy mix consisting of clean subsidies and a carbon price (e.g., Acemoglu et al. 2012; Gerlagh et al. 2009).[7] In contrast, I

[7] For a similar argument in a partial equilibrium context see, for example, Fischer and Newell (2008). In addition, multiple instruments can also be justified if there is a positive probability that the cap is not binding (Lecuyer and Quirion 2013; 2019).

disregard general equilibrium effects and technological change to keep the model simple and focus on the instrument choice to regulate the climate externality. I show that a policy mix can also be justified by the inefficient price path of CAT programs if permit banking (and borrowing) is unconstrained.

By comparing different policy instruments, this paper is also closely related to the partial equilibrium literature on the regulation of pollutants. Hoel and Karp (2001; 2002) and Newell and Pizer (2003) are the first to extend Weitzman's (1974) seminal work to stock pollutants and multiple periods, but without intertemporal permit trading. Similar to my work, Fell et al. (2012) integrate intertemporal trading in this setting and find that such trading improves the performance of CAT programs, but according to Weitzman (2020), banking and borrowing are dominated by either prices or fixed quantities. Pizer and Prest (2020) show that CAT can induce the first best, notably because of intertemporal permit trading, but only if policy-updating rules are in place. However, cap updating is not time-consistent and the Markov policy implies a welfare loss (Karp 2019; Kuusela and Lintunen 2020; Lintunen and Kuusela 2018).[8] I contribute to this literature by examining subsidies within CAT programs as a second-best alternative to carbon taxes or intertemporal trading ratios while considering the commitment and the time-consistent solution.[9]

A related strand of literature analyzes the welfare effects of a combination of price- and quantity-based instruments going back to Roberts and Spence (1976) and Weitzman (1978). In the context of climate change, such hybrid policies are typically analyzed as a CAT program with price collars (Fell and Morgenstern 2010; Fell et al. 2012; Grüll and Taschini 2011; Pizer 2002). I find that subsidies share some properties of a price collar as they reduce the permit price volatility and stabilize the abatement path. Additionally, I show that subsidies affect the regulator's choice of the cap by shifting emissions to the

[8] Other theoretical papers that consider policy updating either assume time-invariant damage (Gerlagh and Heijmans 2018) or ignore explicit damage altogether (Kollenberg and Taschini 2016; Newell et al. 2005). Karp and Traeger (2018) and Karp (2019) study updating of taxes and quantities, but the former ignore intertemporal trading of permits, and the latter models a flow pollutant.

[9] Subsidies to internalize the carbon externality have been analyzed in various contexts (e.g., Abrell et al. 2019; Van Der Ploeg and Withagen 2014), but, to the best of my knowledge, subsidies as a means to correct a dynamically inefficient CAT price have yet not been considered.

future and prolonging the transition phase. Nonetheless, subsidies are only a second-best alternative because they distort prices and therefore lead to overconsumption of energy in my model.

After explaining the general model setup in section 2.1, I derive the social planner solution as an optimal benchmark in section 2.2. In section 2.3, I analyze the decentralized solutions beginning with the carbon tax (section 2.3.1), followed by the CAT program (2.3.2), the clean subsidy (2.3.3), and the combination of CAT and the subsidy (2.3.4). In section 3, I apply the model numerically to the EU electricity sector, where I differentiate between the case of perfect information in section 3.2 and uncertain abatement costs in section 3.3. Section 4 provides a conclusion.

2. The model

In the following subsection, I describe the general model setting and then derive the optimal social planner solution in section 2.2. The policy instruments are analyzed in section 2.3.

2.1. General setup

I consider an energy sector with a representative consumer and two competing energy suppliers in an infinite horizon setting with time steps $t = 0, 1, ..., \infty$. Energy generation causes emissions that result in societal damage. To internalize the damage, a regulator sets the levels of different policy instruments while taking the behavior of the market actors into account. The regulator has access to the same information as the market actors. In section 3, I introduce a shock on the abatement costs that is also symmetric to all actors. In the following section, I describe the market actors and the regulator in more detail.

2.1.1. Energy producers

There are two competitive and representative firms i that generate energy x_{it}. The dirty firm, $i = d$, represents fossil fuel technologies (e.g., coal, gas) with $\phi_d > 0$ emissions per unit of production such that it emits $q_t = \phi_d x_{d,t}$, and the clean firm, $i = c$, is equipped with a carbon-free technology (e.g., wind, nuclear), and thus $\phi_c = 0$. Production costs are given by the continuous and convex functions $C_{it}(x_{it})$ ($C'_{it} > 0$ and $C''_{it} > 0$). To obtain closed-form solutions and to calibrate the model to the EU electricity sector in section 3, I assume the functional form to be

$$C_{it} = \alpha_{it} x_{it} + \frac{\beta_i}{2} x_{it}^2, \tag{2.1}$$

where α_{it} and β_i are cost function parameters. Periodic firm profits are

$$\pi_{it} = w_t x_{it} - C_{it} + \Gamma \tag{2.2}$$

with w_t as energy price and Γ as placeholder for climate policies (tax, CAT, subsidy), which are explained in more detail in section 2.3. The producers' objective is to maximize intertemporal profit

$$\max_{x_{it}} \sum\nolimits_{s=t}^{\infty} \frac{1}{(1+r)^{s-t}} \pi_{is}(x_{is}), \tag{2.3}$$

where r is a discount rate.

2.1.2. Consumer

On the demand side, a representative consumer derives utility $U_t(D_t, N_t)$ from the consumption of D_t units of energy and N_t units of a numeraire (the numeraire's price is normalized to one). The function U_t is continuous and increasing in both arguments, and preferences may change over time. Let $\mathit{\Delta}_t$ be carbon damage and let a_t be an exogenous endowment received by the consumer at the beginning of each period. The endowment reflects the income from the general economy, which increases over time at rate g,

$$a_t = a_0 (1+g)^t. \tag{2.4}$$

Furthermore, the consumer owns the firms (that is, the consumer receives the firm profits $\sum_i \pi_{it}$), and thus the consumer's budget constraint can be written as

$$w_t D_t + N_t = a_t - \Delta_t + \sum_i \pi_{it} + \Gamma. \tag{2.5}$$

The variable Γ appears in equation (2.5) because I assume that the state revenues of the climate policies (tax and CAT) are allocated back to the consumer and that state costs (subsidy) are financed by the consumer on a lump-sum basis.[10] The objective of the consumer is to maximize utility

$$\max_{D_s, N_s} \sum\nolimits_{t=s}^{\infty} \frac{1}{(1+r)^{s-t}} U_s \left(D_s, N_s\right) \tag{2.6}$$

subject to the budget constraint (2.5). Solving this problem yields energy demand, which must be equal to the supply in equilibrium:

$$D_t = \sum_i x_{it}. \tag{2.7}$$

As the focus is on the energy market and to streamline the analysis, I assume the consumer has quasilinear preferences:

$$U_t = \frac{\gamma_t D_t - 0.5 D_t^2}{\epsilon} + N_t, \tag{2.8}$$

where γ_t and ϵ are preference parameters. This specification ignores the wealth effect on the energy demand, which simplifies the analysis. I show in section 2.2 that the assumed functional forms for consumer utility and energy generation costs imply a quadratic abatement cost function, which is standard in the price vs. quantities literature.

[10] Abrell et al. (2019) show that alternative financing rules for the subsidy can significantly increase welfare. However, to streamline the analysis I only consider the case in which the subsidy is financed on a lump sum basis.

2.1.3. Climate damage and the regulator's problem

Carbon emissions accumulate in the atmosphere and cause societal damage, which a regulator seeks to internalize. The stock of emissions in t is

$$\Phi_t = \Phi_{t-1} + q_t. \tag{2.9}$$

To keep the model parsimonious, I ignore the decay of emissions, which is explicitly modeled in the SCC literature and sometimes also considered in similar (partial equilibrium) models like the one presented here.[11]

In line with the SCC literature (Dietz and Venmans 2019; Golosov et al. 2014; Nordhaus 2017), I assume that the carbon damage is multiplicative to the economy,

$$\mathit{\Delta}_t = \delta a_t \Phi_t, \tag{2.10}$$

where δ is a damage function parameter. As δ and a_t are exogenous, equation (2.10) implies constant marginal damage within a period. Constant marginal damage is in contrast to the SCC literature, but a standard assumption in partial equilibrium models (Lintunen and Kuusela 2018; Newell and Pizer 2003). It can be justified by the focus on the regulation of a small sector relative to global emissions; therefore, the sector's emissions do not significantly affect the marginal damage.

The regulator maximizes welfare by choosing the optimal levels of the respective policy instruments reflected by Γ. Welfare is equal to the sum of the discounted utility of the representative consumer. That is, the regulator's problem is

$$\max_{\Gamma} W_0 = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} U_t\left(D_t, N_t\right) \tag{2.11}$$

subject to the dynamic emission stock (2.9) and the market equilibrium conditions that

[11]The decay rate for carbon is low. For example, Newell and Pizer (2003) assume 0.83%, and Karp and Traeger (2018) 0.3%. Moreover, Dietz and Venmans (2019) argue based on results from natural sciences that the effect of the decay rate is more or less offset by other effects such as the saturation of carbon sinks.

solve the producer's and consumer's problem $(w_t^*, x_{it}^*, D_t^*, N_t^*)$. Note that the policy levels are chosen by the regulator at the beginning of the first period for the present and all future periods. However, if a CAT program and clean energy subsidies are implemented, I additionally consider the case where the subsidy level can be reset at the beginning of each period. The reason is that the initially implemented subsidy schedule is dynamically inefficient if subsidies are combined with a CAT program, which is explained in more detail in section 2.3.4.[12]

2.2. Social planner solution

Before the different policy instruments are examined, the optimal solution is derived as a benchmark. The social planner chooses consumption levels and energy production to maximize welfare W_0:

$$\max_{D_t, N_t x_{it}} W_0 = \sum\nolimits_{t=0}^{\infty} \frac{1}{(1+r)^t} U_t\left(D_t, N_t\right) \tag{2.12}$$

subject to the budget constraint (2.5), the equilibrium condition (2.7) and the emission stock (2.9). Solving the problem (see Appendix A.1) yields that the marginal utility from consuming energy must be equal to the marginal production costs plus the cost of the externality:

$$\frac{\partial U_t}{\partial D_t} = \frac{\partial C_{it}}{\partial x_{it}} + \sigma_t \phi_i. \tag{2.13}$$

The variable σ_t reflects the SCC and is defined as follows:

$$\sigma_t = \delta a_0 \sum_{s=t}^{\infty} \frac{(1+g)^s}{(1+r)^{s-t}} = \delta a_0 \left(1+g\right)^t \frac{1+r}{r-g}. \tag{2.14}$$

The SCC is the sum of the discounted marginal damage of all upcoming periods. Intuitively, the SCC increases with damage and decreases with the discount rate because future damage weighs less. The SCC rises at rate g and, given the infinite horizon setting,

[12] While in the case of a tax- and subsidy-only policy, the regulator has no incentive to reset the policy levels, the CAT program is also dynamically inefficient, a finding which has been shown elsewhere (Lintunen and Kuusela 2018; Kuusela and Lintunen 2020). However, to focus attention on the main results, I assume that the initially implemented CAT program is always fixed.

it must hold $r > g$ for the series $\sum_{s=t}^{\infty} \frac{(1+g)^s}{(1+r)^{s-t}}$ to converge. Put differently, if $r \leq g$, then the marginal damage rises as fast or faster than the weight of future periods decreases and thus, $\sigma_t \rightarrow \infty$, implying zero emissions would always be optimal. To avoid such trivial solutions, I concentrate on the case $r > g$. As mentioned in the introduction, a higher interest rate r than the growth rate of the economy g is in line with standard neoclassical models.

Due to the assumptions of quadratic energy generation costs (2.1) and quasilinear and quadratic consumer preferences (2.8), it can be shown (see Appendix A.1.2) that abatement costs are quadratic as in the standard framework for competitive permit markets and more broadly in the price vs. quantities literature beginning with Weitzman (1974). Deriving the parameters of the abatement cost function is useful to simplify the notation throughout the paper. Specifically, abatement costs are

$$AC_t = \frac{\psi}{2} (\varphi_t - q_t)^2 \tag{2.15}$$

with business-as-usual (BAU) emissions φ_t and slope of the marginal abatement costs ψ:

$$\varphi_t = \frac{\phi_d (\beta_c \gamma_t - \alpha_{d,t} (1 + \beta_c \epsilon) + \alpha_{c,t})}{\beta_c + \beta_d + \beta_c \beta_d \epsilon}, \tag{2.16}$$

$$\psi = \frac{\beta_c + \beta_d + \beta_c \beta_d \epsilon}{\phi_d^2 (1 + \beta_c \epsilon)}. \tag{2.17}$$

BAU emissions increase with the energy demand (γ_t), emission factor (ϕ_d) and costs of the clean technology ($\alpha_{c,t}, \beta_c$), and they decrease with the costs of the dirty firm ($\alpha_{d,t}, \beta_d$). The slope of the marginal abatement costs also depends on the technology costs, emission factor and demand reaction to price (ϵ).

2.3. Decentralized solutions: policy instrument analysis

In the decentralized economy, consumption and energy production decisions are made by the utility-maximizing consumer and the profit-maximizing firm, respectively. Without policy intervention, market agents ignore the SCC σ_t and therefore optimality condition (2.13) cannot be fulfilled. To internalize the SCC, the regulator chooses the optimal levels of different policy instruments. In the following, a carbon tax, a CAT program, a clean energy subsidy, and a combination of CAT and subsidy are considered policy instruments. For this purpose, I assume that there is a single finite transition phase $t = 0, 1, ..., \hat{t}$ per policy scenario with $\hat{t}$ as the last period with positive emissions, $q_t > 0$. Specifically, I assume that for $t \leq \hat{t}$ both the clean and the dirty firm supply energy and for $t > \hat{t}$, the transition to carbon neutrality is completed, demand is fully satisfied by the clean firm and thus, $q_t = 0$. This assumption is mild because the stringency of all policy instruments (endogenously) increases over time, so that periodic emissions eventually reach zero.

2.3.1. Carbon tax

The first policy instrument is a carbon tax τ_t. The tax has to be paid by firms per unit of emissions; therefore, the firm profits become:

$$\pi_{it} = w_t x_{it} - \alpha_{it} x_{it} - \frac{\beta_i}{2} x_{it}^2 - \tau_t \phi_i x_{it}. \tag{2.18}$$

Maximizing profits via production x_{it} and considering the energy market equilibrium (2.7) gives production depending only on the tax and parameters (see Appendix A.2.1):

$$x_{d,t} = \begin{cases} \frac{\beta_c \gamma_t + \alpha_{c,t} - \left(\alpha_{d,t} + \tau_t \phi_d\right)(1+\beta_c \epsilon)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} = \frac{\varphi_t}{\phi_d} - \frac{\tau_t}{\psi \phi_d} & \forall\, t \leq \hat{t}_{tax} \\ 0 & \forall\, t > \hat{t}_{tax} \end{cases} \tag{2.19}$$

$$x_{c,t} = \begin{cases} \frac{\beta_d \gamma_t + \alpha_{d,t} - \alpha_{c,t}(1+\beta_d \epsilon) + \tau_t \phi_d}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} & \forall\, t \leq \hat{t}_{tax} \\ \frac{\gamma_t - \alpha_{c,t} \epsilon}{1+\beta_c \epsilon} & \forall\, t > \hat{t}_{tax}. \end{cases} \tag{2.20}$$

During the transition to carbon neutrality, $t \leq \hat{t}_{tax}$, when both firm types are active, production of both increases with demand parameter γ_t and the costs of the other technology, and production decreases with production costs and the demand reaction to price changes (reflected by ϵ). After the transition $t > \hat{t}_{tax}$, dirty production is by definition zero, and the clean firm fully satisfies the energy demand.

The regulator chooses the tax level at the beginning of the first period:

$$\max_{\tau_0,\ldots\tau_t,\ldots\tau_\infty} W_0 = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} U_t\left(x_{d,t}, x_{c,t}\right) \tag{2.21}$$

subject to the emission stock (2.9). Utility can be rewritten to

$$U_t = \frac{\gamma_t \sum_i x_{it} - 0.5\left(\sum_i x_{it}\right)^2}{\epsilon} - \sum_i \left(\alpha_{it} x_{it} + \frac{\beta_i}{2} x_{it}^2\right) - \delta a_t \Phi_t + a_t \tag{2.22}$$

where energy generation x_{it} is given by (2.19) and (2.20). Solving the problem yields (see Appendix A.2.2):

$$\tau_t = \delta a_0 \sum_{s=t}^{\infty} \frac{(1+g)^s}{(1+r)^{s-t}} = \delta a_0 (1+g)^t \frac{1+r}{r-g}. \tag{2.23}$$

Comparing (2.23) to (2.14) shows that the optimal tax is set equal to the SCC, $\tau_t = \sigma_t$. Therefore, the carbon damage is perfectly internalized and the first-best solution is achieved.

It is useful to compare the policy instruments with respect to their implied cumulative emissions, for which I additionally assume that BAU emissions φ are constant. Because periodic emissions can be written as $q_t = x_{d,t}\phi_d = \varphi_t - \frac{\sigma_t}{\psi}$ (see equation (2.19)), cumulative emissions under the tax regime are

$$Q_{tax} = \sum_{t=0}^{\hat{t}_{tax}} q_t = \sum_{t=0}^{\hat{t}_{tax}} \left(\varphi_t - \frac{\sigma_t}{\psi}\right) \tag{2.24}$$

where $q_t = 0 \,\forall\, t > \hat{t}_{tax}$. Equation (2.24) shows that the cumulative emissions depend on the transition length, which is (see Appendix A.2.2)

$$\hat{t}_{tax} = \frac{1}{g} ln \left(\frac{\varphi\psi}{\sigma_0}\right). \tag{2.25}$$

Larger SCC (larger g and σ_0) and lower abatement costs (lower φ and ψ) imply a shorter transition to emission neutrality (and lower cumulative emissions) and vice versa.

2.3.2. Cap-and-trade

Compliance with the CAT program requires that the dirty firm holds one permit for each unit of emission. Permits can be bought at auctions at the beginning of each period. The permit supply of the regulator is z_t, and purchases by the firms are y_{it}, such that in equilibrium it holds

$$z_t = \sum_i y_{it}, \tag{2.26}$$

which is satisfied by permit price p_t. If firms buy more permits than required, $y_{it} - x_{it}\phi_i > 0$, they can bank the permits for later use,

$$b_{it} = b_{it-1} + y_{it} - x_{it}\phi_i, \tag{2.27}$$

with b_t as the banked permits at the end of period t. I assume that a potential permit borrowing constraint, $b_{it} \geq 0$, never binds before all permits are used up, which simplifies the analysis without affecting the main insights. In particular, this assumption implies that the temporal allocation of permits does not affect the results (Salant 2016).[13] The firm's periodic profits become

$$\pi_{it} = w_t x_{it} - \alpha_{it} x_{it} - \frac{\beta_i}{2} x_{it}^2 - p_t y_{it}. \tag{2.28}$$

Maximizing intertemporal profit subject to the banking dynamics (2.27) leads to the following permit price (see Appendix A.3.1):

$$p_t = \psi \left(\varphi_t - z_t - B_{t-1} + B_t \right), \tag{2.29}$$

with $B_t = \sum_i b_{it}$. The permit price reflects the marginal abatement costs, and the term in the brackets is the amount of abatement. It can be shown further that the expected

[13] Alternatively, one could assume that all permits are issued in the first period, which also implies that the borrowing constraint never binds. Ignoring borrowing constraints can be justified by real markets, such as the EU ETS, in which borrowing constraints have played a minor role thus far.

permit price rises at the discount rate as long as the bank is not depleted:

$$\frac{p_{t+1} - p_t}{p_t} = r \quad \forall\, t \leq \hat{t}_{cat}, \tag{2.30}$$

with $\hat{t}_{cat}$ as the last period of the transition to emission neutrality under the CAT regime and thus, in which permits are available. After the transition is completed, that is, $B_t = 0$ for $t > \hat{t}_{cat}$, the energy production of the dirty firm must be zero, and the permit price is at least as high to guarantee this.

By combining (2.29) and (2.30), I derive the permit banking path:

$$B_t = \frac{\sum_{s>t}^{\hat{t}_{cat}} \left((\varphi_s - z_s) + (B_{t-1} + z_t - \varphi_t)\,(1+r)^{s-t} \right)}{\sum_{s=t}^{\hat{t}_{cat}} (1+r)^{s-t}}. \tag{2.31}$$

Higher BAU emissions φ_s in the future increase permit banking because more permits are required for energy generation. Banking is also higher when the future supply z_s is lower, because banked permits are a substitute for the future supply. The bank (2.31) can be used to express production of dirty and clean firms as follows:

$$x_{d,t} = \frac{\varphi_t}{\phi_d} - \frac{\sum_{s=0}^{\hat{t}_{cat}} \varphi_s - Z_{cat}}{\sum_{s=0}^{\hat{t}_{cat}} (1+r)^{s-t}\, \phi_d}\ \forall\, t \leq \hat{t}_{cat}, \tag{2.32}$$

$$x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon - x_{d,t}}{1 + \beta_c \epsilon}\ \forall\, t \leq \hat{t}_{cat}, \tag{2.33}$$

where $Z_{cat} = \sum_{t=0}^{\hat{t}_{cat}} z_t$ is the total supply of permits or the emissions cap. The first term in (2.32) is the BAU production, that is, the energy generation level with a zero carbon price. Because I am interested in the case in which abatement is required, implying that the overall BAU emissions are larger than the cap, $\sum_{s=0}^{\hat{t}_{cat}} \varphi_s - Z_{cat} > 0$, the second term reduces the production of the dirty sector and essentially reflects abatement that is needed for compliance. Note that after the transition, $t > \hat{t}_{cat}$, it holds $x_{d,t} = 0$ and $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c\epsilon}$ as under the tax regime.

The regulator's problem is

$$\max_{Z_{cat}} W_0 = \sum\nolimits_{t=0}^{\infty} \frac{1}{(1+r)^t} U_t \left(x_{d,t}, x_{c,t}\right), \tag{2.34}$$

subject to the emission stock (2.9) where $U_t \left(x_{d,t}, x_{c,t}\right)$ is (2.22) as under the tax regime, but the reaction functions are now given by (2.32) and (2.33). The solution of this problem leads to the following optimal cap (see Appendix A.3.2):

$$Z_{cat} = \sum_{t=0}^{\hat{t}_{cat}} \left(\varphi_t - \frac{\sigma_t}{\psi} \right). \tag{2.35}$$

The optimal cap is the difference between periodic BAU emissions φ_t and the ratio between the SCC and the slope of the marginal abatement costs, $\frac{\sigma_t}{\psi}$, summed over the transition period. Higher BAU emissions and a steeper marginal abatement cost curve (larger ψ) lead to a larger cap because both parameters increase the abatement costs and consequently imply higher optimal emissions. Intuitively, the optimal cap decreases with the SCC σ_t.

To compare the CAT program to the tax, I first write the permit price as a function of the SCC. Using emissions $x_{d,t}\phi_d$ (with $x_{d,t}$ from (2.32)) in the expression for the permit price $p_t = \psi \left(\varphi_t - z_t - B_{t-1} + B_t\right) = \psi \left(\varphi_t - x_{d,t}\phi_d\right)$ gives

$$p_t = (1+r)^t \frac{\sum_{s=0}^{\hat{t}_{cat}} \sigma_s}{\sum_{s=0}^{\hat{t}_{cat}} (1+r)^s} \; \forall \, t \leq \hat{t}_{cat}, \tag{2.36}$$

which implies the following result.

Proposition 1. *The CAT program with free intertemporal trading is inefficient.*

This directly follows from (2.36) because the permit price p_t is generally not equal to the SCC σ_t due to the too-large growth rate of the permit price, $r > g$. The large growth rate also has an implication for the optimal cap Z_{cat} and in turn for the cumulative emissions (see Appendix A.3.3 for proof).

Proposition 2. *Assuming constant BAU emissions φ, the optimal cap of the CAT program with unconstrained intertemporal trading is lower than the cumulative emissions of the tax regime, $Z_{cat} < Q_{tax}$.*

Inspecting equations (2.35) and (2.24) shows that the difference between the cumulative emissions of both instruments depends on the transition lengths ($\hat{t}_{cat}$, $\hat{t}_{tax}$). We can assume for a moment that the lengths and cumulative emissions were equal, and thus, proposition 2 was not true. Then, the higher growth rate of the permit price compared to the tax, $r > g$, implies that the permit price is initially lower and higher in the late phase of the transition (this follows from (2.36)). The initially lower permit price leads to an earlier accumulation of emissions compared to the tax. This implies that if the cumulative emissions are equal at the *end of the transition*, the emission stock, and thus the damage is higher *during the entire transition* under the CAT regime. This higher damage reflects the cost of delay induced by free intertemporal trading in the CAT program. To increase welfare, the regulator has an incentive to lower the cap so that it holds $Z_{cat} < Q_{tax}$, because a lower cap partly compensates for the too-high damage of the CAT program during the transition. Put differently, by lowering the cap, the permit price increases in all periods, and thus, the emission stock and the carbon damage are reduced.

Due to $Z_{cat} < Q_{tax}$, the transition lengths cannot be equal as well. With constant BAU emissions, the optimal transition length under the CAT regime is

$$\hat{t}_{cat} = \frac{1}{r} ln \left(\frac{\varphi\psi}{p_0} \right) = \frac{1}{r} ln \left(\frac{\varphi\psi}{\varphi\psi - \frac{r}{g}\left(\sigma_{\hat{t}_{cat}} - \sigma_0\right)} \right), \tag{2.37}$$

for which the following result holds (see Appendix A.3.4 for proof).

Corollary 1. *Assuming constant BAU emissions φ, the optimal duration of the transition under the CAT program with unconstrained intertemporal trading is shorter than under the tax regime, $\hat{t}_{cat} < \hat{t}_{tax}$.*

Therefore, the failure of the too-high growth rate of the permit price implies less cumulative emissions and a shorter transition compared to the first-best solution.

I show in Appendix A.3.5 that the growth rate of the permit price can be corrected by an intertemporal trading ratio such that the first-best result can also be obtained with the CAT program. However, this does not hold under uncertainty if the regulator does not adjust the permit supply and trading ratio in each period (Pizer and Prest 2020). In the numerical simulation, I show the impact of uncertainty on the size of the trading ratio and the implications for welfare and cumulative emissions if the permit supply and trading ratio are set at the beginning of the first period.

2.3.3. Clean subsidy

In this section, I examine the subsidy separately to show its general effects before I analyze its combined implementation with the CAT program in section 2.3.4. Subsidy ρ_t is paid per unit of clean production and thus, $\rho_t = 0$ for the dirty firm. Profits are given by

$$\pi_{it} = (w_t + \rho_t)\, x_{it} - \alpha_{it} x_{it} - \frac{\beta_i}{2} x_{it}^2. \tag{2.38}$$

I obtain the energy price, the optimal dirty and clean energy generation by following the same steps as under the tax regime (see Appendix A.4.1):

$$w_t = \frac{\beta_c \beta_d \gamma_t + \beta_c \alpha_{d,t} + \beta_d (\alpha_{c,t} - \rho_t)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} \;\forall\, t \leq \hat{t}_{sub}. \tag{2.39}$$

$$x_{d,t} = \frac{\varphi_t}{\phi_d} - \frac{\rho_t}{\psi \phi_d^2 (1 + \beta_c \epsilon)} \;\forall\, t \leq \hat{t}_{sub} \tag{2.40}$$

$$x_{c,t} = \frac{\beta_d \gamma_t + \alpha_{d,t} + (\rho_t - \alpha_{c,t})(1 + \beta_d \epsilon)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} \;\forall\, t \leq \hat{t}_{sub} \tag{2.41}$$

When carbon neutrality is reached, $t > \hat{t}_{sub}$, it holds $x_{d,t} = 0$ and $x_{c,t} = \frac{\gamma_t - (\alpha_{c,t} - \rho_t)\epsilon}{1 + \beta_c \epsilon}$.

The regulator considers the following problem:

$$\max_{\rho_0, \dots \rho_t, \dots \rho_\infty} W_0 = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} U_t (x_{d,t}, x_{c,t}) \tag{2.42}$$

subject to the emission stock (2.9). The utility function is again given by (2.22) and the reaction functions are (2.40) and (2.41). This gives the following optimal subsidy path (see Appendix A.4.2):

$$\rho_t = \begin{cases} \sigma_t \frac{\phi_d}{1+\beta_d \epsilon} & \forall\, t \leq \hat{t}_{sub} \\ \varphi_t \psi \phi_d \left(1 + \beta_c \epsilon\right) & \forall\, t > \hat{t}_{sub}. \end{cases} \tag{2.43}$$

During the transition, $t \leq \hat{t}_{sub}$, the subsidy equals the SCC multiplied by the term $\phi_d \left(1 + \beta_d \epsilon\right)^{-1}$, and thus, the subsidy grows at rate g over time. Emission factor ϕ_d reflects that the subsidy is paid for avoided damage and the term $\left(1 + \beta_d \epsilon\right)^{-1}$accounts for the price distortion of the subsidy. The distortion stems from the price sensitivity of demand due to $\epsilon > 0$, because the subsidy depresses the energy price (see expression (2.39)), leading to overconsumption if consumers are price-sensitive. Therefore, the regulator reduces the subsidy by multiplying with $\left(1 + \beta_d \epsilon\right)^{-1}$ to lower the welfare loss arising from the overconsumption. If the energy demand did not depend on the price, $\epsilon = 0$, clean energy would perfectly replace dirty energy, and thus, the subsidy is $\rho_t = \sigma_t \phi_d$, which induces the first best as under the tax regime. Note that after the transition, $t > \hat{t}_{sub}$, the subsidy is just high enough to set emissions to zero because a higher subsidy would further distort the energy price without having an effect on emissions.

Comparing the subsidy to the first best yields the following result (see Appendix A.4.3 for proof).

Proposition 3. *Assuming constant BAU emissions φ, the optimal transition length of the subsidy regime is longer ($\hat{t}_{sub} > \hat{t}_{tax}$) and cumulative emissions are higher ($Q_{sub} > Q_{tax}$) than under the tax regime if $\epsilon > 0$, and they are equal ($\hat{t}_{sub} = \hat{t}_{tax}$, $Q_{sub} = Q_{tax}$) if $\epsilon = 0$.*

The price-distorting effect of the subsidy directly increases emissions through overconsumption and indirectly increases emissions through the reduction of the subsidy by the regulator. Therefore, the cumulative emissions are higher and the transition is longer if $\epsilon > 0$.

2.3.4. Cap-and-trade and clean subsidy combined

Both, a CAT program with free intertemporal trading and subsidies for clean energy, are imperfect instruments. In this section, I consider whether a combination of both performs better.

The dirty firm faces the same problem as in the CAT-only case in section 2.3.2, and the clean firm has the same problem as in the subsidy-only case in section 2.3.3. Solving the firms' problems yields (see Appendix A.5.1):

$$x_{d,t} = \frac{\varphi_t^{sub}}{\phi_d} - (1+r)^t \frac{\sum_{s=0}^{\hat{t}_{c+s}} \varphi_s^{sub} - Z_{c+s}}{\sum_{s=0}^{\hat{t}_{c+s}} (1+r)^s \phi_d} \; \forall \, t \leq \hat{t}_{c+s}, \tag{2.44}$$

$$x_{c,t} = \frac{\gamma_t - (\alpha_{c,t} - \rho_t)\,\epsilon - x_{d,t}}{1 + \beta_c \epsilon} \; \forall \, t \leq \hat{t}_{c+s}, \tag{2.45}$$

where I adjusted the BAU emissions with the subsidy

$$\varphi_t^{sub} = \frac{\phi_d \left(\beta_c \gamma_t - \alpha_{d,t} (1 + \beta_c \epsilon) + \alpha_{c,t} - \rho_t\right)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon}, \tag{2.46}$$

such that φ_t^{sub} are emissions with the impact of the subsidy but without the effect of the permit price. After the transition to carbon neutrality, $t > \hat{t}_{c+s}$, it holds $x_{d,t} = 0$ and $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c \epsilon}$.

Optimal cap and subsidy levels. The interaction between the subsidy and the CAT program introduces a time-consistency problem. Therefore, I analyze two cases: Under the commitment solution, the regulator is able to commit to the entire subsidy path, $\rho_0, \ldots \rho_t, \ldots \rho_\infty$, determined at the beginning of the first period. Under time consistency, the regulator only sets the subsidy for the current period ρ_t where the level of future subsidies is anticipated. The regulator's problems in the commitment and time-consistent solution, respectively, are

$$\max_{Z_{c+s}, \rho_0, \ldots \rho_t, \ldots \rho_\infty} W_0 = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} U_t (x_{d,t}, x_{c,t}) \tag{2.47}$$

$$\max_{Z_{c+s}, \rho_t} W_t = \sum_{s=t}^{\infty} \frac{1}{(1+r)^{s-t}} U_t (x_{d,t}, x_{c,t}) \tag{2.48}$$

both subject to the emission stock (2.9) and with utility function (2.22) as well as reaction functions (2.44) and (2.45). First, I describe results that apply to both cases and then I show differences (see Appendix A.5.2 for all derivations).

Solving (2.47) or (2.48) with respect to Z_{c+s} shows that for the optimal cap it still holds $Z_{c+s} = \sum_{t=0}^{\hat{t}_{c+s}} \left(\varphi_t - \frac{\sigma_t}{\psi}\right)$, where the transition length $\hat{t}_{c+s}$ is the only difference to the cap of the CAT-only case; see equation (2.35). That is, the subsidy has no direct effect on the choice of the optimal cap but an indirect effect via the transition length, as shown below.

For the subsidy, welfare maximization of either (2.47) or (2.48) leads to the following result:

$$\rho_t = \begin{cases} \frac{\phi_d}{1+\beta_d\epsilon}\left(\sigma_t - p_t\right) & \forall\, t \leq \hat{t}_{c+s} \\ 0 & \forall\, t > \hat{t}_{c+s}. \end{cases} \tag{2.49}$$

During the transition, $t \leq \hat{t}_{c+s}$, the optimal subsidy is the difference between the SCC (σ_t) and the ETS price (p_t) weighted by the factor $\phi_d\left(1+\beta_d\epsilon\right)^{-1}$. Similar to the subsidy-only case, the factor reflects that the subsidy is paid for avoided emissions (ϕ_d) and the term $\left(1+\beta_d\epsilon\right)^{-1}$ accounts for the energy price distortion of the subsidy (cf. section 2.3.3). Put differently, the subsidy reduces the gap between the CAT price and the optimal carbon price while considering the inefficiencies induced by the subsidy. After the transition, $t > \hat{t}_{c+s}$, periodic emissions are zero due to the CAT program. Therefore, the optimal subsidy also becomes zero because a non-zero subsidy would only distort the energy price without having any impact on the emissions after period $\hat{t}_{c+s}$ is reached.

Equation (2.49) shows that the subsidy supports the permit price such that the total (implicit) carbon price $\tilde{p}_t$, which consists of the permit price and the subsidy, is closer to the SCC. The total (implicit) carbon price can be written as

$$\tilde{p}_t = p_t + \tilde{\rho}_t = \left(1-\omega\right)\left(1+r\right)^t p_0 + \omega\left(1+g\right)^t \sigma_0 \tag{2.50}$$

with $\tilde{\rho}_t = \rho_t\left(\phi_d\left(1+\beta_c\epsilon\right)\right)^{-1}$ to convert the subsidy into a carbon price in terms of the effectiveness to reduce emissions.[14] The total carbon price is the weighted average of the permit price and the SCC with weight $\omega = ((1+\beta_c \epsilon)(1+\beta_d \epsilon))^{-1}$ and $0 \leq \omega \leq 1$. The weight reflects the inefficiency of the subsidy: The higher the energy price sensitivity (larger ϵ), the lower ω and thus, the higher the share of the permit price in the total carbon price and vice versa. For instance, with $\omega = 1$ (no energy price sensitivity), the permit price would be zero, because the subsidy would be an ideal instrument and the total carbon price would grow at rate g, and with $\omega = 0$, the energy price distortion is prohibitively high such that $\rho_t = 0$, and the growth rate would be r. For cases between the extremes, $0 < \omega < 1$, there is a non-zero subsidy such that the growth rate of the total carbon price is larger than g but lower than r, meaning it is closer to the optimum compared to the CAT-only case.

Adding subsidies to the CAT program also prolongs the transition to carbon neutrality, which can be expressed as follows:

$$\hat{t}_{c+s} = \frac{1}{r} ln \left(\frac{\varphi\psi - \sigma_{\hat{t}_{c+s}}\omega}{p_0 (1-\omega)} \right) = \frac{1}{g} ln \left(\frac{\varphi\psi - p_{\hat{t}_{c+s}} (1-\omega)}{\sigma_0 \omega} \right). \tag{2.51}$$

Comparing (2.51) to the lengths of the tax and the CAT-only regime, (2.25) and (2.37), shows that if the inefficiency of the subsidy is large, $\omega = 0$, equation (2.51) equals the length of the CAT-only regime, $\hat{t}_{c+s} = \hat{t}_{cat}$, because $\omega = 0$ implies that the optimal subsidy is zero (see previous paragraph). Vice versa, if the inefficiency is low, $\omega = 1$, it holds $\hat{t}_{c+s} = \hat{t}_{tax}$, because the permit price is zero. Thus, for $0 < \omega < 1$, the length of the combined CAT and subsidy program is in between the length of the tax- and CAT-only regimes, $\hat{t}_{tax} \geq \hat{t}_{c+s} \geq \hat{t}_{cat}$. This implies that cumulative emissions are also in between the tax and the CAT regime, $Q_{tax} \geq Z_{c+s} \geq Z_{cat}$, because in all three cases cumulative emissions can be expressed as $\sum_{t=0}^{\hat{t}} \left(\varphi_t - \frac{\sigma_t}{\psi} \right)$. Recall that the reason for lowering the cap in the CAT-only case (compared to the cumulative emissions under the tax regime) is to compensate for the too-low permit price and too-high damage early on (see section

[14] This implicit carbon price of the subsidy is the carbon price that leads to the same abatement as the subsidy, which I obtain by setting dirty production in the tax and subsidy case, equations (2.19) and (2.40), equal and solve for the tax.

2.3.2). As the subsidy counteracts the too-high damage during the transition, the cap can be lifted compared to the cap of the CAT-only regime.

Commitment vs. time consistency. In the following, I examine differences between the commitment and the time-consistent solution. First I consider that under commitment, the subsidy (2.49) can be rewritten as:

$$\begin{aligned}\rho_t \;&=\; \frac{\delta a_0 \phi_d}{(1+\beta_d \epsilon)\sum_{s=0}^{\hat{t}_{c+s}} (1+r)^s} \\ &\times \left(\sum_{t<s\leq \hat{t}_{c+s}} \left((1+r)^s \sum_{s'=0}^{\hat{t}_{c+s}-s} (1+g)^{t+s'} \right) - \sum_{1\leq s\leq t} \left((1+r)^s \sum_{s'=1}^{s} (1+g)^{t-s'} \right) \right) \forall t \leq \hat{t}_{c+s}\end{aligned} \tag{2.52}$$

The first term in the brackets reflects the damage of all upcoming periods until the end of the transition from the perspective of period t, and the second term reflects the damage of the present and past periods. Because the first term decreases and the second term increases over time, the subsidy declines over time and becomes negative at the end of the transition, which is not time-consistent: For example, if the regulator could reset the subsidy in the last transition period $\hat{t}_{c+s}$, it would not be optimal to choose $\rho_{\hat{t}_{c+s}} < 0$ as (2.52) suggests, but instead $\rho_{\hat{t}_{c+s}} = 0$ is optimal. The reason is that from the perspective of $\hat{t}_{c+s}$, there is no value of a non-zero subsidy. In the last period, the problem simplifies to a single period problem with a given cap as permit supply $z_{\hat{t}_{c+s}}$ and initial bank $b_{\hat{t}_{c+s}-1}$ are fixed. Thus, any non-zero subsidy would only distort the energy price but not affect the damage. A similar argument can be made for all periods after the first because the subsidy levels depend on past damage, and thus, the subsidy path is not a Markov policy.

The time-consistent policy yields the following subsidy path:

$$\rho_t = \frac{(1+\beta_c\epsilon)\,\delta a_0 \phi_d \sum_{s=t+1}^{\hat{t}_{c+s}} \left((1+r)^{s-t} \sum_{s'=0}^{\hat{t}_{c+s}-s} (1+g)^{t+s'} \right) + \sum_{s>t}^{\hat{t}_{c+s}} \rho_s}{(\beta_d+\beta_c+\beta_d\beta_c\epsilon)\,\epsilon \sum_{s=0}^{\hat{t}_{c+s}-t} (1+r)^s + \sum_{s=1}^{\hat{t}_{c+s}-t} (1+r)^s} \;\forall\, t < \hat{t}_{c+s}. \tag{2.53}$$

In contrast to the commitment solution, the time-consistent subsidy is a Markov policy and does not depend on the past. Put differently, subsidies in t are set only to maximize welfare in $s \geq t$. Therefore, there is no negative term in (2.53) that would reflect the damage of past periods as in the commitment solution. Without a negative term that grows over time, the Markov subsidy declines at a lower rate compared to the commitment solution and never becomes negative. Negative subsidies increase welfare from the current perspective only when they are set for future periods, because in doing so, the regulator pushes emissions to the future. However, negative subsidies decrease welfare when they are set for the current period because they pull emissions to the present, which never makes sense in this setting.

Whether the regulator is able to commit or not also affects the interaction of the subsidy with the permit price. In general, all transition periods are connected via permit banking, and therefore, setting a subsidy in one transition period affects the permit price in all transition periods, ceteris paribus. This can be shown by writing the permit price as

$$p_t = (1+r)^t \left(\frac{\sum_{s=0}^{\hat{t}_{c+s}} \sigma_s}{\sum_{s=0}^{\hat{t}_{c+s}} (1+r)^s} - \frac{\sum_{s=0}^{\hat{t}_{c+s}} \rho_s}{\phi_d (1+\beta_c \epsilon) \sum_{s=0}^{\hat{t}_{c+s}} (1+r)^s} \right) \forall\, t \leq \hat{t}_{c+s}, \tag{2.54}$$

and thus, the permit price is a negative function of all subsidies. A positive subsidy in any transition period reduces the price in the same way because the subsidy reduces the (anticipated) demand for permits equally, irrespective of the period when the subsidy is paid. However, the opposite is true for negative subsidies, leading to the following result (see Appendix A.5.3 for proof).

Proposition 4. *For a given cap, the optimal subsidy path does not affect the permit price in the commitment solution and the optimal subsidy path reduces the permit price in the time-consistent solution.*

In the commitment solution, positive and negative subsidies are perfectly balanced, $\sum_{s=0}^{\hat{t}_{c+s}} \rho_s = 0$, such that they do not directly affect the permit price (see equation (2.54)). This makes sense from a welfare perspective because the permit price is an efficient abatement instrument (within a period), whereas subsidies distort the energy price. Therefore, an ideal subsidy should not crowd out the permit price but only shift emissions and damage to the future. Under time consistency, in contrast, the subsidy never turns negative, which implies $\sum_{s=0}^{\hat{t}_{c+s}} \rho_s > 0$ such that the subsidy directly reduces the permit price.

However, under both commitment and time consistency, the subsidies have an indirect effect on the permit price: By shifting emissions to the future, the subsidies prolong the transition phase compared to the CAT-only regime such that $\hat{t}_{c+s} > \hat{t}_{cat}$ holds, from which $Z_{c+s} \geq Z_{cat}$ follows (as explained above). As a larger cap implies a lower permit price, the permit price under the combined regime (2.54) is lower than under CAT-only (2.36). Nonetheless, the total carbon price (see equation (2.50)) is larger in the early phase and lower in the late phase of the transition compared to CAT-only, as is shown numerically in the next section.

3. Numerical simulation

In this section, I use a numerical simulation to illustrate the analytical results of the previous section and to investigate the effects of abatement cost uncertainty. Before I present the results, I explain how I calibrate the model to the EU electricity sector.

3.1. Model calibration and scenarios

Table 3.1: Firm data

	α_i (EUR/kWh)	β_i (EUR/kWh^2)	ϕ_i (g/kWh)
Clean firm	0.02821	0.0216	0
Dirty firm	0.0456	0.0040	650

Table 3.1 shows the assumed parameters for clean and dirty firms. The parameters are based on the detailed numerical electricity sector model LIMES-EU. LIMES-EU optimizes the electricity sectors of 29 European countries until 2050 while taking short-term variability of demand and renewable energies (wind and solar), the grid connection between countries and power plant characteristics into account (see Osorio et al. 2018 for more details). For estimating the production cost function of the clean firm, I consider 15 emission-free technologies (renewable energies and nuclear) and for the dirty firm, 21 technologies (mainly coal and gas) of LIMES-EU.[15] The emission intensity of the dirty

[15]To derive the cost functions, I perform several LIMES simulations with increasing clean production shares. Per the simulation, I obtain the total clean and dirty costs (sum of 15 and 21 technologies,

sector ϕ_d is assumed to be 650 g/kWh as a compromise between coal (about 900 g/kWh) and gas (about 350 g/kWh) technologies.

I consider 100 periods in three-year steps starting in the year 2020. The transition (endogenously) ends at the latest after 28 periods, and the remaining periods are included to mimic the infinite horizon. Notice that only the SCC are directly affected by the time horizon, and including more than 100 periods hardly affects the SCC. I set demand parameter γ_t to 4.4 PWh in the first model year (2020) and ϵ (time independent) to 15 (kWh reduction of demand per EUR). This leads to a realized demand between 3.3 PWh and 3.6 PWh in 2020, depending on the scenario, while actual electricity generation in the EU was 3.47 PWh in 2017 according to Eurostat. I further assume that γ_t increases by 0.5% per year. These parameters imply an average (long-run) price elasticity of demand[16] of about -0.2 to -0.3, which is in line with recent empirical estimates (Deryugina et al. 2020). Furthermore, these assumptions lead to BAU emissions φ_t of 1.5189 Gt in 2020, and the slope of the marginal abatement cost curve ψ is 0.0481.

Uncertainty enters the model by an additive shock on energy demand parameter γ_t and thus, on BAU emissions (see equation (2.16)). Specifically, I assume an AR(1) process, $\gamma_t = \eta\gamma_{t-1} + \theta_t$, with $\eta = 0.6$ and $\theta_t \in \{-\vartheta\gamma_0, \vartheta\gamma_0\}$. That is, there is either a positive or negative shock θ_t on demand parameter γ_t equal to $\vartheta\gamma_0$ and the expected value is $E_t[\theta_s] = 0 \; \forall \; s > t$. The properties of the shocks are common knowledge and they are realized at the beginning of a period before the regulator and firms make their decisions. I compute two uncertainty scenarios: one with a relatively weak 8% shock every three years, $\vartheta = 0.08$, and one with a relatively strong 12% shock, $\vartheta = 0.12$. The first shock emerges in $t = 2023$, and to keep the model computable, the last shock is in $t = 2044$.

According to data used by Dietz and Venmans (2019), the growth rate of the social cost of carbon g is about 1.5% to 3.3%, van der Ploeg (2018) applies 2% and Nordhaus (2017) 3%. I assume 2.5% for the growth rate of the social cost of carbon and 5% for the

respectively) as well as the total clean and dirty production levels. The cost function parameters are derived from fitting these costs and production data.

[16]The price elasticity of demand in the model is $\varepsilon_t = \frac{dD_t}{dw_t}\frac{w_t}{D_t} = -\frac{\epsilon w_t}{\gamma_t - \epsilon w_t}$. I take the average over the model periods because the elasticity increases with price w_t and decreases with parameter γ_t.

discount rate. The initial marginal damage is assumed to be $\delta a_{2020} = 0.7174$, such that according to equation (2.14), the social cost of carbon is about 30 EUR/t in 2020. This is roughly in line with SCC estimates from the literature (Nordhaus 2017; van der Ploeg 2018).

My analysis includes the six policies analyzed above: a tax, which serves as optimal benchmark, a CAT program with and without an intertemporal trading ratio (ITR), a subsidy, and a CAT program combined with a subsidy in the commitment (indicated by a C) and time-consistent (indicated by a T) case.[17] I first show the results for the model under perfect foresight, and then analyze the effects of abatement cost uncertainty.

3.2. Transition under certainty

The first-best transition is achieved with the tax and with the CAT program plus the ITR, because in both cases the policy-induced carbon price is always equal to the SCC. All other instruments are imperfect and imply a welfare loss compared to the first best as depicted in Table 3.2 (a). The highest welfare loss is induced if only a subsidy is implemented (250.91 billion EUR), and the second-worst alternative is the CAT-only policy with free intertemporal trading (no ITR). However, CAT-only with a welfare loss of 24.64 billion EUR is much more efficient than the subsidy. Thus, the inefficient permit price growth rate of the CAT program weighs less than the energy-price-distorting effect of the subsidy. Furthermore, combining the CAT program with a subsidy reduces the welfare loss to 9.81 and 13.59 billion EUR under the commitment and time-consistent subsidies, respectively.

The welfare effects can be explained by the different abatement and carbon price paths engendered by the policy instruments as shown in Figure 3.1. In the two policy mix scenarios, the total carbon price (equation (2.50)) is shown in Figure 3.1 (b). Further

[17]The model is solved using the GAMS software as a non-linear program (NLP) while I maximize welfare subject to the first-order conditions of the firms. In the CAT scenarios, the cap is always exogenous and adjusted between model runs until the welfare optimum is found. However, the scenario with CAT combined with the time-consistent subsidy is solved as extended mathematical programming (EMP) in which only the firm profits are maximized. In this case, the subsidy is also exogenous to the model and set according to equation (2.49). The model is iterated until the optimum is found. The code is available upon request.

note that after the transition is completed, carbon prices in all cases are just as high as to guarantee zero emissions. This implies that prices rise linearly after the transition along with the increasing marginal abatement costs due to the growing energy demand.

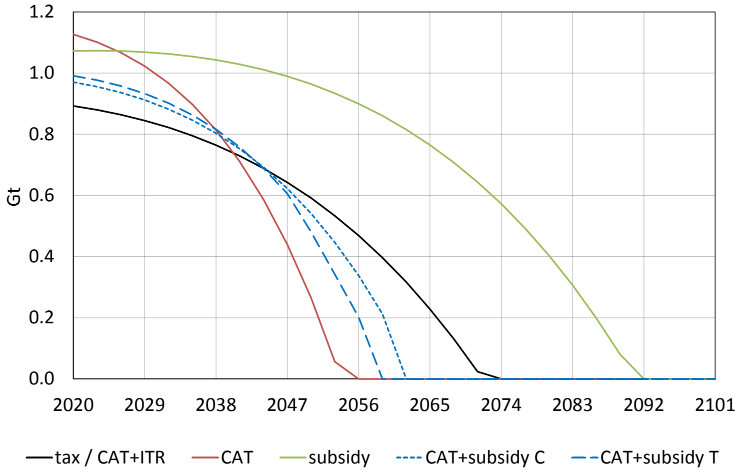
In the CAT-only scenario, the first best cannot be obtained because the permit price rises at the rate of interest, $r = 5\%$, but it should rise with the SCC, $g = 2.5\%$. This implies that the abatement path in CAT-only case is too steep (see Figure 3.1): The emissions are too high early on due to the low permit price, but decline quickly such that the entire transition takes only 36 years compared to 54 years under the first best. Adding subsidies to the CAT program adjusts the growth rate of the total carbon price. Under commitment, the total carbon price rises at 3.2%, and thus, it is significantly closer to the optimum of 2.5%. The Markov subsidy implies a somewhat higher growth rate of 3.5% because time consistency is taken into account. As a result, the abatement and carbon price paths of the policy mixes (CAT+subsidy C and CAT+subsidy T) lie in between the first-best case and the CAT-only case.

The reduction of the growth rate of the total carbon prices in the two policy mix scenarios can be explained by the falling subsidies over time as depicted by the dashed and dotted line in Figure 3.1 (c). The two subsidy paths start at a comparable level, but the commitment subsidy (CAT+subsidy C) declines faster and turns negative in the mid-2040s, implying that the growth rate of the total carbon price is lower than under time consistency (3.2% vs. 3.5%). The negative subsidy explains also why there is no permit price reduction caused by the waterbed effect in the commitment solution. In contrast, the waterbed effect reduces the CAT price by 2.66 EUR/t in 2020 under time consistency.[18]

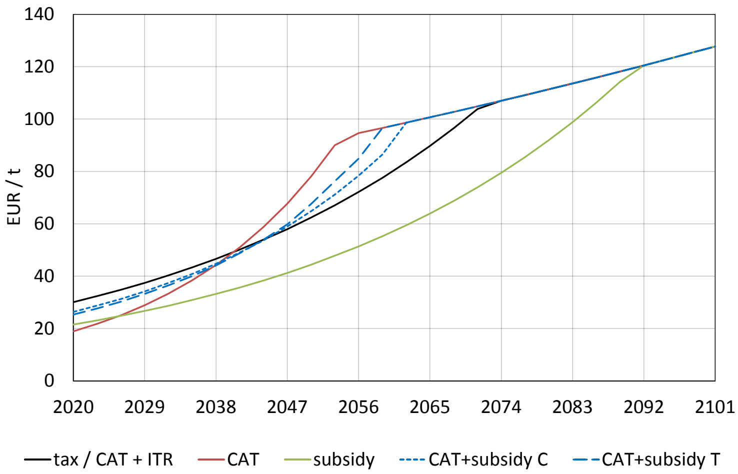
[18] To calculate the permit price reduction of the waterbed effect, I use the cap of the CAT+subsidy T scenario but set the subsidy to zero. The difference in the permit price between this and the CAT+subsidy T scenario is attributed to the waterbed effect of the subsidy.

Figure 3.1: Emission, carbon price and subsidy paths under certainty

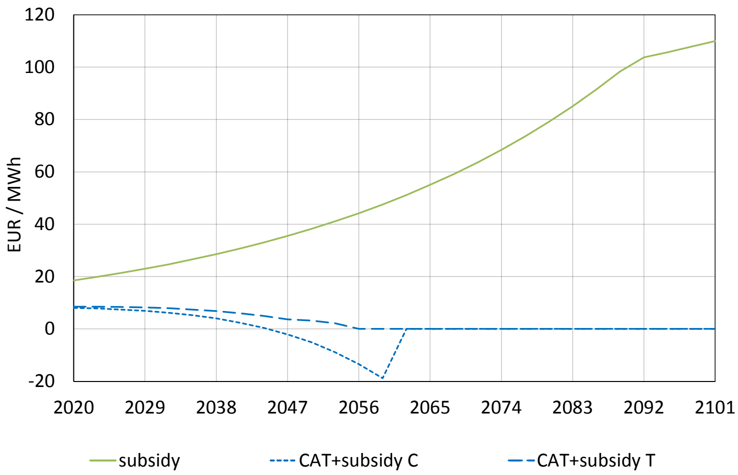
(a) Emissions



(b) Carbon prices



(c) Subsidies



Note: The lines for CAT+subsidy C and CAT+subsidy T in part (b) reflect the total carbon prices (see equation (2.50)).

Cumulative emissions also exhibit large differences between the policy instruments (see Table 3.2 (b)). The largest amount of cumulative emissions is caused by the subsidy-only policy (57.36 Gt) because the subsidy distorts energy prices and leads to overconsumption. In the CAT-only scenario, cumulative emissions (i.e., the cap) are much lower (27.15 Gt), and essentially also significantly lower than under the first-best solution (31.84 Gt). The regulator chooses a lower cap to reduce the inefficient high damage during the transition due to the too low permit price (cf. section 2.3.2). By adding subsidies to the CAT program, the emissions and thus, damage are shifted to the future such that, compared to CAT-only, the cap can be increased to 29.73 Gt and 28.55 Gt under the commitment and time-consistent subsidies, respectively. That is, subsidies do not directly affect cumulative emissions, but they influence the regulator's choice of the cap and therefore, indirectly increase the cumulative emissions.

Table 3.2: Welfare loss, cumulative emissions and transition length

	Tax	CAT	CAT+ITR	Subsidy	CAT+sub C	CAT+sub T
(a) Welfare loss (billion EUR and relative to optimum)						
No shock	-	24.64 (0.21%)	-	250.91 (2.14%)	9.81 (0.08%)	13.59 (0.12%)
8% shock	-	34.65 (0.29%)	11.27 (0.10%)	249.62 (2.11%)	19.92 (0.17%)	23.99 (0.20%)
12% shock	-	46.83 (0.39%)	24.98 (0.21%)	247.89 (2.07%)	31.84 (0.27%)	35.34 (0.30%)
(b) Cumulative emissions (Gt)						
No shock	31.84	27.15	31.84	57.36	29.73	28.55
8% shock	32.53	26.88	32.20	57.95	29.76	29.67
12% shock	33.44	27.57	32.68	58.73	30.30	30.22
(c) Transition length (years)						
No shock	54.00	36.00	54.00	72.00	42.00	39.00
8% shock	52.44	35.20	53.07	71.95	41.57	42.31
12% shock	51.65	35.79	53.12	71.31	41.80	43.36

Note: Welfare is always maximized with the tax because of the assumed constant marginal damage. In the other cases, the welfare loss is the (relative) difference to the tax scenario.

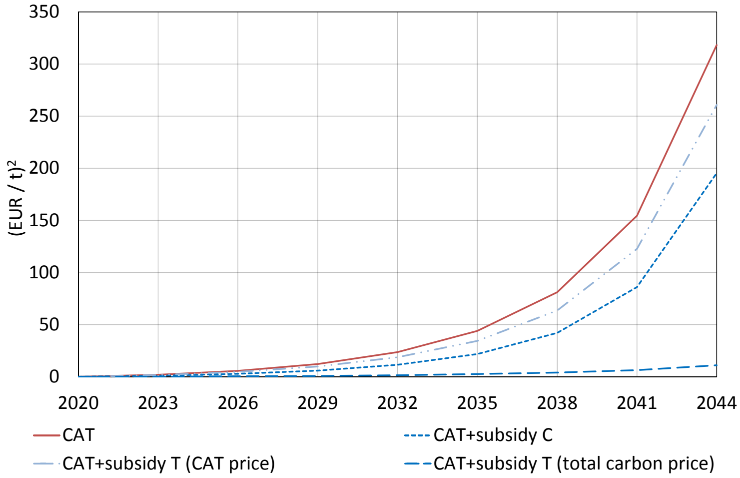
3.3. Transition under uncertainty

Uncertainty generally elevates the welfare advantage of price-based policies because of the assumed constant marginal damage. This is why the tax is always optimal and the welfare disadvantage of the subsidy does not increase with uncertainty, whereas scenarios with CAT exhibit an increasing welfare loss with higher uncertainty (see Table 3.2 (a)). Moreover, the hybrid policies consisting of CAT and a subsidy also do not perform better under uncertainty than the CAT-only scenario (the welfare losses in all CAT scenarios increase at about the same magnitude when a shock is added). Under uncertainty, subsidies lead to welfare advantages, but the subsidies also have adverse effects if added to the CAT program.

An advantage is that subsidies reduce the total carbon price volatility as shown in Figure 3.2. Therefore, the deviations from the optimal carbon price due to shocks are reduced which is welfare enhancing. This even holds for the commitment subsidy (CAT+subsidy C) although the entire subsidy path is determined in the first period, and the subsidy is not state-contingent. That is, the subsidy is a fixed term which would not affect the carbon price volatility if the transition length were given. However, the subsidy shifts emissions to the future and thus, prolongs the transition or banking phase by about six years compared to CAT-only (see Table 3.2 (c)). A longer banking phase reduces the impact of shocks on the permit price because the shocks spread to more periods (Fell et al. 2012), and therefore, the subsidy indirectly reduces the permit price volatility. However, the downside of commitment is lower flexibility because the regulator does not react to new information. This causes a welfare loss which offsets the welfare advantage due to the more stable carbon prices. The initially set subsidy path is a compromise between a potentially shorter or longer transition phase: Compared to the certainty scenario (see Figure 3.3), this explains why the subsidy is less negative before 2060 because non-zero subsidies induce a welfare loss if the emissions are already zero. It also explains why the subsidy is more negative after 2060, because from the perspective of the first period there is a positive probability that the transition is ongoing.

In contrast, under time consistency the regulator has the opportunity to react to

Figure 3.2: Carbon price variance

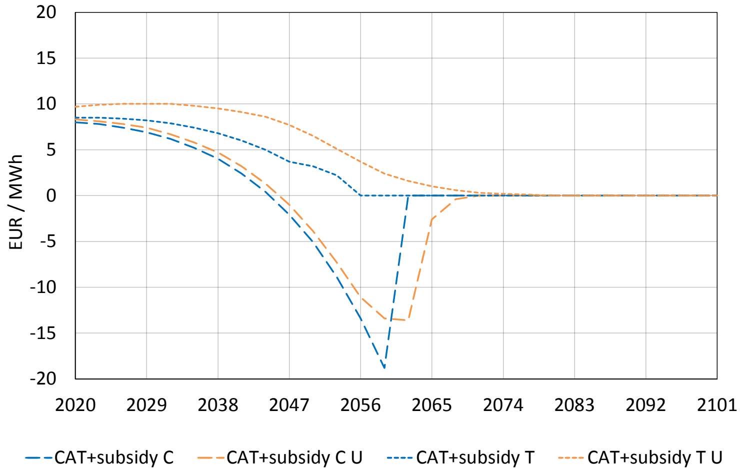


Note: The figure depicts the variance from the perspective of the first period until the period in which the last shock emerges for the 12% shock case. There is no difference between the variance in the CAT price and the total carbon price in the scenario CAT+subsidy C because the subsidy is fixed.

shocks. Therefore, she continues the subsidy program after she learns that the transition is ongoing, and she sets the subsidies to zero after she learns that the transition ends in the next period. The subsidy is increased when new information implies that the transition takes more time and vice versa, because a longer (shorter) transition allows to shift more (less) emissions to the future. This implies that the correlation between the subsidy and the permit price is negative: A positive abatement cost shock increases (expected) emissions in the current and future periods such that more permits are needed, and the expected transition length is shorter than expected before the shock occurred. Therefore, if the abatement costs and thus, the permit prices are high, then the subsidies are low, because the transition phase is expected to be short and vice versa. Put differently, updating subsidies counteracts the permit price shocks such that the total carbon price volatility is much lower (see Figure 3.2) which stabilizes the abatement path. Moreover, the regulator prolongs the expected transition length if she can update the subsidy levels. On average, she raises the subsidy more after the abatement costs decline than she reduces it after they increase. This is welfare enhancing because under CAT the transition length is too short, and thus, on average, higher subsidies bring it closer to the optimum (see Table 3.2 (c)).

However, although updating the subsidies significantly reduces the total carbon price

Figure 3.3: Subsidy paths



Note: U in scenarios CAT+subsidy C U and CAT+subsidy T U refers to uncertainty (12% shock). The path of scenario CAT+subsidy T U reflects the expected subsidy.

volatility and prolongs the transition length, it is not welfare enhancing under uncertainty compared to the commitment solution. The reason is that subsidy updating significantly increases the (expected) subsidy level as shown in Figure 3.3. This reduces the permit price through the waterbed effect: The price-reducing effect of the subsidy is 3.36 EUR/t in 2020 under the 12% shock vs. 2.66 EUR/t under certainty. Thus, the subsidy takes on a stronger role under uncertainty relative to the permit price, which is inefficient because of the energy-price-distorting effect. Therefore, the opportunity to react to shocks may increase or decrease welfare compared to the commitment solution, while in this simulation the effects roughly cancel out, indicated by similarly increasing welfare losses due to uncertainty in Table 3.2 (a). For additional simulation results on the ITR and cumulative emissions under uncertainty, see Appendix B.

4. Conclusion

In this paper, I analyze the interaction of subsidies for clean energy and a CAT program with free intertemporal trading. I show that optimally set subsidies enhance welfare when added to such a CAT program, even if there is only the carbon externality. The reason is that the permit price of the CAT program never equals the SCC because the permit price rises at a higher rate than the SCC. Thus, the permit price is initially too low, implying excessive carbon damage until the transition to carbon neutrality is completed. A subsidy partly corrects this flaw by shifting emissions to the future so that the costs of delayed action are reduced.

The optimal subsidy path strictly declines over time when a CAT program is in place as well. The waterbed effect of the optimal subsidy shifts emissions to the future without directly affecting the permit price. However, this subsidy path is not time-consistent, as the regulator has incentives to deviate from subsidies scheduled in the past. I also derive the time-consistent subsidy path, which partly crowds out the permit price. However, welfare is still significantly higher compared to CAT-only.

In the numerical simulation of the EU electricity sector, I show that subsidies within a CAT program have an additional welfare advantage but also disadvantages if abatement cost uncertainty is considered. On one hand, subsidies reduce the permit price volatility because they prolong the transition or banking phase, implying that shocks spread to more periods. If the regulator additionally reacts to new information in the time-consistent solution, she counteracts the permit price development with higher or lower subsidies such that the total carbon price is relatively stable. On the other hand, I find that the price-reducing effect of the subsidy is exacerbated under uncertainty if the regulator reacts to shocks. In turn, if the regulator does not react to shocks (the commitment solution), uncertainty causes welfare losses because new information is not used to adapt the subsidy. In the current analysis, I do not consider the case of a state-contingent policy with predefined policy rules according to which the subsidy is updated under uncertainty. Such a policy could combine the welfare advantage of commitment and the flexibility to adapt to new information.

However, in the setting of this paper, a combination of clean subsidies and CAT is only a second-best solution. The problem of permit price uncertainty can also be addressed by hybrid instruments (price collars) and the too-high growth rate of the (expected) permit price can be corrected by intertemporal trading ratios. An alternative is to ban banking (and borrowing) and to implement a price responsive permit supply (Traeger et al. 2020). This approach allows for the implementation of an efficient growth rate of the permit price (in expectation) and to reduce the price volatility as supply adjustments counteract shocks.

Nonetheless, the advantage of the CAT and subsidy instrument mix is its high political feasibility, as indicated by its widespread implementation. It can be argued that the prospects of (near-term) implementation of policies should be considered in their evaluation as more feasible policies also reduce the cost of delay (Goulder 2020). In this sense, adding subsidies to a CAT program can be a rational welfare-maximizing decision as long as more efficient alternatives cannot be implemented. This is all the more true if the (perceived) low credibility of the cap of CAT programs, as in the EU's Emission Trading System, and potentially myopic market participants are considered. Both can be reasons for applying inefficiently high discount rates to permit banking (Fuss et al. 2018; Salant 2016), and thus, may depress prices in the early phase of a CAT program, which, in turn, can be counteracted by subsidies. A further argument for a policy mix is technological change. If knowledge stocks or innovation externalities are included in model analyses, the literature finds highly positive research subsidies early on that eventually drop to zero (Acemoglu et al. 2016; Rezai and Van Der Ploeg 2017). Although the subsidy in this paper is not paid for research but rather to shift emissions to the future, the time path is similar to the optimal research subsidies (apart from the period of negative subsidies in the commitment solution). Thus, both arguments - dynamically inefficient carbon price paths and innovation - may justify the addition of clean production subsidies to CAT programs as second-best alternatives.

Having said that, implementing subsidies in a welfare-enhancing way is not without problems. For one, subsidies should reflect damage avoided, in addition to other potential market failures, which hardly seems to be the case in practice. Abrell et al. (2019) show that this requires technology-differentiated subsidies in electricity markets because technologies (e.g., wind and photovoltaics) have different production profiles. However, poorly set subsidies contain the risk of an excessive reduction of the permit price, and therefore, subsidies may undermine the relevance and credibility of CAT programs (Fankhauser et al. 2010). Future research could analyze which specific support schemes (e.g., renewable quotas or feed-in tariffs) are best suited to complement a CAT program and avoid adverse effects on the permit price. In addition, the analysis of this paper builds on several simplifying assumptions as I ignore borrowing constraints in the permit market or other market failures, such as myopia or market power, that may play an important role in real markets, all of which are interesting avenues for future research.

Acknowledgments

The author would like to thank Ottmar Edenhofer, Marina Friedrich, Christian Gambardella, Samuel Okullo, Sebastian Osorio, Michael Pahle and Grischa Perino for valuable comments. This work was funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 730403 (INNOPATHS).

Appendix A. Derivations

Appendix A.1. Social planner problem

Appendix A.1.1. Energy market formulation

The Bellman equation of the planner's problem (2.12) is

$$V_t(\Phi_t) = \max_{D_t,N_t,x_{it}} \; U_t(D_t,N_t) - \lambda_t\left(N_t - a_0(1+g)^t(1-\delta\Phi_t) + \sum_i C_{it}\right) - \mu_t\left(D_t - \sum_i x_{it}\right) - \frac{1}{1+r}V_{t+1}(\Phi_{t+1}), \tag{A.1}$$

subject to the dynamic emission stock (2.9). The first-order conditions are

$$\frac{\partial U_t}{\partial D_t} - \mu_t = 0 \quad (D_t \geq 0), \tag{A.2}$$

$$\frac{\partial U_t}{\partial N_t} - \lambda_t = 0 \quad (N_t \geq 0), \tag{A.3}$$

$$\mu_t - \lambda_t\frac{\partial C_{it}}{\partial x_{it}} - \left(\lambda_t\delta a_0(1+g)^t + \frac{1}{1+r}\frac{\partial V_{t+1}}{\partial \Phi_{t+1}}\right)\phi_i = 0 \quad (x_{it} \geq 0), \tag{A.4}$$

$$\frac{\partial V_t}{\partial \Phi_t} - \lambda_t\delta a_0(1+g)^t - \frac{1}{1+r}\frac{\partial V_{t+1}}{\partial \Phi_{t+1}} = 0 \quad (\Phi_t). \tag{A.5}$$

The first two conditions state that the marginal utility of the respective good must be equal to the shadow value of the energy market clearing and budget constraint, respectively. The third condition reflects that energy is generated until marginal utility (first term) is equal to the marginal production costs (second term) and the marginal environmental damage of production (third term). The marginal environmental damage of emissions (SCC) is given by condition A.5. By inserting $\frac{\partial V_{t+1}}{\partial \Phi_{t+1}}, ..., \frac{\partial V_{\infty}}{\partial \Phi_{\infty}}$ in condition A.5, the SCC can be written as

$$\sigma_t = \frac{\partial V_t}{\partial \Phi_t} = \delta a_0 \sum_{s=t}^{\infty} \frac{(1+g)^s \lambda_s}{(1+r)^{s-t}}. \tag{A.6}$$

If consumer utility is given by equation (2.8), the first two conditions (A.2) and (A.3) become

$$D_t = \gamma_t - \mu_t \epsilon, \tag{A.7}$$

$$\frac{\partial U_t}{\partial N_t} = \lambda_t = 1, \tag{A.8}$$

respectively. Therefore, the SCC become (2.14) and by using (2.14) and (A.2) in (A.4) one obtains (2.13). In addition, quadratic energy generation costs (2.1) imply that equation (2.13) can be written as:

$$x_{it} = \frac{\mu_t - \alpha_{it} - \sigma_t \phi_i}{\beta_i}. \tag{A.9}$$

Inserting (A.9) for both firms and (A.7) into the equilibrium condition (2.7) and solving for μ_t yields

$$\mu_t = \frac{\beta_c \beta_d \gamma_t + \beta_c (\alpha_{d,t} + \sigma_t \phi_d) + \beta_d \alpha_{c,t}}{\beta_c + \beta_d + \beta_c \beta_d \epsilon}. \tag{A.10}$$

Recall that μ_t is the shadow value of the energy equilibrium constraint and thus, can be interpreted as energy price. Inserting μ_t back in the production function (A.9) gives the optimal production levels if both firms are active in equilibrium:

$$x_{d,t} = \frac{\beta_c \gamma_t + \alpha_{c,t} - (\alpha_{d,t} + \sigma_t \phi_d)(1 + \beta_c \epsilon)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon}, \tag{A.11}$$

$$x_{c,t} = \frac{\beta_d \gamma_t + \alpha_{d,t} - \alpha_{c,t}(1 + \beta_d \epsilon) + \sigma_t \phi_d}{\beta_c + \beta_d + \beta_c \beta_d \epsilon}. \tag{A.12}$$

In the social planner solution, the environmental damage is optimally internalized reflected by the term $\sigma_t \phi_d$ in both expressions. The optimal energy consumption D_t is then simply the sum of (A.11) and (A.12), and the optimal consumption of the numeraire N_t can be derived from the budget constraint (2.5).

Appendix A.1.2. Equivalence to quadratic abatement cost problem

In this section, I derive the implied abatement cost function of the energy market problem solved in the previous section. The two parameters of the abatement cost function (2.15) are BAU emissions φ_t and the slope of the marginal abatement costs ψ. I obtain the BAU emissions by multiplying expression (A.11) with ϕ_d (because emissions are $q_t = \phi_d x_{d,t}$) and setting $\sigma_t = 0$, which yields (2.16). To derive the slope of the marginal abatement costs ψ, I set marginal abatement costs equal to the SCC, $\psi\left(\varphi_t - q_t\right) = \sigma_t$, which is required for optimality. Inserting BAU emissions (2.16) and $q_t = \phi_d x_{d,t}$ in $\psi\left(\varphi_t - q_t\right) = \sigma_t$ and solving for ψ yields the slope of the marginal abatement cost curve as shown in (2.17). It can be shown that the problem

$$\min_{q_t} W_0 = \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} \left(\frac{\psi}{2} \left(\varphi_t - q_t\right)^2 + \delta a_0 \left(1+g\right)^t \Phi_t \right) \tag{A.13}$$

subject to the dynamic emission stock (2.9) is equivalent to the social planner problem under the energy market formulation if abatement cost parameters are defined as (2.16) and (2.17).

Appendix A.2. Carbon tax

Appendix A.2.1. Firms' problem

The firms' problem is (2.3) with periodic profits given by (2.18). Note that in this case (as well as under the subsidy-only policy) the firm problem is static. Taking the first-order condition with respect to x_{it} gives the production function $x_{it} = \left(w_t - \alpha_{it} - \tau_t \phi_i\right) \beta_i^{-1}$. Using this function and energy demand[19] $D_t = \gamma_t - w_t \epsilon$ in the energy market equilibrium

[19]The energy demand function is the same in all scenarios. It is obtained from maximizing utility (2.8) subject to the budget constraint (2.5).

(2.7) yields the energy price

$$w_t = \frac{\beta_c \beta_d \gamma_t + \beta_c \left(\alpha_{d,t} + \tau_t \phi_d\right) + \beta_d \alpha_{c,t}}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} \,\forall\, t \leq \hat{t}_{tax}, \tag{A.14}$$

which is equal to the optimal energy price (A.10) if $\tau_t = \sigma_t$ holds. By inserting (A.14) in the production function, $x_{it} = (w_t - \alpha_{it} - \tau_t \phi_i)\, \beta_i^{-1}$, one obtains (2.19) and (2.20) for $t \leq \hat{t}_{tax}$. After the transition is completed, $t > \hat{t}_{tax}$, only the clean generates energy. Therefore, the energy market equilibrium becomes $x_{c,t} = (w_t - \alpha_{c,t})\, \beta_c^{-1} = \gamma_t - w_t \epsilon$. Solving for w_t and inserting back into the production function gives $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c \epsilon}$ for $t > \hat{t}_{tax}$.

Appendix A.2.2. Regulator's problem

Substituting the firm profits (2.18) and the carbon damage (2.10) into the budget constraint (2.5), allows to write the consumption of the numeraire as follows:

$$N_t = a_t - \sum_i \left(\alpha_{it} x_{it} + \frac{\beta_i}{2} x_{it}^2 \right) - \delta a_t \Phi_t. \tag{A.15}$$

By using (A.15) and the energy market equilibrium condition (2.7) in the utility function (2.8), one obtains utility (2.22). Next, utility (2.22) can be inserted in the regulator's problem (2.21):

$$\max_{\tau_0, \ldots \tau_t, \ldots \tau_\infty} W_0 = \sum\nolimits_{t=0}^{\infty} \frac{1}{(1+r)^t} \left(\frac{\gamma_t \sum_i x_{it} - 0.5 \left(\sum_i x_{it}\right)^2}{\epsilon} - \sum_i C_{it} - \delta a_t \Phi_t + a_t \right). \tag{A.16}$$

Considering the emission stock constraint (2.9) and taking the first-order conditions with respect to τ_t yields the optimal tax (2.23).

To derive the optimal transition length, I consider infinitely small time steps $\Delta t \approx 0$ (continuous time) and constant BAU emissions φ. Then, the optimal transition length can be derived from the fact that at the end of the transition emissions are zero and it holds:

$$\begin{aligned} q_{\hat{t}_{tax}} &= \varphi - \tfrac{\tau_{\hat{t}_{tax}}}{\psi} &= 0 \\ \Rightarrow \tau_{\hat{t}_{tax}} &= \tau_0 e^{g \hat{t}_{tax}} &= \varphi \psi, \end{aligned} \tag{A.17}$$

which can be rearranged to (2.25).

Appendix A.3. Cap-and-trade

Appendix A.3.1. Firms' problem

The firms solve problem (2.3) via production decisions x_{it} and permit purchases y_{it} where the periodic profit π_{it} is given by (2.28). The Bellman equation is

$$V_{it}(b_{it}) = \max_{x_{it},y_{it}} \left(w_t x_{it} - \alpha_{it} x_{it} - \frac{\beta_i}{2} x_{it}^2 - p_t y_{it} \right) + \frac{1}{1+r} V_{it+1}(b_{it+1}), \tag{A.18}$$

subject to the banking constraint (2.27). The first-order conditions are:

$$w_t - \alpha_{it} - \beta_i x_{it} - \frac{\phi_d}{1+r} \frac{\partial V_{it+1}}{\partial b_{it+1}} = 0 \quad (x_{it} \geq 0), \tag{A.19}$$

$$p_t - \frac{1}{1+r} \frac{\partial V_{it+1}}{\partial b_{it+1}} = 0 \quad (y_{it}), \tag{A.20}$$

$$\frac{\partial V_{it}}{\partial b_{it}} - \frac{1}{1+r} \frac{\partial V_{it+1}}{\partial b_{it+1}} = 0 \quad (b_{it}). \tag{A.21}$$

Using (A.19) and (A.20) yields the production function $x_{it} = (w_t - \alpha_{it} - p_t \phi_i) \beta_i^{-1}$. Inserting the production functions of both firms and energy demand $D_t = \gamma_t - w_t \epsilon$ into the energy market equilibrium (2.7) gives the energy price $w_t(p_t)$ as in (A.14) but with p_t instead of τ_t. Substituting the energy price into the production functions, $x_{it} = (w_t - \alpha_{it} - p_t \phi_i) \beta_i^{-1}$, yields production depending on the permit price:

$$x_{d,t} = \frac{\beta_c \gamma_t + \alpha_{c,t} - (\alpha_{d,t} + p_t \phi_d)(1 + \beta_c \epsilon)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} \,\forall\, t \leq \hat{t}_{cat}, \tag{A.22}$$

$$x_{c,t} = \frac{\beta_d \gamma_t + \alpha_{d,t} + p_t \phi_d - \alpha_{c,t}(1 + \beta_d \epsilon)}{\beta_c + \beta_d + \beta_c \beta_d \epsilon} \,\forall\, t \leq \hat{t}_{cat}. \tag{A.23}$$

These expressions can be inserted in the permit equilibrium, $z_t = \sum_i y_{it} = \phi_d x_{d,t} + B_t - B_{t-1}$, which results in the permit price (2.29), for which I have used (2.16) and (2.17).

Combining (A.20) and (A.21) yields the intertemporal price dynamics as shown in (2.30) and thus, the permit price rises at rate r. Inserting the permit price (2.29) in the expression for the intertemporal price dynamics (2.30) gives the permit bank

$$B_t = \frac{(1+r)(z_t + B_{t-1} - \varphi_t) + \varphi_{t+1} + B_{t+1}}{1+1+r} \; \forall \, t \leq \hat{t}_{cat}. \tag{A.24}$$

Once the budget is used up the bank is empty, that is, $B_t = 0$ for all $t \geq \hat{t}_{cat}$. Inserting $B_{\hat{t}_{cat}} = 0$ for B_{t+1} in (A.24), gives the bank one period before the transition ends, $B_{\hat{t}_{cat}-1}$, which only depends on parameters. Inserting $B_{\hat{t}_{cat}-1}$ in $B_{\hat{t}_{cat}-2}$ and so forth until the first period is reached results in the whole banking path as shown in (2.31).

I use the permit price (2.29) in (A.22) to rewrite dirty energy generation:

$$x_{d,t} = (z_t - B_t + B_{t-1})\, \phi_d^{-1}. \tag{A.25}$$

Substituting the bank (2.31) shows that dirty production is (2.32) for $t \leq \hat{t}_{cat}$. For the clean production, I use permit price (2.29) in (A.23) and insert the bank (2.31), which results in (2.33) for $t \leq \hat{t}_{cat}$.

After the transition is completed, $t > \hat{t}_{cat}$, dirty generation is by definition zero and for the clean firm it holds $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c \epsilon}$, which I obtain by applying the same steps as under the tax regime, see Appendix A.2.1.

Appendix A.3.2. Regulator's problem

The regulator's problem is

$$\max_{Z_{cat}} W_0 = \sum\nolimits_{t=0}^{\infty} \frac{1}{(1+r)^t} \left(\frac{\gamma_t \sum_i x_{it} - 0.5 \left(\sum_i x_{it}\right)^2}{\epsilon} - \sum_i C_{it} - \delta a_t \Phi_t + a_t \right), \tag{A.26}$$

subject to emission stock (2.9) and with reaction functions (2.32) and (2.33). To solve the problem I take the first-order condition of (A.26) with respect to Z_{cat}, which results in (2.35). In doing so, I make use of the fact that for $t > \hat{t}_{cat}$ it holds $x_{d,t} = 0$ and $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c \epsilon}$ so that Z_{cat} does not affect the outcome. Moreover, the damage is $\delta a_t \Phi_t = \delta a_t Z_{cat}$ for $t > \hat{t}_{cat}$.

As in the tax case, I derive the optimal transition length while assuming constant

BAU emissions φ and small time steps $\Delta t \approx 0$. The optimal cap (2.35) then becomes

$$Z_{cat} = \int_0^{\hat{t}_{cat}} \left(\varphi - \frac{\sigma_0 e^{gt}}{\psi}\right) dt = \hat{t}_{cat}\varphi - \frac{\sigma_0 \left(e^{g\hat{t}_{cat}} - 1\right)}{g\psi}. \tag{A.27}$$

To derive the optimal transition length, an additional equation is required because (A.27) contains two unknowns. I obtain an additional equation from the firm problem (see Schennach 2000 for a more detailed derivation): The permit price and emissions at the end of the transition are $p_{\hat{t}_{cat}} = \psi\left(\varphi - q_0\right) e^{r\hat{t}_{cat}}$ and $q_{\hat{t}_{cat}} = \varphi - \frac{p_{\hat{t}_{cat}}}{\psi} = 0$. Combining the two yields

$$\varphi - \left(\varphi - q_0\right) e^{r\hat{t}_{cat}} = 0. \tag{A.28}$$

Furthermore, cumulative emissions must be equal to the cap

$$\begin{aligned} \hat{t}_{cat}\varphi - \textstyle\int_0^{\hat{t}_{cat}} e^{rt}\left(\varphi - q_0\right) dt &= Z_{cat} \\ \Rightarrow \hat{t}_{cat}\varphi - \left(\varphi - q_0\right) \frac{e^{r\hat{t}_{cat}} - 1}{r} &= Z_{cat}, \end{aligned} \tag{A.29}$$

for which I have used periodic emissions $q_t = \varphi - \frac{p_t}{\psi}$ with $p_t = \psi\left(\varphi - q_0\right) e^{rt}$. Substituting (A.28) in (A.29) gives

$$\tilde{Z}_{cat} = \hat{t}_{cat}\varphi + \varphi \frac{e^{-r\hat{t}_{cat}} - 1}{r}, \tag{A.30}$$

where the tilde indicates that $\tilde{Z}_{cat}$ is a result of the firm problem, whereas (A.27) reflects the optimal cap as a result of the social planner problem. Now I have two functions that relate the transition length to the cap. Combining (A.30) and (A.27) yields (2.37).

Appendix A.3.3. Proof of proposition 2

The difference between the cap (2.35) and the cumulative emissions of the tax regime (2.24) is determined by the transition lengths ($\hat{t}_{cat}$, $\hat{t}_{tax}$). A shorter ($\hat{t}_{cat} < \hat{t}_{tax}$) and a longer ($\hat{t}_{cat} > \hat{t}_{tax}$) transition both imply fewer cumulative emissions in the CAT program compared to the tax, $Z_{cat} < Q_{tax}$. To see why, consider that periodic emissions under the tax regime are $q_t = \varphi_t - \frac{\sigma_t}{\psi}$ and thus, it holds $q_{\hat{t}_{tax}} = \varphi_{\hat{t}_{tax}} - \frac{\sigma_{\hat{t}_{tax}}}{\psi} = 0$ at the end of the transition. Because it holds $\varphi_t - \frac{\sigma_t}{\psi} \leq 0$ for $t \geq \hat{t}_{tax}$ (otherwise emissions were positive, which is ruled out by definition of $\hat{t}_{tax}$), the sum over more time steps in (2.35) compared to (2.24) implies $Z_{cat} < Q_{tax}$. Vice versa, if $\hat{t}_{cat} < \hat{t}_{tax}$ the sum over less time steps in (2.35) compared to (2.24) implies $Z_{cat} < Q_{tax}$ as well because for $t < \hat{t}_{tax}$ it holds $\varphi_t - \frac{\sigma_t}{\psi} > 0$.

Therefore, it suffices to show that the transition lengths are not equal. If transition lengths (2.25) and (2.37) were equal, it would hold

$$\hat{t}_{cat} = \hat{t}_{tax} = \frac{1}{g} ln \left(\frac{\varphi\psi}{\sigma_0} \right) = \frac{1}{r} ln \left(\frac{\varphi\psi}{\varphi\psi - \frac{r}{g}\left(\varphi\psi - \sigma_0\right)} \right), \tag{A.31}$$

for which I have inserted $\sigma_{\hat{t}_{cat}} = \varphi\psi$ in (2.37). However, (A.31) only holds if $g = r$, which is ruled out by assumption and therefore $\hat{t}_{cat} \neq \hat{t}_{tax}$. It follows that the cap is lower than the cumulative emissions under the tax regime $Z_{cat} < Q_{tax}$.

Appendix A.3.4. Proof of corollary 1

Consider expression (A.30) to see that the relationship between the cap and the transition length is positive:

$$\frac{d\tilde{Z}_{cat}}{d\hat{t}_{cat}} = \varphi \left(1 - e^{-r\hat{t}_{cat}}\right) > 0. \tag{A.32}$$

Therefore, proposition 2 implies $\hat{t}_{cat} < \hat{t}_{tax}$, because the lengths $\hat{t}_{tax}$ and $\hat{t}_{cat}$ are only equal if $Z_{cat} = Q_{tax}$, but it holds $Z_{cat} < Q_{tax}$.

Appendix A.3.5. Extension by intertemporal trading ratio

The only difference to the previous CAT problem is the adaption of the permit banking equation by the intertemporal trading ratio $(1 + r_b)$,

$$b_{it} = y_{it} - x_{it}\phi_i + b_{it-1}\left(1 + r_b\right). \tag{A.33}$$

Instead of an one-to-one exchange of permits over time, firms receive for each banked permit $(1 + r_b)$ permits in the next period. The first-order conditions for the firms remain unchanged apart from the intertemporal condition, which becomes

$$\frac{\partial V_{it}}{\partial b_{it}} - \frac{1+r_b}{1+r}\frac{\partial V_{it+1}}{\partial b_{it+1}} = 0 \quad (b_{it}), \tag{A.34}$$

and therefore, the growth rate of the permit price is

$$\frac{r-r_b}{1+r_b} = \frac{p_{t+1}-p_t}{p_t}\, \forall\, t \leq \hat{t}_{itr}, \tag{A.35}$$

where itr stands for intertemporal trading ratio. I derive the permit price again by inserting the production functions in the permit equilibrium, $p_t = \psi\left(\varphi_t - z_t - B_{t-1}\left(1+r_b\right) + B_t\right).$ Using this price in (A.35) yields the permit bank:

$$B_t = \frac{\sum_{s>t}^{\hat{t}_{itr}}\left(\left(1+r_b\right)^{2\hat{t}_{itr}-s-t}\left(\varphi_s - z_s\right) + \left(B_{t-1}\left(1+r_b\right) + z_t - \varphi_t\right)\left(1+r_b\right)^{2\left(\hat{t}_{itr}-s\right)}\left(1+r\right)^{s-t}\right)}{\sum_{s=t}^{\hat{t}_{itr}}\left(1+r_b\right)^{2\left(\hat{t}_{itr}-s\right)}\left(1+r\right)^{s-t}}. \tag{A.36}$$

This expression can again be used in the production functions (A.22) and (A.23) where p_t is replaced by $p_t = \psi\left(\varphi_t - z_t - B_{t-1}\left(1+r_b\right) + B_t\right)$. Based on the resulting reaction functions and assuming that all permits are issued in the first period the regulator maximizes welfare via the cap[20], which yields

$$z_0 = \sum_{t=0}^{\hat{t}_{itr}} \frac{1}{\left(1+r_b\right)^t}\left(\varphi_t - \frac{\sigma_t}{\psi}\right). \tag{A.37}$$

Thus, the optimal issuance of permits is the discounted sum of optimal periodic emissions with intertemporal trading ratio $(1+r_b)$ as discount factor. This restores the first-best outcome if the trading ratio is set to $(1+r_b) = (1+r)(1+g)^{-1}$, which implies that the permit price grows at the optimal rate g: Inserting the expressions for the bank (A.36), the cap (A.37) and the trading ratio $(1+r_b) = (1+r)(1+g)^{-1}$ into the permit price, $p_t = \psi\left(\varphi_t - z_t - B_{t-1}\left(1+r_b\right) + B_t\right)$, shows that the permit price is equal to the SCC,

[20]Note that because the available permits depend on the bank level, the temporal issuance of permits matters. For simplicity, I assume that all permits are issued in the first period. However, other allocations that also lead to the first best are possible.

$p_t = \sigma_t$ and therefore, the CAT program with intertemporal trading ratio is optimal.

Appendix A.4. Clean subsidy

Appendix A.4.1. Firms' problem

The firms maximize profits (2.38) via production x_{it} similar to the tax case. The first-order conditions result in production functions $x_{c,t} = (w_t + \rho_t - \alpha_{c,t})\,\beta_c^{-1}$ and $x_{d,t} = (w_t - \alpha_{d,t})\,\beta_d^{-1}$. The production functions and energy demand $D_t = \gamma_t - w_t\epsilon$ are again inserted in the energy market equilibrium (2.7) to derive the energy price (2.39). The energy price can then be used in the production functions to derive dirty and clean production (2.40) and (2.41) for $t \leq \hat{t}_{sub}$, respectively. For $t > \hat{t}_{sub}$, clean production is obtained via the energy market equilibrium, $x_{c,t} = (w_t + \rho_t - \alpha_{c,t})\,\beta_c^{-1} = \gamma_t - w_t\epsilon$. The resulting equilibrium energy price is substituted into $x_{c,t} = (w_t + \rho_t - \alpha_{c,t})\,\beta_c^{-1}$ to derive $x_{c,t} = \frac{\gamma_t - (\alpha_{c,t} - \rho_t)\epsilon}{1+\beta_c\epsilon}$ for $t > \hat{t}_{sub}$.

Appendix A.4.2. Regulator's problem

The regulator's problem becomes:

$$\max_{\rho_0,\dots\rho_t,\dots\rho_\infty} W_0 = \sum\nolimits_{t=0}^{\infty} \frac{1}{(1+r)^t} \left(\frac{\gamma_t \sum_i x_{it} - 0.5\left(\sum_i x_{it}\right)^2}{\epsilon} - \sum_i C_{it} - \delta a_t \Phi_t + a_t \right), \tag{A.38}$$

subject to the emission stock (2.9). Under consideration of the firms' reaction functions (2.40) and (2.41) for $t \leq \hat{t}_{sub}$ and $x_{d,t} = 0$ and $x_{c,t} = \frac{\gamma_t - (\alpha_{c,t} - \rho_t)\epsilon}{1+\beta_c\epsilon}$ for $t > \hat{t}_{sub}$, solving the problem yields (2.43). Setting (2.40) to zero and solving for ρ_t yields the subsidy for $t > \hat{t}_{sub}$ as presented in (2.43).

Multiplying dirty production (2.40) with ϕ_d gives emissions

$$q_t = \varphi_t - \frac{\rho_t}{\psi\phi_d\,(1+\beta_c\epsilon)}. \tag{A.39}$$

Considering small time steps and constant BAU emissions allows to derive the transition length again. Because the subsidy rises at rate g until the transition ends, it holds

$$\begin{aligned} q_{\hat{t}_{sub}} &= \varphi - \frac{\rho_0 e^{g\hat{t}_{sub}}}{\psi\phi_d(1+\beta_c\epsilon)} = 0 \\ \Rightarrow \hat{t}_{sub} &= \frac{1}{g} ln\left(\frac{\varphi\psi(1+\beta_c\epsilon)(1+\beta_d\epsilon)}{\sigma_0}\right), \end{aligned} \tag{A.40}$$

for which I inserted (2.43) for ρ_0.

Appendix A.4.3. Proof of proposition 3

Comparing the transition length (2.25) and (A.40) directly shows that $\hat{t}_{sub} > \hat{t}_{tax}$ if $\epsilon > 0$ and $\hat{t}_{sub} = \hat{t}_{tax}$ if $\epsilon = 0$. Cumulative emissions under the subsidy regime can be written as follows:

$$Q_{sub} = \int_0^{\hat{t}_{sub}} q_t dt = \hat{t}_{sub}\varphi - \frac{1}{g}\frac{\sigma_0\left(e^{g\hat{t}_{sub}} - 1\right)}{\psi\left(1+\beta_c\epsilon\right)\left(1+\beta_d\epsilon\right)}, \tag{A.41}$$

for which I have used (A.39). Inserting the optimal transition lengths (2.25) and (A.40) in (2.24) and (A.41), respectively, yields

$$Q_{tax} = \frac{1}{g}\left(\varphi ln\left(\frac{\varphi\psi}{\sigma_0}\right) - \varphi + \frac{\sigma_0}{\psi}\right), \tag{A.42}$$

$$Q_{sub} = \frac{1}{g}\left(\varphi ln\left(\frac{\varphi\psi\left(1+\beta_c\epsilon\right)\left(1+\beta_d\epsilon\right)}{\sigma_0}\right) - \varphi + \frac{\sigma_0}{\psi\left(1+\beta_c\epsilon\right)\left(1+\beta_d\epsilon\right)}\right), \tag{A.43}$$

from which directly follows that $Q_{sub} > Q_{tax}$ if $\epsilon > 0$ and $Q_{sub} = Q_{tax}$ if $\epsilon = 0$.

Appendix A.5. Clean subsidy and cap-and-trade

Appendix A.5.1. Firms' problem

The Bellman equation of the dirty firm's problem is again (A.18) subject to the banking constraint (2.27) such that the first-order conditions (A.19) to (A.21) still apply and the dirty firm's production function is $x_{it} = (w_t - \alpha_{it} - p_t\phi_i)\,\beta_i^{-1}$. The clean firm maximizes profits (2.38) as in the subsidy-only case and thus, its production function is again $x_{c,t} = (w_t + \rho_t - \alpha_{c,t})\,\beta_c^{-1}$. Inserting both production functions and demand

$D_t = \gamma_t - w_t\epsilon$ in the energy market equilibrium (2.7), yields the energy price:

$$w_t = \frac{\beta_c\beta_d\gamma_t + \beta_c\left(\alpha_{d,t} + p_t\phi_d\right) + \beta_d\left(\alpha_{c,t} - \rho_t\right)}{\beta_c + \beta_d + \beta_c\beta_d\epsilon} \,\forall\, t \leq \hat{t}_{c+s}. \tag{A.44}$$

Following the same steps as under the CAT program (see Appendix A.3.1) results in (2.44) and (2.45) for $t \leq \hat{t}_{c+s}$, for which I have redefined the BAU emissions to (2.46). After carbon neutrality is reached, $t > \hat{t}_{c+s}$, dirty energy generation is zero by definition, $x_{d,t} = 0$, and clean generation is $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c\epsilon}$.

Appendix A.5.2. Regulator's problem

First, I consider the commitment case. The problem is:

$$\max_{Z_{c+s},\rho_0,\ldots\rho_t,\ldots\rho_\infty} W_0 = \sum\nolimits_{t=0}^{\infty} \frac{1}{(1+r)^t}\left(\frac{\gamma_t\sum_i x_{it} - 0.5\left(\sum_i x_{it}\right)^2}{\epsilon} - \sum_i C_{it} - \delta a_t\Phi_t + a_t\right), \tag{A.45}$$

subject to the emission stock (2.9) and with reaction functions (2.44) and (2.45). Recall that for $t > \hat{t}_{c+s}$ it holds $x_{d,t} = 0$, $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c\epsilon}$ and $\rho_t = 0$. That is, for $t > \hat{t}_{c+s}$ the cap Z_{c+s} only affects the damage $\delta a_t\Phi_t = \delta a_t Z_{c+s}$ but no other variables. To derive the optimal cap, I take the first-order condition with respect to Z_{c+s}, which results in $Z_{c+s} = \sum_{t=0}^{\hat{t}_{c+s}}\left(\varphi_t - \frac{\sigma_t}{\psi}\right)$.

Due to binding cap Z_{c+s} the problem of finding the optimal subsidies turns the infinite into a finite horizon setting until the end of transition $\hat{t}_{c+s}$ as it holds $\rho_t = 0$ for $t > \hat{t}_{c+s}$. Taking the first-order conditions of (A.45) with respect to $\rho_0, \ldots\rho_t, \ldots\rho_{\hat{t}_{c+s}}$ gives the subsidies depending on all other subsidies $\rho_t\left(\rho_s\right)$ with $t \neq s$ and for all $t \leq \hat{t}_{c+s}$. Inserting the subsidies in each other yields (2.52), which can be simplified to

$$\rho_t \;=\; \frac{\phi_d\sigma_t}{1+\beta_d\epsilon}\left(1 - e^{t(r-g)}\frac{r}{g}\frac{e^{g\hat{t}_{c+s}} - 1}{e^{r\hat{t}_{c+s}} - 1}\right) \tag{A.46}$$

by assuming small time steps such that $e^{rt} \approx (1+r)^t$ and $e^{gt} \approx (1+g)^t$.

The permit price p_t can be derived by using emissions, $x_{d,t}\phi_d$ (with $x_{d,t}$ from (2.44)) in the expression for the permit price $p_t = \psi\left(\varphi_t^{sub} - z_t - B_{t-1} + B_t\right) = \psi\left(\varphi_t^{sub} - x_{d,t}\phi_d\right)$ while considering that $Z_{c+s} = \sum_{t=0}^{\hat{t}_{c+s}}\left(\varphi_t - \frac{\sigma_t}{\psi}\right)$, which yields (2.54). Now I can use the permit price (2.54) and the SCC (2.14) to reformulate (A.46) to (2.49).

To derive the optimal transition length I consider the total carbon price in the terminal transition period $\tilde{p}_{\hat{t}_{c+s}}$ as expressed in equation (2.50). Assuming constant BAU emissions φ, the duration of the transition satisfies

$$\tilde{p}_{\hat{t}_{c+s}} = e^{r\hat{t}_{c+s}} p_0 (1-\omega) + e^{g\hat{t}_{c+s}} \sigma_0 \omega = \varphi\psi, \tag{A.47}$$

which can be rearranged to (2.51).

Next, I consider the time-consistent solution. In this case, the regulator maximizes welfare via the subsidy ρ_t at the beginning of each period:

$$\max_{Z_{c+s},\rho_t} W_t = \sum\nolimits_{s=t}^{\infty} \frac{1}{(1+r)^{s-t}} \left(\frac{\gamma_t \sum_i x_{it} - 0.5 \left(\sum_i x_{it}\right)^2}{\epsilon} - \sum_i C_{it} - \delta a_t \Phi_t + a_t \right), \tag{A.48}$$

subject to the emission stock (2.9) and the firm production decisions given by (2.44) and (2.45). It again holds $x_{d,t} = 0$ and $x_{c,t} = \frac{\gamma_t - \alpha_{c,t}\epsilon}{1+\beta_c\epsilon}$ for $t > \hat{t}_{c+s}$, and by taking the derivative with respect to Z_{c+s} one obtains again $Z_{c+s} = \sum_{t=0}^{\hat{t}_{c+s}} \left(\varphi_t - \frac{\sigma_t}{\psi}\right)$.

Concerning the optimal subsidy, a difference is that under time consistency the subsidy is already set to zero in the terminal transition period $\rho_{\hat{t}_{c+s}} = 0$, whereas under commitment it is negative in the last transition period $\rho_{\hat{t}_{c+s}} < 0$ (see expression (2.52)) and zero only thereafter (see above). The reason for this difference is that under time consistency the regulator determines the subsidy in $\hat{t}_{c+s}$ rather than in $t = 0$ as under commitment. From the perspective of $\hat{t}_{c+s}$ a non-zero subsidy cannot increase welfare $W_{\hat{t}_{c+s}}$ because the emission level in $\hat{t}_{c+s}$ is already determined by the remaining permits of the CAT program. The first-order conditions of (A.48) with respect to $\rho_0, ...\rho_t, ...\rho_{\hat{t}_{c+s}-1}$ yield (2.53) or assuming small times steps $e^{rt} \approx (1+r)^t$ and $e^{gt} \approx (1+g)^t$:

$$\rho_t = \frac{\phi_d}{1+\beta_d\epsilon} \left(\sigma_t \left(1 - \frac{r}{g} \frac{e^{g\left(\hat{t}_{c+s}-t\right)} - 1}{e^{r\left(\hat{t}_{c+s}-t\right)} - 1} \right) + \frac{\phi_d r \int_t^{\hat{t}_{c+s}} \rho_s ds}{(1+\beta_c\epsilon)\left(e^{r\left(\hat{t}_{c+s}-1\right)} - 1\right)} \right). \tag{A.49}$$

As in the commitment solution, (A.49) can be rewritten to (2.49) by using the permit price (2.54) and the SCC (2.14). As expression (2.50) for the total carbon price still applies, the transition length (2.51) can be derived in the same way as in the commitment case.

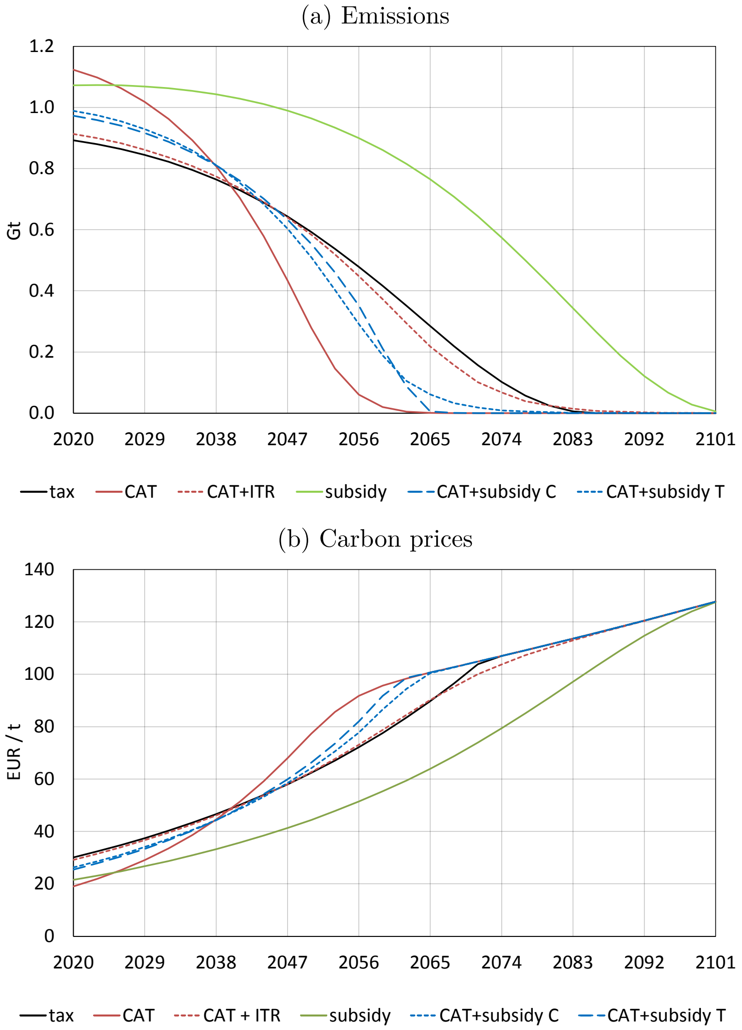
Appendix A.5.3. Proof of proposition 4

Consider the expression of the permit price (2.54). For a given cap Z_{c+s}, the duration of the transition phase $0, ...t, ...\hat{t}_{c+s}$ is fixed because the cap can be expressed as $Z_{c+s} = \sum_{t=0}^{\hat{t}_{c+s}} \left(\varphi_t - \frac{\sigma_t}{\psi}\right)$. For a given cap, the subsidies therefore only have an impact on the permit price if the second term in brackets of equation (2.54) is non-zero. In the commitment solution, the subsidy has no impact on the price as the sum of the subsidies over the transition periods is zero $\sum_{s=0}^{\hat{t}_{c+s}} \rho_s = 0$, which follows from (2.52). Under time consistency, it holds $\sum_{s=0}^{\hat{t}_{c+s}} \rho_s > 0$ due to the fact that the subsidy can be written as (2.53), which is always positive. It follows that the time-consistent subsidy reduces the permit price (2.54) for a given cap.

Appendix B. Additional simulation results for uncertainty scenarios

Appendix B.1. Emission and carbon price path

Figure B.1: Emission and carbon price paths under uncertainty (12% shock)



Note: The lines for CAT+subsidy C and CAT+subsidy T in part (b) reflect the total carbon prices (see equation (2.50)).

Appendix B.2. Cumulative emissions

Cumulative emissions are in all cases but one (CAT with an 8% shock) higher due to uncertainty (see Table 3.2), which can be explained by the zero lower bound of the emissions at the end of the transition. To see this, first consider a price policy (tax or subsidy) in which the bound implies that negative shocks, which would lead to fewer emissions, have no or a weaker effect if the emissions are already zero or close to zero.

In contrast, positive shocks have the full emission increasing effect if the emissions are (close to) zero, thus inducing larger overall (expected) emissions.

Under CAT, there are two effects on the cumulative emissions: First, for a given cap, the expected transition length is always shorter than the transition length under certainty (not shown; note that the values in Table 3.2 (c) are for adjusted caps as shown in panel (b) of the table). This can be traced back to the concavity of the emission paths that results from exponentially increasing permit prices, implying an asymmetric impact of shocks on the transition length. Because the optimal cap decreases with a shorter transition length (cf., equation (2.35)), this first effect leads to less cumulative emissions. The second effect increases the optimal cap: If there is a positive probability that the bank is depleted in the next period, the expected permit price generally rises at a lower rate than the interest rate because of the convenience yield (Schennach 2000). Thus, from the first period onward in which the transition could end (i.e., the bank is depleted), the expected growth rate of the price is lower than under certainty, that is, lower than r. For a given cap, this implies lower expected prices and more emissions later. In turn, prices are higher early due to intertemporal trading. The resulting lower emissions (and damage) in the beginning allow to increase the cap similar to the case when subsidies are added to a CAT program. The first effect dominates for the 8% shock in the CAT-only scenario, and therefore, the cumulative emissions are lower (see Table 3.2 (b)), while for the 12% shock the second effect dominates.

Appendix B.3. Intertemporal trading ratio

Although a CAT program with an ITR restores the first best under certainty, the program performs worse with increasing uncertainty, and essentially, the welfare loss increases faster than in the CAT and subsidy cases. Thus, subsidies may outperform the ITR as a complement to CAT with increasing uncertainty, but not given the parameter assumptions in this simulation. The reason for the increasing inefficiency of the ITR is that uncertainty causes the expected permit price to grow at a rate that declines over time because of the convenience yield. In response to the lower average permit price growth rate, the regulator also reduces the ITR slightly to 2.3% and 2.2% (compared to 2.5% under certainty) given the 8% and 12% shock, respectively, because a lower ITR increases the growth rate (see Appendix A.3.5). This, however, cannot restore the optimal abatement path, because the permit price growth rate is time-variable. Because the fixed ITR is too low for the early phase and too high for the later phase, the permit price is too low in the beginning and too high later (see Figure B.1).

References

Abrell, J., S. Rausch, and C. Streitberger (2019). The economics of renewable energy support. *Journal of Public Economics 176*, 94–117.

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton, NJ, USA: Princeton University Press.

Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hemous (2012). The Environment and Directed Technical Change. *American Economic Review 102*(1), 131–166.

Acemoglu, D., U. Akcigit, D. Hanley, and W. Kerr (2016, feb). Transition to Clean Technology. *Journal of Political Economy 124*(1), 52–104.

Böhringer, C. and K. E. Rosendahl (2010). Green promotes the dirtiest: On the interaction between black and green quotas in energy markets. *Journal of Regulatory Economics 37*(3), 316–325.

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

Daniel, K. D., R. B. Litterman, and G. Wagner (2019). Declining CO2 price paths. *Proceedings of the National Academy of Sciences of the United States of America 116*(42), 20886–20891.

Deryugina, T., A. MacKay, and J. Reif (2020). The Long-Run Dynamics of Electricity Demand: Evidence from Municipal Aggregation. *American Economic Journal: Applied Economics 12*(1), 86–114.

Dietz, S., C. Gollier, and L. Kessler (2018). The climate beta. *Journal of Environmental Economics and Management 87*, 258–274.

Dietz, S. and F. Venmans (2019). Cumulative carbon emissions and economic policy: In search of general principles. *Journal of Environmental Economics and Management 96*, 108–129.

Fankhauser, S., C. Hepburn, and J. Park (2010). Combining Multiple Climate Policy Instruments: How not to do it. *Climate Change Economics 1*(3), 209–225.

Fell, H., D. Burtraw, R. D. Morgenstern, and K. L. Palmer (2012). Soft and hard price collars in a cap-and-trade system: A comparative analysis. *Journal of Environmental Economics and Management 64*(2), 183–198.

Fell, H., I. A. MacKenzie, and W. A. Pizer (2012). Prices versus quantities versus bankable quantities. *Resource and Energy Economics 34*(4), 607–623.

Fell, H. and R. D. Morgenstern (2010). Alternative Approaches to Cost Containment in a Cap-and-Trade System. *Environmental and Resource Economics 47*(2), 275–297.

Fischer, C. and R. G. Newell (2008). Environmental and technology policies for climate mitigation. *Journal of Environmental Economics and Management 55*(2), 142–162.

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.

Gerlagh, R. and R. J. R. K. Heijmans (2018). Regulating Stock Externalities. *CESifo Working Paper No. 7383*.

Gerlagh, R., S. Kverndokk, and K. E. Rosendahl (2009). Optimal timing of climate change policy: Interaction between carbon taxes and innovation externalities. *Environmental and Resource Economics 43*(3), 369–390.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica 82*(1), 41–88.

Goulder, L. H. (2020). Timing Is Everything: How Economists Can Better Address the Urgency of Stronger Climate Policy. *Review of Environmental Economics and Policy*, 1–15.

Grüll, G. and L. Taschini (2011). Cap-and-trade properties under different hybrid scheme designs. *Journal of Environmental Economics and Management 61*(1), 107–118.

Hoel, M. and L. Karp (2001). Taxes and quotas for a stock pollutant with multiplicative uncertainty. *Journal of Public Economics 82*(1), 91–114.

Hoel, M. and L. Karp (2002). Taxes versus quotas for a tax pollutant. *Resource and Energy Economics 24*, 367–384.

Hotelling, H. (1931). The Economics of Exhaustible Resources. *Journal of Political Economy 39*(2), 137–175.

ICAP (2019). *Emissions Trading Worldwide: Status Report 2019.* Berlin, Germany: International Carbon Action Partnership.

Interagency Working Group on the Social Cost of Greenhouse Gases (2016). *Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866.* Washington, DC: Interagency Working Group on the Social Cost of Carbon.

Jarke, J. and G. Perino (2017). Do renewable energy policies reduce carbon emissions? On caps and inter-industry leakage. *Journal of Environmental Economics and Management 84*, 102–124.

Karp, L. and C. Traeger (2018). Prices versus Quantities Reassessed. *CESifo Working Paper No. 7331*.

Karp, L. S. (2019). Welfare Ranking of Emissions Taxes, Cap & Trade, and Banking & Borrowing. *Working paper*.

Kling, C. and J. Rubin (1997). Bankable permits for the control of environmental pollution. *Journal of Public Economics 64*(1), 101–115.

Kollenberg, S. and L. Taschini (2016). Emissions trading systems with cap adjustments. *Journal of Environmental Economics and Management 80*, 20–36.

Kuusela, O.-p. and J. Lintunen (2020). A Cap-and-Trade Commitment Policy with Allowance Banking. *Environmental and Resource Economics 75*, 421–455.

Lecuyer, O. and P. Quirion (2013). Can uncertainty justify overlapping policy instruments to mitigate emissions? *Ecological Economics 93*, 177–191.

Lecuyer, O. and P. Quirion (2019). Interaction between CO2 emissions trading and renewable energy subsidies under uncertainty: feed-in tariffs as a safety net against over-allocation. *Climate Policy 19*(8), 1002–1018.

Leiby, P. and J. Rubin (2001). Intertemporal permit trading for the control of greenhouse gas emissions. *Environmental and Resource Economics 19*(3), 229–256.

Lintunen, J. and O. P. Kuusela (2018). Business cycles and emission trading with banking. *European Economic Review 101*, 397–417.

Newell, R., W. Pizer, and J. Zhang (2005). Managing Permit Markets to Stabilize Prices. *Environmental & Resource Economics 31*(2), 133–157.

Newell, R. G. and W. A. Pizer (2003). Regulating stock externalities under uncertainty. *Journal of Environmental Economics and Management 45*(2, Supplement), 416–432.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences of the United States of America 114*(7), 1518–1523.

Osorio, S., P. Nahmmacher, E. Schmid, and B. Knopf (2018). Documentation of LIMES-EU - A long-term electricity system model for Europe. *PIK model documentation.*

Perino, G. (2018). New EU ETS Phase 4 rules temporarily puncture waterbed. *Nature Climate Change 8*(4), 260–261.

Pizer, W. A. (2002). Combining price and quantity controls to mitigate global climate change. *Journal of Public Economics 85*(3), 409–434.

Pizer, W. A. and B. C. Prest (2020). Prices versus Quantities with Policy Updating. *Journal of the Association of Environmental and Resource Economists 7*(3), 483–518.

REN21 (2019). *Renewables 2019 Global Status Report.* Paris: REN21 Secretariat.

Rezai, A. and F. Van Der Ploeg (2017). Abandoning Fossil Fuel: How Fast and How Much. *Manchester School 85*(14373), e16–e44.

Roberts, M. J. and M. Spence (1976). Effluent charges and licenses under uncertainty. *Journal of Public Economics 5*(3-4), 193–208.

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

Traeger, C., G. Perino, K. Pittel, T. Requate, and A. Schmitt (2020). The Flexcap. An Innovative CO2 Pricing Method for Germany. *ifo DICE Report 18*(1), 42–48.

van der Ploeg, F. (2018). The safe carbon budget. *Climatic Change 147*(1-2), 47–59.

Van Der Ploeg, F. and C. Withagen (2014). Growth, renewables, and the optimal carbon tax. *International Economic Review 55*(1), 283–311.

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

Weitzman, M. L. (1978). Optimal rewards for economic regulation. *American Economic Review 68*(4), 683–691.

Weitzman, M. L. (2020). Prices or Quantities Can Dominate Banking and Borrowing. *The Scandinavian Journal of Economics 122*(2), 437–463.

Yates, A. J. and M. B. Cronshaw (2001). Pollution permit markets with intertemporal trading and asymmetric information. *Journal of Environmental Economics and Management 42*(1), 104–118.

Chapter 4

Adverse effects of rising interest rates on sustainable energy transitions[1]

Tobias S. Schmidt
Bjarne Steffen
Florian Egli
Michael Pahle
Oliver Tietjen
Ottmar Edenhofer

[1]Published in *Nature Sustainability* as Schmidt, T. S., Steffen, B., Egli, F., Pahle, M., Tietjen, O., Edenhofer, O. (2019). Adverse effects of rising interest rates on sustainable energy transitions. Nature Sustainability, Vol. 2, pages 879-885, available online: https://doi.org/10.1038/s41893-019-0375-2.

nature sustainability

ANALYSIS
https://doi.org/10.1038/s41893-019-0375-2

Adverse effects of rising interest rates on sustainable energy transitions

Tobias S. Schmidt[1]*, Bjarne Steffen[1]*, Florian Egli[1]*, Michael Pahle[2], Oliver Tietjen[2,3] and Ottmar Edenhofer[2,3,4]

Increasing the use of renewable energy (RE) is a key enabler of sustainable energy transitions. While the costs of RE have substantially declined in the past, here we show that rising interest rates (IRs) can reverse the trend of decreasing RE costs, particularly in Europe with its historically low IRs. In Germany, IRs recovering to pre-financial crisis levels in 5 years could add 11% and 25% to the levelized cost of electricity for solar photovoltaics and onshore wind, respectively, with financing costs accounting for about one-third of total levelized cost of electricity. As fossil-fuel-based electricity costs are much less and potentially even negatively affected by rising IRs, the viability of RE investments would be markedly deteriorated. On the basis of these findings, we argue that rising IRs could jeopardize the sustainable energy transition and we propose a self-adjusting thermostatic policy strategy to safeguard against rising IRs.

Replacing fossil fuel (FF)-based with renewable energy (RE)-based electricity generation technologies has multiple societal benefits, such as climate change mitigation or improved air quality and thus better health[1–3]. Doubling the global share of RE is therefore one of the targets of Sustainable Development Goal (SDG) 7 of the 2030 United Nations Agenda for Sustainable Development, a goal whose attainment is also highly important to reach several other SDGs[4,5]. This particularly holds for climate action (SDG 13): all emission pathways that reach the Paris Agreement's target of limiting global warming to well below 2 °C assume strong increases of RE[6,7].

Achieving such a sustainable energy transition has become easier since RE, especially solar photovoltaics (PV) and wind, have experienced substantial cost reductions in the past decades[8]. This dynamic has been enabled mainly by RE deployment policies, particularly in countries of the European Union (EU), a front-runner in large-scale RE deployment. These policies induced technological and organizational innovation and contributed to the formation of a global RE industry that exploited economies of scale in production and thereby allowed RE technologies to progress down their cost learning curves[9,10]. Today, in many European countries, the levelized costs of electricity (LCOE) of RE investments are comparable with the marginal costs of gas- and coal-based electricity plants[8,11]. In line with these developments, recent auctions for RE in Europe were concluded at wholesale market prices. Since April 2017, such subsidy-free auction results have appeared in Denmark, Germany, the Netherlands, Spain, Portugal and Sweden[12]. In autumn 2018, for the first time, a large PV plant that relies only on income from the wholesale electricity market was commissioned in Spain[13]. By mid-2019, similar projects were announced and constructed in Germany[13,14].

These developments have beguiled scholars, industry experts, policy makers and the media into believing that the trend of decreasing RE costs is irreversible[15,16] and claiming that the times of subsidizing RE are over[12,17–19]. As a result, countries particularly in the EU—yet again being front runners, although in reverse direction—are considering abandoning RE subsidies and leaving RE deployment to market forces[20,21] and the EU Emissions Trading System (ETS). While phase-outs in Europe have not been implemented yet, the trend towards phasing out RE deployment policies is apparent from the EU's recent decision no longer to impose legally binding RE targets and respective deployment policies in its member states.

However, it should not be taken for granted that the strong downward trend of RE costs observed in the past is going to continue. New data for Germany show that the past RE cost reductions not only stem from technological innovation but also, to a substantial extent, arise from improved financing conditions for RE power plants, particularly lowered long-term interest rates (IRs)[22]. Lower IRs translate directly into lower cost of debt and equity[22,23], which lowers the LCOE of capital-intensive RE investments[24,25]. Thus far, the potential effects of rising IRs on the viability of RE investments are unexplored.

To address this gap, first, we analyse the effects of IR increases on the LCOE of large-scale solar PV and onshore wind investments, finding that their LCOE might increase by 11% (PV) and 25% (wind), should IRs reach pre-financial crisis levels over the next 5 years. Second, we compare these LCOE with the marginal cost of installed FF plants, as these typically set the wholesale market prices. We find that the viability of RE investments solely relying on income from the wholesale market is drastically reduced by rising IRs. Third, based on these findings, we argue that solely relying on wholesale markets and the EU ETS is a risky strategy and we propose an alternative policy strategy, relying on RE auctions in the short run and an ETS price floor in the longer run.

IR effects on the cost of RE

The low IRs observed in recent years in Europe (and beyond) are mostly a consequence of monetary policy. In the wake of the financial crisis of 2007–2008, the European Central Bank lowered IRs overnight and in 2015 started purchasing large amounts of sovereign and corporate bonds—an approach termed 'quantitative easing'. This contributed to low levels of long-term IRs (see Fig. 1a), allowing RE plants to

[1]Energy Politics Group, ETH Zurich, Zürich, Switzerland. [2]Potsdam-Institute for Climate Impact Research, Potsdam, Germany. [3]TU Berlin, Berlin, Germany. [4]Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany. *e-mail: tobiasschmidt@ethz.ch; bjarne.steffen@gess.ethz.ch; florian.egli@gess.ethz.ch

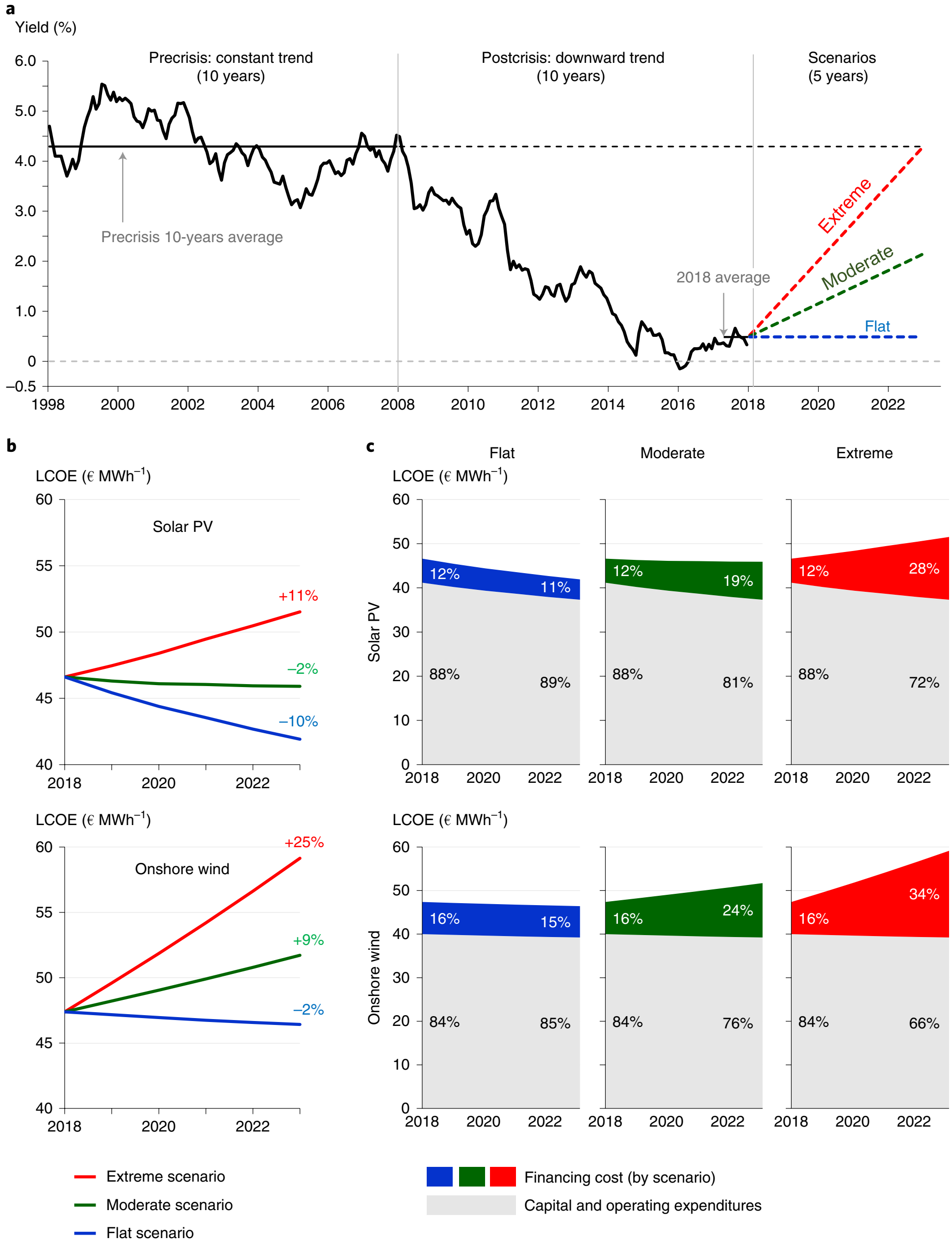


Fig. 1 | IR dynamics and their effects on the levelized cost of RE-based electricity generation. a, Historical development of long-term IRs (black solid line) in Germany and future scenarios. IR recovery scenarios are based on historical estimates. The moderate scenario (green dashed line) features the same upward slope as the downward slope of the postcrisis trend. The extreme scenario (red dashed line) doubles that slope. The flat scenario (blue dashed line) assumes constant IRs. **b**, Solar PV (top) and onshore wind (bottom) LCOE developments for 2018 and 5 years into the future in the three IR scenarios (using the same colour codes as in **a**). **c**, Share of financing cost in the LCOE of solar PV (top) and onshore wind (bottom) across all three scenarios (using the same colour coding as in **a** and **b**).

borrow capital at very low rates[22]. While long-term IRs in Europe are still historically low, financial expert estimates suggest that they might rise again, as they already have in the United States since 2016[26].

To quantify the effect of possible future IR increases, we calculate the LCOE for solar PV and onshore wind plants in Germany in three steps (see Methods for details). First, we use historical government bond yields to develop three future IR scenarios from 2019 to 2023 (Fig. 1a). Second, we project technology- and time-specific future costs of capital (CoC) in each of the three scenarios. Third, we calculate LCOEs for each scenario using the scenario-specific CoC (Fig. 1b) and derive the part of the LCOE attributable to financing costs (Fig. 1c).

We assume that the cost of debt is composed of a long-term IR component (yield of 10-year government bond) and a debt margin. We assume that the cost of equity equals the cost of debt plus an equity premium. To represent the changes in CoC due to IR changes, we vary the long-term IR component according to the IR scenario while using technology-specific debt margins and equity premiums from Egli et al.[22]. The three IR scenarios are (see Fig. 1a): (1) a 'flat' scenario, where IRs remain at the 2018 average (0.49%); (2) a 'moderate' scenario, where IRs increase at the speed at which they decreased in the postcrisis period to attain 2.15% in 2023; (3) an 'extreme' scenario, where IRs increase at twice the speed of the postcrisis decrease to attain 4.29% in 2023. Note that the moderate scenario is in line with the 2018 forecasts of major financial institutions (see Methods).

Our results (Fig. 1b) show that, in the moderate scenario, the LCOE-reducing learning curve effects of solar PV are almost entirely offset by the LCOE-increasing effects of the rising IR. In the extreme IR recovery scenario, the LCOE rise by 11%, with financing cost contributing up to 28% to the LCOE (Fig. 1c). For wind, the IR effects are even larger (Fig. 1b), outweighing the learning effects and resulting in an increased LCOE of 9% (moderate) and 25% (extreme recovery scenario). In this scenario, financing costs contribute over one-third of the LCOE (Fig. 1c). In other words, while the window of extremely low IRs helped RE to become cheap, rising IRs could mean that the decreasing cost trend of RE might be reversed—also because cost reductions along the learning curve through continued incremental technological innovation are becoming less important (especially in the case of wind).

Viability of investments in subsidy-free renewables

To realize the sustainable energy transition in time to meet the Paris targets and SDG 13, key scenarios[6,7] show that it is necessary that RE are deployed rapidly, displacing FF-based electricity in the mid- to long-term, and eventually stranding FF-assets. For this to happen, investments in subsidy-free new RE capacity (relying on income from the wholesale market) need to remain attractive. To this end, the LCOE of new RE plants need to be lower or equal to the short-run marginal costs of the price-setting plants in wholesale markets. Whether rising IRs deteriorate the viability of subsidy-free RE investments therefore also depends on the effect of IR hikes on the short-run marginal costs of price-setting plants, that is, typically FF plants using lignite, hard coal or natural gas[27] (for more details see Methods).

Theoretical analysis suggests that variations of the IR could have the following two effects on commodity prices[28–33]. First, higher IRs could increase the supply of commodities, since investors earn more interest with the revenues they receive from selling resources compared to leaving them in the ground. Second, commodity supply could also be increased because holding commodity inventories has higher opportunity costs and thus inventory levels are reduced. A related point is that speculators leave commodity markets if the IR is high since alternative investments offer higher returns at low risk (treasury bills). Empirical estimates tend to support theoretical findings that higher IRs reduce commodity prices in general and energy commodities (most studies focus on oil) in particular, although estimated parameters are for the most part not significant or small[30,34–38]. For this article, the impact on natural gas and hard coal prices is decisive. To the best of our knowledge, there is no estimate for coal. The only estimate for natural gas[35] finds a negative relationship but estimated parameters are not significant. Further, one study[36] also considers a fuel commodity price index that includes oil, gas and coal, detecting a negative impact of an IR increase on this index. The difficulties to find evidence for an IR effect on commodities might be explained by complex interactions between the IRs and commodity prices[32]. Not only do shocks of the IR affect commodity prices but exogenous shocks of commodity prices also affect the economic activity and, due to market interactions, the IR as well[39]. Moreover, central banks may respond to commodity price shocks (or the implied inflation) which further complicates the interaction between fuel prices and the IR[40,41]. Whether central banks systematically respond to oil price shocks is, however, controversial as more recent work does not support this relationship[42].

Given the lack of clear evidence we resort to theory and assume a generally negative effect of IRs on the price of fossils. To account for uncertainty and divergent results of the empirical literature, we cover a range of fuel price changes from −7.5% to +2.5% for hard coal and natural gas for each 1%-point of IR increase (see Methods). We keep the marginal cost of lignite that is used in mine-mouth plants constant. This serves as input for modelling the effect of IR developments on the marginal cost of FF-based electricity generation in the three IR scenarios. We further consider the range of thermal efficiencies in Germany's electricity generation park for plant types using lignite, hard coal and gas in combined cycle gas turbine setups, as in most cases one of these technologies sets the price. ETS emission prices are held constant at 2018 levels (see Methods).

The following three observations can be made based on our cost projections (Fig. 2). First, assuming constant carbon prices, lignite plants' marginal costs remain out of reach for both solar PV and wind LCOE across all scenarios. Second, the LCOE of solar PV is lower than the marginal cost of gas plants and of almost all hard-coal plants in a flat IR scenario (Fig. 2a). With rising IRs, the viability of solar PV investments in a wholesale market-based setting deteriorates. In the moderate IR rise scenario, solar PV LCOE remain lower than the cost range for running gas plants and the upper-half of the cost range for hard-coal plants (Fig. 2b). In the extreme IR rise scenario, solar PV LCOE cease to be comparable with running hard-coal power plants and only undercut the upper-half of the marginal cost range of gas plants (Fig. 2c). Third, onshore wind LCOE are below the marginal costs of gas and less efficient hard-coal plants only if IRs remain at today's levels (Fig. 2a). With moderately rising IRs (Fig. 2b), wind investments become less viable. LCOE are above the marginal cost of even the least efficient hard-coal plants. In the extreme IR rise scenario, onshore wind LCOE are even higher than the marginal costs of most gas plants (Fig. 2c). In sum, the comparisons show that an increase in IRs would substantially deteriorate the economic viability of RE investments that need to earn their LCOE from wholesale market prices set by FF plants.

Sustainable energy transition at risk

Our findings have important ramifications for the policy mix driving the sustainable energy transition in Europe. If governments of EU countries abandon their RE deployment policies, comparatively higher costs of RE would need to be absorbed by the EU ETS. However, the long-term economic performance of the EU ETS may be hampered by distortions such as myopic decision-making by investors and the limited credibility of the government-imposed emissions cap[43]. Hence, while allowance prices in the EU ETS are currently recovering, it is possible that history will repeat itself (C. Flachsland et al., manuscript in preparation) and prices may collapse again as observed several times in the past.

From the perspective of ensuring a continuous transition, relying on the EU ETS in its current state alone might be a risky strategy. Importantly, even short-term slumps of RE deployment due to deteriorated economic viability might have negative long-term consequences for sustainable developments and related goals. Industry slumps and consequential layoffs would have a negative impact on decent jobs and economic growth (SDG 8). This often results in the loss of hard-earned technological capabilities and tacit knowledge in technology development, manufacturing, project development and financing (SDG 9), in turn resulting in increased technology adjustment costs[44]. Given the global importance of the European RE technology industry, the effect could have worldwide implications.

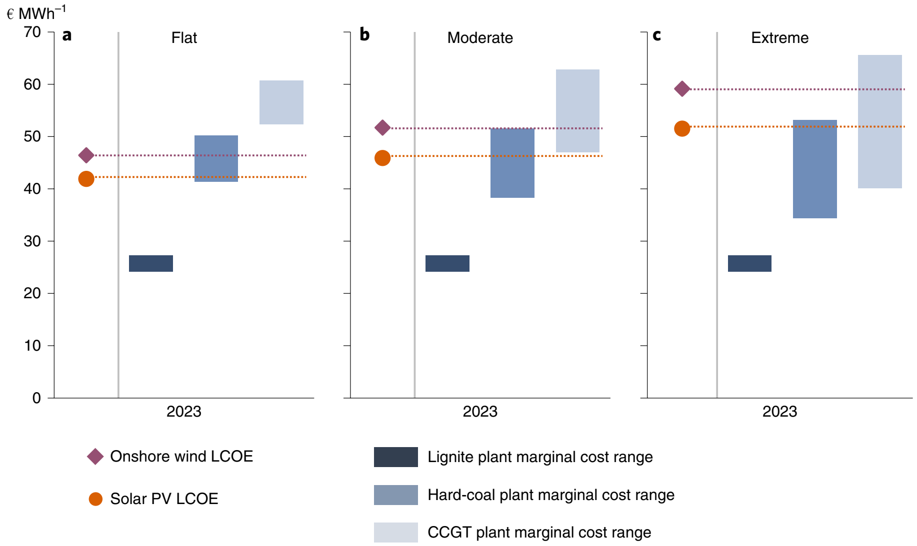


Fig. 2 | Projected range of short-run marginal cost of FF-based vis-à-vis LCOE of RE-based electricity generation in 2023. a, Projected ranges assuming a flat IR, where ranges represent different thermal efficiencies. **b**, Projected ranges assuming a moderate rise in IRs, where ranges relate to the combination of thermal efficiencies and IR–fuel cost elasticities. **c**, Projected ranges assuming an extreme rise in IRs, where ranges relate to the combination of thermal efficiencies and IR–fuel cost elasticities. CCGT, combined cycle gas turbine.

In consequence, abandoning RE deployment policies in the face of rising IR is bad timing and could jeopardize the sustainable energy transition. To prevent this, we recommend a 'thermostatic policy' strategy[45], which—like a thermostat—automatically counterbalances potential increases in IR. The most direct policy measure to address rising IRs would be to couple the provision of subsidized loans with IRs: for example, financed via issuing green bonds[46]. For instance, the European Investment Bank, which has played a major role in financing RE plants[47,48], could provide such subsidized loans in response to rising IRs. However, subsidized loans could crowd out private finance and create problems in calibrating subsidy levels to avoid over- or under-installations of RE (for example, in case of unexpected changes of fuel cost of competing FF-based electricity generation).

Recommendations for policy makers and researchers

In view of these results, we recommend a two-stage policy strategy, supported by new energy models. In the short run, RE auction policies could work like a thermostat: as long as financing conditions remain favourable, bids will continue to yield market prices (or premiums of zero), representing a zero-cost policy. If IRs increase, bids will go up again, reflecting worsened financing conditions. Competitive auctions for RE could thus ensure continued RE capacity additions while avoiding the above-mentioned calibration problems in a cost-effective way. Hence, instead of phasing out RE policies, governments should keep or—where not yet in place—introduce such auction policies. Sophisticated auction policies, for example, using contracts for differences, could further improve the financing conditions[49]. Importantly, while the past role of RE deployment policies was to induce innovation, the new role would be to safeguard RE deployment—and thus the SDGs and the Paris targets—against the negative consequences of rising IRs.

In the longer run, a durable price floor could be introduced in the EU ETS to remedy the current above-mentioned distortions (as is the case in the Californian ETS). Ideally, such a floor would over time become high enough to ensure RE deployment even in times of high IRs, allowing RE auctions to be eventually phased out. The continuous deployment of RE assured by the near-term auction policies, would probably increase the political feasibility of an ETS price floor due to positive feedback effects, such as increased political support for RE and decreased influence of the FF industry[10,50,51]. New models are required to understand which price floor levels could sustain renewable deployment and how energy market and policy risks not considered here would factor into financing costs. Such models should also consider broader general equilibrium effects not accounted for in the partial equilibrium perspective taken here.

This line of research should extend into a broader research stream, bringing this issue to policy makers' attention and exploring the full scope of implications. The reason most policy makers are unaware of the implications of higher IRs is that the models they typically use for energy- and climate-related decision-making ignore IR dynamics. Future model-based research should therefore incorporate IR dynamics and explicitly cover aspects of thermostatic policy strategies, including how to deal with the potential trade-off between rising policy costs and adjustment costs resulting from potential RE industry slumps. Moreover, political scientists should consider the fact that IRs can change the cost dynamics of RE, which in turn affects the dynamics of energy and climate politics. In addition, researchers should explore potential new roles of public and central banks in addressing macroeconomic risks to the clean-energy transition that could jeopardize the Paris targets, attainment of SDG 7 and sustainable development more broadly.

In parallel, a second stream of research needs to consider the lost revenue from stranded FFs. The question of whether to compensate asset owners or not is an ethical (and political) one that is difficult to answer. On the one hand, according to the 'polluter pays principle', the costs of avoiding pollution need to be carried by the polluters, implying no compensation. On the other hand, from a political economy perspective, displacement of workers creates a strong incentive for politicians to compensate workers or regions. In the case of the German coal phase-out, heavy compensations along with structural aid were brokered by a coal phase-out commission[52]. A sustainable energy transition requires considering all stakeholders and making tradeoffs, including between the affected SDGs, such as clean energy (SDG 7), climate action (SDG 13), decent work (SDG 8) and reduced inequalities (SDG 10)[4]. Importantly, the effectiveness of other policies addressing these SDGs might also be affected by IR changes. Further research analysing the effect of IRs on attaining these SDGs and on their tradeoffs could facilitate the debate.

Methods

We calculate the LCOE for solar PV and onshore wind plants in Germany in three steps. First, we use historical government bond yields to develop three future IR scenarios for 5 years (2019 to 2023). Second, we project technology- and time-specific future CoC in each of the three scenarios. Third, we calculate LCOEs for each scenario using the CoCs and other parameters and derive the part of the LCOE attributable to financing costs. The approach, data sources and assumptions for each step are described next.

IR scenarios. In the first step, we use monthly data on 10-year German government bond yields from July 1998 to June 2018 (20 years)[53]. We define a 'pre-financial crisis' and a 'post-financial crisis' period of 120 months (10 years) each, using the month with the highest yield (June 2008, at 4.52%) as the separator. The pre-financial crisis period covers July 1998 to June 2008 and the post-financial crisis period covers July 2008 to June 2018. The separator month is 3 months before Lehman Brothers filed for Chapter 11 bankruptcy protection, which marks the peak of the financial crisis in the United States that subsequently spread to the Euro-zone[54]. We calculate the average over the 10-year precrisis period and use it as a reference point for the level that long-term IRs could reach again in a rebound. Specifically, we define three scenarios: (1) a flat scenario in which long-term IRs stay constant at the 2018 average of 0.49%; (2) a moderate scenario in which long-term IRs rise at the same rate at which they decreased in the 10 years during and after the financial crisis, reaching 50% of the precrisis average (2.15%) in 2023; (3) an extreme scenario in which long-term IRs rise at twice the rate at which they previously declined, reaching the precrisis average (4.29%) in 2023. For comparison, the moderate scenario (0.97% in December 2019) is in line with several financial institutions' 2019 outlooks on the long-term IR (10-year German government bond), which project a level of 0.8% in July 2019 and 1.0% in the fourth quarter of 2019 or the end of 2019, respectively[26,55,56].

Cost of capital of RE plants. In the second step, we use the three IR scenarios and project-level data on PV and wind financing conditions to calculate technology- and time-specific costs of capital[22]. We build on the methodology of Egli et al.[22], calculating the after-tax CoC using equation (1), in which E and D denote equity and debt investment, respectively; V signifies the total investment sum; K_D and K_E refer to the cost of debt and the cost of equity, respectively; and T represents the corporate tax rate. The leverage ratio is equal to D/V.

$$\mathrm{CoC} = K_\mathrm{D}\frac{D}{V}(1-T) + K_\mathrm{E}\frac{E}{V} \quad (1)$$

Again following Egli et al.[22], we split the cost of debt into a long-term IR component (IR) and the debt margin (DM), as shown in equation (2). Furthermore, we follow the energy-finance literature[57] in defining the cost of equity as the cost of debt plus an equity premium (EP), as shown in equation (3).

$$K_\mathrm{D} = \mathrm{IR} + \mathrm{DM} \quad (2)$$

$$K_\mathrm{E} = \mathrm{IR} + \mathrm{DM} + \mathrm{EP} \quad (3)$$

The cost of debt and the cost of equity change over time, depending on the IR scenario. All other indicators, namely the debt margin, the equity premium, the leverage ratio and the tax rate, are held constant. We use the technology-specific 2017 average values from Egli et al.[22] for the first three indicators and the German corporate tax rate in 2017 for the last one. All parameters are summarized in Supplementary Tables 1 and 2.

LCOE model for RE plants. In the third step, we parametrize an LCOE model for both technologies in each year (2019 to 2023) using equation (4). Note that this formulation of LCOE represents a cash-flow perspective and hence does not account for depreciation[58].

$$\mathrm{LCOE} = \frac{C^{\mathrm{CAPEX}} + \sum_{t=1}^{t=25}\frac{C_t^{\mathrm{OPEX}}}{(1+\mathrm{CoC})^t}}{\sum_{t=1}^{t=25}\frac{\mathrm{FLH}_t}{(1+\mathrm{CoC})^t}} \quad (4)$$

C^{CAPEX} denotes the initial investment cost per MW (CAPEX) at $t=0$, C_t^{OPEX} represents the operation and maintenance costs per MW per year (OPEX) from $t=1$ to $t=25$ (constant) and FLH_t signifies the full-load hours of the asset per year from $t=1$ to $t=25$ (constant). The discount rate CoC is the technology- and time-specific cost of capital. We calculate future investment costs (€ MW^{-1}) using global cumulative installed capacity by combining global capacity data in 2017 from IRENA[59] with deployment scenarios for the years 2018 to 2023 from IEA[60]. As a starting point, we use German investment costs in 2018[61]. The future investment costs are then calculated using a one-factor learning curve commonly used in the literature[62] and learning rates specific to the German context[61]. We further parametrize the LCOE model by using data for Germany in 2018 for full-load hours, operation and maintenance cost (€ $\mathrm{MW}^{-1}\,\mathrm{year}^{-1}$) and asset lifetime[61]. We use solar PV full-load hours for central Germany and onshore wind values for northern Germany. We assume that full-load hours, operation and maintenance costs and asset lifetime stay constant for both technologies from 2018 to 2023. All parameters are summarized in Supplementary Tables 1 and 2.

Finally, we follow the approach of Egli et al.[22] in splitting the LCOE into a CAPEX/OPEX component and a financing-cost component (see Fig. 1c and Main). The latter consists of debt service (principal repayment and IR payment) and returns to equity. We do so by estimating an LCOE with 0% cost of capital for both technologies in each year. We define the difference between the LCOE estimated using our technology- and time-specific cost of capital and the LCOE estimated using 0% cost of capital as the financing-cost share δ, according to equation (5).

$$\delta = \mathrm{LCOE} - \mathrm{LCOE}_{\mathrm{CoC}=0} \quad (5)$$

Comparison of RE LCOE with marginal cost of FF plants. In the final step, we compare the projected LCOE of solar PV and onshore wind with the short-run marginal cost of FF-based plants. Generally, the viability of RE investments depends on a comparison of their LCOE with expected electricity wholesale market prices, both of which could be affected by IR changes. While near-zero marginal cost will allow RE plants to produce electricity before FF plants are dispatched, investments in new RE capacity are only viable if market prices allow them to earn their full LCOE. Following the merit order principle, the wholesale price that subsidy-free RE plants can earn is an average of prices set by the different marginal plants depending on the market situation (given that investment costs of these marginal plants are sunk and thus are neither considered for dispatch nor decommissioning decisions). In Europe, wholesale price levels dropped over the last decade due to higher renewable penetration and lower cost of emission certificates and fuels[27,63]. However, Germany and its neighbouring countries still have overcapacities from an investment boom in the 2000s[63,64] and only recently consider capacity payments to incentivize new dispatchable capacity once it will be needed[65]. Thus, short-run marginal costs of existing FF plants will probably continue setting market prices. In the German case, price-setting plants are typically either lignite-based plants, hard-coal-based plants or combined cycle gas turbine plants[27]. For our analysis we do not know which of these technologies sets the price during how much of the time where the RE plants produce electricity but compare RE LCOE to the marginal cost of all three technologies.

The marginal costs of FF plants (MC) in € MWh^{-1} are calculated as

$$\mathrm{MC} = \frac{C^{\mathrm{fuel}}}{\eta} + \frac{C^{\mathrm{emissions}} \times \mathrm{EF}}{\eta} + C^{\mathrm{VOM}} \quad (6)$$

where C^{fuel} are fuel costs (in € per $\mathrm{MWh}_{\mathrm{thermal}}$), η is the thermal efficiency (in $\mathrm{MWh}_{\mathrm{electric}}$ per $\mathrm{MWh}_{\mathrm{thermal}}$), $C^{\mathrm{emissions}}$ are the CO_2 ETS emission certificate costs (€ t^{-1}), EF is the CO_2 emissions factor (t per $\mathrm{MWh}_{\mathrm{thermal}}$) and C^{VOM} are variable operations and maintenance costs (€ per $\mathrm{MWh}_{\mathrm{electric}}$). Parameters are taken from recent studies on electricity generation cost in Germany and held constant across scenarios (except for fuel prices, see later). Values and sources are summarized in Supplementary Table 3. The ranges for thermal efficiencies reflect the fact that in each technology group, power plants of different age and efficiency currently exist in Germany, yielding a range of marginal costs per FF-based plant type.

For CO_2 emissions cost, we assume constant EU ETS costs at the average market price for in 2018 (closing European Climate Exchange European Emission Allowances futures prices, continuous contract no. 1: non-adjusted price based on spot-month continuous contract calculations). For fuel costs, the flat scenario is based on price projections for 2023 based on multiple market studies as compiled in ref. [11]. For the moderate and extreme scenarios, we model a change of fuel costs with changes in the general IR level. The exact effect size depends on the temporal dimension. Typically, models estimate the reaction to a shock over time[33]. Estimated effects of a 100 basis points decrease of the IR on the oil price range from around 0 to 7% for months 1 to 47 after the shock. Depending on the model specification, the maximum effect ranges from 2.1 to 14.4% and occurs 4–6 months after the shock. For a commodity index, the range is 0.7 to 6.0% for the maximum effect (3–15 months after the shock). Given that the theoretical and empirical literature describes a negative fuel price per IR elasticity but is inconclusive regarding the precise quantification of the elasticity (see section 'Viability of investments in subsidy-free renewables' above), we model a fuel price change between −7.5% and +2.5% for each 100 basis points of IR increase for hard coal and natural gas (which are commodities). For our argument, this range is a conservative reading of the empirical literature as it even allows for a positive effect, which is typically not empirically observed, and as it limits the negative effect to the baseline estimation of Anzuini et al.[33], while alternative estimations produce a twice-as-high negative effect size.

In contrast to hard coal and natural gas, lignite is generally not traded but mined on-site (a non-commodity), so the fuel price is considered independent of the IR level. Overall, it should be kept in mind that the present analysis depicts the situation in Europe where an installed base of FF-based power plants competes with newly erected RE plants—which is why marginal costs are compared to LCOE of the latter.

Data availability

All data and the models used for this paper are provided in the Supplementary Dataset.

Received: 9 January 2019; Accepted: 7 August 2019;
Published online: 9 September 2019

References

1. Giang, A. & Selin, N. E. Benefits of mercury controls for the United States. *Proc. Natl Acad. Sci. USA* **113**, 286–291 (2016).
2. Peng, W. et al. Managing China's coal power plants to address multiple environmental objectives. *Nat. Sustain.* **1**, 693–701 (2018).
3. IPCC *Special Report on Renewable Energy Sources and Climate Change Mitigation* (eds Edenhofer, O. et al.) (Cambridge Univ. Press, 2011).
4. Fuso Nerini, F. et al. Mapping synergies and trade-offs between energy and the sustainable development goals. *Nat. Energy* **3**, 10–15 (2018).
5. Fuso Nerini, F. et al. Connecting climate action with other Sustainable Development Goals. *Nat. Sustain.* **2**, 674–680 (2019).
6. Rogelj, J. et al. in *Global Warming of 1.5 °C* (eds Masson-Delmotte, V. et al.) 93–174 (IPCC, 2018).
7. Luderer, G. et al. Residual fossil CO_2 emissions in 1.5–2 °C pathways. *Nat. Clim. Change* **8**, 626–633 (2018).
8. Ralon, P., Taylor, M., Ilas, A., Diaz-Bone, H. & Kairies, K.-P. *Electricity Storage and Renewables: Costs and Market to 2030* (IRENA, 2017).
9. Trancik, J., Brown, P., Jean, J., Kavlak, G. & Klemun, M. *Technology Improvement and Emissions Reductions as Mutually Reinforcing Efforts* (MIT, 2015).
10. Schmidt, T. S. & Sewerin, S. Technology as a driver of climate and energy politics. *Nat. Energy* **2**, 17084 (2017).
11. Kost, C., Shammugam, S., Jülich, V., Nguyen, H.-T. & Schlegel, T. *Levelized Cost of Electricity Renewable Energy Technologies* (Fraunhofer Institute for Solar Energy Systems, 2018).
12. Wronski, M. *Renewables 2.0: the Subsidy-free Revolution* (Aurora Energy Research, 2018).
13. *BayWa: First Subsidy-Free Solar Park Under Construction in Germany* (Marketscreener, accessed 8 July 2019); https://www.marketscreener.com/BAYWA-AG-435730/news/BayWa-First-Subsidy-Free-Solar-Park-Under-Construction-in-Germany-28666080/
14. *EnBW Developing 175MW Subsidy-free Solar Project in Germany* (PVtech, accessed 8 July 2019); https://www.pv-tech.org/news/enbw-developing-175mw-pv-project-in-germany
15. Obama, B. The irreversible momentum of clean energy. *Science* **355**, 126–129 (2017).
16. Clark, P. The big green bang: how renewable energy became unstoppable. *Financial Times* (18 May 2017).
17. Bassi, S., Carvalho, M., Doda, B. & Fankhauser, S. *Credible, Effective and Publicly Acceptable Policies to Decarbonise the European Union* (LSE/Grantham Research Institute, 2017).
18. Pfeifer, S. Subsidy-free renewable projects on 'cusp of breakthrough'. *Financial Times* (28 March 2018).
19. *Managing Merchant Risk in Renewables* (Aurora Energy Research, 2018).
20. *Green Finance: Mobilising Investment in Clean Energy and Sustainable Development* (UK Parliament, 2018).
21. *Koalitionsvertrag: Ein neuer Aufbruch für Europa Eine neue Dynamik für Deutschland Ein neuer Zusammenhalt für unser Land* (CDU/CSU/SPD, 2018).
22. Egli, F., Steffen, B. & Schmidt, T. S. A dynamic analysis of financing conditions for renewable energy technologies. *Nat. Energy* **3**, 1084–1092 (2018).
23. Steffen, B. The importance of project finance for renewable energy projects. *Energy Econ.* **69**, 280–294 (2018).
24. Schmidt, T. S. Low-carbon investment risks and de-risking. *Nat. Clim. Change* **4**, 237–239 (2014).
25. Hirth, L. & Steckel, J. C. The role of capital costs in decarbonizing the electricity sector. *Environ. Res. Lett.* **11**, 114010 (2016).
26. *Global Economic Outlook* (Swiss Re Institute, 2018).
27. Kallabis, T., Pape, C. & Weber, C. The plunge in German electricity futures prices—analysis using a parsimonious fundamental model. *Energy Policy* **95**, 280–290 (2016).
28. Frankel, J. A. & Hardouvelis, G. A. Commodity prices, money surprises and Fed credibility. *J. Money, Credit Bank* **17**, 425–438 (1985).
29. Frankel, J. A. Expectations and commodity price dynamics: the overshooting model: reply. *Am. J. Agric. Econ.* **69**, 856–856 (1986).
30. Frankel, J. A. Effects of speculation and interest rates in a 'carry trade' model of commodity prices. *J. Int. Money Finance* **42**, 88–112 (2014).
31. Barsky, R. B. & Kilian, L. in *NBER/Macroeconomics Annual* (eds Bernanke, B. S. & Rogoff, K.) 137–183 (MIT Press, 2002).
32. Kilian, L. & Zhou, X. *Oil Prices, Exchange Rates and Interest Rates* CEPR Discussion Paper No. DP13478 (SSRN, 2019).
33. Anzuini, A., Lombardi, M. & Pagano, P. *The Impact of Monetary Policy Shocks on Commodity Prices* Bank of Italy Temi di Discussione Working Paper No. 851 (SSRN, 2012).
34. Arango, L. E., Arias, F. & Flórez, A. Determinants of commodity prices. *Appl. Econ.* **44**, 135–145 (2012).
35. Chiou-Wei, S. Z., Chen, S. H. & Zhu, Z. Energy and agricultural commodity markets interaction: an analysis of crude oil, natural gas, corn, soybean, and ethanol prices. *Energy J.* **40**, 265–296 (2019).
36. Hammoudeh, S., Nguyen, D. K. & Sousa, R. M. US monetary policy and sectoral commodity prices. *J. Int. Money Finance* **57**, 61–85 (2015).
37. Huang, W., Mollick, A. V. & Nguyen, K. H. Dynamic responses and tail-dependence among commodities, the US real interest rate and the dollar. *Empir. Econ.* **53**, 959–997 (2017).
38. Ratti, R. A. & Vespignani, J. L. Oil prices and global factor macroeconomic variables. *Energy Econ.* **59**, 198–212 (2016).
39. Bodenstein, M., Erceg, C. J. & Guerrieri, L. Oil shocks and external adjustment. *J. Int. Econ.* **83**, 168–184 (2011).
40. Bodenstein, M., Guerrieri, L. & Kilian, L. Monetary policy responses to oil price fluctuations. *IMF Econ. Rev.* **60**, 470–504 (2012).
41. Bernanke, B. S., Gertler, M. & Watson, M. Systematic monetary policy and the effects of oil price shocks. *Brook. Pap. Econ. Act.* **1**, 91–142 (1997).
42. Kilian, L. & Lewis, L. T. Does the Fed respond to oil price shocks? *Econ. J.* **121**, 1047–1072 (2011).
43. Fuss, S. et al. A framework for assessing the performance of cap-and-trade systems: insights from the European Union emissions trading system. *Rev. Environ. Econ. Policy* **12**, 220–241 (2018).
44. Vogt-Schilb, A., Meunier, G. & Hallegatte, S. When starting with the most expensive option makes sense: optimal timing, cost and sectoral allocation of abatement investment. *J. Environ. Econ. Manag.* **88**, 210–233 (2018).
45. Cashore, B. & Howlett, M. Punctuating which equilibrium? Understanding thermostatic policy dynamics in Pacific Northwest forestry. *Am. J. Pol. Sci.* **51**, 532–551 (2007).
46. Edenhofer, O. Klima, Kohle, Kapital. Ökonomische Hebel in der internationalen Klimapolitik. *Aus Politik und Zeitgeschichte* (18 May 2018); http://www.bpb.de/apuz/269302/klima-kohle-kapital-oekonomische-hebel-in-der-internationalen-klimapolitik?p=all
47. Geddes, A., Schmidt, T. S. & Steffen, B. The multiple roles of state investment banks in low-carbon energy finance: an analysis of Australia, the UK and Germany. *Energy Policy* **115**, 158–170 (2018).
48. Steffen, B. & Schmidt, T. S. A quantitative analysis of 10 multilateral development banks' investment in conventional and renewable power generation technologies from 2006 to 2015. *Nat. Energy* **4**, 75–82 (2019).
49. May, N., Neuhoff, K. & Richstein, J. C. Affordable electricity supply via contracts for difference for renewable energy. *DIW Weekly Report* **8**, 251–259 (2018).
50. Pahle, M. et al. Sequencing to ratchet up climate policy stringency. *Nat. Clim. Change* **8**, 861–867 (2018).
51. Meckling, J., Kelsey, N., Biber, E. & Zysman, J. Winning coalitions for climate policy. *Science* **349**, 1170–1171 (2015).
52. Edenhofer, O. & Pahle, M. The German coal phase out: buying out polluters, not (yet) buying into carbon pricing. *EAERE Mag.* **5**, 11–14 (2019).
53. *Main Economic Indicators* (OECD, 2016); https://doi.org/10.1787/data-00052-en
54. Gorton, G. & Metrick, A. Getting up to speed on the financial crisis: a one-weekend-reader's guide. *J. Econ. Lit.* **50**, 128–150 (2012).
55. *Macro Research Economic Forecasts* (Deka, 2018).
56. *Commerzbank Forecasts* (Commerzbank, 2018).
57. Donovan, C. & Nuñez, L. Figuring what's fair: the cost of equity capital for renewable energy in emerging markets. *Energy Policy* **40**, 49–58 (2012).
58. Comello, S., Reichelstein, S. & Sahoo, A. The road ahead for solar PV power. *Renew. Sustain. Energy Rev.* **92**, 744–756 (2018).
59. *Renewable Capacity Statistics 2017* (IRENA, 2017).
60. *Renewables 2018 Market Analysis and Forecast from 2018 to 2023* (IEA, 2018).
61. *Stromgestehungskosten Erneuerbare Energien* (Fraunhofer ISE, 2018).
62. Rubin, E. S., Azevedo, I. M. L., Jaramillo, P. & Yeh, S. A review of learning rates for electricity supply technologies. *Energy Policy* **86**, 198–218 (2015).
63. Keles, D., Bublitz, A., Zimmermann, F., Genoese, M. & Fichtner, W. Analysis of design options for the electricity market: the German case. *Appl. Energy* **183**, 884–901 (2016).
64. *Definition und Monitoring der Versorgungssicherheit an den Europäischen Strommärkten* Projekt No. 047/16 (Fraunhofer ISI, TEP, Consentec & r2b, 2019); https://www.r2b-energy.com/uploads/media/BMWi_VS-Monitoring_Ber1_deu_20190520_03.pdf

65. Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C. & Fichtner, W. A survey on electricity market design: insights from theory and real-world implementations of capacity remuneration mechanisms. *Energy Econ.* **80**, 1059–1078 (2019).

Acknowledgements
This research was conducted as part of the EU Horizon 2020 research and innovation programme project INNOPATHS (grant no. 730403). As such, it was partly supported by the Swiss State Secretariat for Education, Research and Innovation (contract no. 16.0222). The opinions expressed and arguments used herein do not necessarily reflect the official views of the Swiss Government. M.P. was also supported by Stiftung Mercator Foundation under the research project AHEAD. We gratefully acknowledge all funding.

Author contributions
T.S.S., B.S., F.E., M.P., O.T. and O.E. developed the research idea. B.S., F.E. and T.S.S. compiled the data and developed the model. T.S.S., B.S., F.E., M.P., O.T. and O.E. interpreted the results. T.S.S., together with B.S., F.E., M.P., O.T. and O.E., wrote the paper. M.P. and T.S.S. secured project funding.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41893-019-0375-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.S.S., B.S. or F.E.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Chapter 5

The risk of softening the cap in emissions trading systems[1]

Michael Pahle
Oliver Tietjen
Sebastian Osorio
Florian Egli
Bjarne Steffen
Tobias S. Schmidt
Ottmar Edenhofer

[1]Submitted to *Proceedings of the National Academy of Sciences of the United States of America (PNAS).*

Abstract

Mitigating climate change requires a fundamental low-carbon transition, and emissions trading is widely thought to be a fail-safe instrument for achieving it. However, if prices become hockey stick-shaped, the resulting political pressure may soften the cap, implying the transition will not be followed through. Combing economics and political science, we conceptualise the feedback between allowance price and politics, underline the importance of the discount rate, and infer indicators for the risk of softening the cap. We then quantify these indicators for the EU, assessing a scenario with market risk exposure for renewables and rising interest rates. Compared to a scenario extrapolating current conditions, we find that allowance prices double in the long-term. Moreover, renewable capacity deployment is substantially delayed in the mid-term, and fossil fuel plant profits surge rather than decline. This suggests a considerable risk that the cap will be softened.

Significance Statement

Emissions trading systems are adopted by more and more jurisdictions worldwide. Their particular appeal is that emissions never exceed the capped amount. This led to a widespread belief that emissions trading is “fail-safe” and emissions targets will always be reached. We challenge this belief and provide a framework for how to assess the risk that the emissions cap will eventually be softened. Combining economics and political science, we conceptualize the feedback between allowance price and politics, and infer suitable indicators to assess the risk. Using an energy-system model, we then quantify these indicators for the EU-ETS for two different scenarios to underline that a hockey-stick price path implies a considerable risk that the cap will be softened.

Main Text

Introduction

The appeal of emissions trading – also referred to as cap-and-trade – is to achieve a given emission reduction target at the lowest cost (1, 2). The cap acts as a backstop that ensures emissions never exceed a certain level, and allowance prices automatically adjust to changes in economic and policy conditions. This particular feature has led to a widespread belief that emissions trading is “fail-safe”: once a cap is in place, it will always be reached. Yet, economic and political research challenges this belief. Economic work has hypothesised that if prices reach unanticipated levels, this may lead to a retrospective change in the cap (3). Specifically, if prices rise too sharply, the cap might be revoked (4). Correspondingly, a certain level may exist beyond which allowance prices are not politically acceptable (5). Relatedly, political science has become concerned with the stickiness of cap-and-trade programmes – analysing political resilience across election cycles, the ability to adapt (6), and specifically in Europe, the politics of ratcheting up (7). From all this work, it transpires that important feedback of allowance prices on politics exists, which induces a risk that the cap will be softened. However, this feedback has not been made explicit so far, and the feedback mechanisms have not been analysed.

Our work addresses this gap by combining the two disciplines. First, we conceptualise how alternative allowance price paths for the same cap provoke different feedback on politics. Specifically, we argue that firms’ discount rates are crucial for such feedback in emission trading systems since they can lead to a hockey stick-shaped price path. Such a path raises and concentrates the economic adjustment costs of fossil technologies and decelerates the deployment of clean technologies. These are indicators for a shifting balance of the political power of respective industries, which is at the core of the risk that the cap will be softened.

Second, using the Long-term Investment Model for the Electricity Sector (LIMES-EU), we quantify these indicators for the European Union Emission Trading System (EU ETS), responding to a recent call to expand the application of numerical economic models to inform policy makers on the political feasibility of policies (8). We find that, compared to currently low discount rates due to existing monetary and renewable policies, a higher rate induces a substantial risk of softening the cap by mid-2030 in terms of the indicators: allowance prices double by 2055, renewable capacity deployment is delayed eight years by 2035, and fossil fuel plant profits surge to their peak by 2025. This is valuable information for EU policy makers, who just started deliberating about the post-2030 policy framework: If they take a wait-and-see approach and ignore the effect of other policies on allowance prices, their inaction may backfire with a softening of the cap in the future. Furthermore, general insights can also help policy makers in other jurisdictions design future emissions trading programmes.

Dynamic feedback from allowance prices on politics

In this section, we conceptualise how the risk of a future softening of the cap arises from the feedback of allowance prices on politics. To that end, we draw on both the political science and economics literature to answer the following questions: First, what are the specific mechanisms that may trigger future softening or strengthening of the cap? Second, which policy settings (e.g. socioeconomic parameters, design features) are crucial for determining the balance of feedback and the likelihood of future softening?

To begin with, the political science literature argues that policy feedback is an important determinant for policy change (9, 10). More specifically, the stickiness of a policy is largely determined by the feedback this policy – once enacted – creates on future politics. Both negative and positive feedback may exist: If positive feedback prevails over negative feedback, the likelihood of policy stickiness increases. If negative feedback prevails over positive feedback, the likelihood of policy dismantling, e.g. softening the cap, increases (11, 12). Whether positive or negative feedback dominates depends largely on the costs and benefits a policy creates and how they are distributed (12). The cost and benefits are strongly related to the industry structure. Structural change, in turn, affects related political coalitions and feedback (13).

Specifically for emissions trading, how the feedback develops over time depends on the allowance price path. According to the economic literature on emissions trading, allowance prices rise exponentially over time (14, 15). When allowances can be banked, the rate of price increase over time equals the discount rate due to intertemporal arbitrage. Accordingly, the general shape of the price path is determined by the discount rate. A high rate implies that the price starts at a relatively low level but rises sharply towards the end. Recent economic work alluding to the problems of such paths (4, 5) refers to them as hockey stick-shaped, which we also do in the following.

Combining political science and economic perspectives, as shown in Figure 1, suggests that emission trading is particularly prone to the risk of future dismantling when the discount rate is high. To illustrate this, we make use of two stylised scenarios: “sticky cap” and “softened cap”. The “sticky cap” can be thought of as the scenario underlying the view that the cap is fail-safe, and it is thus the appropriate starting point to make our case. In the early phase, when prices are relatively low, the industry structure is dominated by carbon-intensive fossil-fuel (FF) capacity. The corresponding brown political coalition dominates policy choices. However, with rising prices, green firms benefit because this increases renewable energy (RE) technologies’ competitiveness and generates profits. This, in turn, fosters the growth of green industries. The higher the industries' profits, the more strongly they engage in a supporting coalition to sustain the policy – similar to the effect of green industrial policies (16). In parallel, the FF capacity becomes less competitive and is pushed out of the market, and the corresponding brown coalition becomes smaller and smaller. Accordingly, in the late phase when prices are high, overall feedback is positive working towards a sticky cap.

This is different in the "softened cap" scenario. A high discount rate implies lower prices and, therefore, less green investments in the early phase, which postpones the transformation of the industry structure and the related expansion of the green coalition. At the same time, the sharper rise of prices in the late phase provokes stronger opposition from the brown coalition because of higher adjustment costs. More specifically, within a short time, considerable FF capacity becomes devaluated (17). To avoid this, the still-dominant brown political coalition opposes the policy (18). Since the green coalition is still relatively nascent, overall feedback is negative working towards a softening of the cap.

Notably, this situation is only exacerbated when firms anticipate future dismantling and adapt their demand for allowances and intertemporal trading behaviour accordingly. This is because it effectively leads to a higher discount rate (19). Accordingly, anticipating future intervention is a sort of self-fulfilling prophecy in the sense that it makes a future dismantling more likely.

To quantitatively analyse specific cases, we propose the following indicators to capture the relative risk of dismantling between different scenarios: First, the intersection of allowance price paths marks the onset of the blade of the hockey stick, from which point one can expect the risk of softening to become increasingly severe. Second, the delay in renewable capacity deployment captures the industry structure and particularly the relative strength of the positive feedback from the supporting green coalition. A larger delay implies weaker support. Third, the profit dynamics of fossil fuel-based plants capture the adjustment costs and strength of the negative feedback from the brown coalition. A sharper decline implies stronger opposition.

Application to the EU ETS

In this section, we apply this framework to the EU ETS and quantify the above indicators using LIMES-EU. The motivation for selecting this case is that, in the EU, the two main components of the discount rate – the general risk-free interest rate (IR) and the market risk premium – are currently particularly low. As we will explain, this is mainly due to energy and macroeconomic policies. While in the turmoil of the COVID-19 pandemic the prospects of these policies are difficult to assess, it certainly merits investigation as a contingency scenario (20).

The persistently low IR in the EU has been the result of monetary policies (including quantitative easing) over recent years. If monetary policy is reversed, the IR may revert to the "old normal". This affects the allowance price path directly, as explained in the previous section, and also indirectly by altering the costs of capital. Green technologies, namely renewables, are typically more capital intensive than brown technologies (21). Because a higher interest rate increases the cost of capital, green technologies become relatively more expensive (22). This, in turn, raises the initial allowance price level, implying that the price grows faster in absolute terms and becomes more bent in the later phase. Likewise, higher market risks have the same effect. Currently, dedicated support policies in the EU practically shield renewables from market risk. But political support is waning, and there is an ongoing discussion about phasing them out to rely solely on the EU-ETS for incentivising investments (22). This would expose renewables to the full market risk, implying a risk premium that increases the costs of capital even further.

Against this background, we analyse two scenarios using LIMES-EU, a long-term cost optimisation model of the sectors regulated under the EU-ETS (see Methods for a more detailed description): The first is a low discount rate (LoDR) scenario resembling the current situation to establish a reference case for comparison. The second is a high discount rate (HiDR) scenario that could arise under the policy choices described above. Accordingly, we break the discount rate down into the general IR and a risk premium depending on the market risk that power plants face: (1) In the LoDR scenario, we assume that monetary and energy policies remain as they currently are. Hence, the general IR remains at the current level of around 0%. FF-based power plants remain exposed to market risk, which we assume translates into a cost of capital of 5%. RE plants continue to be supported by dedicated policies, which effectively nullify their market risk exposure (23–26). Accordingly, we assume a cost of capital of 0%. (2) In the HiDR scenario, we assume that the general IR rises to 5%, and RE policies will be phased out, implying that RE technologies also face the full market risk. Consequently, investments into all technologies – FF and RE – are discounted with a uniform rate of 10%, resembling the market risk premium plus the increased general IR.

Analysing the scenarios using LIMES-EU, we first look at the two allowance price paths (Figure 2) to establish the extent to which the price path is more hockey stick-shaped in the HiDR scenario. We find that the allowance price in the HiDR scenario is lower until 2035 despite the higher cost of capital for renewable energies. Yet, in the long term, prices take a hockey stick-shape because of the high growth rate of the allowance price. Prices rise to around €160 per ton of CO2 (€/t) in 2055 compared to only around 80 €/t in the LoDR scenario. The fact that prices in the HiDR scenario surpass prices in the LoDR scenario only after 2035 suggests political pressure to soften the cap will mount two or three decades from now and, hence, seems to be a distant concern. However, looking at the specific indicators, it may turn out that the critical period will start in the nearer future, around 2030.

Turning to RE capacity deployment, which indicates the strength of the green coalition, we find that lower allowance prices early on and higher costs of capital for REs imply significantly lower deployment (Figure 3). By 2035, the renewable capacity in the LoDR scenario (1,363 Gigawatt [GW]) is around 430 GW higher than in the HiDR scenario (932 GW). In other words, the level reached in the HiDR scenario by 2035 is already reached by 2027 in the LoDR scenario, implying a “deployment time lag” of around eight years between the two scenarios. To put this into perspective, eight years corresponds to approximately two election cycles in most EU countries, providing an opportunity for campaigning, thereby asserting influence on the composition of the political coalitions in parliaments and governments, i.e. two feedback loops, as shown in Figure 1. It is worth noting that increasing either the IR or market risk, keeping the other fixed at the LoDR level, already decreases deployment by around 280 GW (not shown in Figure 3). This corresponds to a deployment lag of around 5 years. In other words, if just one of the two effects would eventually unfold as assumed in the scenario, the deployment lag would still be substantial.

Finally, we turn to the third indicator, the profits of fossil fuel-based plants. Figure 4 shows that overall short-run profits are substantially higher in the HiDR scenario compared to the LoDR scenario. Even after 2035, when carbon prices are higher, profits remain higher. The main reason is the deployment lag of renewable energy in HiDR, which increases market shares and, thus, profits for fossil technologies. With a view on how negative political feedback unfolds, the situation is somewhat ambiguous since profits also decline in the LoDR scenario between 2020 and 2030. However, it is unlikely that this will result in a backlash that endangers the cap for two reasons. First, in the period through 2030, allowance prices are still relatively low and do not exceed 30 €/t. Thus, the brown coalition cannot credibly point to a high carbon price burden to lobby against the ETS. Second, since profits are relatively low, the funds that could be used for lobbying are also relatively low in contrast to the HiDR scenario, in which considerably more funds would be available due to the higher and more prolonged profits (27). More broadly, whereas profits in the LoDR scenario indicate a phase-out trajectory, in the HiDR scenario, profits still rise and thus “peak fossil” still lies ahead. The former pattern discourages lobbying (“lost cause”), whereas the latter encourages it (“playing for time”).

In sum, this analysis of the indicators suggests that the risk of softening the cap is considerably higher in the HiDR scenario than in the LoDR scenario for two intertwined reasons. First, in the HiDR scenario, the green coalition expands much slower due to the delay in RE capacity, and correspondingly, the extant brown coalition remains dominant for a longer time. Second, the brown coalition's lobbying power is fuelled by higher and prolonged profits and the prospect that the fossil phase-out can be postponed. Overall, it seems likely that by mid-2030 the cap will come under pressure: By then the lag in deployment would have accumulated to eight years and fossil profits would have started to decline, which triggers a political pushback from respective generation asset owners.

Discussion and conclusion

We argue that emissions trading is not a fail-safe policy instrument to achieve the low-carbon transition. To that end, we establish how political feedback arising from allowance prices induces risk that the cap might be softened: a hockey stick-shaped price path arising from a high discount rate delays the formation of a strong green coalition in the short term and strengthens the lobbying power of the brown coalition in the longer term. These two dynamically interlinked effects determine the inherent risk of a future softening of the cap. Suitable indicators to quantify this risk are the allowance price path (onset of hockey stick blade), installed green capacity (relative strength of green coalition), and profits of fossil technologies (relative strength of brown coalition).

The particular value of our approach lies in combining political science and economics, allowing us to overcome the respective blind spots in the two disciplines. On the one hand, economics is concerned with how policy design and market conditions influence the allowance price path. However, while the risk of regulatory intervention is acknowledged, its drivers are rarely explicitly considered and analysed. On the other hand, the focus of political science is on political feasibility and the stickiness of the cap. However, little attention is given to long-term feedback effects from allowance prices and the role of related market conditions. It is only by combining the two disciplines that this important interlinkage – policy feeds back on politics, which in turn feeds back on policy – can be unveiled. This interlinkage is what questions emissions trading as a fail-safe policy instrument.

The results of our quantitative application are of high relevance for EU policy makers, who are about to initiate the process of designing the post-2030 policies. The indicators suggest that in the HiDR scenario, negative feedback will have mounted to exert substantial pressure to soften the cap by mid-2030. As for example Germany's coal phase-out shows, this pressure can generally be dealt with at the time it arises – but only with high compensatory payments (28). Accordingly, taking measures to prevent a softening of the cap will pay out in the future, and several options for doing so exist. A continuation of the renewable support and keeping the general interest rates low, as in the LoDR scenario, is the apparent choice. However, climate policy makers have little say in that, so it should be addressed directly in the emissions trading design. A suitable mechanism would be a price collar that effectively "flattens" the hockey stick. This could reduce the risk of softening and would make the policy maker's commitment to the cap more credible (29). Accordingly, whether emissions trading is fail-safe is a matter of policy design. Having a better understanding of the risk, to which we hope this work contributes, allows policy makers to take precautionary measures now to avoid that the cap will be softened in the future.

Methods

The LIMES-EU is a linear dynamic cost-optimisation model with a focus on the electricity sector. It simultaneously optimises investment and dispatch decisions for generation, storage and transmission technologies in five-year time steps from 2010 to 2070. Each year is modelled using six representative days, including eight blocks of three hours. The representative days are estimated using a clustering algorithm (30), which allows capturing the short-term variability of supply (namely wind and solar) and demand. The model contains 32 generation and storage technologies, including different vintages for lignite, hard coal and gas. The energy-intensive industry is also covered and represented by a step-wise linear marginal abatement cost curve for each country. The EU ETS is implemented in line with the recent 2018 reform, including the Market Stability Reserve (MSR) and cancellation of allowances. A comprehensive description of the model (parameters, equations and assumptions) is provided in the LIMES-EU documentation available from the model's website (31).

Like all optimisation models of this type, LIMES-EU allows for a single discount rate. Accordingly, technology-specific discount rates need to be implemented indirectly. To this end, we convert the market risk premium (δ [%]) into a monetary value expressing the net present value of the technology-specific risk premium (RP [€/MW]) and deduct it from the model's default investment costs. For this purpose, we make use of the fact that LIMES-EU results resemble a competitive market equilibrium, i.e. the NPV of all investments equals zero. To determine the risk premium, we define

$$(1)\ NPV_A = -IC + \sum_t \frac{1}{(1+r+\delta)^t} CF_t = 0 \ \Rightarrow\ IC = \sum_t \frac{1}{(1+r+\delta)^t} CF_t$$

$$(2)\ NPV_B = -(IC + RP) + \sum_t \frac{1}{(1+r)^t} CF_t = 0 \ \Rightarrow\ IC + RP = \sum_t \frac{1}{(1+r)^t} CF_t$$

where CF_t are the cash flows from selling the plant's production, and IC are the investment costs. Assuming that the cash flows are constant over time, the risk-adjusted investment costs ($IC + RP$) can be derived by expanding the right-hand side of (2) and inserting (1), which leads to the following expression:

$$(3)\ IC + RP = \frac{IC \sum_t \frac{1}{(1+r)^t}}{\sum_t \frac{1}{(1+r+\delta)^t}}$$

Acknowledgments

This research was conducted as part of the EU Horizon 2020 research and innovation programme project INNOPATHS (grant no. 730403). As such, it was partly supported by the Swiss State Secretariat for Education, Research and Innovation (contract no. 16.0222). The opinions expressed and arguments used herein do not necessarily reflect the official views of the Swiss Government. Paste your acknowledgments here.

References

1. J.-M. Neuhaus, L. Sticher, F. Meins, Jr., T. Boller, A short C-terminal sequence is necessary and sufficient for the targeting of chitinases to the plant vacuole. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10362–10366 (1991).

2. E. van Sebille, M. Doblin, Data from "Drift in ocean currents impacts intergenerational microbial exposure to temperature." Figshare. Available at https://dx.doi.org/10.6084/m9.figshare.3178534.v2. Deposited 15 April 2016.
3. A. V. S. Hill, "HLA associations with malaria in Africa: Some implications for MHC evolution" in Molecular Evolution of the Major Histocompatibility Complex, J. Klein, D. Klein, Eds. (Springer, 1991), pp. 403–420.

1. International Emissions Trading Association, Cap And Trade: The Basics (2020) (May 6, 2020).

2. German Council of Economic Experts, "Setting Out For A New Climate Policy" (2019).

3. D. Burtraw, K. Palmer, D. Kahn, A symmetric safety valve. *Energy Policy* **38**, 4921–4932 (2010).

4. O. Edenhofer, *et al.*, "Decarbonization and EU ETS Reform: Introducing a price floor to drive low-carbon investments." (2017).

5. S. Borenstein, J. Bushnell, F. A. Wolak, M. Zaragoza-Watkins, Expecting the Unexpected: Emissions Uncertainty and Environmental Market Design. *Am. Econ. Rev.* **109**, 3953–3977 (2019).

6. B. G. Rabe, The Durability of Carbon Cap-and-Trade Policy. *Governance* **29**, 103–119 (2016).

7. J. Wettestad, T. Jevnaker, Smokescreen Politics? Ratcheting Up EU Emissions Trading in 2017. *Rev. Policy Res.* **36**, 635–659 (2019).

8. L. H. Goulder, Timing Is Everything: How Economists Can Better Address the Urgency of Stronger Climate Policy. *Rev. Environ. Econ. Policy*, 1–15 (2020).

9. P. Pierson, Increasing Returns, Path Dependence, and the Study of Politics. *Am. Polit. Sci. Rev.* **94**, 251–267 (2000).

10. D. Béland, E. Schlager, Varieties of Policy Feedback Research: Looking Backward, Moving Forward. *Policy Stud. J.* **47**, 184–205 (2019).

11. G. Skogstad, Policy feedback and self-reinforcing and self-undermining processes in EU biofuels policy. *J. Eur. Public Policy* **24**, 21–41 (2017).

12. A. M. Jacobs, R. K. Weaver, When Policies Undo Themselves: Self-Undermining Feedback as a Source of Policy Change. *Governance* **28**, 441–457 (2015).

13. N. Schmid, S. Sewerin, T. S. Schmidt, Explaining Advocacy Coalition Change with Policy Feedback. *Policy Stud. J.*, psj.12365 (2019).

14. J. D. Rubin, A Model of Intertemporal Emission Trading, Banking, and Borrowing. *J. Environ. Econ. Manage.* **31**, 269–286 (1996).

15. C. Kling, J. Rubin, Bankable permits for the control of environmental pollution. *J. Public Econ.* **64**, 101–115 (1997).

16. J. Meckling, N. Kelsey, E. Biber, J. Zysman, Winning coalitions for climate policy. *Science*

(80-.). **349**, 1170–1171 (2015).

17. J. Rozenberg, A. Vogt-Schilb, S. Hallegatte, Instrument choice and stranded assets in the transition to clean capital. *J. Environ. Econ. Manage.*, 102183 (2018).

18. B. J. Cook, Arenas of Power in Climate Change Policymaking. *Policy Stud. J.* **38**, 465–486 (2010).

19. S. W. Salant, What ails the European Union's emissions trading system? *J. Environ. Econ. Manage.* **80**, 6–19 (2016).

20. B. Steffen, F. Egli, M. Pahle, T. S. Schmidt, Navigating the Clean Energy Transition in the COVID-19 Crisis. *Joule*, 1–5 (2020).

21. T. S. Schmidt, Low-carbon investment risks and de-risking. *Nat. Clim. Chang.* **4**, 237–239 (2014).

22. T. S. Schmidt, *et al.*, Adverse effects of rising interest rates on sustainable energy transitions. *Nat. Sustain.* **2**, 879–885 (2019).

23. M. Pahle, H. Schweizerhof, Time for tough love: Towards gradual risk transfer to renewables in Germany. *Econ. Energy Environ. Policy* **5** (2016).

24. F. Egli, B. Steffen, T. S. Schmidt, A dynamic analysis of financing conditions for renewable energy technologies. *Nat. Energy* **3**, 1084–1092 (2018).

25. F. Egli, Renewable energy investment risk: An investigation of changes over time and the underlying drivers. *Energy Policy* **140**, 111428 (2020).

26. L. Kitzing, C. Mitchell, P. E. Morthorst, Renewable energy policies in Europe: Converging or diverging? *Energy Policy* **51**, 192–201 (2012).

27. L. C. Stokes, *Short Circuiting Policy: Interest Groups and the Battle Over Clean Energy and Climate Policy in the American States* (Oxford University Press, 2020).

28. S. Osorio, R. C. Pietzcker, M. Pahle, O. Edenhofer, How to deal with the risks of phasing out coal in Germany. *Energy Econ.* **87**, 104730 (2020).

29. S. Fuss, *et al.*, A Framework for Assessing the Performance of Cap-and-Trade Systems: Insights from the European Union Emissions Trading System. *Rev. Environ. Econ. Policy* **12**, 220–241 (2018).

30. P. Nahmmacher, E. Schmid, L. Hirth, B. Knopf, Carpe diem: A novel approach to select representative days for long-term power system modeling. *Energy* **112**, 430–442 (2016).

31. S. Osorio, R. C. Pietzcker, O. Tietjen, "Documentation of LIMES-EU -A long-term electricity system model for Europe" (2020).

Figures and Tables

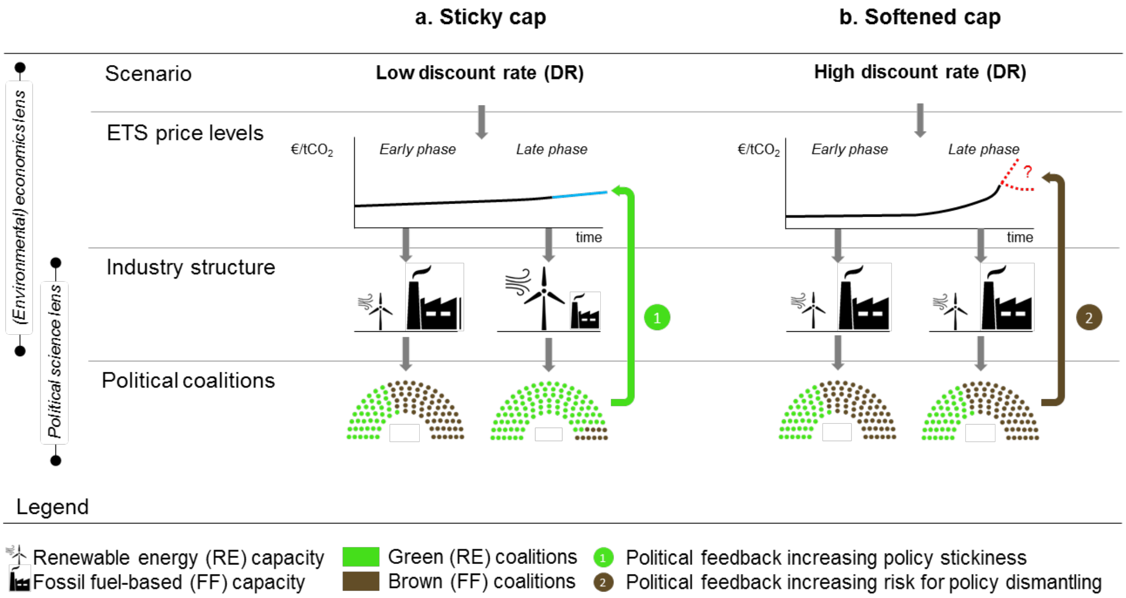


Figure 1. Dynamic policy feedback from allowance prices on politics. a) In case of a low discount rate, prices in the early phase are relatively high, inducing investment that “greens” the industry structure early on. When the late phase is reached, predominantly positive feedback ensures that the cap is sticky. b) The opposite holds for a high discount rate, where negative feedback is predominant in the late stage and softens the cap. Note that adjustment costs (not shown) are smoothly distributed over time in a) and concentrated in the late phase in b).

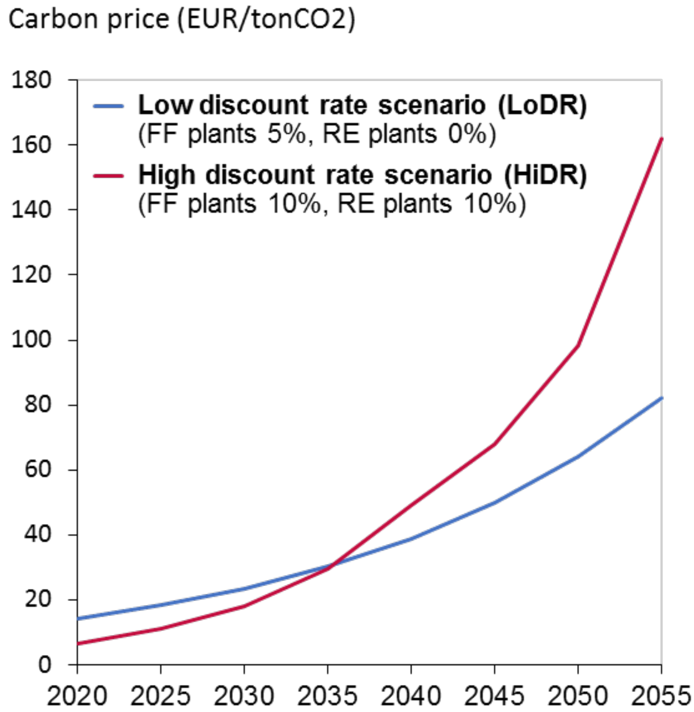


Figure 2. ETS allowance price paths in the two scenarios. Prices are shown up until the year 2055 when, according to current regulations, the ETS cap will reach zero. Price paths intersect in the year 2035, after which prices in the HiDR scenario become markedly more hockey stick-shaped.

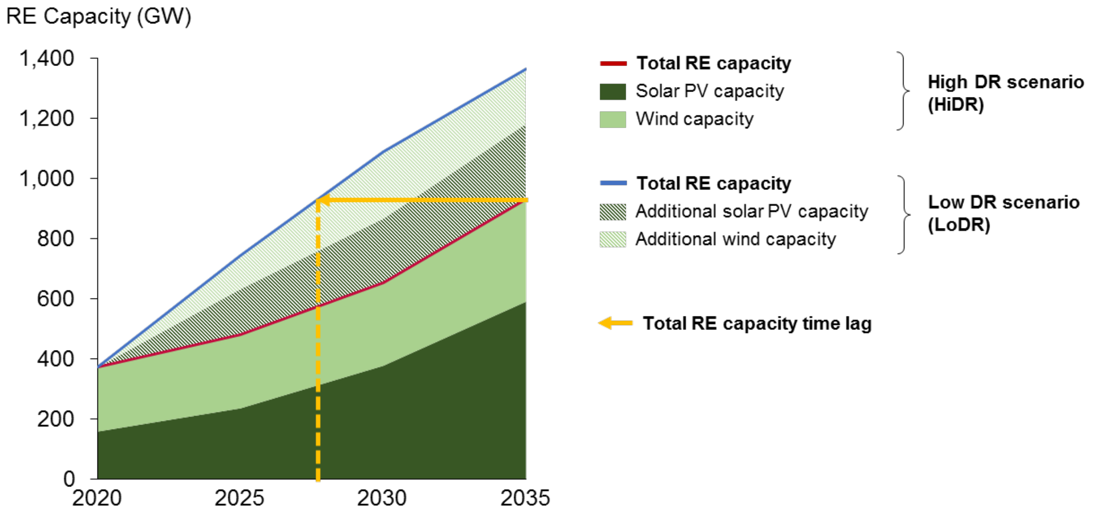


Figure 3. RE capacity deployment in both scenarios. Total RE capacity in the HiDR scenario (red line) by 2035 is equal to the level already reached eight years earlier in the LoDR scenario (blue line), as indicated with the yellow arrow. By 2035, RE capacity in the LoDR scenario is around 430 GW higher than in the HiDR scenario.

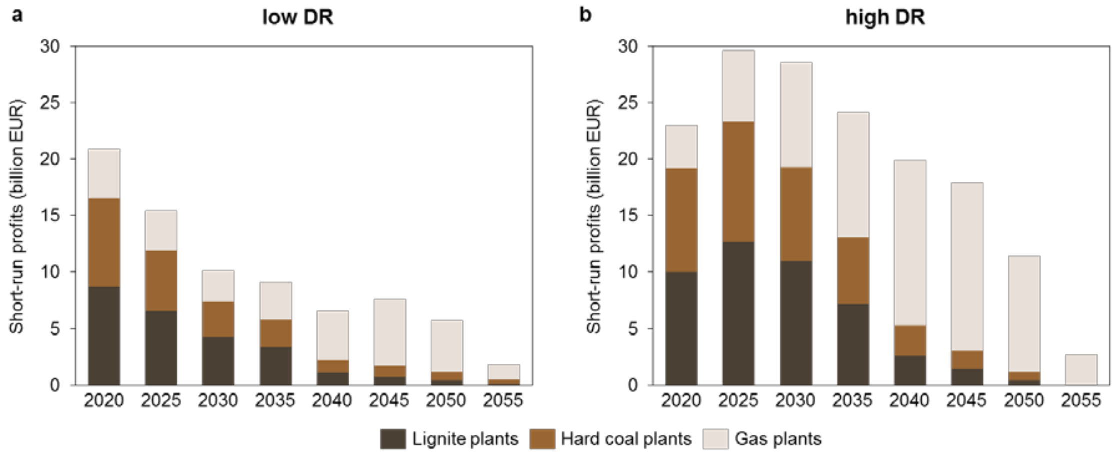


Figure 4. Short-run profits of fossil technologies in both scenarios. Profits in the LoDR scenario show a marked phase-out trajectory and are lower than in the HiDR scenario, even through 2035, when the allowance price is higher. In contrast, profits in the HiDR scenario describe a trajectory where "peak fossil" still lies ahead and is particularly prone to lobbying.

Chapter 6

Hedging and temporal permit issuances in cap-and-trade programs: the Market Stability Reserve under risk aversion[1]

Oliver Tietjen
Kai Lessmann
Michael Pahle

[1]Submitted to *Resource and Energy Economics*. The final version is available under `https://doi.org/10.1016/j.reseneeco.2020.101214`.

Hedging and temporal permit issuances in cap-and-trade programs: the Market Stability Reserve under risk aversion

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

[a] *Technical University of Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany*
[b] *Potsdam Institute for Climate Impact Research, Telegraphenberg A 31, 14473 Potsdam, Germany*

Abstract

Cap-and-trade programs such as the European Union's Emissions Trading System (EU ETS) expose firms to considerable risks, to which the firms can respond with hedging. We develop an intertemporal stochastic equilibrium model to analyze the implications of hedging by risk-averse firms. We show that the resulting time-varying risk premium depends on the size of the permit bank. Applying the model to the EU ETS, we find that hedging can lead to a U-shaped price path, because prices initially fall due to negative risk premiums and then rise as the hedging demand declines. The Market Stability Reserve (MSR) reduces the permit bank and thus, increases the hedging value of the permits. This offers an explanation for the recent price hike, but also implies that prices may decline in the future due to more negative risk premiums. In addition, we find higher permit cancellations through the MSR than previous analyses, which do not account for hedging.

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

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

*Corresponding author. e-mail: oliver.tietjen@pik-potsdam.de, phone: +49 331 288 20796

November 2020

1. Introduction

The European Union's Emission Trading System (EU ETS) is the flagship policy for the EU's aspiration to reach climate neutrality by 2050 (European Council 2019). However, observers are puzzled by the ETS permit[1] price development and question whether the EU ETS works efficiently (Ellerman et al. 2016; Friedrich et al. 2020). In particular, the permit price declined from 30 EUR/t in 2008 to well below 10 EUR/t in 2012 where the price stayed until early 2018.[2] One explanation for the price drop is lower-than-expected emissions because of, among others, the financial crisis in 2007–2009 and corresponding lower economic growth rates (Hintermann et al. 2016). According to the European Parliament and Council of the European Union (2015), the resulting "supply-demand imbalance" has destabilized the market. In response, the EU implemented the Market Stability Reserve (MSR), which has two main mechanisms: First, the issuance of permits is postponed, and they are placed in a reserve instead. Second, permits are ultimately canceled when the reserve becomes too large (European Parliament and Council of the European Union 2018). Since this mechanism was announced at the end of 2017, the price has increased to about 25 EUR/t.

However, it remains controversial whether the MSR fixes the EU ETS's problems (Flachsland et al. 2020; Gerlagh et al. 2020). For one, the EU ETS might be plagued by fundamental failures such as myopia, regulatory uncertainty and excessive discounting – all of which distort the intertemporal permit price development (Fuss et al. 2018). In this article, we also consider an intertemporal price distortion, which is affected by the temporal permit issuance: That is, the time schedule when the regulator supplies the permits to the regulated firms. In idealized cap-and-trade programs, the temporal permit issuance is irrelevant as long as permits can be freely banked between periods so that permit holders can decide when to use their permits (Salant 2016). As firms exploit intertemporal arbitrage, free banking implies that the (expected) permit price rises at

[1]In cap-and-trade programs, tradable permits allow firms to release emissions. In the EU ETS, permits are called European Union Allowances (EUAs) where one EUA permits emission of one ton of carbon dioxide equivalent.

[2]Data are publicly available, for example, at https://www.quandl.com.

the discount rate over time (Cronshaw and Kruse 1996; Rubin 1996).[3] Although banking is free in most emission trading programs, such as the EU ETS (ICAP 2020),[4] we show in this paper that the growth rate of the permit price nevertheless can be distorted by the temporal issuance of permits.

The underlying market failure in our approach is a distortion of the permit price due to hedging by risk-averse firms when the markets for risk are incomplete. Specifically, firms reduce their risk exposure (i.e., hedging) by banking permits, as the value of the banked permits (negatively) covaries with the firm profits such that the overall profits are stabilized. However, a limited number of permits are available, and as a result, hedging opportunities are constrained implying a risk premium as part of the permit price. As an alternative to hedging via permits, firms may also trade derivatives of permits (futures contracts) with financial traders (speculators), which, however, only reduces and does not eliminate the risk premium. Our analysis comprises two steps: First, we analyze theoretically how hedging affects prices in intertemporal cap-and-trade programs, such as the EU ETS, and show that the size of the permit bank becomes an important price driver. Second, we apply this theory to assess the price effects of the MSR. Analyzing the MSR is a relevant application, because it shifts permits to the future and thus, reduces the number of permits available for hedging.

Our theoretical approach regarding hedging is based on long-standing literature in financial economics that focuses on the interaction of producers and speculators in commodity (permits, in this case) futures markets (Keynes 1930; Hicks 1939; Anderson and Danthine 1979; 1981; Bessembinder and Lemmon 2002; Goldstein et al. 2014; Ekeland et al. 2019). In this hedging pressure theory, risk-averse producers reduce their profit risk exposure by trading futures contracts with speculators. The demand for futures by the producers raises the price by the risk premium, which indicates the costs of hedging for producers. Hirshleifer (1990) shows that risk premiums arise only from hedging demand

[3] This price path is known as the Hotelling price path (Hotelling 1931).

[4] In the EU ETS and many other ETS programs, banking is free, but borrowing from future periods is not allowed. However, as, for instance, in the EU ETS the actual bank levels are highly positive (European Commission 2019), the borrowing constraint does not play a large role.

in general equilibrium if there is market friction, as otherwise, speculators eliminate the risk premium through diversification. Although several frictions may cause such "limits to arbitrage" (Shleifer and Vishny 1997), we follow Acharya et al. (2013) in assuming that liquidity constraints limit speculators' risk-taking capacity. Therefore, speculators cannot fully satisfy producers' hedging demand implying a non-zero risk premium in the permit price.[5]

Several papers find empirical evidence for such risk premiums in different commodity markets (e.g., Acharya et al. 2013; Hamilton and Wu 2014; Kang et al. 2020), and in particular in the EU ETS (Pinho and Madaleno 2011; Chevallier 2013; Trück and Weron 2016). Furthermore, a survey among market participants of the EU ETS indicates that hedging is the most important motive for trading (KfW and ZEW 2016). Interviews conducted by Schopp and Neuhoff (2013) reveal that electricity producers follow risk management procedures and hold permits for hedging profits several years ahead.

Against that background, we develop a stochastic intertemporal model that comprises dirty (coal) and relatively clean (gas) firms that generate electricity and are regulated by a cap-and-trade program. Firms build up capacity stocks, which constrain electricity generation and amplify the impact of hedging. The risk premium that affects the level and growth rate of the permit price is a function of the firms' hedging demand for permits, the permit price variability and the size of the permit bank. This gives rise to a distinct intertemporal permit price profile. Initially, the dominant hedging demand of dirty coal firms creates a negative risk premium, and thus, they apply a lower discount rate compared to a risk-neutral reference firm. Over time, the market becomes cleaner, implying declining hedging demand by dirty firms. In addition, firms build up a permit bank which allows them to hedge. Thus, the risk premium becomes less negative and

[5]In addition to the hedging pressure theory, our work contains elements of the theory of storage (Kaldor 1939; Working 1949; Brennan 1958; Deaton and Laroque 1992), similar to Acharya et al. (2013) and Ekeland et al. (2019) who also combine both perspectives. The theory of storage explains the relationship between commodity spot and futures prices by the non-negative constraint for commodity inventories and storage costs (e.g., Deaton and Laroque 1992), where permit markets are a special case with negligible storage costs. Moreover, in our model, there is also a non-negative constraint for banking, but it is not our focus. However, the possibility of banking links permit prices over time, and therefore, (expected) risk premiums in the future affect current spot prices.

may turn positive. However, the price path strongly depends on the permits available for hedging purposes, which, in turn, depend on the regulator's time plan (schedule) for issuing permits.

In our stylized simulation of the EU ETS, we find a declining price until 2025–2030 and then a rising price in the counterfactual case without the MSR. Accordingly, hedging results in a U-shaped price path. The MSR amplifies the U-shape as prices are higher in early years than without the MSR, but also decline at a higher rate because the MSR reduces the permit bank level leading to a more negative risk premium. Therefore, the recently observed price hike in the EU ETS presumably due to the MSR may imply that prices in the future will rise only very slowly or even decline.

These findings stand in contrast to previous work on the originally proposed MSR without cancellation of permits (European Parliament and Council of the European Union 2015). An important result of these studies is that the temporal issuance is irrelevant as long as the overall cap remains unchanged, and banking and borrowing constraints do not bind (Salant 2016). Perino and Willner (2016) accordingly find that a cap-neutral MSR lifts the (short-term) permit price only if the borrowing constraint binds earlier due to the MSR. As long-term prices are lower, the authors also conclude that low-carbon investments with long lifetimes may decline (see also Perino and Willner 2019). We find that investments in relatively clean gas capacities are hardly affected, and investments in coal capacity significantly decline in the short-term and are slightly higher in the long-term even when the MSR is cap-neutral. This result can be traced back to worse hedging conditions for dirty firm capacity in the early years and price-level effects related to the risk premium.

Kollenberg and Taschini (2019) go a step further and relate price variability positively to the risk premium for banking permits. Because the MSR raises price variability, the MSR may even lead to lower prices in the short-term, as firms want to use more permits early 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 the risk premium. In doing so, we find the differing result that even the cap-neutral MSR raises short-term prices substantially because the hedging value of the 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 the MSR is also analyzed by Schopp and Neuhoff (2013) and Schopp et al. (2015). Their approach does not explicitly account for risk and implies inconsistent price jumps (cf. Salant 2016). We overcome these drawbacks by explicitly including a risk factor (permit supply risk) and risk aversion.[6]

Furthermore, several papers[7] analyze the cancellation mechanism of the new MSR. However, these papers assume given discount rates and ignore uncertainty. An exception is provided by Quemin and Trotignon (2019) who analyze the impact of firms' limited planning horizons and limited responsiveness to the MSR. They find a relatively high number of permanent permit withdrawals (5 to 10 Gt) compared to the literature (1.7 to 6.0 Gt) especially if the firms have a limited horizon. We also find a relatively high number of MSR cancellations (8.6 Gt) due to the negative risk premiums, which reduce the applied discount rates in the early years. Therefore, the permit bank is larger, which, in turn, leads to a larger influx in the MSR and thus, more cancellations.

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. In Section 3, we apply the model numerically to the EU ETS for multiple periods to assess the MSR regarding its price and investment effects. Finally, we discuss the results and conclude in Section 4.

2. The model

In our model, we consider firms competing in an electricity market. Emissions are a byproduct of electricity generation, and firms are heterogeneous in how clean or dirty their generation is. Emissions are limited by a cap-and-trade program, but the number of

[6] Several other papers consider risk aversion (Baldursson and von der Fehr 2004; 2012; Colla et al. 2012; Haita-Falah 2016) and ambiguity aversion (Quemin 2017) in permit markets. However, all of them have a different focus than we do.

[7] Beck and Kruse-Andersen (2018), Bocklet et al. 2019, Bruninx et al. (2018), Carlén et al. (2019), Gerlagh et al. 2019, Quemin and Trotignon (2019) and Perino and Willner (2017).

permits issued in the future is uncertain, and so are firms' profits. This creates a demand for hedging profits when firms are risk averse. Electricity generators hedge by banking permits and by trading permit futures contracts. In the futures market, we also model a speculator who serves as trading counterparty to the generators. In the following, we describe the model in detail.

2.1. General model setup

We consider N competitive firms, indexed by i, that produce a homogeneous and non-storable good (electricity) x_{it} at T periods, indexed by t. Demand is given by $D(w_t)$ with $D^{'} < 0$ and price w_t. The equilibrium condition,

$$\sum\nolimits_{i}^{N} x_{it} = D(w_t), \tag{2.1}$$

is always fulfilled. Firms use production technologies that differ in emission intensity and in how costly they are to install and operate (capacity and production costs), for example, coal and gas plants. We model production costs as a function $C_{Xi}(x_{it})$ with $C_{Xi}^{'} > 0$. To produce x_{it} units, firms also need at least k_{it} units of capacity, for which the capacity costs are given by $C_{Ki}(k_{it})$ with $C_{Ki}^{'} > 0$. We assume that the production and capacity costs are separable, which is a standard assumption in electricity market modeling (Stoft 2002). Defining $\zeta_{it} \equiv \frac{x_{it}}{k_{it}}$ as the capacity utilization rate, production is constrained by

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

Although the utilization rates can be immediately adjusted within a period, investments in (plant) capacity, $I_{Kit} \geq 0$, are added to the existing capacity stock with a lag of one period,

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

where δ is the rate of depreciation.

The firm-specific emission intensity is captured by a time-invariant emission factor ϕ_i;

i.e., the production of each unit of x_{it} causes ϕ_i units of emissions.[8] The heterogeneity in emission factors ϕ_i is important for our analysis, because permit supply uncertainty affects dirty coal firms (high ϕ_i) differently from relatively clean gas firms (low ϕ_i). Overall emissions are capped, because emissions are regulated by a cap-and-trade program. To comply with the regulations, firms need at least as many permits as emissions $x_{it}\phi_i$ at the end of each period t. At the beginning of each period, S_t permits are auctioned[9] by the regulator at price p_t that clears the permit market, that is,

$$S_t = \sum_i^N y_{it}, \tag{2.4}$$

where y_{it} is the number of purchased permits. Uncertainty enters the model through permit supply risk in the following way: Initially, the regulator announces a permit auction schedule for the entire lifetime of the cap-and-trade program beginning in the first period and ending in the last period $t = T$. However, in each period the regulator may deviate from her previous announcement and in addition, may announce a new permit supply schedule for future periods τ, $S_\tau \; \forall \, \tau > t$. Thus, uncertainty about the permit supply in period t is resolved at the beginning of t, but the supply in future periods τ remains uncertain. Therefore, in firm expectation the overall supply, or cap, in any period t is

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

Furthermore, a reserve mechanism similar to the MSR in the EU ETS affects the permit supply as well. We explain and implement the MSR in detail in the numerical simulation

[8] Constant and time-invariant emission factors are standard assumptions for electricity plants because each unit of fossil fuel (coal, gas) leads to the same amount of emissions and electricity. We ignore technological progress which could improve the conversion efficiency from fossil fuel to electricity. Given the maturity of fossil fuel plants, this is a mild assumption.

[9] Throughout the paper, we assume that the initial allocation of permits is through auctioning; that is, there is no free allocation. Although the allocation method, in general, can affect market outcomes (e.g., Böhringer and Lange 2005), and in particular, if firms are risk averse (Baldursson and von der Fehr 2004; 2012), we omit this to streamline the analysis. In this paper, we focus on the EU ETS and on the electricity sector in the program, where, in principle, all permits are auctioned (European Parliament and Council of the European Union 2018).

in Section 3, but we consider a stylized representation in this analytical part. Recall that the MSR has two effects: First, it shifts the permit supply to the future, and second, it reduces the overall cap by the cancellation of permits. We separate these two mechanisms and model the permit shift as

$$\Delta E_t \left[\bar{S}\right] = \Delta S_t + \sum\nolimits_{\tau>1}^{T} \Delta E \left[S_\tau\right] = 0. \tag{2.6}$$

That is, any change in permit supply ΔS_t in any period is fully compensated by the (announced) supply in other periods such that the total expected cap is always the same from the perspective of period t. Thus, only the temporal permit issuance is affected which corresponds to the first mechanism of the MSR. In addition, the cancellation mechanism leads to an overall lower supply if too many permits are in the reserve. Below, we discuss how hedging affects the number of canceled permits, and we model the entire MSR explicitly in the numerical simulation.

If firms buy more permits y_{it} than they have emissions $x_{it}\phi_i = k_{it}\zeta_{it}\phi_i$, additional permits can be transferred to the next period (banking). Let b_{it} be the banked permits at the end of period t. Then, the dynamic banking constraint is

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

while borrowing from the future is not allowed:

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

Moreover, firms can also trade futures contracts on permits denoted by f_{it}. We consider only futures contracts that expire in the next period. That is, the buyer of $f_{it} > 0$ units of futures bought at price p_t^f in period t receives $f_{it}p_{t+1}$ in period $t+1$, and the seller ($f_{it} < 0$) receives $f_{it}p_t^f$ in period t and has to pay $f_{it}p_{t+1}$ in period $t+1$. For both, the expected payoff of the futures is $\left(E\left[p_{t+1}\right] - p_t^f\right) f_{it}$. As further shown below, futures are a hedging instrument for the electricity-generating firms because the futures' payoff negatively covaries with the plant profits. In addition, a representative speculator is active in the futures market who seeks to gain speculative profits given (in expectation) by the futures' payoff $\left(E\left[p_{t+1}\right]-p_t^f\right) f_{sp,t}$ where $f_{sp,t}$ is the number of futures bought ($f_{sp,t} > 0$) or sold ($f_{sp,t} < 0$) by the speculator. In the futures market equilibrium, positive and negative positions must be balanced:

$$\sum\nolimits_i^N f_{it} + f_{sp,t} = 0. \tag{2.9}$$

Furthermore, firms invest in a risk-free asset stock l_{it}, providing a safe return r. This serves as an alternative investment opportunity, allowing for risk-free allocation of wealth over time. Denoting investments in the risk-free asset as I_{Lit}, the risk-free asset stock is

$$l_{it} = (1+r)\, l_{it-1} + I_{Lit}. \tag{2.10}$$

Given this setup, the profits of the electricity-generating firms in period t are

$$\pi_{it} = w_t k_{it} \zeta_{it} - C_{Xi}\left(k_{it}\zeta_{it}\right) - C_{Ki}\left(k_{it}\right) - p_t y_{it} - I_{Lit} - p_t^f f_{it} + p_t f_{it-1}, \tag{2.11}$$

where $w_t k_{it} \zeta_{it}$ describes the revenue for selling electricity, $C_{Xi}\left(k_{it}\zeta_{it}\right)$ and $C_{Ki}\left(k_{it}\right)$ are costs for producing and for plant capacities,[10] respectively. The terms $p_t y_{it}$ and I_{Lit} are costs (> 0) or revenues (< 0) for trading permits and the risk-free asset, respectively. The term $p_t^f f_{it}$ denotes investments in futures contracts, and the term $p_t f_{it-1}$ reflects profits from futures contracts invested in the previous period. We further assume that firms have concave preferences regarding profits described by a von Neumann–Morgenstern utility function $U_{it}\left(\pi_{it}\right)$ with $U_{it}' > 0$ and $U_{it}'' < 0$. This implies that firms have a preference for a more stable profit, meaning they behave in a risk-averse manner which causes the desire to hedge.[11] The problem of the electricity-generating firms is

[10]Note that we assume for simplicity that there are no costs for investing in plant capacity I_{Kit}. Instead, investment costs are allocated to the capacity costs.

[11]There are several reasons why firms behave as if they are risk averse (Froot et al. 1993; Acharya

$$\max_{\zeta_{it},y_{it},I_{Kit},I_{Lit},f_{it}} \sum_{t=1}^{T} \frac{1}{(1+r)^{t-1}} E\left[U_{it}\left(\pi_{it}\right)\right] \tag{2.12}$$

subject to

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

$$k_{it} = (1-\delta)\, k_{it-1} + I_{Kit-1} \tag{2.13}$$

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

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

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

$$\pi_{it} = \pi_{it}^{plant} + p_t\left(b_{it-1} - b_{it}\right) - I_{Lit} - p_t^f f_{it} + p_t f_{it-1}, \tag{2.14}$$

with $\pi_{it}^{plant} = w_t k_{it}\zeta_{it} - C_{Xi}\left(k_{it}\zeta_{it}\right) - C_{Ki}\left(k_{it}\right) - p_t k_{it}\zeta_{it}\phi_i$.

The speculator is not active in the electricity market and trades only the risk-free asset and futures contracts. The speculator's profits are

$$\pi_{sp,t} = -I_{L,sp,t} - p_t^f f_{sp,t} + p_t f_{sp,t-1}, \tag{2.15}$$

and the speculator's maximization problem is

$$\max_{I_{L,sp,t},f_{sp,t}} \sum_{t=1}^{T} \frac{1}{(1+r)^{t-1}} E\left[U_{sp,t}\left(\pi_{sp,t}\right)\right] \tag{2.16}$$

subject to

$$l_{sp,t} = (1+r)\, l_{sp,t-1} + I_{L,sp,t}. \tag{2.17}$$

et al. 2013): for example, costs associated with financial distress or principal agent issues that result in higher utility from more stable profit.

Similar to the electricity generators, the speculator evaluates profits based on a concave function $U_{sp,t}(\pi_{sp,t})$ with $U'_{sp,t} > 0$ and $U''_{sp,t} < 0$. In doing so, we follow the financial economics literature (Acharya et al. 2013), by interpreting this as capital constraint. For instance, the constraint could be due to value-at-risk (VaR) limits, and thus, taking risky positions is constrained. Therefore, even if the speculator is risk neutral, she behaves in a risk-averse manner.

2.2. Two-period model

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 electricity demand is linear, $D(w_t) = A - aw_t$, the firms' production costs are quadratic $C_{Xi}(x_{it}) = \frac{\beta_i}{2}x_{it}^2$ and their capacity costs are linear $C_{Ki}(k_{it}) = \gamma_i k_{it}$.[12] There are only two firms, $i = c, d$, a relatively clean gas firm and a dirty coal firm with $\phi_d > \phi_c$. Moreover, to arrive at closed-form results we assume a quadratic utility function in some cases, for both electricity generators $U_{it}(\pi_{it}) = \pi_{it} - \pi_{it}^2$ and the speculator $U_{sp,t}(\pi_{sp,t}) = \pi_{sp,t} - \pi_{sp,t}^2$.[13] 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. Period 2 equilibrium

We solve the model backward and start in period 2. Note that all derivations can be found in Appendix A.

As period 2 is the final period, no further investments in plant capacity and futures contracts are made, $I_{Ki,2} = f_{i,2} = f_{sp,2} = 0$, and all available permits are used or sold

[12] These assumptions are motivated by electricity sectors: First, marginal production costs increase with production implying an upward sloping electricity supply curve as with the merit-order curve in electricity markets. Second, capacity costs typically exhibit constant marginal costs per unit of capacity (e.g., costs for coal plant capacity do not increase with the number of installed plants). However, the specific functional forms have only minor relevance for the analytical results as long as one firm type has costs which imply that the firm benefits from a higher permit price (in terms of higher profits), while another firm type has costs so that it loses from a higher price (and vice versa). We show below that this is the case given these assumptions.

[13] Assuming quadratic utility is akin to mean-variance optimization which has a long tradition in financial economics, and especially in the hedging pressure literature (Anderson and Danthine 1979; Ekeland et al. 2019).

(assuming a strictly positive permit price p_2). Similarly, the stock of the risk-free asset is depleted, implying $b_{i,2} = l_{i,2} = l_{sp,2} = 0$. As a result, the speculator has no decision to make. Uncertainty has been resolved, and the generating firms' problem, thus, is to maximize $U_{i,2}(\pi_{i,2})$ over $\zeta_{i,2}$ and $y_{i,2}$ subject to the constraints in (2.13). Taking the first-order conditions (see Appendix A.1), the utilization rate and the permit purchases can be written as

$$\zeta_{i,2} = \frac{w_2 - p_2\phi_i}{\beta_i k_{i,2}} - \frac{\mu_{i,2}}{U'_{i,2}\beta_i k_{i,2}^2}, \tag{2.18}$$

$$y_{i,2} = \phi_i \left(\frac{w_2 - p_2\phi_i}{\beta_i} - \frac{\mu_{i,2}}{U'_{i,2}\beta_i k_{i,2}} \right) - b_{i,1}, \tag{2.19}$$

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:

$$\mu_{i,2} = \begin{cases} U'_{i,2}\left(k_{i,2}w_2 - \beta_i k_{i,2}^2 - p_2\phi_i k_{i,2}\right) & if\ \zeta_{i,2} = 1 \\ 0 & if\ 1 \geq \zeta_{i,2} \geq 0 \end{cases}. \tag{2.20}$$

The shadow value $\mu_{i,2}$ indicates the scarcity of capacity k_{it}, which cannot be increased within a period due to the time lag for investments. For (2.19), 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 in period 2 (in (2.18), either $\mu_{i,2} = 0$ or $U'_{i,2}$ is canceled out due to (2.20)). It adjusts only the shadow value of the capacity which, however, triggers no changes in the firm behavior, because the firm cannot change its capacity level within a period.

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

$$w_2 = \frac{1}{(\beta_d + \beta_c + \beta_c\beta_d a)} \left(A\beta_c\beta_d + p_2(\beta_d\phi_c + \beta_c\phi_d) + \beta_d\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}} + \beta_c\frac{\mu_{d,2}}{U'_{d,2}k_{d,2}} \right). \tag{2.21}$$

Similarly, the permit price can be derived from using (2.19) in the permit equilibrium condition, $S_t = \sum_i^N y_{it}$. By additionally considering (2.21), we get:

$$\begin{aligned} p_2 &= \frac{A\left(\beta_d\phi_c + \beta_c\phi_d\right) - \left(\beta_d + \beta_c + \beta_c\beta_d a\right)\left(b_{c,1} + b_{d,1} + S_2\right)}{\left(\phi_c - \phi_d\right)^2 + a\left(\beta_c\phi_d^2 + \beta_d\phi_c^2\right)} \\ &+ \frac{\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}}\left(\phi_d - \phi_c\left(1 + \beta_d a\right)\right) + \frac{\mu_{d,2}}{U'_{d,2}k_{d,2}}\left(\phi_c - \phi_d\left(1 + \beta_c a\right)\right)}{\left(\phi_c - \phi_d\right)^2 + a\left(\beta_c\phi_d^2 + \beta_d\phi_c^2\right)}. \end{aligned} \tag{2.22}$$

Intuitively, the electricity price is a positive function of demand, reflected by A (the intercept of the demand function), and the permit price p_2. Shocks in the permit price, therefore, are transferred to consumers via the electricity price. The only source of uncertainty (from the perspective of period 1) is the permit supply in period 2, S_2.

To examine the effect of a permit supply shock, we assume for a moment that the plant capacity constraints (reflected by the last two terms in (2.21) and (2.22)) do not bind, $\mu_{i,2} = 0$. In this case, a positive shock on S_2 (a less ambitious policy) leads to a lower permit price and vice versa, as can be seen directly from (2.22). 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) higher capacity utilization by the clean firm, $\frac{d\zeta_{c,2}}{dS_2} > 0$, iff $\phi_d > \phi_c\left(1 + \beta_d a\right)$ holds, and (2) lower capacity utilization by the dirty firm, $\frac{d\zeta_{d,2}}{dS_2} < 0$. For a negative permit price shock, the opposite holds.*

While the dirty firm always produces less when the permit price increases and vice versa, for the clean firm, it depends on the 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 electricity market is strong enough, reflected by a high a, or the clean firm is not clean enough (i.e., ϕ_c is too large) such that the inequality is violated, the clean firm produces more if the permit price is low. 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, which reflects the fuel switch in electricity markets. A higher permit price leads to less coal (dirty) and more gas (clean) production.[14] Note that the assumption of non-binding

[14]The fuel switch from coal to gas plants is one of the most important abatement options in the EU

capacity constraints is innocuous for this result. For one, capacity constraints would not switch the sign of the effect on the utilization rates, but instead, restrict the effect size, as the constraints limit or even prevent how firms change their production after a shock. Moreover, typically capacity constraints do not bind in expectation in electricity markets, as power plants are not always fully utilized. Demand varies on a short time scale, and plants have to be ramped up and down. In this sense, the utilization rate in the 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 opposite 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, the clean firm 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.

Although 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 sensitivity to permit supply shocks, which we consider as the measure of price variability. Specifically, it can be shown (see Appendix A.2) that if electricity generation is constrained by plant capacity, the permit price variability is higher. The intuition is that capacity partly locks in 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 the capacity constraints bind. To comply with the cap, the permit price must rise to a higher level than without capacity constraints, because abate-

ETS when the permit price is above approximately 20-30 EUR/t (depending on coal and gas prices; Friedrich et al. 2020). As the permit price has been above 20 EUR/t since 2019 (with the exception of a short period due to the COVID-19 shock), modeling the fuel switch is very relevant.

ment is achieved with more expansive technologies (i.e., via lower electricity demand in the model). As a result, the higher price variability also leads to higher profit variability, and thus, it amplifies the effect of hedging which we analyze further below.

2.2.2. Period 1 equilibrium

In period 1, electricity-generating firms have to make decisions under uncertainty about the permit supply by maximizing utility in (2.12) for $T = 2$ and subject to (2.13). While the capacity utilization rate $\zeta_{i,1}$ and the permits trades $y_{i,1}$ must fulfill the same condition as in period 2, generators additionally decide about the optimal permit bank level $b_{i,1}$, capacity level for period 2 $k_{i,2}$[15] and the amount invested in the risk-free asset $l_{i,1}$ and futures contracts $f_{i,1}$. In addition, the speculator maximizes (2.16) subject to (2.17) via investments in the risk-free asset $l_{sp,1}$ and futures contracts $f_{sp,1}$ (see Appendix A.1 for all first-order conditions). First, we analyze how the generators hedge via the permit bank, and how this affects the permit price while we ignore futures markets and plant capacities. Then, we add the futures market to the analysis and show that its main effect is to reduce risk premiums. In the last two parts of this section, we examine the capacity effects and discuss the impact of the MSR.

Banking and hedging. The number of permits firms buy is equal to their period 1 emissions plus the desired bank at the end of period 1,

$$y_{i,1} = \phi_i \zeta_{i,1} k_{i,1} + b_{i,1}, \tag{2.23}$$

where the banking demand can be written as follows:

$$b_{i,1} = \frac{E\left[p_2\right] - p_1\left(1+r\right)}{\lambda_i Var\left[p_2\right]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var\left[p_2\right]} - \frac{\left(1+r\right)\varphi_{i,1}}{U'_{i,1} Var\left[p_2\right]}, \tag{2.24}$$

for which we assume quadratic utility with $\lambda_i = -\frac{U''_{i,1}}{U'_{i,1}}$ as the coefficient for absolute risk aversion. The third term on the right side in (2.24) includes the shadow price of

[15]Due to the time lag for investments in capacity, we assume for simplicity that there are sufficient initial capacities $k_{i,1}$ such that the capacity constraints do not bind in period 1.

the borrowing constraint $\varphi_{i,1}$ (due to inequality (2.8)), which is positive if firms want to borrow ($b_{i,1} < 0$) 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 to hold a positive bank for purely speculative reasons and vice versa. The second term is the hedging demand, determined by the covariance of plant profits with the period 2 permit price. It reflects the number of permits that firms want to bank to reduce their risk exposure. For this hedging demand, we have the following proposition.

Proposition 1. *For a pure hedging purpose, the dirty firm wants to hold a positive number of permits $b_{d,1} > 0$ (banking), and the clean firm holds no permits, $b_{c,1} = 0$. The clean firm holds a positive number of permits only 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 electricity market; for clean firms, the opposite holds (see Lemma 2). This is reflected by the hedging demand, the second term in Equation (2.24), 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, because we assume that borrowing is not allowed, clean firms cannot hedge their electricity 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 bank, because the expected profit for banking compensates for the higher risk exposure due to banking.

The implications of the hedging demand for the permit price can be analyzed by decomposing the price dynamics into three parts:

$$\frac{E[p_2] - p_1}{p_1} = r + \frac{(1+r)\varphi_{i,1}}{p_1 E\left[U'_{i,2}\right]} + q_1. \tag{2.25}$$

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 present only 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 in period $t = 1$ which emerges endogenously due to the firms' hedging demand. With a general utility function, it is $q_1 = -\frac{Cov\left[U'_{i,2}, p_2\right]}{E\left[U'_{i,2}\right]p_1}$, and thus, it depends on the firms' risk preferences, reflected by the marginal utility $U'_{i,t}$ 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:

$$q_1 = \frac{\Lambda}{p_1}\left(Cov\left[\pi^{plant}_{d,2}, p_2\right] + Cov\left[\pi^{plant}_{c,2}, p_2\right] + Var\left[p_2\right]B_1\right), \tag{2.26}$$

where $B_1 = b_{d,1} + b_{c,1}$ is the total bank, and $\Lambda \geq 0$ is a parameter that reflects the risk-taking capacity of the market. The risk-taking capacity if both firms bank is $\Lambda = \left(\lambda_d^{-1} + \lambda_c^{-1}\right)^{-1}$. If only one firm banks, the risk-taking capacity is $\Lambda = \lambda_i$ (recall that λ_i is the coefficient of absolute risk aversion). A large Λ implies a low risk-taking capacity, and in the case of risk neutrality, $\Lambda = 0$, the risk-taking capacity is infinitely large, and the risk premium would disappear, $q_1 = 0$. Equation (2.26) further shows that the price variability, $Var\left[p_2\right]$, has a positive effect on the risk premium, because price variability increases the risk of permit banking, and thus, firms require a higher return for banking. Similarly, a higher overall bank, B_1, in isolation increases the volume of risky permits for which firms require a larger risk premium. In contrast, the hedging demand may have a positive or negative effect on the risk premium. The clean firm's hedging demand increases, and the dirty firm's hedging demand decreases, the risk premium, because $Cov\left[\pi^{plant}_{c,2}, p_2\right] > 0$, and $Cov\left[\pi^{plant}_{d,2}, p_2\right] < 0$ (see Lemma 2).

However, the clean firm banks only if the risk premium is positive (cf. Proposition 1), and thus, the sign of the risk premium depends only on the strength of the dirty firm's hedging demand and the risk of banking permits $\left|Cov\left[\pi^{plant}_{d,2}, p_2\right]\right| \lesseqqgtr Var\left[p_2\right]b_{d,1}$. If the former exceeds the latter, the risk premium is negative. This is because banking has the additional benefit of 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 are too high so 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 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.*

In the absence of capacity constraints, a positive risk premium always leads to a lower price and higher emissions in period 1, and a higher (expected) price and lower emissions in period 2. By rewriting (2.25) as $p_1 = \frac{E[p_2]}{(1+r+q_1)}$, it becomes obvious that the risk premium has the same effect as the risk-free rate. Thus, a positive risk premium increases the applied discount rate and leads to a steeper price path, and vice versa. The size of the risk premium hinges on the risk-taking capacity of the market reflected by Λ (see Equation (2.26)). Next, we show how futures markets reduce risk premiums by increasing the risk-taking capacity (lower Λ).

The effect of futures markets. In this section, we add the futures market and speculators to the model as described in Section 2.1. Assuming quadratic utility, and maximizing (2.16) via $f_{sp,1}$ subject to (2.17), yields the speculator's futures trades:

$$f_{sp,1} = \frac{E\left[p_2\right] - p_1^f\left(1+r\right)}{\lambda_{sp} Var\left[p_2\right]}. \tag{2.27}$$

The coefficient of absolute risk aversion $\lambda_{sp} \geq 0$ reflects the severity of the speculator's capital constraint (cf. Acharya et al. 2013). If $\lambda_{sp} = 0$, the constraint does not bind, and the speculator can fully exploit intertemporal arbitrage implying that she invests in the futures market until it holds $E\left[p_2\right] = p_1^f\left(1+r\right)$. If $\lambda_{sp} \to \infty$, the speculator has no liquid funds to invest in the futures market implying $f_{sp,1} = 0$. For $0 < \lambda_{sp} < \infty$, the speculator increases the funds invested in the futures market with the expected profit for this investment, $E\left[p_2\right] - p_1^f\left(1+r\right)$.

Dirty and clean firms maximize (2.12) via $f_{i,1}$ subject to (2.13), which gives the demand for futures:

$$f_{i,1} = \frac{E[p_2] - p_1^f (1+r)}{\lambda_i Var[p_2]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var[p_2]} - b_{i,1}. \tag{2.28}$$

The expression shows that an increase in the permit bank $b_{i,1}$, reduces the demand for futures by the same number (all else equal). The reason is that buying a permit instead of a futures contract is a perfect substitute in terms of hedging: Buying one permit or one futures contract in $t = 1$ both yields the same random profit p_2 in $t = 2$ implying that they have the same hedging effect reflected by $Cov\left[\pi_{i,2}^{plant}, p_2\right]$. As long as there is a positive bank in equilibrium (i.e., the borrowing constraint does not bind, $\varphi_{i,1} = 0$), the permit price and the futures price must be equal, $p_1 = p_1^f$, due to arbitrage. This can be seen by using (2.28) and (2.27) in the equilibrium condition of the futures market (2.9) to derive the futures price:

$$p_1^f = \frac{E[p_2]}{(1+r)} - \frac{\Lambda^f}{(1+r)} \left(Cov\left[\pi_{d,2}^{plant}, p_2\right] + Cov\left[\pi_{c,2}^{plant}, p_2\right] + Var[p_2] B_1\right), \tag{2.29}$$

where Λ^f is the risk-taking capacity of the market if a futures market exists, as opposed to Λ without a futures market. Similarly, the permit price p_1 can be derived by using the demand for permits (2.23) in the ETS market equilibrium (2.4) (see Appendix A.3.1), which yields the same expression (if the borrowing constraint does not bind) implying $p_1 = p_1^f$. However, a difference from permits is that futures allow the clean firm to hedge as well, because short positions ($f_{it} < 0$) are possible, which is not allowed in the permit market due to the borrowing constraint ($b_{it} \geq 0$).

The main implication of the futures market is that the speculator increases the risk-taking capacity because $\Lambda^f = \left(\lambda_d^{-1} + \lambda_c^{-1} + \lambda_{sp}^{-1}\right)^{-1} < \Lambda = \left(\lambda_d^{-1} + \lambda_c^{-1}\right)^{-1}$ holds. That is, if a speculator is active in the futures market, the risk premium becomes smaller as can be shown by replacing Λ by Λ^f in the expression for the risk premium (2.26). The strength of this effect depends on the speculator's capital constraint: If the constraint does not bind ($\lambda_{sp} = 0$), the speculator eliminates the risk premium, and if the constraint is too binding ($\lambda_{sp} \to \infty$,), the risk-taking capacity becomes $\Lambda = \left(\lambda_d^{-1} + \lambda_c^{-1}\right)^{-1}$, because the speculator does not trade futures.

Capacity effects. In this section, we look at the effect of plant capacity, which we have ignored thus far. Optimal capacity investments can be decomposed into three parts:

$$k_i \quad = \quad \frac{1}{\gamma_i}\left(E\left[\zeta_{i,2}\right]E\left[\mu_{i,2}^{RN}\right] + Cov\left[\zeta_{i,2}, \mu_{i,2}^{RN}\right] + \frac{1}{U'_{i,1}}Cov\left[U'_{i,2}, \mu_{i,2}^{RN}\right]\right), \qquad (2.30)$$

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 (2.20)). The first two terms on the right side in (2.30) reflect the 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 $Cov\left[\zeta_{i,2}, \mu_{i,2}^{RN}\right]$. Because $Cov\left[\zeta_{i,2}, \mu_{i,2}^{RN}\right]$ 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. Thus, $\mu_{i,2}^{RN}$ is bounded at zero but has no upper bound. Therefore, capacity constraints induce an asymmetric impact of symmetric shocks if the shocks are large enough. This leads to higher expected profits reflected by a higher capacity value implying more investments in capacity.

The third term represents the effect of risk aversion. If firms do not bank permits, then $Cov\left[U'_{i,2}, \mu_{i,2}^{RN}\right] \leq 0$ holds, and thus, risk aversion has a negative impact on investments, ceteris paribus. This is intuitive, as capacity investments are risky, and firms are risk averse. However, the effect of banking permits on $Cov\left[U'_{i,2}, \mu_{i,2}^{RN}\right]$ is positive for the dirty firm and negative for the clean firm, and we have the following result.

Proposition 3. *Banking has a positive effect on investments in dirty capacity and a negative effect on investments in clean capacity, ceteris paribus.*

The intuition is that banking hedges dirty plant profits, but increases the risk for clean firms, and the investment incentives change accordingly. For hedging purposes, clean plants require a futures market that allows them to take short positions akin to permit borrowing as explained in the previous section.

The impact of the MSR. Next, we consider the effect of the MSR on the permit price path and investments in plant capacities. We analyze the effect of shifting permits to the future before we discuss permit cancellations.

We model shifts of permits to the future with a cap-neutral permit reallocation in the sense of Equation (2.6). Issuing more permits in period 2, rather than in period 1, reduces the permit bank of all firms with a positive bank at the end of period 1. By using the first-order conditions, we get the following relation between 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 and exploiting the concavity of the utility function ($U'_{i,1} > 0$ and $U''_{i,2} < 0$) yields

$$\frac{\partial p_1}{\partial b_{i,1}} = \frac{E\left[U''_{i,2}p_2^2\right]U'_{i,1} + E\left[U'_{i,2}p_2\right]U''_{i,1}p_1}{U'^2_{i,1}} < 0. \tag{2.31}$$

Thus, if the bank volume decreases, p_1 increases. This is because firms require a lower return for holding fewer permits (lower risk premium, see (2.26)), which is achieved with a higher price in period 1. Intuitively, a higher permit price in period 1 leads to less emissions in period 1. If the total (expected) number of permits is given, this implies that the expected emissions in period 2 must increase, and in turn, the expected period 2 permit price must decline. We summarize this in the following proposition.

Proposition 4. *A temporal reallocation of permits by the regulator to period 2 in the sense of Equation* (2.6) *such that bank B_1 decreases leads to a higher permit price and lower emissions in period 1 and a lower expected permit price and higher expected emissions in period 2, and vice versa.*

Thus, the regulator's decision about the temporal issuance of permits has real production effects even if the borrowing constraint is not affected.[16] The reason is that it matters who owns the permits: If firms bank permits in private accounts, the firms bear the risk of a changing permit price. However, this also allows firms to hedge their profits

[16]In the risk-neutral case, the shift of permits to the future affects the price only via the borrowing constraint (Perino and Willner 2016). We exclude the effect of the borrowing constraint as it never binds before all permits are used up after the second period. However, we account for this effect in the numerical simulation in the next section.

by exploiting the covariance of the permit price and plant profits (see above). In contrast, if the permits are issued later and are transferred into the MSR instead, the firms cannot use the permits for hedging purposes. Thus, if not enough permits are available for hedging purpose, dirty firms are willing to pay for holding a bank (negative risk premium) to reduce their risk exposure. If instead, too many permits are available, the firms require a positive risk premium for holding permits. These hedging or risk costs are incorporated into the permit price, such that firms emit less in the first period if the number of permits available is reduced through the shifting mechanism of the MSR.

The implications for the investment incentives in dirty and clean capacity are ambiguous. On one hand, a lower expected permit price in the future increases (decreases) investments in dirty (clean) capacity. On the other, a lower permit bank level raises the costs of hedging dirty plants (see Proposition 3) implying weaker incentives to invest in them.

The second mechanism of the MSR cancels permits if too many of them are stored in the reserve. The main effect of this measure is that the overall cap is reduced such that the entire price path is lifted upward. As the number of permits in the reserve depends on the size of the bank B_1, ultimately, the number of canceled permits depends on B_1. Compared to the risk-neutral reference case, hedging may increase or decrease the bank, and thus, cancellations: If the hedging demand of dirty firms outweighs the available permits and the hedging demand of clean firms, the bank is larger due to hedging and vice versa (see above). We analyze the implications of hedging for permit cancellations in more detail in the following section.

3. Numerical application to the EU ETS

In this section, we apply the model to the EU ETS to (1) demonstrate the impact of hedging in a multi-period setting and (2) assess the effects of the explicitly implemented MSR rather than the stylized MSR version of the previous section. As a reminder, the MSR was introduced to stabilize the permit price on a higher level and spur cleaner investments. However, as the model is highly stylized, the numerical outcomes should be interpreted as qualitative results rather than numerical estimates. In the following section, we explain the model implementation and important assumptions.

3.1. Model implementation

The main sectors of the EU ETS are the electricity sector and the energy-intensive industry. However, we explicitly consider only the electricity sector for which dirty (coal) and relatively clean (gas) plants can be identified, and hedging behavior is also observed in practice (Schopp and Neuhoff 2013). That is, we solve the firms' problems given by Equations (2.12) and (2.13) for $i = c, d$, a representative gas and coal firm. In principle, the analysis carried out in this paper should also hold for firms in other sectors because the permit price affects their profits in a similar way.

We focus on the time period between 2018 and 2057, but solve the model until 2102 to set investment incentives beyond 2057. 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 five-year period and $y = 2018, 2019, ...2102$ for every year.

Due to the more detailed approach of modeling the MSR compared to the analytical section, we adapt the notation slightly. At the beginning of the first year, the regulator announces it will issue $\hat{S}_y$ permits each year (for $\hat{S}_t$, we take the average of the respective five years). The parameter $\hat{S}_y$ corresponds to the (announced) permits to be auctioned in the EU ETS between 2018 and 2057, with a linear reduction factor[17] of 1.74% until 2020 and then 2.2% (European Parliament and Council of the European Union 2018), which implies that the last permits are issued in 2057. In line with current regulation, we assume that permits issued to the electricity sector are auctioned, while the auction share of all issued permits is 57% (European Parliament and Council of the European Union 2018). The remaining 43% of the permits are freely allocated to the other ETS sectors. We assume that the freely allocated permits cover the emissions from these sectors such that the expected net permit demand of these other sectors is zero.

We consider an additive shock θ_t to the permit supply such that the permits available

[17]The linear reduction factor determines by how much the annually issued permits are reduced (excluding the impact of the MSR).

to the electricity sector (excluding the MSR effects) are $S_t = \hat{S}_t + \theta_t$. Regulatory supply uncertainty is one rationale for this shock. As an alternative interpretation, the shock θ_t may also include the uncertain permit demand of the other ETS sectors, such that S_t would reflect the permit supply net of the other sectors' demand. Specifically, we assume the following shock process:

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

with $\epsilon_t \in \{-0.35\hat{S}_t, 0.35\hat{S}_t\}$ where the positive and negative shocks have the same probability.[18] Lacking real-world guidance, we assume that the shocks are a proportion (35%) of the initially announced permits ($\hat{S}_t$). This yields a price volatility that is close to the actual observed volatility,[19] but it is clearly only a rough representation of the actual shocks in the EU ETS.

We model the MSR close to its actual implementation (European Parliament and Council of the European Union 2015; 2018). If the aggregate firm bank in the previous year B_{y-1} is larger than 0.833 Gt, a share ω_y of that bank is deducted from the auctioned permits in year y (if there are enough permits to be auctioned). The share is $\omega_y = 0.24$ until $y = 2023$ and $\omega_y = 0.12$ thereafter. Permits that are not auctioned due to this mechanism are denoted by M_y^{in} and go into the reserve denoted by M_y. If the banked permits in the previous year are lower than 0.4 Gt, the number of M_y^{out} is released from the reserve and added to the auctioned permits. This number is equal to 0.1 Gt (if there are enough permits in the MSR). If the bank in the previous year lies within the corridor, $0.4 < B_{y-1} < 0.833$, the permit supply is not adjusted. Therefore, the actual permits

[18] Note that we assume that the last shock emerges in period $t = 2045$ (2043–2047) due to computational constraints.

[19] In the model, the price volatility (measured as the relative standard deviation) is about 58% (excluding the MSR) in 2025 from the perspective of 2020 (see Figure C.5 in Appendix C). The actual price volatility of the EU ETS price (2008 until the end of 2019) is 60%. Further note that to avoid a negative auction supply, we set potential negative auction values due to the shocks to zero.

issued after the impact of the MSR S_y^M reads

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

and the number of permits in the reserve 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 the permits. 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 with $M_{2019} = 1.525$ Gt permits.[20] Under the current regulation, the number of permits in the MSR would only slowly decline in some scenarios, and therefore, a positive reserve could remain in the terminal model period. Because we focus on the time until 2057, but a positive reserve in the terminal model period effectively reduces the cap (and thus, affects the permit price in all periods), we assume that from 2058 onward the outtake of the MSR increases from 0.1 to 1 Gt.

Based on the European Commission (2019), we set the initial bank volume to $B_{2018} = 1.655$ Gt and assume that initially all permits are held by the dirty coal firm because the gas firm has no incentive to bank for hedging purposes (cf. Proposition 1). The risk-free rate is assumed to be $r = 3\%$. Additional details of the assumed parameters are in Appendix B. To solve the model with the MSR, we initially run the model with the auction schedule S_t without the MSR. 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

[20] The MSR is initially filled with permits that were backloaded between 2014 and 2016 (0.9 Gt) and other unallocated permits that are estimated to be between 0.55 to 0.7 Gt (European Commission 2015). Taking the sum of the arithmetic mean of this estimate and of the 0.9 Gt backloaded permits yields $M_{2019} = 1.525$ Gt.

with the adjusted permit issuance S_t^M. This procedure is iterated until it converges.[21]

3.2. Results

To disentangle the two effects of the MSR (permit shifting over time and cancellation), we consider three base scenarios: a scenario without the MSR, one with the MSR but without cancellation, and a scenario with the MSR and cancellation. Each base scenario is run in two variations, with risk aversion (RA) and with risk neutrality (RN), to show the effects of hedging. Thus, we have six scenarios in total denoted by *RN*, *RA* (both without the MSR), *RN MSR*, *RA MSR* (the MSR without cancellation), *RN MSR + cancel,* and *RA MSR + cancel* (the MSR with cancellation). In scenarios without the MSR the initial MSR bank is added to the initial bank level of the coal firm. In Section 3.2.1, we focus on the price effects of hedging and on the effect of permit shifting due to the MSR, ignoring cancellation. Then in Section 3.2.2 we examine the actual MSR as implemented in the EU ETS including cancellation. We show only results until 2055, for the full time horizon, see Appendix C.[22]

3.2.1. Hedging effects and the MSR without cancellation

Figure 3.1 (a) shows the development of the expected permit price for all six scenarios. First, we focus on the differences between the scenarios without the MSR (*RA*, *RN*, black lines) which reveal the effect of hedging. Initially, the price is higher with risk aversion; then the price declines and drops below the risk-neutral case from 2035 onward. Deviations between *RA* and *RN* are driven by the firms' hedging demand, as reflected by risk premiums, shown in Figure 3.1 (b). In the early years, the risk premium at -5% is highly negative, and because the permit price grows at the sum of the risk-free rate (3%) and the risk premium, the price actually declines. The negative risk premium can be explained by the high coal production level and thus, the coal firm's high hedging demand. The available permits for banking do not suffice to cover the coal firm's high

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

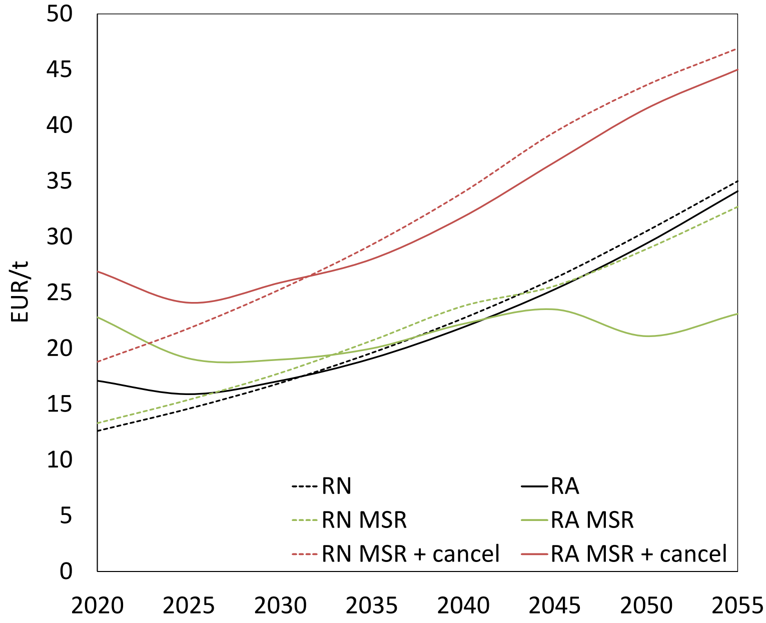
[22]Note that we concentrate on scenarios without a futures market as they do not affect the main insights (see Section 2.2.2). However, we briefly compare the results to the case with the futures market in Section 3.2.2 and show results for futures markets in Appendix C.

hedging demand. As a result, the firm accepts a reduced return for holding permits reflected by the negative risk premium. Over time, the hedging demand declines as coal production is reduced in the market, and the bank volume rises (see Figure 3.2 (a)). Consequently, the risk premium declines. Note that the lower growth rate and the price decline after 2045 are due to binding borrowing constraints.

Next, we consider the impact of the MSR if cancellation is not active by comparing the green and black lines in Figure 3.1 (a). The figure shows that the MSR raises near-term prices but lowers long-term prices under risk neutrality and risk aversion. The effect is small for risk neutrality and 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 firms' hedging demand is considered. However, the short-term price increase in the case of risk aversion implies a lower growth rate such that the price level in *RA MSR* in 2040 is only as high as in 2020. The reason for the strong effects of the MSR even without cancellation is the firm's reduced bank level as shown in Figure 3.2 (a). Instead of firms holding permits, a large number of permits are transferred into the MSR bank (see Figure 3.2 (b)) where they cannot cover the firms' hedging demand. Note that even without the MSR there are not enough permits to cover the hedging demand reflected by the negative risk premium. Because the MSR reduces the permit availability further, it implies an even more negative risk premium (see Figure 3.1 (b)), 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

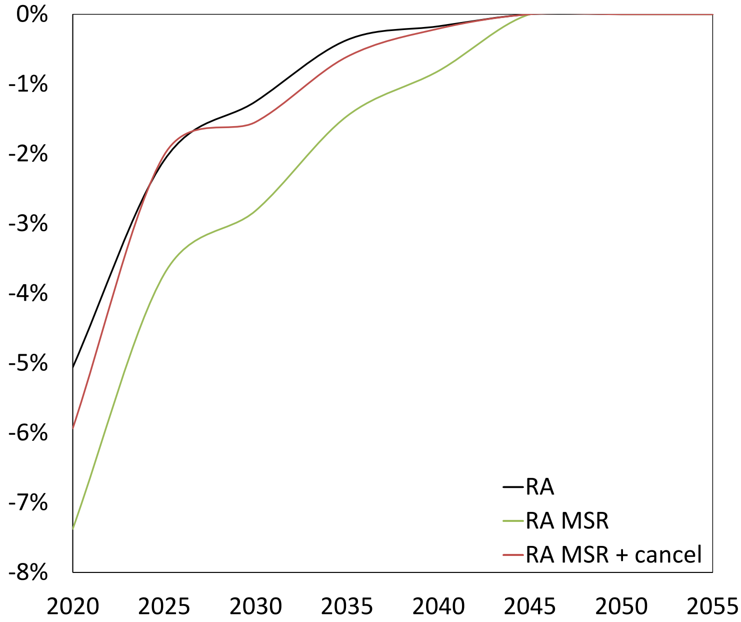
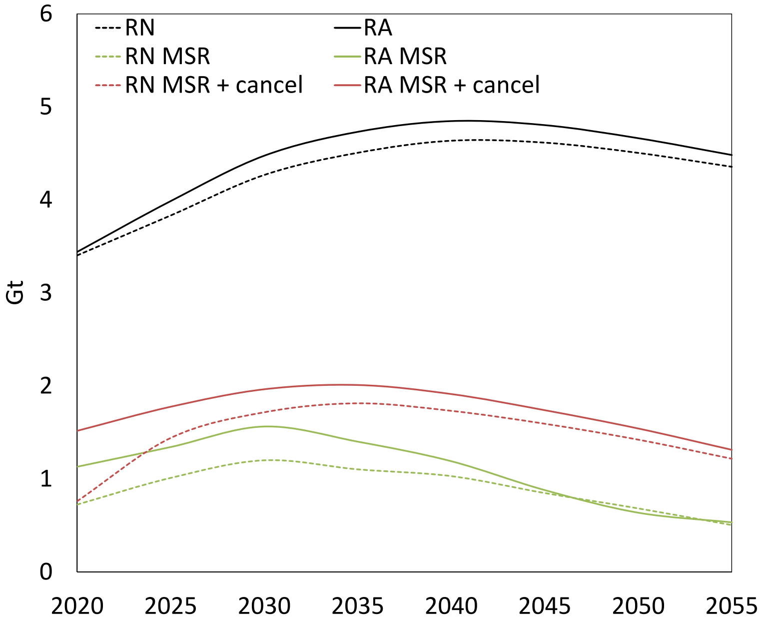
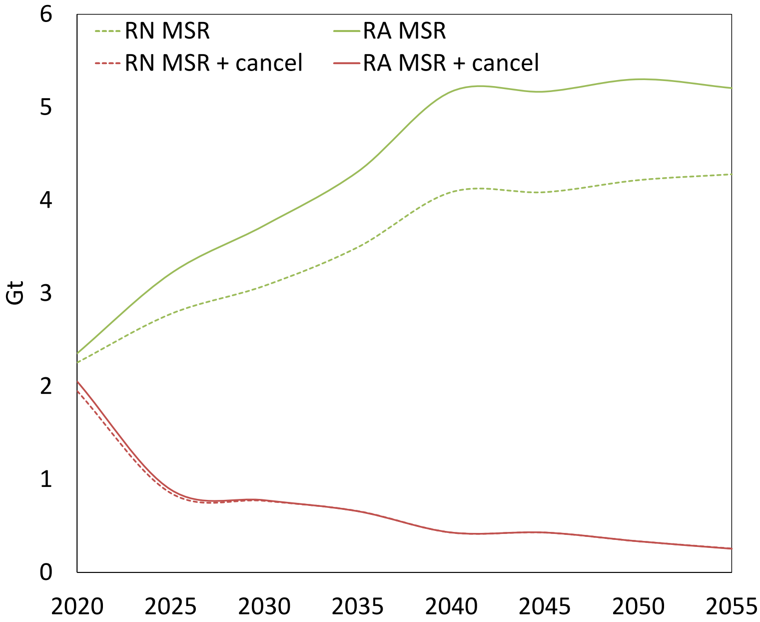


Figure 3.2: Expected firm bank and MSR level

(a) Aggregate firm bank



(b) MSR level



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 level. 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

We consider how the cancellation mechanism affects the price pattern and how hedging affects the number of cancellations. Then, we briefly discuss the impact of the MSR on plant investments.

In scenarios with cancellation (red lines in Figure 3.1 (a)), a similar price pattern to scenarios without cancellation can be observed but at a higher level. Moreover, cancellation mitigates the price drop after 2020, and thus, the price level of 2020 is reached in 2035 instead of 2040 as without cancellation. This can be traced back to the higher price level induced by cancellations: First, higher prices imply less coal production, and thus, a reduced need to hedge dirty profits. Second, less coal production also implies a higher bank level (see Figure 3.2 (a)). Therefore, the mismatch between hedging demand and permit availability is lower compared to *RA MSR,* and in turn, the risk premium is less negative as well (see Figure 3.1 (b)). Overall, the price starts at a higher level and declines less. As a result, prices are strictly higher than without the MSR.

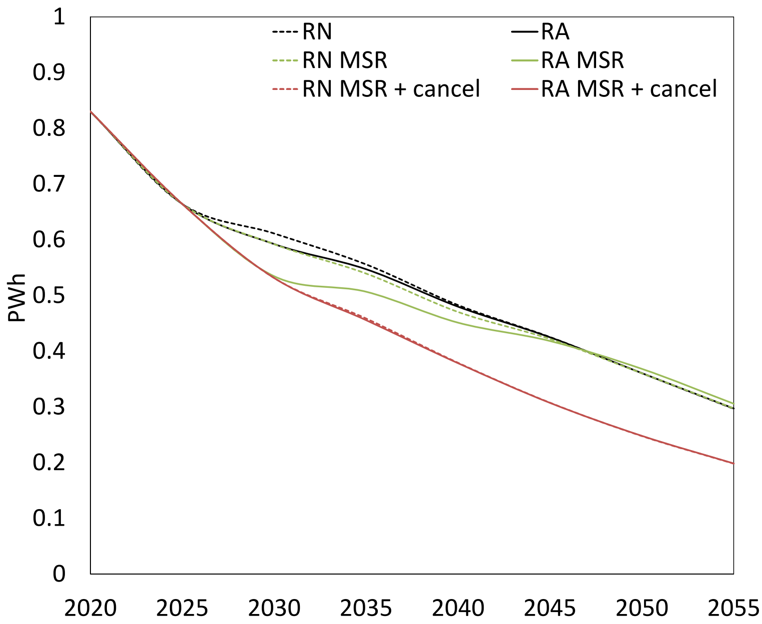
The cancellation totals to 7.60 Gt and 8.59 Gt in the case of *RN MSR + cancel* and *RA MSR + cancel*, respectively. Thus, if the hedging demand is considered, cancellation is about 1 Gt higher. This can be explained by the higher value of the permits in the early years due to firms' hedging demand. Specifically, the hedging value raises the price in 2020 significantly in *RA MSR + cancel* compared to *RN MSR + cancel* (see Figure 3.1 (a)) leading to less emissions and a larger bank (see Figure 3.2 (a)). In turn, the influx into the MSR is higher, and thus, more permits are canceled. The lower prices after 2030 (implying opposite effects) cannot outweigh this effect because the cancellations mainly take place before 2030.

Two modifications illustrate how the numerical findings are affected by the risk-free rate and the futures market. First, we run a scenario with a risk-free rate of 5% in which cancellations are lower (6.23 Gt in *RN MSR + cancel* and 6.81 Gt in *RA MSR + cancel*), which is in line with results from the literature (e.g., Bocklet et al. 2019). Second, adding futures contracts and speculators to the model reduces the risk premium, because the risk-taking capacity of the market increases (see Section 2.2.2). The cancellations are somewhat lower with the futures market (8.27 Gt in *RA MSR + cancel* compared to 8.59 Gt; note in the case of risk neutrality, futures markets have no effect), because the effect of risk aversion becomes weaker. Both modifications do not change the nature of the results because the price pattern is similar (see Appendix C).

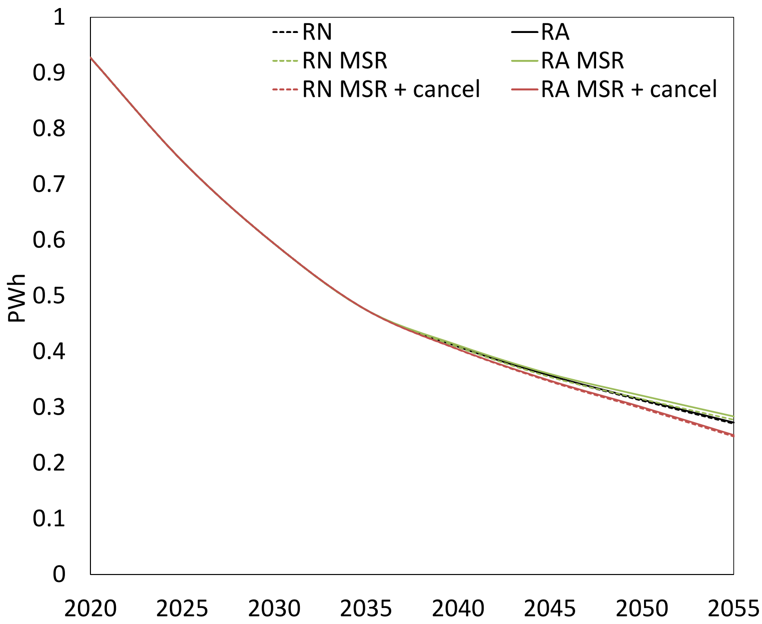
Finally, we consider the impact of the MSR on investments in capacity. Figure 3.3 shows that the higher permit price path due to the cancellation mechanism significantly reduces investments in coal capacity, while gas capacity is only slightly affected. If the cancellation mechanism is not active, the MSR also has a significant effect on coal capacities if firms are risk averse (*RA MSR*): Due to the higher permit price until about 2040 and the worse hedging opportunities (see Section 2.2.2), there is less coal capacity compared to *RA*. However, from 2045 onward there is slightly more coal capacity, because the MSR leads to lower prices in the long-term. Overall, the effect of the cancellation mechanism is significantly stronger than the effect of shifting permits to the future. However, a potential disadvantage of the cancellation mechanism that deserves more attention in future research is that it may increase the price variability (see Figure C.5 in Appendix C).

Figure 3.3: Expected capacity

(a) Coal capacity



(b) Gas capacity



4. Conclusion

We analyze the impact of 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 a permit supply risk 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 permit borrowing is not allowed, which is typically the case, 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. As the hedging demand of dirty firms is typically high in the early years (implying price growth at a low rate) of a cap-and-trade program and low in later years (implying a higher growth rate), the expected growth rate of the permit price may has a U-shape.

We numerically apply the model to the EU ETS to investigate price effects of the MSR. The core mechanisms of the MSR are shifting permits to the future and canceling 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 induces even more negative risk premiums because it reduces the size of the permit bank. The results offer an explanation for the recent permit price hike in the EU ETS because more negative risk premiums lead to higher short-term prices. An additional consequence is that prices may grow only at a low rate or even decline in the coming years, which is also in line with analysts' forecasts for the coming decade (see Carbon Pulse 2019 for a poll).

In addition to the higher hedging value of permits due to the MSR, an important reason for the recent price increase is the cancellation of permits from 2023 onward. We find that cancellations may be higher than previous analyses suggest. The hedging demand and the associated negative risk premium imply that firms use a lower discount rate for banking permits and build up a larger bank. This, in turn, increases the MSR cancellations. We also stress the role of capacity constraints, which prevail in electricity markets. Specifically, we show that they increase the permit price variability and therefore, amplify the effects of risk aversion.

However, this study also has limitations. First, we consider a highly stylized model with only two electricity generators and one speculator. Considering more firm types would affect how risks can be allocated as, for instance, firms may pursue a plant portfolio approach by investing in clean and dirty plants to lower their overall risk exposure (Roques et al. 2008). In reality, there are also more, and essentially more complex, derivatives, such as options that allow to improve hedging opportunities. The main effect of including more derivatives and more complex firm structures is more efficient risk allocation implying lower risk premiums which could be further analyzed in future research. Similarly, we assume simple functional forms for electricity costs and demand, and ignore certain aspects of the EU ETS, such as grandfathering of permits and the explicit modeling of non-electricity sectors. In general, our model is only roughly calibrated to the EU ETS so that the numerical results should be understood only as stylized illustrations. A more detailed and calibrated modeling of financial aspects such as hedging and capital constraints in emission trading systems is an interesting avenue for future research.

Our work also raises other issues for further research. At the time of writing, actual discount rates in futures markets of the EU ETS are about 1.5%,[23] far below the typically assumed rate of 3% to 10% in the theoretical and numerical ETS literature. Given the high degree of uncertainty in this market, our analysis suggests that such a low rate can be explained by negative risk premiums. For future research, it would be interesting to empirically investigate the risk premium, ideally with a dedicated proxy for the firms' hedging demand (see Acharya et al. 2013 for a similar analysis for other commodity markets). Another promising research field would be to examine the impact of the MSR on the permit price variability, given that previous work considering uncertainty examines only the original MSR without the cancellation mechanism (Richstein et al. 2015; Fell 2016; Perino and Willner 2016; Kollenberg and Taschini 2019). In particular, our results

[23]See https://www.barchart.com/futures/quotes/CK*0/futures-prices (18-07-2019)

indicate that the permit price variability may increase due to the cancellation mechanism.

Acknowledgments

The authors would like to thank Dallas Burtraw, Ottmar Edenhofer, Christian Gambardella, Andrew McConnell, Samuel Okullo, Simon Quemin, Rick van der Ploeg and participants of the WCERE conference 2018 in Gothenburg for valuable comments. Oliver Tietjen's contribution was funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 730403 (INNOPATHS). Kai Lessmann's contribution was funded by the Federal Ministry of Education and Research (BMBF) program "Global Change 5+1" as part of the grant agreement 01LN1703A (FINFAIL). Michael Pahle was funded by Stiftung Mercator Foundation under the research project AHEAD.

Appendix A. Derivations

Appendix A.1. First-order conditions

The first-order conditions of the generators' problem in period 2 are

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

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

The first-order conditions of the generators' problem in period 1 are

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

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

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

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

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

$$U'_{i,1}p_1^f-\frac{E\left[U'_{i,2}p_2\right]}{1+r}=0 \quad (f_{i,1}). \tag{A.8}$$

The first-order conditions of the speculator's problem are

$$U'_{sp,1}-E\left[U'_{sp,2}\right]=0 \quad (l_{sp,1}), \tag{A.9}$$

$$U'_{sp,1}p_1^f-\frac{E\left[U'_{sp,2}p_2\right]}{1+r}=0 \quad (f_{sp,1}). \tag{A.10}$$

Note that for a risk-neutral firm that maximizes its expected profits the optimality conditions are the same but with constant marginal utility; i.e., $U'_{i,1}=E\left[U'_{i,2}\right]=1$.

Appendix A.2. Period 2 equilibrium

Lemma 1. Inserting the electricity market price (2.21) in the utilization rate (2.18), if the capacity constraints do not bind, $\mu_{c,2}=\mu_{d,2}=0$, yields

$$\zeta_{i,2}=\frac{A\beta_c\beta_d+p_2\left(\beta_d\phi_c+\beta_c\phi_d\right)}{\left(\beta_d+\beta_c+\beta_c\beta_d a\right)\beta_i k_i}-\frac{p_2\phi_i}{\beta_i k_i}. \tag{A.11}$$

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

$$\frac{d\zeta_{c,2}}{dp_2}=\frac{\phi_d-\phi_c\left(1+\beta_d a\right)}{\left(\beta_d+\beta_c+\beta_c\beta_d a\right)k_c}>0, \tag{A.12}$$

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

$$\frac{d\zeta_{c,2}}{dp_2} = \frac{\phi_c - \phi_d\left(1+\beta_c a\right)}{\left(\beta_d + \beta_c + \beta_c\beta_d a\right)k_d} < 0, \tag{A.13}$$

as, by definition, $\phi_d > \phi_c$.

Lemma 2. Using the electricity price (2.21) and the utilization rate (2.18), the plant profit can be written as

$$\pi_{i,2}^{plant} = \frac{\left(A\beta_c\beta_d + p_2\left(\beta_d\phi_c + \beta_c\phi_d\right)\right)^2}{2\beta_i\left(\beta_d + \beta_c + \beta_c\beta_d a\right)^2} - \frac{\left(A\beta_c\beta_d + p_2\left(\beta_d\phi_c + \beta_c\phi_d\right)\right)p_2\phi_i}{\beta_i\left(\beta_d + \beta_c + \beta_c\beta_d a\right)} + \frac{p_2^2\phi_i^2}{2\beta_i}, \tag{A.14}$$

for which we assume 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

$$A\beta_d + p_2\left(\phi_d - \phi_c\left(1+\beta_d a\right)\right) > 0, \tag{A.15}$$

which is always the case if $\phi_d - \phi_c\left(1+\beta_d a\right) > 0$ 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

$$A\beta_c + p_2\left(\phi_c - \phi_d\left(1+\beta_c a\right)\right) > 0 \tag{A.16}$$

holds. If the price is

$$p_2 = \frac{A\beta_c}{\left(\phi_d\left(1+\beta_c a\right) - \phi_c\right)}, \tag{A.17}$$

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.17), the dirty firm does not produce, and thus, profits also are not affected. This can be seen by inserting (A.17) in the utilization rate (A.11) which yields $\zeta_{d,2} = 0$. The same is true for higher prices because of Lemma 1. For lower prices than in (A.17), condition (A.16) is fulfilled, and thus, $\frac{d\pi_{d,2}^{plant}}{dp_2} < 0$ and $Cov\left[\pi_{d,2}^{plant}, p_2\right] < 0$ hold.

The effect of plant capacity constraints on permit price variability. The effect of capacity constraints on the ETS price is given by the second line in (2.22), which we replicate for convenience

$$\frac{\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}}\left(\phi_d-\phi_c\left(1+\beta_d a\right)\right)+\frac{\mu_{d,2}}{U'_{d,2}k_{d,2}}\left(\phi_c-\phi_d\left(1+\beta_c a\right)\right)}{\left(\phi_c-\phi_d\right)^2+a\left(\beta_c\phi_d^2+\beta_d\phi_c^2\right)}. \tag{A.18}$$

The first line in (2.22) is the same as without capacity. 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 (2.22) and (2.21) in (2.20) such that we get

$$\begin{aligned}\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}} &= \frac{A\phi_d\left(\phi_d-\phi_c\right)-k_{c,2}\left(\left(\phi_d-\phi_c\right)^2+a\left(\beta_d\phi_c^2+\beta_c\phi_d^2\right)\right)}{\phi_d^2 a}\\ &\quad+\frac{\left(b_{c,1}+b_{d,1}+S_2\right)\left(\phi_c\left(1+\beta_d a\right)-\phi_d\right)}{\phi_d^2 a}.\end{aligned} \tag{A.19}$$

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

$$\frac{d\left(\frac{\mu_{c,2}}{U'_{c,2}k_{c,2}}\right)}{dS_2}=\frac{\phi_c\left(1+\beta_d a\right)-\phi_d}{\phi_d^2 a}<0, \tag{A.20}$$

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, $\mu_{c,2}=0$ and $\mu_{d,2}>0$ hold. 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 no longer binds, $\mu_{d,2}=0$ (Lemma 1). To see that a declining $\mu_{d,2}$ and a rising $\mu_{c,2}$ lead to a stronger ETS price increase in (2.22), 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, $\mu_{c,2}>0$ and $\mu_{d,2}>0$ hold. As in case 2, the dirty constraint cannot bind after a negative shock has emerged, which 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, $\mu_{c,2} = \mu_{d,2} = 0$ holds. As in case 1, a negative shock implies that $\mu_{d,2} = 0$ still holds and $\mu_{c,2} \geq 0$. Thus, 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. Period 1 equilibrium

Appendix A.3.1. Banking and hedging

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

$$\frac{E\left[p_2\right] - p_1}{p_1} = r + \frac{\left(1+r\right)\varphi_{i,1}}{p_1 E\left[U'_{i,2}\right]} - \frac{Cov\left[U'_{i,2}, p_2\right]}{E\left[U'_{i,2}\right] p_1}. \tag{A.21}$$

Assuming quadratic utility, $U_t\left(\pi_{it}\right) = \pi_{it} - \pi_{it}^2$, we can write the covariance as $Cov\left[U'_{i,2}, p_2\right] = -2\left(Cov\left[\pi_{i,2}^{plant}, p_2\right] + Var\left[p_2\right] b_{i,1}\right)$, and inserting it in (A.21) yields

$$b_{i,1} = \frac{E\left[p_2\right] - p_1\left(1+r\right)}{\lambda_i Var\left[p_2\right]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var\left[p_2\right]} - \frac{\left(1+r\right)\varphi_{i,1}}{U'_{i,1} Var\left[p_2\right]}. \tag{A.22}$$

Assuming a risk premium of zero, $E\left[p_2\right] - p_1\left(1+r\right) = 0$, the pure banking or borrowing demand is due only 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 $b_{i,1} > 0$ permits, 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 the risks of this action, $E\left[p_2\right] - p_1\left(1+r\right) > \lambda_i Cov\left[\pi_{c,2}^{plant}, p_2\right]$.

Next, we turn to the price effects of hedging. Consider that the permit demand can be written as

$$
\begin{aligned}
y_{i,1} \quad = \quad & \frac{\phi_i A\beta_c\beta_d}{(\beta_d+\beta_c+\beta_c\beta_d a)\,\beta_i} + p_1\left(\frac{(\beta_d\phi_c+\beta_c\phi_d)\,\phi_i}{(\beta_d+\beta_c+\beta_c\beta_d a)\,\beta_i} - \frac{\phi_i^2}{\beta_i} - \frac{(1+r)}{\lambda_i Var\,[p_2]}\right) \qquad \text{(A.23)}\\
& + \frac{E\,[p_2]}{\lambda_i Var\,[p_2]} - \frac{Cov\left[\pi_{i,2}^{plant}, p_2\right]}{Var\,[p_2]},
\end{aligned}
$$

for which we used (2.24) and (A.11) (but for period 1) in (2.23), and we assumed non-binding capacity constraints. Inserting this permit demand for clean and dirty firms in the permit equilibrium condition (2.4) yields the permit price,

$$
p_1 = \frac{E\,[p_2]}{(1+r)} - \frac{\Lambda}{(1+r)}\left(Cov\left[\pi_{d,2}^{plant}, p_2\right] + Cov\left[\pi_{c,2}^{plant}, p_2\right] + Var\,[p_2]\,B_1\right), \qquad \text{(A.24)}
$$

The whole term in brackets is the absolute risk premium, and dividing by p_1 this term becomes the relative risk premium as shown in (2.26). Repeating the steps in this section for the case with the futures market shows that Equation (A.24) is still valid, but $\Lambda = \left(\lambda_d^{-1}+\lambda_c^{-1}\right)^{-1}$ must be replaced by $\Lambda^f = \left(\lambda_d^{-1}+\lambda_c^{-1}+\lambda_{sp}^{-1}\right)^{-1}$.

Appendix A.3.2. Capacity effects

The first-order condition for $k_{i,2}$ (Equation (A.6)) can be reformulated as

$$
\begin{aligned}
E\,[\zeta_{i,2}]\,(E\,[w_2] - E\,[\zeta_{i,2}]\,\beta_i k_i - E\,[\zeta_{i,2}]\,\phi_i) - \gamma_i & \\
+Cov\,[\zeta_{i,2}, w_2 - \zeta_{i,2}\beta_i k_i - p_2\phi_i] & \qquad \text{(A.25)}\\
+\frac{1}{U'_{i,1}} Cov\left[U'_{i,2}, \zeta_{i,2}w_2 - \zeta_{i,2}^2\beta_i k_i - p_2\zeta_{i,2}\phi_i\right] &= 0,
\end{aligned}
$$

for which we 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 $\mu_{i,2}^{RN} = k_{i,2}w_2 - x_{i,2}k_{i,2}\beta_i - p_2 k_{i,2}\phi_i$ (see Equation (2.20) and consider that in the case of risk neutrality $U'_{i,2} = 1$ holds), we can rewrite this further and finally get (2.30).

Effect of uncertainty if firms are risk neutral. Risk neutrality implies $U'_{i,2} = 1$, and thus, $Cov\left[U'_{i,2}, \mu^{RN}_{i,2}\right] = 0$. Therefore, only the first two terms in (2.30) matter in the risk-neutral case. Moreover, in the deterministic case $Cov\left[\zeta_{i,2}, \mu^{RN}_{i,2}\right] = 0$ holds, and investments are determined by only the first term in (2.30) (ignoring the expectation operator). Thus, given risk neutrality, the effect of uncertainty compared to the deterministic case is given by the second term, $Cov\left[\zeta_{i,2}, \mu^{RN}_{i,2}\right]$. We can rewrite this term as $Cov\left[\zeta_{i,2}, \mu^{RN}_{i,2}\right] = E\left[\mu^{RN}_{i,2}\right](1 - E[\zeta_{i,2}])$, for which we used $E\left[\mu^{RN}_{i,2}\right] = E\left[\zeta_{i,2}\mu^{RN}_{i,2}\right]$, because $\mu^{RN}_{i,2}$ is positive only if $\zeta_{i,2} = 1$ and zero otherwise. As $0 < E[\zeta_{i,2}] < 1$, and $E\left[\mu^{RN}_{i,2}\right] > 0$, $Cov\left[\zeta_{i,2}, \mu^{RN}_{i,2}\right] > 0$ holds. Thus, uncertainty has a positive effect on investments if firms are risk neutral.

Effect of risk aversion without banking. Next, we consider the effect of risk aversion given by the third term in (2.30). As $\frac{1}{U'_{i,1}} > 0$, the sign of the effect of risk aversion depends on $Cov\left[U'_{i,2}, \mu^{RN}_{i,2}\right]$. Assuming that the permit bank is zero, marginal utility $U'_{i,2}$ depends only on plant profit $\pi^{plant}_{i,2}$ and risk-free asset returns. The latter do not affect the covariance because they are nonrandom. Due to the concavity of $U_{i,2}$, it follows $\frac{dU_{i,2}}{d\pi^{plant}_{i,2}} < 0$. Thus, the sign of $Cov\left[U'_{i,2}, \mu^{RN}_{i,2}\right]$ is inversely related to $Cov\left[\pi^{plant}_{i,2}, \mu^{RN}_{i,2}\right]$ which is positive,

$$Cov\left[\pi^{plant}_{i,2}, \mu^{RN}_{i,2}\right] = Cov\left[\zeta_{i,2}\left(w_2 - \frac{\beta_i}{2}x_{i,2} - p_2\phi_i\right), \zeta_{i,2}\left(w_2 - \beta_i x_{i,2} - p_2\phi_i\right)\right] k^2_{i,2} \geq 0, \tag{A.26}$$

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

Effect of risk aversion with banking. We again consider $Cov\left[\pi_{i,2}, \mu^{RN}_{i,2}\right]$ which becomes $Cov\left[\pi_{i,2}, \mu^{RN}_{i,2}\right] = Cov\left[\pi^{plant}_{i,2}, \mu^{RN}_{i,2}\right] + Cov\left[p_2 b_{i,1}, \mu^{RN}_{i,2}\right]$ when firms bank. Compared to the case without banking, there is an additional effect given by $Cov\left[p_2 b_{i,1}, \mu^{RN}_{i,2}\right]$. Firms invest, ceteris paribus, more if $Cov\left[p_2 b_{i,1}, \mu^{RN}_{i,2}\right] < 0$ and less if $Cov\left[p_2 b_{i,1}, \mu^{RN}_{i,2}\right] > 0$, because a lower $Cov\left[\pi_{i,2}, \mu^{RN}_{i,2}\right]$ implies a higher $Cov\left[U'_{i,2}, \mu^{RN}_{i,2}\right]$ 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. Therefore, $Cov\left[p_2 b_{i,1}, \mu^{RN}_{i,2}\right] < 0$ holds. For clean firms, the opposite holds.

If capacity constraints are strictly binding such that firms cannot produce more in the case of positive (clean firm) or negative (dirty firm) price shocks, or stick to the full capacity utilization in the opposite case, we get $Cov\left[p_2 b_{i,1}, \mu_{i,2}^{RN}\right] = \left(Cov\left[p_2, w_2\right] - Var\left[p_2\right]\phi_i\right) b_{i,1} k_{i,2}$. Thus, if permit price shocks are disproportionately transferred to the electricity 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 B. Parameter assumptions for the numerical simulation

Cost function parameters are chosen in line with coal and gas power plants for the representative dirty and clean firms, respectively (for the parameters see Table B.1). For the electricity demand function, $D\left(w_t\right) = A - a_t w_t$, we assume the intercept is $A = 3,462$ TWh, which is the total electricity generation in the EU28, Iceland and Norway (the EU ETS countries except Liechtenstein) 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 larger 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 increases at a 9% rate every five years. The increase in a_t mainly reflects market entry of renewable energies (e.g., due to support programs).

These parameter assumptions lead to an ETS price of 26.9 EUR/t and 0.78 Gt emissions in the first model period of the scenario *RA MSR + cancel*. The price is in line with actual (futures) prices between 2018 and 2022 (26.15–27.44 EUR/t).[24] Our emission level is somewhat lower than the emissions due to combustion in the EU ETS in 2018,

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

which are 1.1 Gt.[25] 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 the 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 the higher ETS prices and the growing renewable energy output.

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 the empirical estimates, we set relative risk aversion to $\eta = 1.5$ (cf. 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 first period, which is between 23 and 38 billion EUR for the coal firm 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.

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

Table B.1: Firm data

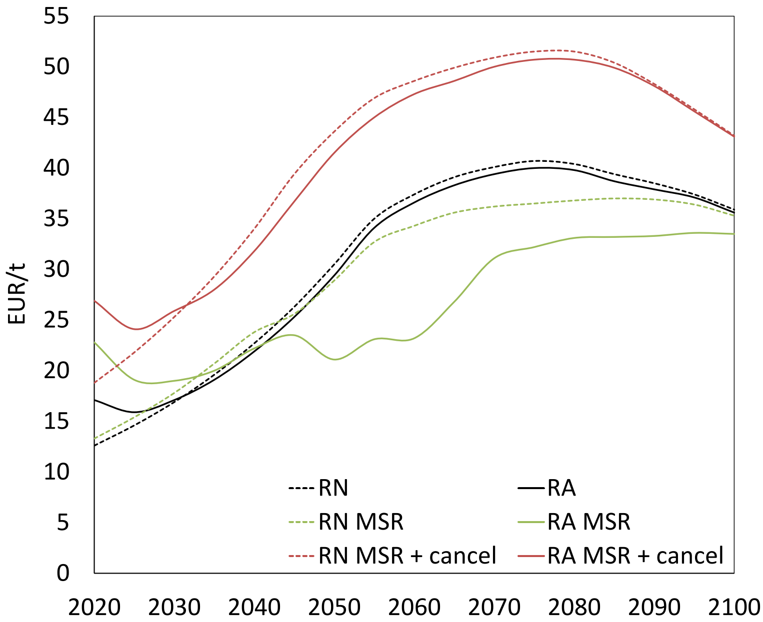
	Clean	Dirty
Production costs (EUR/GWh):β_i	0.050	0.020
Capacity costs (EUR/GWh): γ_i	0.0049	0.0084
Emission factor (t/GWh): ϕ_i	333	950
Initial capacities (TWh): $k_{i,2020}$	830.2	927.5
Capacity depreciation: δ	0.2	0.2

Note: Emission factors are based on UBA (2014) and divided by conversion efficiencies (fuel to electricity) of 60% for gas and 40% for coal. Capacity costs are based on the 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%.

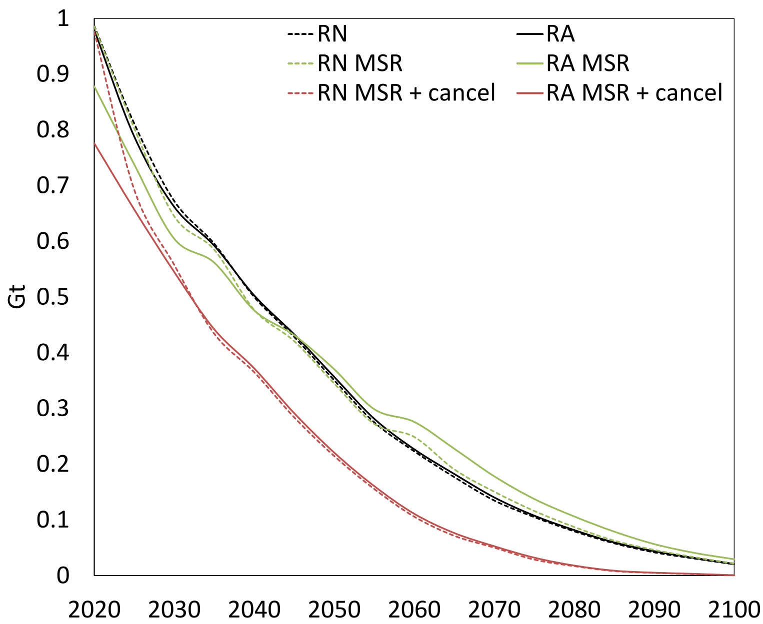
Appendix C. Additional simulation results

Figure C.1: Results for the full time horizon: price and emissions

(a) Price



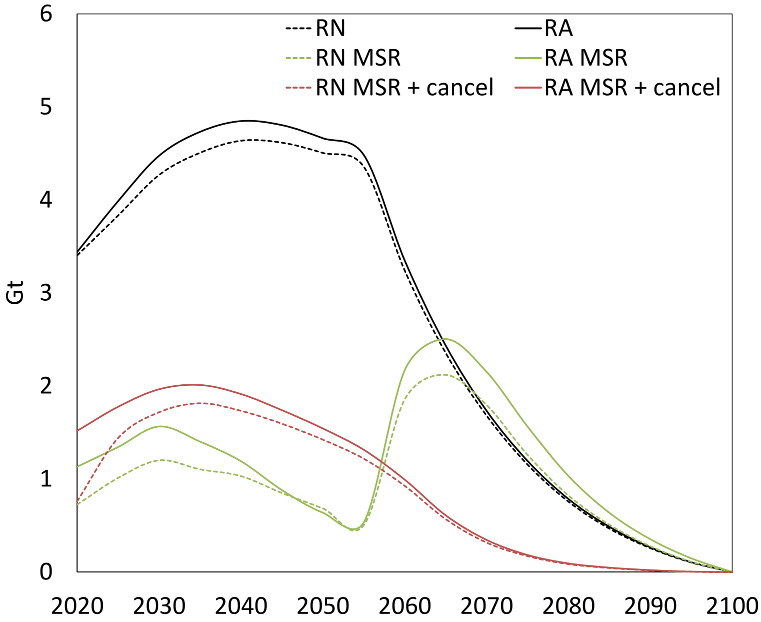
(b) Emissions



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).

Figure C.2: Results for the full time horizon: firm bank and MSR level

(a) Aggregate firm bank



(b) MSR level

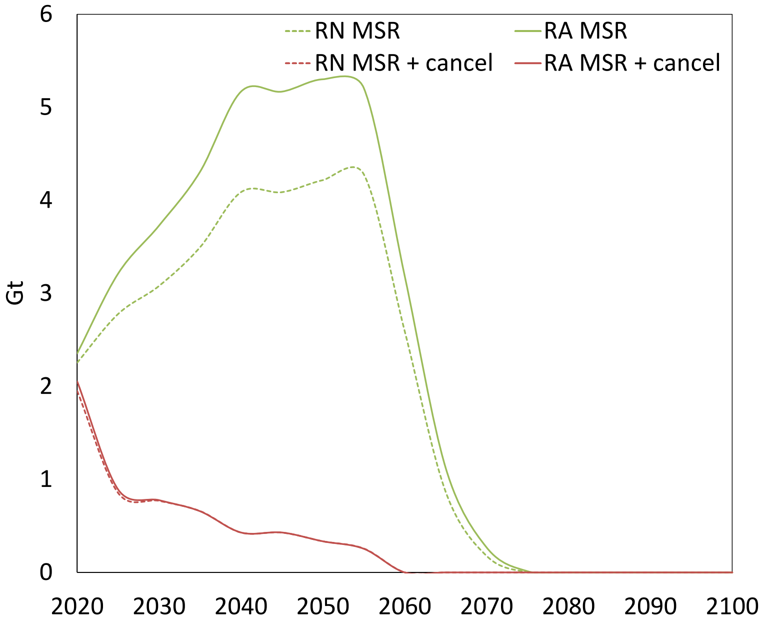
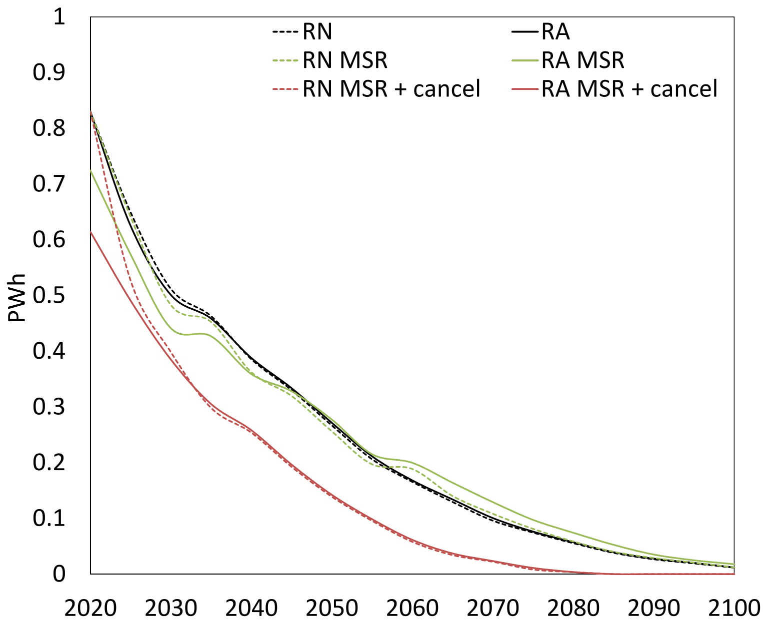


Figure C.3: Results for the full time horizon: production

(a) Coal production



(b) Gas production

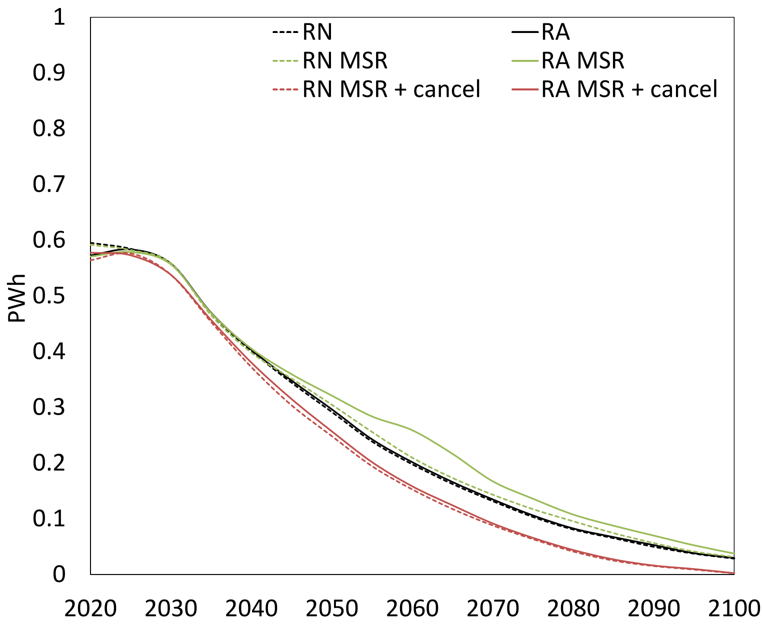
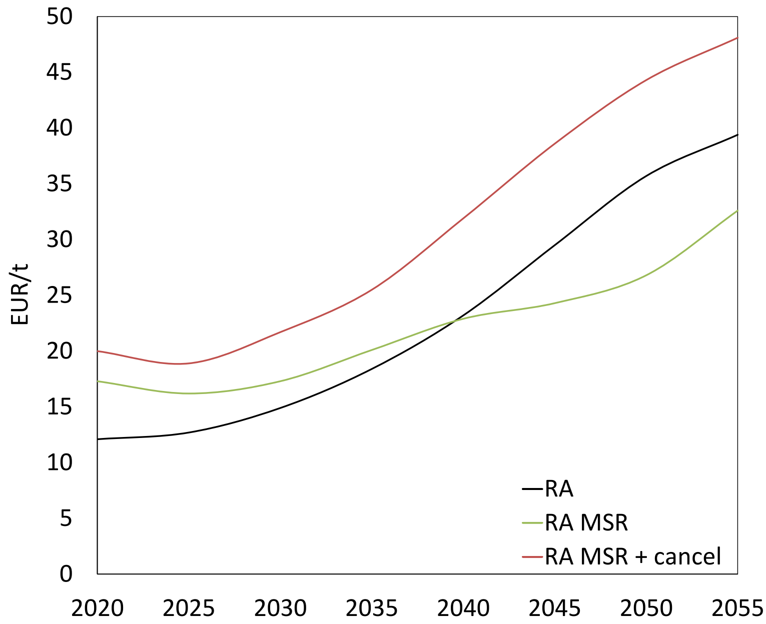
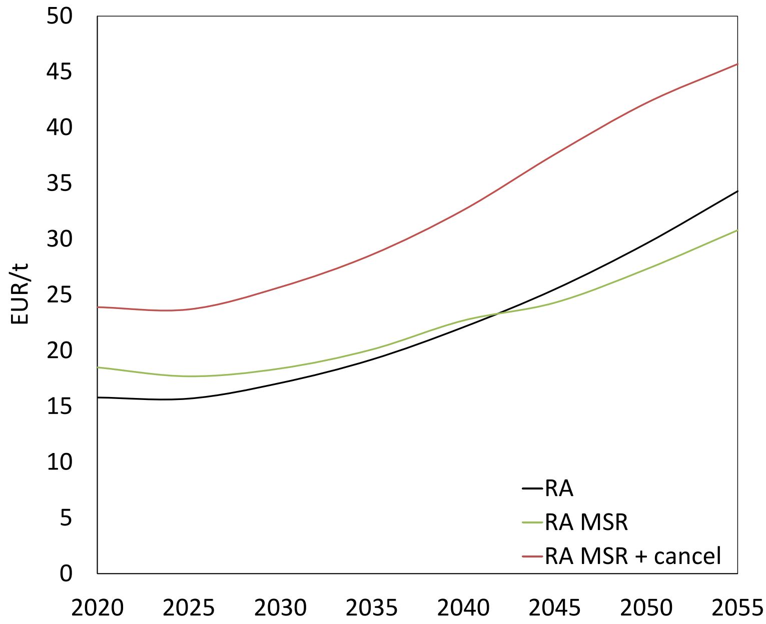


Figure C.4: Expected permit price with risk-free rate of 5% and futures market

(a) Risk-free rate of 5%



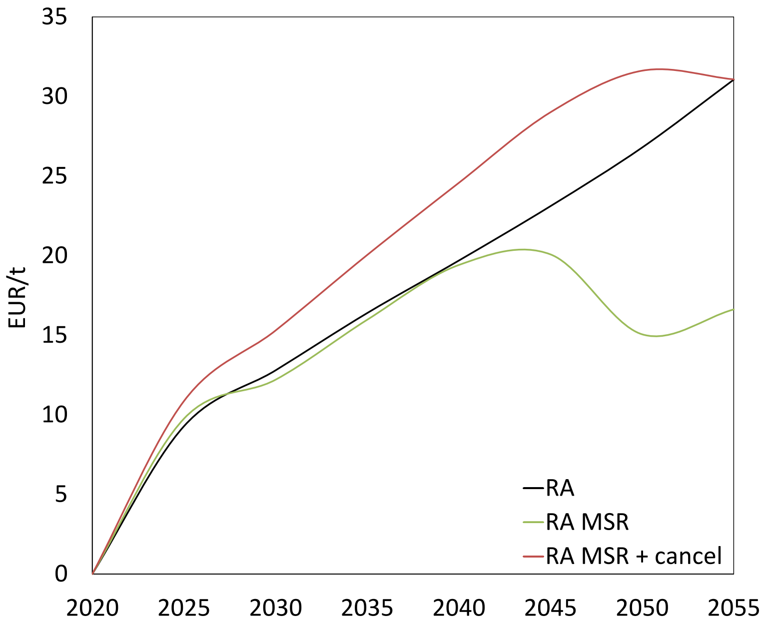
(b) Futures market



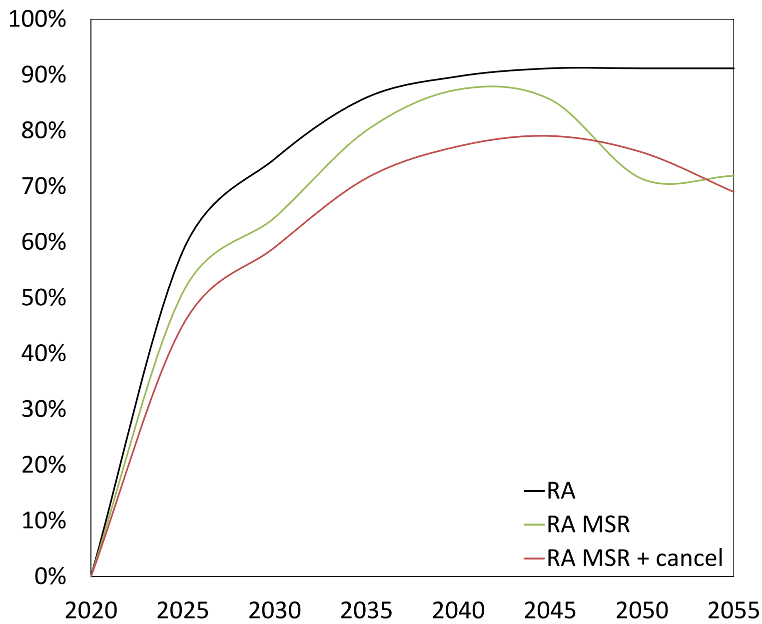
Note: For part (b), we assume a risk-free rate of 3% and that speculators have the same initial endowment and level of risk aversion as the electricity generators.

Figure C.5: Permit price variability

(a) Standard deviation



(b) Relative standard deviation



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.

Anderson, R. W. and J.-P. Danthine (1979). Hedging and Joint Production: Theory and Illustrations. *The Journal of Finance 35*(2), 487–498.

Anderson, R. W. and J.-P. Danthine (1981). Cross Hedging. *Journal of Political Economy 89*(6), 1182–1196.

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. *Working Paper 2018:2*.

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. *Energy Economics 84*.

Böhringer, C. and A. Lange (2005). On the design of optimal grandfathering schemes for emission allowances. *European Economic Review 49*(8), 2041–2055.

Brennan, M. J. (1958). The Supply of Storage. *The American Economic Review 48*(1), 50–72.

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

Carbon Pulse (2019). POLL: Analysts trim near-term EUA price forecasts as auctions put a damper on market. https://carbon-pulse.com/67157/.

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

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*, 152–182.

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

Deaton, A. and G. Laroque (1992). On the Behaviour of Commodity Prices. *The Review of Economic Studies 59*(1), 1–23.

Ekeland, I., D. Lautier, and B. Villeneuve (2019). Hedging pressure and speculation in commodity markets. *Economic Theory 68*(1), 83–123.

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 (2015). Impact Assessment Accompanying the document 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 Council (2019). European Council conclusions, 12 December 2019. *https://www.consilium.europa.eu/media/41768/12-euco-final-conclusions-en.pdf*.

European Parliament and Council of the European Union (2015). Decision (EU) 2015/1814. *Official Journal of the European Union 58*.

European Parliament and Council of the European Union (2018). Directive (EU) 2018/410. *Official Journal of the European Union 61*.

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

Flachsland, C., M. Pahle, D. Burtraw, O. Edenhofer, M. Elkerbout, C. Fischer, O. Tietjen, and L. Zetterberg (2020). How to avoid history repeating itself: the case for an EU Emissions Trading System (EU ETS) price floor revisited. *Climate Policy 20*(1), 133–142.

Friedrich, M., S. Fries, M. Pahle, and O. Edenhofer (2020). Understanding the explosive trend in EU ETS prices - fundamentals or speculation? *Working paper. https://arxiv.org/pdf/1906.10572.pdf*.

Friedrich, M., E.-M. Mauer, M. Pahle, and O. Tietjen (2020). From fundamentals to financial assets: the evolution of understanding price formation in the EU ETS. *EconStor 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.

Gerlagh, R., R. Heijmans, and K. E. Rosendahl (2019). Endogenous Emission Caps Always Induce a Green Paradox. *CESifo Working Paper No. 7862*.

Gerlagh, R., R. J. Heijmans, and K. E. Rosendahl (2020). COVID-19 Tests the Market Stability Reserve. *Environmental and Resource Economics 76*(4), 855–865.

Goldstein, I., Y. Li, and L. Yang (2014). Speculation and hedging in segmented markets. *Review of Financial Studies 27*(3), 881–922.

Haita-Falah, C. (2016). Uncertainty and speculators in an auction for emissions permits. *Journal of Regulatory Economics 49*(3), 315–343.

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

Hicks, J. (1939). *Value and Capital.* Cambridge: Oxford University Press.

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.

Hirshleifer, D. (1990). Hedging Pressure and Futures Price Movements in a General Equilibrium Model. *Econometrica 58*(2), 411–428.

Hotelling, H. (1931). The Economics of Exhaustible Resources. *Journal of Political Economy 39*(2), 137–175.

ICAP (2020). *Emissions Trading Worldwide: Status Report 2020.* Berlin, Germany: International Carbon Action Partnership.

IEA (2016). World Energy Outlook 2016.

Kaldor, N. (1939). Speculation and Economic Stability. *The Review of Economic Studies 7*(1), 1–27.

Kang, W., K. G. Rouwenhorst, and K. Tang (2020). A Tale of Two Premiums: The Role of Hedgers and Speculators in Commodity Futures Markets. *The Journal of Finance 75*(1), 377–417.

Keynes, J. (1930). *A treatise on money, Vol 2.* London: Macmillan.

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.

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.

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

Perino, G. and M. Willner (2019). Rushing the Impatient: Allowance Reserves and the Time Profile of Low-Carbon Investments. *Environmental and Resource Economics 74*(2), 845–863.

Pinho, C. and M. Madaleno (2011). Links between spot and futures allowances: ECX and EEX markets comparison. *International Journal of Global Energy Issues 35*, 101–131.

Quemin, S. (2017). Intertemporal Abatement Decisions underAmbiguity Aversion in a Cap and Trade. *FAERE Working Paper, 2017.06*.

Quemin, S. and R. Trotignon (2019). Emissions Trading with Rolling Horizons. *Working paper*.

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.

Roques, F. a., D. M. Newbery, and W. J. Nuttall (2008). Fuel mix diversification incentives in liberalized electricity markets: A Mean-Variance Portfolio theory approach. *Energy Economics 30*(4), 1831–1849.

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. *DIW Discussion Papers 1483*.

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

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

Stoft, S. (2002). *Power System Economics: Designing Markets for Electricity*. IEEE Press & Wiley-Interscience.

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.

Working, H. (1949). The Theory of Price of Storage. *The American Economic Review 39*(6), 1254–1262.

Chapter 7

Reviewing the Market Stability Reserve in light of more ambitious EU ETS emission targets[1]

Sebastian Osorio
Oliver Tietjen
Michael Pahle
Robert Pietzcker
Ottmar Edenhofer

[1]Submitted to *Energy Policy*. The final version is available under https://doi.org/10.1016/j.enpol.2021.112530.

Reviewing the Market Stability Reserve in light of more ambitious EU ETS emission targets

Sebastian Osorio*[a], sebastian.osorio@pik-potsdam.de

Oliver Tietjen[a,b], oliver.tietjen@pik-potsdam.de

Michael Pahle[a], michael.pahle@pik-potsdam.de

Robert Pietzcker[a], pietzcker@pik-potsdam.de

Ottmar Edenhofer[a,b,c], ottmar.edenhofer@pik-potsdam.de

[*] Corresponding author

[a] Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, P.O. Box 60 12 03, D-14412 Potsdam, Germany

[b] Technical University Berlin, Germany

[c] Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany

This version: May 2020

Abstract

The stringency of the EU's Emission Trading System (ETS) is bound to be ratcheted-up to deliver on more ambitious goals as put forth in the EU's Green Deal. Tightening the cap needs to consider the interactions with the Market Stability Reserve (MSR), which will be reviewed in 2021. Against that background, we employ the detailed model LIMES-EU to analyse options for the upcoming reforms. First, we examine how revising MSR parameters impacts allowance cancellations through the MSR. We find that under current regulation, the MSR cancels 5.1 Gt of allowances. Varying MSR parameters leads to cancellations in the range of 2.6 and 7.9 Gt, with the intake/outtake thresholds having the highest impact. Intake rates above 12% only have a limited effect but cause oscillatory intake behaviour. Second, we analyse how the 2030 targets can be achieved by adjusting the linear reduction factor (LRF). We find that the LRF increases MSR cancellations substantially (up to 10.0 Gt). This implies that increasing the LRF from currently 2.2% to 2.6% could already be consistent with the 55% EU-wide emission reduction target in 2030. However, we highlight that the number of MSR cancellations is subject to large uncertainty. Overall, the MSR increases the complexity of the market. In face of that, we suggest to develop the MSR into a Price Stability Reserve.

Keywords: EU climate policy; EU ETS reform; linear reduction factor (LRF); market stability reserve (MSR); EU ETS Phase IV.

Key policy insights

- We estimate that the MSR cancels 5.1 Gt of allowances under current regulation.
- MSR cancellations are more sensitive to the upper than to the lower threshold.
- A high intake rate could increase EUA price uncertainty.
- Cancellations are sensitive to the LRF due to a reinforcement effect with the MSR.
- A LRF of 2.4% and 2.6% could be in line with a 50% and 55% 2030 target.

1. Introduction

Being reformed only recently, the Emission Trading System (ETS) of the European Union (EU) is yet again bound for another major reform. In 2018, the EU strengthened the ETS cap in order to deliver on the 40% emission reduction target by 2030. However, this target will likely be ratcheted-up in the near future: the EU Commission aims for a reduction of 50% or 55% by 2030 to eventually reach emission neutrality in 2050 (European Commission, 2019). As the EU ETS covers more than 40% of total EU emissions, its stringency needs to be ramped up to reach this target. The regulatory entry point is the review of the Market Stability Reserve (MSR) planned for 2021. The MSR started operating in 2019 and is a mechanism that reduces the total number of allowances in circulation (TNAC[1]) and ultimately cancels allowances based on a complex mechanism. As such it affects the overall cap and therefore should be considered when increasing the stringency of the EU ETS.

The purpose of this paper is twofold: (*i*) to analyse which MSR parameters have a significant effect on the cap size by affecting MSR cancellations; (*ii*) to show which linear reduction factor (LRF[2]) would achieve a given 2030 emission targets when considering the interaction with the MSR. This is of importance for the MSR review in 2021 and in particular for reforming the EU ETS towards higher stringency. We conduct our analysis in four steps.

First, we provide the policy background and briefly review the main results of the quickly growing literature on the MSR. We find that there is a broad range of MSR cancellation estimates from the literature (from 1.7 Gt to 13 Gt, making up 4% to 32% of the total pre-MSR

[1] In literature one could also find the terms "allowances bank" or "surplus", but we use TNAC as it is used in the official EU reports.
[2] The LRF sets the level of ambition of the EU ETS. It is the percentage of the average total quantity of allowances issued annually in 2008-2012 at which the cap decreases each year. In phase 3, the LRF being 1.74%, this amounts to a reduction of 38,264,246 allowances each year (European Commission, 2016).

budget[3]), these results being driven mainly by the assumed discount rate and the marginal abatement cost (MAC) curves considered.

In a second step, we conduct our own analysis relying on the highly detailed electricity sector and industry model LIMES-EU with an endogenous representation of the MSR mechanism. In our reference scenario, we find moderate cancellations of about 5.1 Gt.

Third, we analyse the effect of a broad range of key MSR parameters which potentially are adjusted in the upcoming MSR review in 2021: intake and outtake rates of allowances into the MSR, TNAC thresholds that determine when in- and outtake begins, and auction shares for newly issued allowances. We show that cancellation is more sensitive to the upper than to the lower TNAC threshold. Moreover, increasing the intake rate has a rather limited effect on cancellations but it could induce an oscillatory behaviour on the TNAC due to a discontinuous MSR intake. Increasing the auction share as envisaged by policy makers reduces cancellations and thus also needs to be considered when calibrating the MSR.

Fourth, we analyse the effect of increasing the LRF in order to comply with the more ambitious 2030 targets that arise from the EU Green Deal. We find that a higher LRF not only directly decreases the cap, but also leads to significantly more MSR cancellations. Under our default assumptions, the LRF would therefore only need to be increased from currently 2.2% to 2.4% and 2.6% to bring 2030 emissions in line with an overall 50% and 55% emission reduction target by 2030, respectively.

To the best of our knowledge the only other work in this direction is Quemin (2020). In the same vein as our analysis it assesses both potential changes in the MSR parameters in light of the 2021 review, and how to raise ambition through the LRF and the MSR. The main difference

[3] We calculate the total budget (ignoring MSR cancellations) from 2018 until the last allowances are issued (2057 if the LRF stays 2.2% after 2030) to be 40.1 Gt.

to our work is that Quemin (2020) assumes that firms have a rolling horizon rather than an infinite horizon as in our model, and conducts simulations based on stylized cost function rather than a detailed sectoral as ours. Notably because of the former he finds that combining the MSR and the LRF is more efficient than solely relying on the LRF, which stands in contrast to our recommendation that assess regulatory complexity more broadly.

In the next section we review the latest two ETS reforms and the scientific literature that analyses them. In Section 3 we describe the model, show the results of the reference scenario that uses current ETS parameters and examine alternative parameters. In Section 4 we discuss the results and conclude.

2. The EU ETS and its recent reforms

In this section we shortly present the policy background, review previous work that has analysed the effects of the MSR and provide an overview about cancellation estimates.

The EU ETS covers the power sector, energy intensive industry and aviation which made up more than 40% of GHG emissions in 2017 in the regulated regions – the EU, Iceland, Liechtenstein and Norway (EEA, 2019a, 2019b). Firms are allowed to bank EU allowances (EUA) between years without restriction.[4] This implies that allowances prices are linked over time and, according to theory, they should rise at the discount rate due to intertemporal arbitrage (Rubin, 1996). Yet, in practice prices remained very low (between 3 and 9 €/t) until the beginning of 2018. The low price is attributed to several reasons as for example the economic downturn (e.g. Ellerman et al., 2016; Fuss et al., 2018; van Renssen, 2018). At the same time the TNAC grew continuously between 2008 and 2013 up to 2.1 Gt, which was interpreted as

[4] However, borrowing from future periods is not allowed.

"structural supply-demand imbalance" (European Parliament and Council of the European Union, 2015).

In order to tackle these imbalances, increase the resilience regarding future imbalances and to bring the ETS on track to reach the 2030 emission targets, the MSR was implemented in 2015 to start its operation in 2019 (European Parliament and Council of the European Union, 2015). Basically, the MSR reduces the supply of allowances if the TNAC reaches a certain upper threshold and transfers them into the MSR instead. Allowances are released from the MSR if the TNAC drops below a certain lower threshold. Since this initial version of the MSR was cap-neutral it was expected to have only a weak effect on the EUA price (Perino and Willner, 2016) and may even deter clean long-term investments (Perino and Willner, 2019). In addition, the MSR might also raise the price volatility (Kollenberg and Taschini, 2019; Mauer et al., 2019; Perino and Willner, 2016; Richstein et al., 2015), although Fell et al. (2016) find the opposite.

Facing this criticism and since the price indeed did not increase significantly, the European Parliament and Council of the European Union (2018) agreed to reform the MSR even before it came into effect in 2019: first, if the amount of allowances in the MSR exceeds the amount of auctioned allowances, allowances are permanently cancelled from 2023 onwards and, second, the intake into the MSR is increased until 2023. In addition, it was agreed to raise the LRF from 1.74% to 2.2% for phase IV of the EU ETS (2021-2030) which significantly reduces the cap. The price surged and stabilized in the range of 20-30 €/t since 2019, which suggests that the reform indeed created the expectations of a more stringent ETS.

This reform has evoked a wave of studies on the new MSR version and especially on the cancellation mechanism.[5] Table 1 provides an overview of cancellation results from the

[5] In doing so, several papers also examine how MSR cancellations are affected by additional policies such as RES support (Beck and Kruse-Andersen, 2018; Burtraw et al., 2018; Carlén et al., 2019; Gerlagh et al., 2019; Pahle et al., 2019; Perino et al., 2019; Quemin and Trotignon, 2019; Silbye and Sørensen, 2018).

literature ordered from highest to lowest. A crucial reason for the large range from less than 2 to 13 Gt is the variety of assumed discount rates. Since allowance banking is a provision to reduce costs in the future, firms bank less if they discount at a higher rate. Put differently, if firms have a higher discount rate they put a lower weight on the future and thus bank less. A lower bank in turn implies that fewer allowances go into the MSR and therefore also cancellations are lower. Table 1 indeed indicates that cancellations tend to go up if the discount rate is low.

However, the discount rate can only partly explain the cancellations as is in particular shown by the result of Bruninx et al. (2019a): they find the by far highest cancellations (13 Gt) despite a high discount rate. In contrast to the other papers, they use a relative detailed electricity sector model rather than stylized MAC curves.[6] This may explain some of the differences because it affects the timing of the abatement path which again affects the TNAC and thus the inflow into the MSR. For example, a detailed model would capture that abatement may come in "blocks" instead of a continuous decline in emission as implied by simple MAC curves applied in the other papers. Yet, probably the most important reason for the huge amount of cancellations in Bruninx et al. (2019a) is the slope of their implied MAC curve: since they assume strongly increasing MAC, firms bank a large amount of allowances in order to prevent having to pay high costs for deep emissions reductions later.

Table 1. Comparison of certificate cancellations in the literature

Source	Cancellation (Gt)	Discount rate
Bruninx et al. (2019a)	13	10%
Quemin and Trotignon (2019)[a]	10	3%
Quemin (2020)[a]	8.7	3%
Tietjen et al. (2020)[b]	7.6	3%

[6] Mauer et al. (2019) and Tietjen et al. (2020) additionally consider stylized capacity stocks in the electricity sector. However, their models are still stylized.

Source	Cancellation (Gt)	Discount rate
Beck and Kruse-Andersen (2018)	6	5%
Gerlagh et al. (2019)	5.5	5%
Silbye and Sørensen (2018)	5	7.4%
Quemin and Trotignon (2019)[a]	5	7%
Quemin (2020)[a]	4.2	7%
Carlén et al. (2019)	3.4	3.5%
Bocklet et al. (2019)	2	8%
Perino and Willner (2017)[c]	1.7	10%
Mauer et al. (2019)[d]	1.2	10%

Notes: these values correspond to the central, standard or reference scenario of the respective study. Some numbers are taken from figures and thus might not be perfectly accurate. We only include scientific papers with a model horizon longer than 2030 to increase the comparability.

[a] No single standard scenario. 10 and 5 Gt (Quemin and Trotignon, 2019) and 8.7 and 4.2 Gt (Quemin, 2020) refer to scenarios with rolling horizon and infinite horizon of market agents, respectively. Both scenarios include anticipation of MSR effects though.

[b] Tietjen et al. (2020) consider only the electricity sector. They assume that the not covered sectors (mainly energy-intensive industry) receive all permits they require for free as approximately happened in the past. The numbers correspond to the risk-neutral case.

[c] Perino and Willner (2017) assume that the cap decreases exponentially by 2.2% instead of using the 2.2% as a linear reduction factor (as determined by the EU) resulting in a total emission budget of 53.8 Gt, i.e. 33% higher than in our assumptions.

[d] Mauer et al. (2019) consider only the electricity sector. They multiply all MSR parameters by the electricity sector share.

In the following section we conduct our own cancellation estimation. In doing so, we rely on a model that exhibits significantly more power sector details than the model used by Bruninx et al. (2019a). In combination with our assumed MAC curve for the industry our implied total MAC curve is much flatter and therefore we find significantly fewer cancellations despite assuming a discount rate of only 5%.

3. Model analysis

In this section we examine the ETS and, in particular, the MSR. We first describe the model and scenarios and then show the results of the reference scenario which includes the current regulation. Thereafter the impact of parameters that might be adjusted during the upcoming MSR review is analysed. Finally, we assess in section 3.4 the cancellations triggered by a tighter cap and under which LRF the 2030 emission targets of the EU can be reached.

3.1. Model and scenario description

We use the long-term model for the EU electricity sector (LIMES-EU). It simultaneously optimizes investment and dispatch decisions for generation, storage and transmission technologies, and abatement alternatives for the energy-intensive industry in a 5-year time step, from 2010 to 2070 and covers all EU ETS countries except Cyprus, Iceland, Liechtenstein and Malta, but includes Switzerland and the Balkan region. The model captures the variability of supply (namely wind and solar) and demand by modelling each year through 6 representative days, which are calculated through a clustering algorithm (Nahmmacher et al., 2016). For each day, eight blocks of three hours are assumed. The model contains 32 generation and storage technologies, including different vintages for lignite, hard coal and gas. The energy-intensive industry is included through a MAC curve, which is derived from (Gerbert et al., 2018). We implement the EU ETS with intertemporal banking according to Rubin (1996). This implies that the ETS price grows at the interest rate (assumed to be 5%) as long as the TNAC is positive. More detail on data sources, parameters and the model equations is available in the LIMES-EU documentation (Osorio et al., 2020).

In the reference scenario, we set all ETS parameters to their current values and assume that they remain at these values after 2030 (current regulation only defines values until 2030). The LRF determines by how much the issued allowances are reduced each year. The LRF is 1.74% until

2020 and increases to 2.2% as of 2021 which implies that allowances would be supplied until 2057. See Appendix A for an elaborated description of the cap estimation. Due to a lack of real world guidance we assume in all scenarios that allowances cannot be banked after 2057 and thus we constrain emissions from the EU ETS sector to zero after 2057. We feel this assumption is appropriate when policy targets are analysed since the EU aims for emission neutrality in 2050. Hence it seems to be implausible that firms bank certificates for decades after 2050 when emission should be zero. We elaborate on the implications of this assumption in Appendix C. The initial TNAC (end of 2017) is 1.65 $GtCO_2$ (EEA, 2018) and we assume that the MSR has an extra intake in 2019 and 2020 of 1.55 Gt in total[7]. In the reference scenario the share of allowances to be auctioned is set to be 57% over the entire model horizon, while the remaining 47% are allocated for free[8].

The MSR is modelled based on its operation rules: (*i*) allowances are withheld from auctioning and transferred to the MSR when the TNAC of the previous year, is higher than 833 Mt, the intake to the MSR equalling a share of the TNAC level (24% until 2023 and 12% afterwards); (*ii*) allowances are transferred back from the MSR to the market when the TNAC of the previous year is lower than 400 Mt; the outtake from the MSR (available through auctions) equals 100 Mt (unless the level of the MSR is lower); and (*iii*) when the size of the MSR stock is higher than the number of certificates auctioned in the previous year, the difference between both is cancelled from the MSR. Given the non-linearity of the MSR conditions, it is not possible to embed such equations directly in LIMES-EU. In addition, embedding the MSR into LIMES-EU would be inconsistent with the perfect competitiveness assumption in the model. We thus

[7] This corresponds to 900 Mt that were not auctioned between 2014 and 2016 (backloading) which go directly into the MSR (European Parliament and Council of the European Union, 2015). The remaining 650 $MtCO_2$ (350 $MtCO_2$ in 2017 and 300 $MtCO_2$ between 2018 and 2020) are the estimated unallocated certificates until 2020 (European Commission, 2015). We assume that 250 $MtCO_2$ are transferred in 2019 and 1300 $MtCO_2$ in 2020, as suggested by Burtraw et al. (2018).

[8] The targeted auction share from 2021 onwards is 57% (European Parliament and Council of the European Union, 2018). Notice that the auction share before 2021 is not relevant for our analysis as the difference between auctioning and free allocation only affects the functioning of the MSR.

couple LIMES-EU with a simulation of the MSR, following an iterative approach described in detail in Appendix B.

Table 2 summarizes the parameter values used in the reference scenario (current values) and additionally shows the range used in our analysis. All the variations are implemented after 2023 because we consider this as a plausible first year for new parameters since the MSR review is in 2021.

Table 2. Overview of analysed ETS parameters

Parameter	Current values (reference)	Analysed range (after 2023)
Linear Reduction Factor (LRF)	Until 2020: 1.74% After 2020: 2.2%	1.7-6.0% (step of 0.1%)
Thresholds[a]	Lower threshold: 400 Mt Upper threshold: 833 Mt	0-1500 $MtCO_2$ (step of 100 $MtCO_2$)
Intake rate	Until 2023: 24% After 2023: 12%	0-100% (step of 2%)
Outtake parameter	100 Mt per year	0-1000 $MtCO_2$ (step of 100 $MtCO_2$)
Auction share	57%	0-100% (step of 10%)

[a] We evaluate all possible combinations within that range (in step of 100 $MtCO_2$)

In the next sections we present the results focussing on cumulative emissions and MSR cancellations. In our model cumulative emissions are always equal to the pre-MSR cap (resulting from the LRF) minus the MSR cancellations. Hence for a given LRF, cancellations determine cumulative emissions. However, from policy perspective annual targets are often of greater importance. We thus also present emissions and carbon prices for 2030 as this year is the current focal point of EU climate policy.

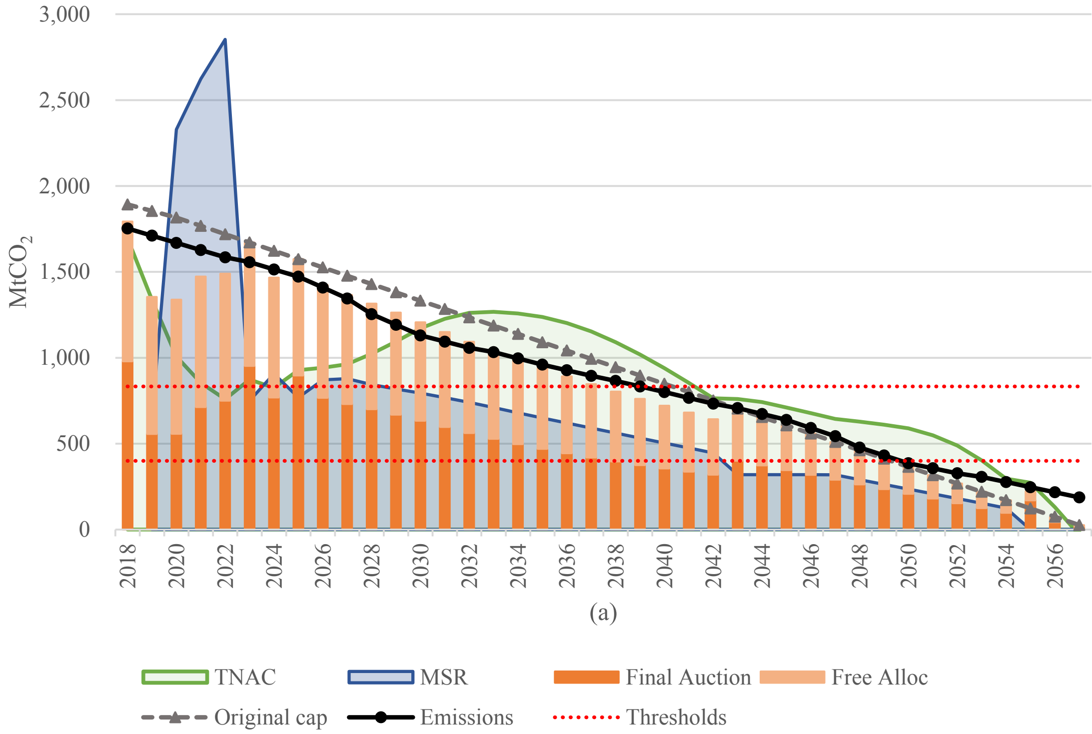
3.2. Reference scenario

Figure 1 shows the main variables determining the long-term dynamics of the EU ETS, including the MSR. While (a) shows the TNAC and MSR levels as well as parameters and variables that influence them, (b) shows the MSR level and the flows that determine it. The TNAC changes as a result of the annual difference between emissions and supply of certificates as well as the TNAC level in the previous time step. The supply consists of freely allocated allowances (43% of the original pre-MSR cap) and auctioned allowances. The actual auctioned volume in turn depends on the TNAC level: fewer certificates are auctioned and instead flow into MSR (intake) when the TNAC is higher than the upper threshold, while additional certificates from the MSR (outtake) are auctioned when the TNAC is lower than the bottom threshold. The two thresholds are indicated by the dotted lines in part (a) of Figure 1.

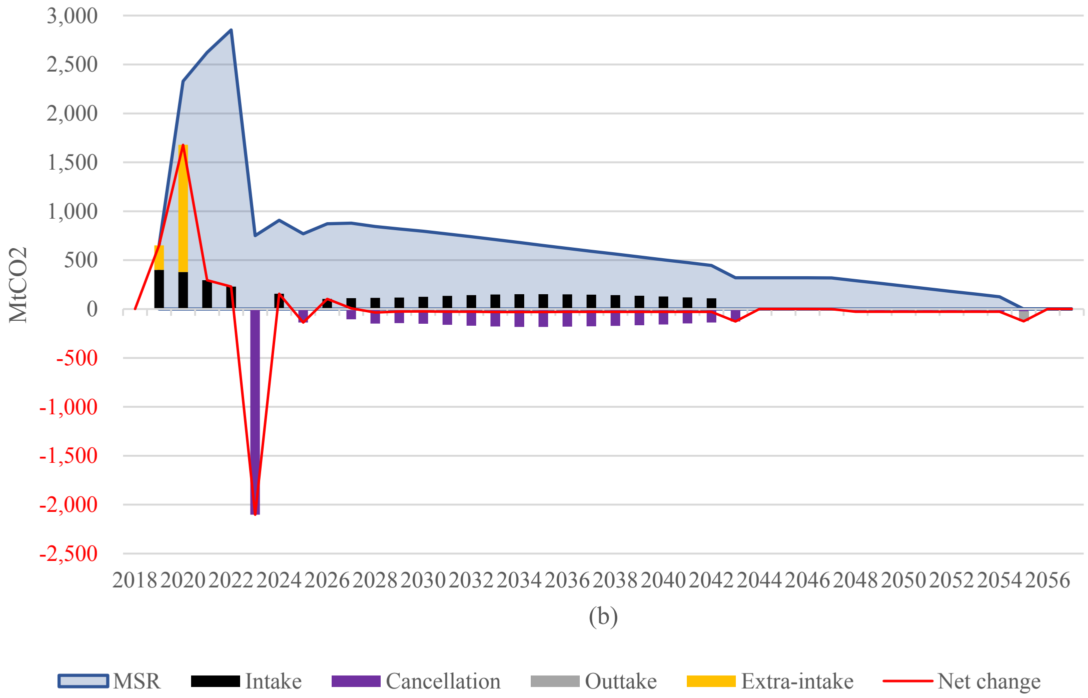
While the TNAC decreases until 2022, the MSR level quickly rises, achieving a maximum of 2853 MtCO_2 in 2022, mainly explained by the extra intake of the not issued allowances before the MSR has started (1550 Mt). There is ongoing intake to the MSR between 2019 and 2042 (except for 2023 and 2025). During the same period, the MSR still progressively decreases due to the higher cancellation of certificates, which takes place from 2023 to 2043 (except for 2024 and 2026) and later between 2047 and 2055.

Such a prolonged cancellation can be partly explained by the MSR rules itself. Since cancellation is determined by the difference between the MSR level and the auction volume of the previous year, cancellation reinforces itself: cancellation implies a lower total cap and thus higher allowance prices. Consequently, emissions are lower and the TNAC is higher which in turn increases the inflow into the MSR. If more allowances flow into the MSR, first, the MSR level is higher and, second, the auction volume is lower, while both imply more cancellations. In addition, the TNAC increase between 2023 and 2033 is caused by a faster decrease of emissions in this period (see Figure 1). As a result, the TNAC remains above the upper threshold (833 $MtCO_2$), which in turn triggers the intake to the MSR and later the cancellation.

The TNAC remains between both thresholds from 2042 to 2053, i.e., there are no transfers from or to the MSR. When the bank falls below the bottom threshold (400 $MtCO_2$) in 2054 and triggers the outtake from the MSR, the MSR level is already at a very low level (125 $MtCO_2$), which limits the reinjection of certificates into the market (outtake only takes place in 2055). In total, from the 5243 $MtCO_2$ certificates withdrawn from the market (including the extra intake in the beginning), 5143 $MtCO_2$ are cancelled, i.e., 98%, with the majority of the cancellation occurring before 2030 (2787 $MtCO_2$, i.e., 54% of total cancellation). As a result, cumulative EU ETS emissions from 2018 until 2057 amount to 34.9 Gt.



(a)



(b)

Figure 1. EU ETS dynamics. (a) TNAC and MSR levels as well as parameters and variables that determine influence them; (b) MSR level and its flows.

We perform a sensitivity analysis on key assumptions in Appendix C. We show that the amount of cancellation is relatively robust towards changes in the abatement costs (fuel prices, electricity demand, technology costs). However, we confirm the strong effect of the discount rate found in the literature, and show the relevance of the banking horizon as already mentioned above.

3.3. Analysis of MSR parameters

In this section we show how the MSR parameters, namely the thresholds, intake rate, outtake rate and share of auctions, affect cancellations.

3.3.1. Thresholds

Figure 2 shows the impact of the lower (currently 400 Mt) and upper (currently 833 Mt) thresholds on EUA cancellation. Cancellation is highest when the lower and upper thresholds are lowest and vice versa. For the evaluated thresholds, cancellation remains within the range of 3.1 to 7.9 Gt, meaning that between 8% and 20% of the total pre-MSR budget since 2018 is cancelled.

It can be observed that the cancellation is more sensitive to the upper threshold than to the lower threshold: cancellations decrease on average 93 $MtCO_2$ for a 100 $MtCO_2$ lower threshold increase, where it is 226 $MtCO_2$ for a 100 $MtCO_2$ upper threshold increase. To understand why, recall that the upper threshold mainly affects the intake to the MSR and the lower threshold the outtake from the MSR. Since the inflow into the MSR occurs in the cancellation phase (see Figure 1(b)), almost all allowances that go into the MSR are cancelled. The reason is that cancellation is the difference between the MSR level (higher due to more inflow) and auctioned certificates (lower due to more inflow) if it is positive. Hence additional inflow (due to a lower upper threshold) is more or less immediately cancelled.

The lower threshold, in contrast, plays only a role later on when the TNAC is low enough (recall Figure 1(a)). However, at this time many allowances have already been cancelled such that not many allowances actually can leave the MSR. Essentially, the MSR level cannot be higher than the auction volume because of the cancellation mechanism. Therefore there is only little room for the lower threshold to have an effect on cancellations.

In addition, price effects reinforce the effect of varying the thresholds. With more intake (lower upper threshold), there is more cancellation and hence a higher price. This in turn leads to a higher TNAC, and thus to more intake and eventually more cancellation. A lower bottom threshold has in principle a similar effect as it leads to lower prices and eventually to less cancellation and vice versa. Yet, again after cancellation takes place not many allowances are left such that outtake generally is relative low in our model.

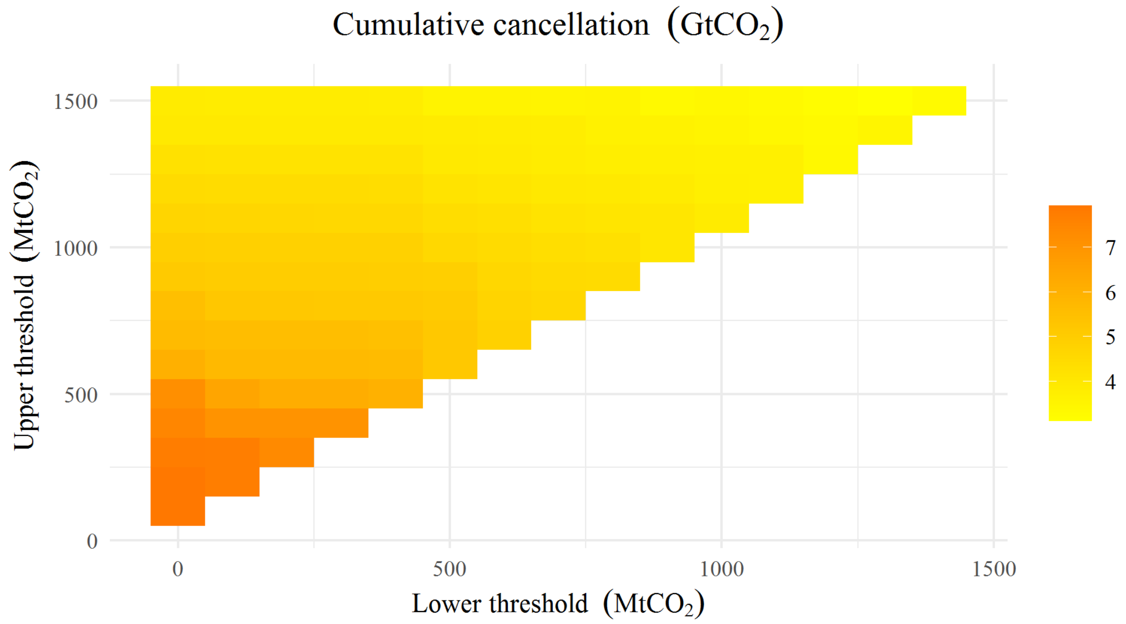


Figure 2. Impact of MSR thresholds on certificate cancellation.

3.3.2. Intake and Outtake rates

Figure 3 shows the impact of the intake rate on cancellations. Notice that even when the intake rate after 2023 is 0%, cancellation still equals 2.6 Gt because 2.7 Gt are transferred to the MSR before 2023 and outtake equals 0.1 Gt. While cancellation increases sharply for rates between 0 and 12% (from 2.6 to 5.1 Gt), it is hardly affected between 12% and 50% (5.1 ± 0.1 $GtCO_2$) and only slightly increases when the rates are higher than 50% and remains within the range of 5.0 to 5.9 Gt. The maximum cancellation (5.9 $GtCO_2$) occurs when the intake rate is 58%. Hence compared to current regulation (12% after 2023) that leads to a cancellation of 5.1 Gt (reference scenario) a higher intake rate only has a moderate effect on the overall cap.

Moreover, for rates higher than 12% the effect on cancellation is non-monotonic. On the one hand, a higher intake rate raises the transfers to the MSR in years with intake. On the other hand, the TNAC cut-backs are severer and thus a higher rate leads to fewer years in which certificates are transferred to the MSR because the upper threshold is less often reached. If the first effect dominates a higher rate leads to more cancellations and if the second effect is stronger cancellations go down.

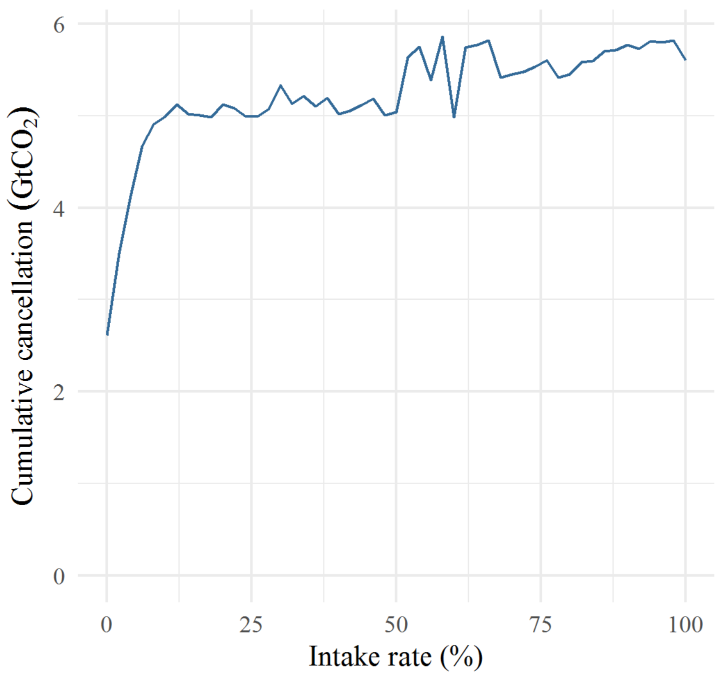


Figure 3. Impact of MSR intake rate after 2023 on certificates cancellation.

Taking a closer look to the intake volumes and TNAC levels in Figure 4 shows these effects. When the rates are 12%, 24% and 36% the maximum annual intakes are respectively 157, 231 and 340 Mt during the period 2024-2043. Within the same period there are respectively 1, 9 and 12 years in which transfers to the MSR do not occur. This is because the TNAC oscillates around the upper threshold (red dashed line at 833 $MtCO_2$) more frequently when the intake rate is high whereas it is constantly above the threshold when the rate is low (see Figure 4).

This unveils a potential risk posed by high intake rates. While aggregate cancellation it not much affected (see above), moderately higher intake rates may induce some instability. Although we do not model this explicitly, it is plausible that under uncertainty a TNAC that is closely below or above the upper threshold may cause some additional price jumps or higher price volatility because only a very little change in emissions and thus TNAC levels, may imply that the threshold is reached or not. If it is reached significant fewer allowances are issued in the next year and potentially cancelled implying a higher price and vice versa. Moreover, even small firms relative to the market size could try to affect the market outcome because only a relative small amount of allowances is needed. For example, increasing emissions may reduce the TNAC such that the upper threshold is not reached, implying that more allowances are issued in the future.

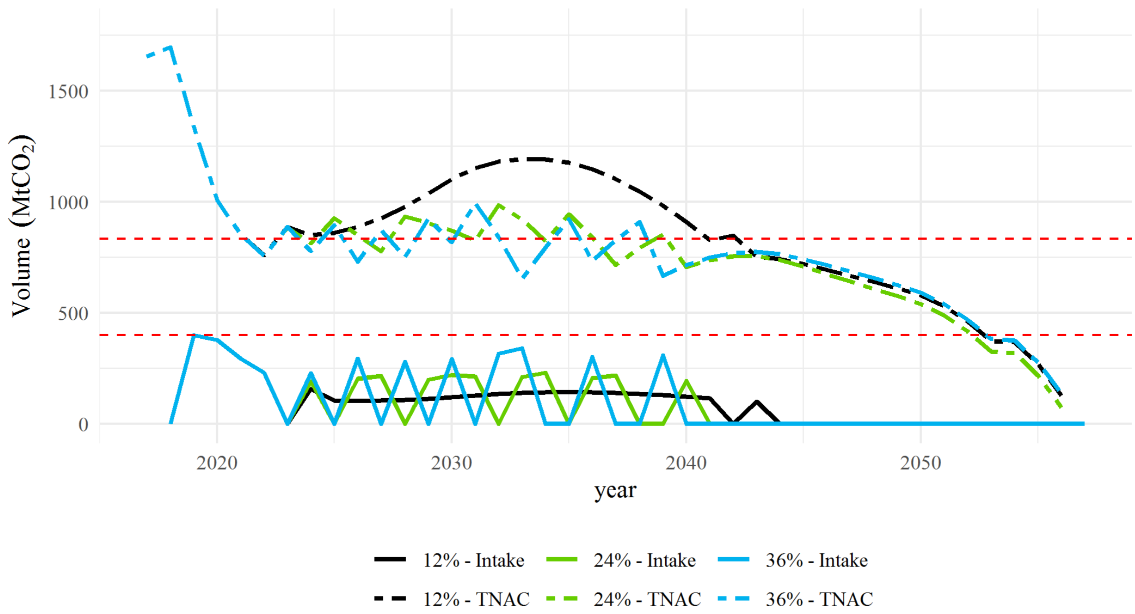


Figure 4. Effect of the intake rates on transfers to the MSR. The red dashed horizontal lines represent the MSR thresholds.

The outtake parameter is an absolute value (currently 100 Mt) that determines the outflow from the MSR when the TNAC is below the lower threshold. However, as explained in the previous section, the cancellation mechanism leaves only a small amount in the MSR. Accordingly we find that the outtake parameter plays only a very limited role (not shown). When the outtake rate is 0 Mt, cancellation reaches 5.3 Gt, while outtake rates higher than 200 Mt lead to cancellations slightly lower than 5.1 Gt.

3.3.3. Auction share

Lastly, we analyse the effect of the auction share which is under current regulation targeted to be 57%. The auction volume is relevant for two reasons: first, cancellation is determined as the difference between the MSR level and the auction volume of the previous year when this difference is positive; second, the intake to the MSR is subtracted from the auction volume and thus the auction volume constrains the annual intake. Therefore, increasing the auction share allows for more intake to the MSR eventually increasing cancellations, but it also leads to less cancellations for a given MSR level.

As shown in Figure 5, the relationship between cancellation and auction share is non-monotonic. In the hypothetical case when the auction share is zero (i.e. all certificates are freely allocated instead), cancellation is limited to the total intake to the MSR of the 2018 – 2023 period. Increasing the auction share to 20% raises cancellations because it softens the constraint on the annual intake (see above) and therefore more allowances flow into the MSR. For auction shares above 20% the other effect dominates: the difference between the MSR level and the auction volume is lower in many years, implying fewer cancellations. Overall, cancellation is highest when the auction share is 20% (5.3 $GtCO_2$) and lowest when it is 0% (2.6 $GtCO_2$). Moderate deviations from the current auction share of 57%, however, only have small impact on cancellations.

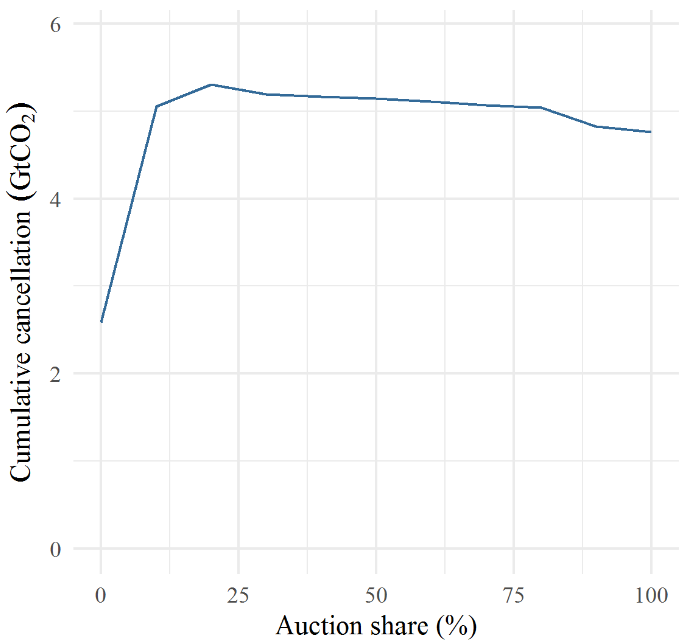


Figure 5. Cumulative cancellations (2018-2057) as a function of the auction share.

3.4. Achieving more ambitious climate targets: The interaction of the MSR with increased LRFs

The EU Green Deal contains a tightening of the 2030 GHG emission targets to -50 or -55% vs. 1990 compared to current -40%. This will in turn require an update of the ETS cap, determined by the LRF. However, the effective cap not only depends on the LRF, but also on cancellations through the MSR. Accordingly, the interplay between LRFs and MSR needs to be considered, which we do in the next section. Based on this, we subsequently analyse how the LRF should be adjusted to reach the 2030 targets.

3.4.1. Interaction between LRF and MSR

The net budget of allowances depends on MSR cancellations in a non-trivial manner. This can be seen in Figure 6, which depicts cumulative emissions and cancellations as a function of the LRF. The sum of both reflects the total gross (pre-MSR) budget of allowances. Looking at cancellations, they increase substantially when the LRF increases from 1.7% (2.6 $GtCO_2$) to 2.6% (9.8 $GtCO_2$). The reason is a reinforcement mechanism between the LRF and the cancellations, which can be disentangled into two effects. First, a higher LRF implies lower supply of certificates, and thus higher prices, with an equal percentage-wise increase in each time step. At the same time, the changes in the LRF have only a small effect on near-term caps, but a large effect on caps in 2040 and 2050 due to the basic linear rule for calculating the cap for any year. Thus, emissions in the first decade decline due to increased prices but annual caps are hardly reduced, which leads to an increase of TNAC, which in turn increases the inflow into the MSR and results in more cancellations. Second, a higher LRF raises cancellations because they depend on the number of auctioned allowances: in each year allowances in the MSR above the auction volume of the previous year are cancelled.

Assuming the continuation of the MSR as currently implemented, the interaction effect is particularly sensitive to changes in the LRF in the range of 2% to 2.6%. For example, increasing the LRF from 2.2 to 2.6% reduces the overall net allowances budget by 4.6 Gt due to the lower LRF itself and additionally by 4.7 Gt due to more cancellations, leading to overall 9.3 Gt lower cumulative emissions.

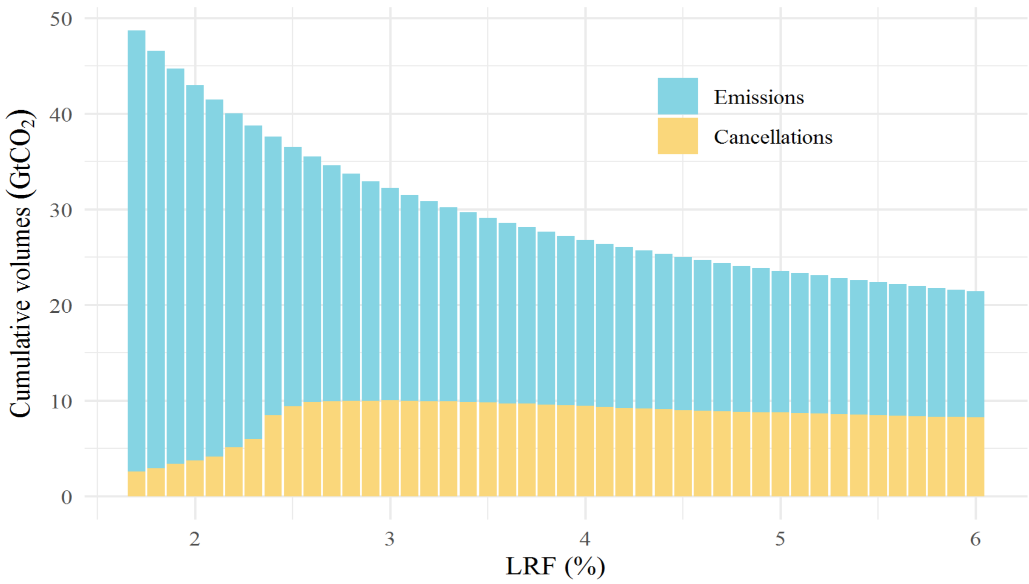


Figure 6. Impact of the LRF on certificate cancellations and cumulative 2020-2055 emissions.

However, for LFRs larger than about 2.6% the reinforcing effect of the LRF and the MSR cancellation becomes significantly weaker and even reverses the sign: the highest cancellation is reached when the LRF is 3.0% (10.0 Gt), and afterwards cancellation decreases to 8.2 Gt when the LRF is 6.0%. The reason for this declining effect is that the transfers to the MSR are constrained by the certificates to be auctioned. Put differently, if the cap becomes smaller, auctions decline, and therefore the amount of certificates that could potentially be cancelled also decreases. Still, the share of cancellations from the total pre-MSR budget (the sum of cumulative emissions and cancellations) increases from 5% to 39%.

We also analyse a simultaneous modification of the LRF, TNAC thresholds and intake rate. We find that the intake rate has a larger impact and the TNAC thresholds a lower impact when the

LRF is increased. However, the degree of interaction between the LRF and the MSR parameters is limited and the main qualitative insights do not change. We elaborate on these mechanisms in Appendix D.

3.4.2. Readjusting the LRF to achieve more ambitious climate targets

We contrast two types of approaches for setting the LRF in order to achieve more ambitious emission targets. In the "conservative approach", policy makers simply calculate the LRF that would be needed to make the annual cap in 2030 equal to the 2030 target as derived from the Green Deal targets. This approach ignores the effects of banking and MSR cancellations on resulting 2030 emissions: if firms use allowances in 2030 banked from previous years, the actual emissions could exceed the target. Vice versa, if firms bank allowances in 2030 in expectation of higher decarbonization challenges after 2030, emission would be lower than the 2030 cap. Moreover, the MSR endogenously adjusts the issued allowances and cancels an unknown number. Since cancellations are ignored we consider this as a conservative approach as emissions very likely will not exceed the target. In the "expected emissions approach", the expected effects of banking and MSR cancellations are included in the calculation of the 2030 emissions. This approach minimizes the LRF required to achieve a given emission target, while at the same time increasing the risk that the 2030 target will be missed because cancellations turn out smaller than expected.

For each LRF, we calculate the resulting emissions using our model (keeping the current MSR parameters). First of all, note that in our model there is a unique optimal carbon price path for a particular amount of cumulative emissions.[9] Since we assume the same abatement costs in all scenarios (apart from the sensitivity analysis in Appendix C) there is thus a one-to-one

[9] This is a well-known result from the literature, e.g. Rubin (1996) and Salant (2016). The reason is that we assume perfect intertemporal markets and that the EU ETS allows to freely bank allowances when the TNAC never drops to zero before the total budget of allowances is exhausted which is the case in all scenarios of this paper.

correspondence between the price and emissions in any year and the cumulative emissions, which in turn are determined by the cap set via the LRF and the cancellations. Figure 7 shows emissions in 2030 and cumulative emissions for LRF values ranging from 1.7% to 6.0%. The figure also includes the different 2030 targets: the current -40% target, and the potential new -50% and -55% targets which translate to -43% (1352 $MtCO_2$), -56% (1044 $MtCO_2$) and -63% (878 $MtCO_2$) for the EU ETS, respectively, if the current split of efforts between ETS and Effort Sharing Regulation is kept constant[10].

First note that the current 43% target is reached even with a LRF of 1.7% (minimum evaluated), i.e., it is clearly achieved with the current cap. Recall that under the current EU ETS and MSR configuration (reference scenario, see Section 3.1), cancellations amount to 5.1 Gt, leading to 35 Gt cumulative emissions. This implies 2030 emissions of 1109 Mt, much lower than the current 2030 target of 43% reduction, i.e., 1352 Mt.

[10] The current target establishes an EU-wide reduction of 40% with respect to 1990 emission level. Accordingly, the EU sets a target of 43% reduction for the EU ETS (i.e., 1018 $MtCO_2$) and of 30% reduction for the Effort Sharing Regulation (ESR) (i.e., 857 $MtCO_2$) with respect to 2005 emissions (2368 and 2855 $MtCO_2$, respectively). This implies that the ETS is expected to contribute 54% of emissions reduction by 2030. If the EU-wide target increases to 55% (i.e., 15% more), then 859 $MtCO_2$ additional reductions are required in 2030. Assuming the same contribution as for the current policy, we estimate that emissions in the EU ETS would need to reduce additionally by 467 $MtCO_2$, i.e., 1485 $MtCO_2$ in total. Such volume implies a 63% reduction compared to the 2005 value. Likewise, an EU-wide reduction of 50% would imply a reduction of 1249 $MtCO_2$ in the EU ETS in 2030, i.e., 56% reduction with respect to 2005.

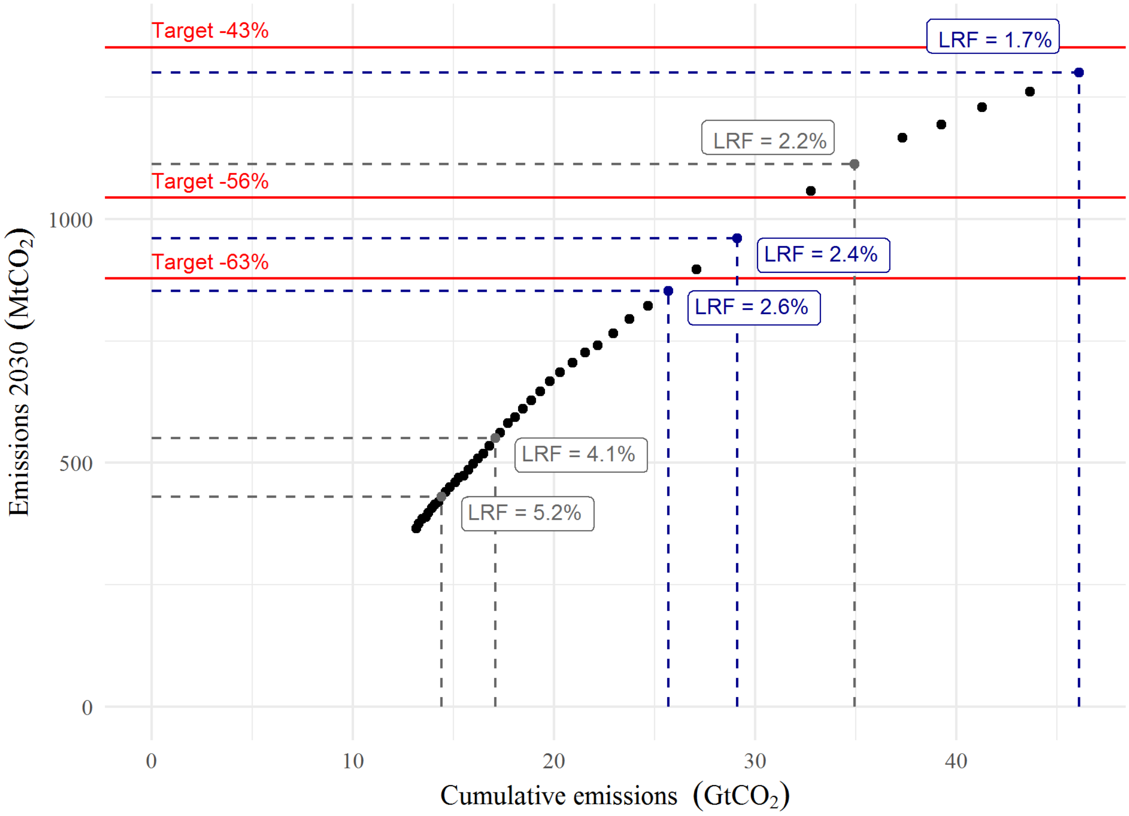


Figure 7. Correspondence between cumulative emissions and emissions in 2030 under various LRF values. The required level of emissions in 2030 to reach the target is shown in red lines. The labels show the resulting emissions in 2030 and cumulative values for the 'expected emissions' (blue) and 'conservative' (grey) approaches.

What happens when tightening the 2030 targets? Under the "conservative approach", policymakers would choose a cap equivalent to the desired reduction, implying LRFs of 4.1% and 5.2% for -50% and -55%, respectively. Due to banking and the MSR cancellations, the emissions in 2030 resulting from such LRFs would be substantially lower than the targets, as Figure 7 shows. The effective emission reduction in 2030 would be 77% and 82%, respectively, with respect to 2005.

For the "expected emissions" approach, LRF increases are much lower. The -56% target (1044 Mt) is reached when cumulative emissions are lower than 29.1 Gt, namely by increasing the current LRF from 2.2% to 2.4%. The -63% target (878 Mt) is reached when cumulative emissions are lower than 25.7 Gt, i.e., only when LRF is at least 2.6%. Hence, due to the positive effect on cancellations, only a relative modest increase of the LRF is necessary to reach significantly more ambitious targets.

Table 3 summarizes the emission targets in 2030, the LRFs derived from the two approaches, as well as the resulting emissions and CO_2 prices. Note that the required EUA price in 2030 to reach the 63% target (accounting for MSR cancellations) more than doubles compared to the price required to achieve the current target, but still remains far below the 2030 economy-wide CO_2 price levels of 61-169 €/tCO_2 that were found by Knopf et al. (2013) in a multi-model comparison study as being in line with a 40% reduction vs. 1990 of economy-wide EU emissions in 2030. In contrast, the 2030 EUA price connected to the "conservative" implementation of the -63% target would be in the lower half of this range.

Table 3. Emissions and prices implied by 2030 targets

Approach	2030 ETS reductions	Target 2030 emissions	Implied LRF	Resulting 2030 cap	Resulting 2030 emissions with MSR & banking	Resulting 2018-2057 budget with MSR & banking	CO2 Price in 2030
		MtCO_2/yr		MtCO_2/yr	MtCO_2/yr	GtCO_2	€/tCO_2
Conservative	-43%	1353	2.2%	**1353**	1112	40.1	27
	-56%	1044	4.1%	**1044**	550	26.4	67
	-63%	878	5.2%	**878**	429	23.1	76
Expected emissions	-43%	1353	1.7%[a]	1410	**1300**	48.6	16
	-56%	1044	2.4%	1302	**960**	37.6	37
	-63%	878	2.6%	1271	**853**	35.5	43

[a] The lowest LRF examined is 1.7%, for which -43% target is largely achieved. Hence, the implied LRF for achieving such target could be lower.

Note: cumulative emissions refer to the period after 2018. Numbers in bold and underlined highlight the values that are brought as close to the target emissions as possible under a given approach by varying the LRF. For the "conservative" approach, the LRF can be directly calculated; in the "expected emissions" approach, the implied LRF correspond to that of the scenarios in which the resulting emissions in 2030 are closest (and below) the corresponding target, thus the resulting emissions do not exactly match the target emissions.

4. Conclusion and Policy Implications

In this paper we analyse key EU ETS parameters with a view on the upcoming MSR review and a potentially broader reform of the EU ETS to reach more ambitious emission targets. We find that under the current regulation the reduction of the cap through cancellations of allowances amounts to 5.1 Gt.

Analysing the MSR parameters we find that especially the upper threshold of the TNAC has a significant impact on cancellations and thus on the cap: when the threshold is decreased from the current 833 Mt to 100 Mt, about 2.8 Gt more certificates are cancelled because a lower threshold implies more inflow to the MSR. Since a high share of the allowances in the MSR is always cancelled, more inflow means more cancellations. This also implies that the bottom threshold which determines the outflow from the MSR is of lower relevance: since a high share of allowances that go into the MSR is cancelled anyway, only a low share can actually leave the MSR regardless the bottom threshold level.

Furthermore, we find that cancellation would strongly decrease if intake rate were decreased from 12% to 0%, from 5.1 Gt to 2.6 Gt. However, intake rates above 12% only have a small additional effect but may lead to discontinuous cancellation and intake because the TNAC fluctuates around the upper threshold relevant for intake. This may have undesirable side effects because small deviations in the TNAC decide whether the threshold is reached or not, potentially implying higher price volatility and a larger impact of firms applying market power. A dedicated analysis of uncertainty and market power in the context of the MSR is, however, left to future research.

We further show that the MSR has another so far under acknowledged side effect, as the share of auctioned certificates has an impact on the cap. In total we find that increasing the auction share raises cancellations up to an auction share of 20% and reduces cancellations in the range of 20% to 100% (currently targeted to be 57%). However, in the practically most relevant range above 20% the effects are relative weak (less than 0.5 Gt difference in cancellations).

We further show that cancellations may vary significantly when the LRF is increased. Beyond the direct cap decreasing effect, a higher LRF also indirectly affects the cap through MSR cancellations. Up to a LRF of about 2.6% (currently 2.2%) we find a strong positive feedback between the LRF and cancellations. However, this effect declines with higher LRF and becomes negative from a LRF of about 3% onwards, though the negative effect is weak. We additionally find that – keeping all other parameters fixed – a LRF higher than 2.4% and 2.6% could be in line with the potential new 2030 EU emission targets of 50% and 55%, respectively, if banking and cancellations turn out as our model suggests.

However, the actual number of cancelled allowances can vary considerably depending on key design parameters set by policy makers, but also on market actors' time horizons and discount rates, as well as their expectations about the future costs of abatement. For instance, increasing the discount rate from 5% to 7% leads only to 2.1 Gt cancellations and decreasing it to 3% leads to 10.0 Gt compared to 5.1 Gt in our reference scenario, and the banking horizon proves to be a critical assumption for cancellations under low discount rates (+5 Gt). Put more succinctly, the (unpredictable) *expectations* of market actors about future CO_2 prices and costs will – via the MSR – influence the size of the cap. Investors expecting higher future abatement costs will bank certificates, thereby increasing cancellations and thus increasing abatement costs. This feedback effect to expectations makes it hardly possible to tune the ETS and MSR parameters to reach a certain emission target.

In light of this observation, and given that the aim of the MSR was to stabilize the ETS, make it more resilient against shocks and increase planning certainty, a more profound reform of the MSR seems to be recommendable. A promising way forward would be to trigger in- and outtake from the MSR by prices rather than emissions, developing it into what could be called a "Price Stability Reserve". Such a reserve would turn the ETS into a classical hybrid instrument, which is typically considered to be more efficient (e.g. Roberts and Spence, 1976; Weitzman, 1978). In particular it would consolidate expectations about future CO_2 prices and thus increase planning security for development of and investments into decarbonization technologies.

Acknowledgments

We thank Brigitte Knopf, Christian Gambardella and Samuel Okullo for their helpful comments. This work received funding from the German Federal Ministry of Education and Research (BMBF) under grant agreement No 01LA1810C (FFF). Oliver Tietjen's and Robert Pietzcker's contribution was funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 730403 (INNOPATHS).

References

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

Bocklet, J., Hintermayer, M., Schmidt, L., Wildgrube, T., 2019. The Reformed EU ETS - Intertemporal Emission Trading with Restricted Banking (No. 2019–4), EWI Working Papers. Energiewirtschaftliches Institut an der Universitaet zu Koeln (EWI).

Bruninx, K., Ovaere, M., Delarue, E., 2019a. The Long-Term Impact of the Market Stability Reserve on the EU Emission Trading System.

Bruninx, K., Ovaere, M., Gillingham, K., Delarue, E., 2019b. The unintended consequences of the EU ETS cancellation policy.

Burtraw, D., Keyes, A., Zetterberg, L., 2018. Companion Policies under Capped Systems and Implications for Efficiency—The North American Experience and Lessons in the EU Context (RFF report). Resources for the Future.

Carlén, B., Dahlqvist, A., Mandell, S., Marklund, P., 2019. EU ETS emissions under the cancellation mechanism - Effects of national measures. Energy Policy 129, 816–825. https://doi.org/10.1016/j.enpol.2019.02.060

EEA, 2019a. EEA greenhouse gas - data viewer [WWW Document]. Eur. Environ. Agency. URL https://www.eea.europa.eu/data-and-maps/data/data-viewers/greenhouse-gases-viewer (accessed 1.13.20).

EEA, 2019b. EU Emissions Trading System (ETS) data viewer [WWW Document]. Eur. Environ. Agency. URL https://www.eea.europa.eu/data-and-maps/dashboards/emissions-trading-viewer-1 (accessed 1.13.20).

EEA, 2018. Trends and projections in the EU ETS in 2018. The EU Emissions Trading System in numbers (No. 14/2018). European Environment Agency, Copenhagen.

Ellerman, A.D., Marcantonini, C., Zaklan, A., 2016. The European Union Emissions Trading System: Ten Years and Counting. Rev. Environ. Econ. Policy 10, 89–107. https://doi.org/10.1093/reep/rev014

European Commission, 2019. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE

EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS The European Green Deal COM/2019/640 final [WWW Document]. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1588580774040&uri=CELEX:52019DC0640 (accessed 5.14.20).

European Commission, 2016. Emissions cap and allowances [WWW Document]. Clim. Action - Eur. Comm. URL https://ec.europa.eu/clima/policies/ets/cap_en (accessed 5.29.18).

European Commission, 2015. COMMISSION STAFF WORKING DOCUMENT IMPACT ASSESSMENT Accompanying the document 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 SWD/2015/0135 final - 2015/0148 (COD) (Commission staff working document impact assessment No. SWD(2015) 135 final). Brussels.

European Parliament, Council of the European Union, 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. Off. J. Eur. Union 76, 3–27.

European Parliament, Council of the European Union, 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. Off. J. Eur. Union 264, 1–5.

Eurostat, 2019. Energy balance sheets - 2017 data (2019 edition). Luxembourg.

Fell, H., 2016. Comparing policies to confront permit over-allocation. J. Environ. Econ. Manag., The economics of the European Union Emission Trading System (EU ETS) market stability reserve 80, 53–68. https://doi.org/10.1016/j.jeem.2016.01.001

Fuss, S., Flachsland, C., Koch, N., Kornek, U., Knopf, B., Edenhofer, O., 2018. A Framework for Assessing the Performance of Cap-and-Trade Systems: Insights from the European Union Emissions Trading System. Rev. Environ. Econ. Policy 12, 220–241. https://doi.org/10.1093/reep/rey010

Gerbert, P., Herhold, P., Burchardt, J., Schönberger, S., Rechenmacher, F., Kirchner, A., Kemmler, A., Wünsch, M., 2018. Klimapfade für Deutschland (Study commissioned by the the Federation of German Industries (Bundesverband der Deutschen Industrie – BDI) to the Boston Consulting Group (BCG) and Prognos).

Gerlagh, R., Heijmans, R.J.R.K., Rosendahl, K.E., 2019. Endogenous Emission Caps Always Induce a Green Paradox.

Gomez, D., Watterson, J.D., Americano, B.B., Ha, C., Marland, G., Matsika, E., Namayanga, L., Osman-Elasha, B., Kalenga Saka, J.D., Treanton, K., Quadrelli, R., 2006. 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Volume 2. Energy. Chapter 2: Stationary combustion. IPCC.

ICAO, 2016. On board. A sustainable future. 2016 Environmental report. International Civil Aviation Authority (ICAO).

Knopf, B., Chen, Y.-H.H., De Cian, E., Förster, H., Kanudia, A., Karkatsouli, I., Keppo, I., Koljonen, T., Schumacher, K., Van Vuuren, D.P., 2013. Beyond 2020 — strategies and costs for transforming the european energy system. Clim. Change Econ. 04, 1340001. https://doi.org/10.1142/S2010007813400010

Kollenberg, S., Taschini, L., 2019. Dynamic supply adjustment and banking under uncertainty in an emission trading scheme: The market stability reserve. Eur. Econ. Rev. 118, 213–226. https://doi.org/10.1016/j.euroecorev.2019.05.013

Mauer, E.-M., Okullo, S.J., Pahle, M., 2019. Evaluating the performance of the EU ETS MSR.

Nahmmacher, P., Schmid, E., Hirth, L., Knopf, B., 2016. Carpe diem: A novel approach to select representative days for long-term power system modeling. Energy 112, 430–442. https://doi.org/10.1016/j.energy.2016.06.081

Osorio, S., Pietzcker, R.C., Tietjen, O., 2020. Documentation of LIMES-EU - A long-term electricity system model for Europe (v2.36) [WWW Document]. Model Doc. V236. URL https://www.pik-potsdam.de/research/transformation-pathways/models/limes/limes-documentation-may-2020

Pahle, M., Edenhofer, O., Pietzcker, R., Tietjen, O., Osorio, S., Flachsland, C., 2019. Die unterschätzten Risiken des Kohleausstiegs. Energiewirtschaftliche Tagesfragen 69, 1–4.

Perino, G., Ritz, R.A., van Benthem, A., 2019. Understanding Overlapping Policies: Internal Carbon Leakage and the Punctured Waterbed (Working Paper No. 25643). National Bureau of Economic Research. https://doi.org/10.3386/w25643

Perino, G., Willner, M., 2019. Rushing the Impatient: Allowance Reserves and the Time Profile of Low-Carbon Investments. Environ. Resour. Econ. 74, 845–863. https://doi.org/10.1007/s10640-019-00350-x

Perino, G., Willner, M., 2017. EU-ETS Phase IV: allowance prices, design choices and the market stability reserve. Clim. Policy 17, 936–946. https://doi.org/10.1080/14693062.2017.1360173

Perino, G., Willner, M., 2016. Procrastinating reform: The impact of the market stability reserve on the EU ETS. J. Environ. Econ. Manag., The economics of the European Union Emission Trading System (EU ETS) market stability reserve 80, 37–52. https://doi.org/10.1016/j.jeem.2016.09.006

Quemin, S., 2020. Using Supply-Side Policies to Raise Ambition: The Case of the 2021 EU ETS Review.

Quemin, S., Trotignon, R., 2019. Emissions Trading with Rolling Horizons.

Richstein, J.C., Chappin, É.J.L., de Vries, L.J., 2015. The market (in-)stability reserve for EU carbon emission trading: Why it might fail and how to improve it. Util. Policy 35, 1–18. https://doi.org/10.1016/j.jup.2015.05.002

Roberts, M.J., Spence, M., 1976. Effluent charges and licenses under uncertainty. J. Public Econ. 5, 193–208. https://doi.org/10.1016/0047-2727(76)90014-1

Rubin, J.D., 1996. A Model of Intertemporal Emission Trading, Banking, and Borrowing. J. Environ. Econ. Manag. 31, 269–286. https://doi.org/10.1006/jeem.1996.0044

Salant, S.W., 2016. What ails the European Union′s emissions trading system? J. Environ. Econ. Manag., The economics of the European Union Emission Trading System (EU ETS) market stability reserve 80, 6–19. https://doi.org/10.1016/j.jeem.2016.06.001

Silbye, F., Sørensen, P.B., 2018. Towards a more efficient european carbon market.

Tietjen, O., Lessmann, K., Pahle, M., 2020. Hedging and temporal permit issuances in cap-and-trade programs: The Market Stability Reserve under risk aversion (SSRN Scholarly Paper No. ID 3436736). Social Science Research Network.

van Renssen, S., 2018. The inconvenient truth of failed climate policies. Nat. Clim. Change 8, 355. https://doi.org/10.1038/s41558-018-0155-4

Weitzman, M., 1978. Optimal Rewards for Economic Regulation. Am. Econ. Rev. September, 68, 683–691.

Appendix

A. Construction of annual caps

The cap (before the impact of the MSR) for the stationary sector (i.e. all but aviation) has decreased since 2013 (beginning of Phase III) at a rate of 1.74% (of the average cap during phase II, i.e. 38.3 $MtCO_2$ per year). The resulting cap in 2020 is 1,816 $MtCO_2$ and is set to decrease at a rate of 2.2% (i.e. 48.4 $MtCO_2$ per year) until 2030. We assume the cap keeps decreasing afterwards at the same rate.

Since we do not model the heating-related and aviation emissions explicitly in LIMES-EU, we assume exogenous emissions as follows:

Heating: The combustion sector emissions added up to 1163 $MtCO_2$ (66% of stationary sector emissions) in 2017, accounting mainly for power plants. To differentiate electricity- and heat-related emissions, we estimate them using the primary energy consumed from power plants (Eurostat, 2019) and the emission factors from the IPCC's guidelines (Gomez et al., 2006). We allocate the emissions from cogeneration heat and power (CHP) plants according to the power plants output. We estimate that heating-related emissions added up to 11% of the total stationary sector emissions in 2017. We thus assume exogenous emissions accounting for 11% of the cap for the entire modelling period.

Aviation: this sector has its own cap (about 37 $MtCO_2$ per year have been allocated since 2013), but is allowed to buy certificates from the stationary sector. Emissions have increased from 53 $MtCO_2$ in 2013 to 64 $MtCO_2$ in 2017, the sector having always a negative balance of EU allowances for aviation (EUAA), i.e. airlines have had to buy allowances from the stationary sector to cover their emissions. The EU forecasts aviation emissions (under the current scope of the EU ETS, i.e. only covering intra-

European Economic Area (EEA) flights) to be between 65 and 70 $MtCO_2$ in 2030 (EEA, 2018). However, it is not clear whether the scope will remain, as the current derogation from the EU ETS obligations for flights to and from third countries is extended until 31 December 2023, subject to review. There is also significant uncertainty about the future demand and technical improvements as well as on feasibility of implementing alternative fuels on a large scale (ICAO, 2016). We assume that emissions from aviation remain at 60 $MtCO_2$ per year and the cap – starting in 37 $MtCO_2$ per year in 2020 – decreases at the same pace as the stationary cap. The difference between emissions and the aviation cap are thus subtracted from the stationary cap.

B. Coupling the MSR simulation with LIMES-EU

Since LIMES-EU is a linear model, including the MSR rules as part of the optimisation problem would not be possible. Converting the model into a non-linear one risk the non-convergence of the runs given the size of the model. In addition, the MSR rules are stated on an annual basis, while LIMES-EU runs in a 5-year basis. To reconcile these issues, we couple LIMES-EU with a simulation of the MSR through an iterative process, which we summarize in the flow diagram presented in Figure B.1.

We estimate the cap on an annual basis (p_cap_{t2}), based on the assumed LRF. We 'translate' this cap into a 5-year value (v_cap_t), averaging the corresponding 5 year values to each year in LIMES-EU[11]. For instance, the cap in LIMES-EU in 2020 equals the average of the annual cap

[11] To distinguish the variables computed in LIMES from those computed in the MSR simulation, we name as v_* for the former and p_* for the latter. In addition, the index t is only used for input from or variables used in LIMES-EU ($t = 2015,2020 \ldots .2055$), while $t2$ is only used for those related to the MSR simulation ($t2 = 2017,2018 \ldots .2057$).

between 2018 and 2022. In a first iteration, the certificates supply (v_supEUA_t) equals the cap (v_cap_t).

From the LIMES-EU results, we use the total emissions for the EU ETS (v_emi_t) and the bank at the end of 2015 (v_TNAC_{2015}) as input for the MSR. These 5-year-based inputs nonetheless have to be 'translated' into annual values for the MSR simulation. This is necessary because of the MSR operation criteria, e.g., use TNAC from year *t2-1* to estimate the intake into the MSR in *t2*, works on an annual basis. Recall that each year in LIMES-EU corresponds to the 5 years around it. To smoothen the input, we interpolate the emission volumes between LIMES-EU years and then normalize them to ensure that the 5-years average equals the LIMES-EU value. Unlike emissions, which are a flow, the TNAC in 2015 from LIMES-EU (v_TNAC_{2015}) is a stock. This corresponds to the initial TNAC used in the MSR simulation, p_TNAC_{2017} (TNAC at the end of 2017). From the annual cap, we estimate the preliminary auctions ($p\_prelaucEUA_{t2}$, see Eq. (B.1)) and certificates to be freely allocated ($p\_freeEUA_{t2}$, see Eq. (B.2)).

Other parameters such as the thresholds ($p\_lower\_threshold_{t2}$ and $p\_upper\_threshold_{t2}$), the intake rate ($p\_rateintakeMSR_{t2}$), the outtake rate ($p\_rateoutakeMSR_{t2}$) and the additional intake ($p\_extraintake_{t2}$) are required to simulate the MSR. Once the MSR is simulated, we are able to estimate the intake (Eq. (B.3)), outtake (Eq. (B.4)), cancellation (Eq. (B.5)), MSR level (Eq. (B.6)), certificates to be auctioned (Eq. (B.7)) and TNAC (Eq. (B.8)) on an annual basis as of 2019.

$$p\_prelaucEUA_{t2} = p\_cap_{t2} \times (1 - p\_sharefreeEUA_{t2}) \quad \text{(B.1)}$$

$$p\_freeEUA_{t2} = p\_cap_{t2} \times p\_sharefreeEUA_{t2} \quad \text{(B.2)}$$

$$\text{If } p\_TNAC_{t2-1} > p\_upper\_threshold_{t2},$$
$$p\_intake_{t2} = min\big({}^{2}/_{3}\, p\_TNAC_{t2-2} \times p\_rateintakeMSR_{t2-2} + {}^{1}/_{3}\, p\_TNAC_{t2-1} \times p\_rateintakeMSR_{t2-1}, p\_prelaucEUA_{t2}\big), \tag{B.3}$$
$$\text{in other case } p\_intake_{t2} = 0$$

$$\text{If } p\_TNAC_{t2-1} < p\_lower\_threshold_{t2},$$
$$p\_outtake_{t2} = min(p\_MSR_{t2-1}, p\_rateouttakeMSR_{t2}), \tag{B.4}$$
$$\text{in other case } p\_outtake_{t2} = 0$$

$$p\_cancellation_{t2} = 0 \quad \forall t2 \leq 2023$$
$$p\_cancellation_{t2} = max(p\_MSR_{t2-1} - p\_prelaucEUA_{t2-1}, 0) \ \forall t2 > 2024 \tag{B.5}$$

$$p\_MSR_{t2} = p\_MSR_{t2-1} + p\_extraintake_{t2} + p\_intake_{t2} - p\_outtake_{t2} - p\_cancellation_{t2} \tag{B.6}$$

$$p\_aucEUA_{t2} = p\_prelaucEUA_{t2} - p\_intake_{t2} + p\_outtake_{t2} \tag{B.7}$$

$$p\_TNAC_{t2} = p\_TNAC_{t2-1} + p\_aucEUA_{t2} + p\_freeEUA_{t2} - p\_emi_{t2} \tag{B.8}$$

The intake to the MSR (Eq. (B.3)) is modelled in detail, i.e., the exact time in which allowances are removed from the auctions is considered. The European Commission informs each May about the TNAC by the end of the previous year and about the volume of certificates to be transferred to the MSR. A volume calculated on the basis of the TNAC of a year *t-1* is removed from the auctions between September in year *t* and August of year *t+1*. Since the MSR only starts absorbing certificates in January 2019, 16% of the TNAC in 2017, informed in May 2018

(1.65 $GtCO_2$), i.e., 264 $MtCO_2$, will be transfer to the MSR between January and August 2019[12]. Likewise, the TNAC at the end of 2018, informed in May 2019 (1.65 $GtCO_2$), determined the amount of certificates being removed from auctions between September 2019 and August 2020[13] and transferred to the MSR. Accordingly, it can be assumed that the intake for each year *t* amounts to two thirds of the volume calculated on the basis of the TNAC by the end of *t-2* and one third of the volume calculated on the basis of the TNAC by the end of the year *t-1*, such volume depending on the intake rate.

[12] Communication from the Commission C(2018) 2801 final, available at https://ec.europa.eu/clima/sites/clima/files/ets/reform/docs/c_2018_2801_en.pdf

[13] Communication from the Commission C(2019) 3288 final, available at https://ec.europa.eu/clima/sites/clima/files/ets/reform/docs/c_2019_3288_en.pdf

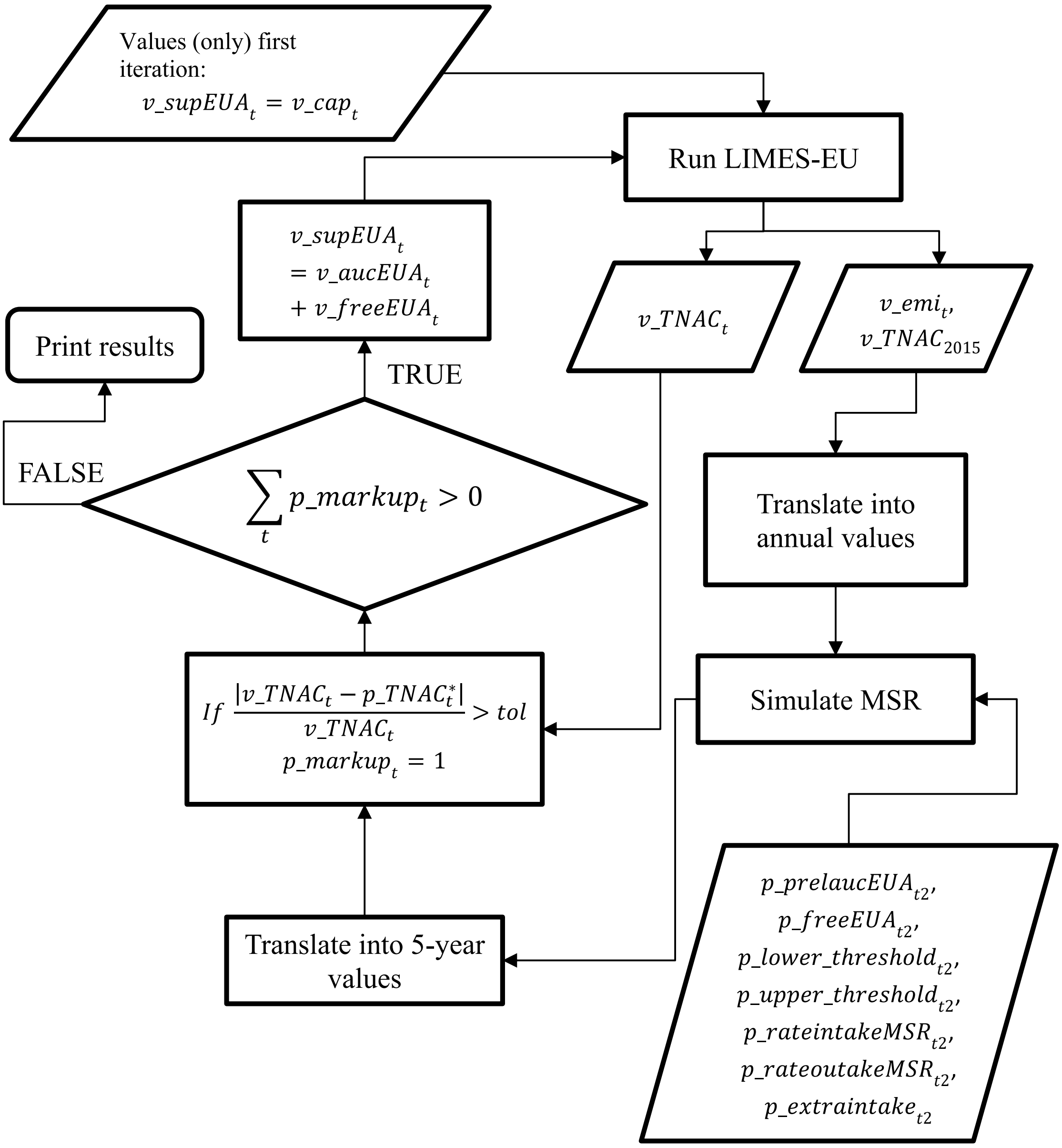


Figure B.1. Iterative process to couple LIMES-EU with the MSR simulation.

This output is 'translated' into 5-year data. For flow-type variables we compute the average for the 5-corresponding years. For instance, the average EUA auctioned (p_aucEUA_{t2}) between 2018 and 2022 is used for the 2020 volume in LIMES-EU (v_aucEUA_{2020}). For stock-type variables, p_TNAC_{t2} and p_MSR_{t2}, we use the value from the last corresponding year. For instance, their level in 2022 would correspond to 2020 in LIMES-EU years. We compute the error between the 'translated' TNAC from the MSR simulation ($p\_TNAC_t^*$) and that from LIMES (v_TNAC_t). If the error is higher than the tolerance margin ($tol = 0.05$) for any t, LIMES-EU is run again with an updated supply of certificates (v_supEUA_t). This equals the sum between the 'translated' free allocated EUA ($v\_freeEUA_t$), which does not change across iterations, and the 'translated' final auctioned EUA (v_aucEUA_t), estimated through the MSR simulation. This process is followed until the TNAC from both LIMES-EU and the MSR simulation converge.

C. Sensitivity analysis

We perform a sensitivity analysis to our main assumptions. First we show the impact of fuel prices (±50% by 2050), capital costs of variable renewable energy sources (vRES) as photovoltaics and wind mills (±30% by 2050), electricity demand (+50% by 2050) and the industry abatement costs (between -50% and +100%)[14]. Second, we provide more details on the impact of the interest rate on cancellations. Finally, we evaluate the impact of a longer banking horizon which we restrict to 2057 in the main scenarios.

Table C.1 shows cancellations for variations in the first set of parameters. Cancellations lie within a range of 4.3 to 7.3 Gt, i.e., 17% lower and 42% higher than in the reference scenario, only low gas price and low industry abatement costs having a significant effect on cancelations. Considering the large variations assumed, this highlights the robustness of our results.

When it comes to fossil fuel prices, cancellation (4.3-6.4 $GtCO_2$) is more sensitive to changes in gas prices (independently of coal prices) because investments in gas plans depend heavily on their marginal costs. When gas prices are low, gas-fired generation displaces that from hard

[14] For all these parameters, except the industry MACC, we assume that they grow linearly between 2020 and 2050 up to the value specified, e.g., electricity demand is 34% higher than in BAU by 2040.

coal, increasing the TNAC and thus cancellation. On the contrary, high gas prices are high result in lower ETS prices, from which industry profits to abate less. With overall higher emissions, the TNAC is lower and thus cancellation too.

A higher electricity demand triggers a higher amount of cancellations (5.9 $GtCO_2$). EUA price increases as a result of the higher electricity demand. As a consequence, industry emissions decrease. However, the rise in electricity emissions, due to larger demand requirements, do not offset such drop. There is thus an overall decrease in emissions, that leads to higher TNAC, and thus to higher cancellations. A similar effect is observed when vRES investment costs vary: more expensive vRES increase the ETS price, which leads to less emissions and thus a larger TNAC in the near-term and overall more cancellations (and vice versa). Hence the MSR tends to amplify the effect of higher abatement costs. From an economic perspective, this is usually not desirable as higher abatement costs should imply a softer cap (see also Bruninx et al. (2019b)).

However, in case of the energy-intensive industry this effect is non-monotonic, despite carbon prices showing a monotonic behaviour, i.e., they are higher when industry MAC are higher. When the industry MAC are 50% cheaper, cancellations are also higher (7.3 $GtCO_2$) than in the reference scenario. Similarly, higher industry MAC lead to more cancellations (e.g., 6.1 $GtCO_2$ when +100%). In the former case, industry abate more but the electricity sector profits from lower carbon prices and emit more. In the latter, higher carbon prices encourage more abatement in the electricity sector, but industry emits more. The overall effect is less emissions, and thus more cancellations.

Table C.1. Sensitivity analysis. Impact of fuel prices, vRES capital costs, electricity demand and the industry abatement costs on cancellations.

Scenario		**Cancellation ($GtCO_2$)**
Reference scenario		5.1
Fossil fuel prices	Low	5.9
	High	4.5
	High gas /low coal	4.3
	Low gas /high coal	6.4
	Low gas	6.0
	High gas	4.6
	Low coal	5.0
	High coal	5.1
High electricity demand		5.9
vRES capital costs	Cheap	4.3
	Expensive	5.8
Energy intensive MAC	-50%	7.3
	-25%	5.1
	25%	5.0
	50%	5.3
	100%	6.1

The second part of the sensitivity examines the effect of the discount rate. As explained in Section 2, the discount rate is one of the driving parameters for the wide range of cancellation estimations in the literature. Figure C.1 shows that the discount rate also has a huge impact on our results as cancellations lie within a range of 1.4 and 15.1 Gt, i.e., 73% lower and 196% higher than in reference scenario, respectively. Note that the effect of the interest rate is very strong when the discount rate is lower than 7%. Higher discount rates almost have no effect on cancellations because in the short-term the TNAC can hardly fall below a certain level as emissions are almost fixed until the first cancellation happens in 2023 (if only the discount rate is varied).

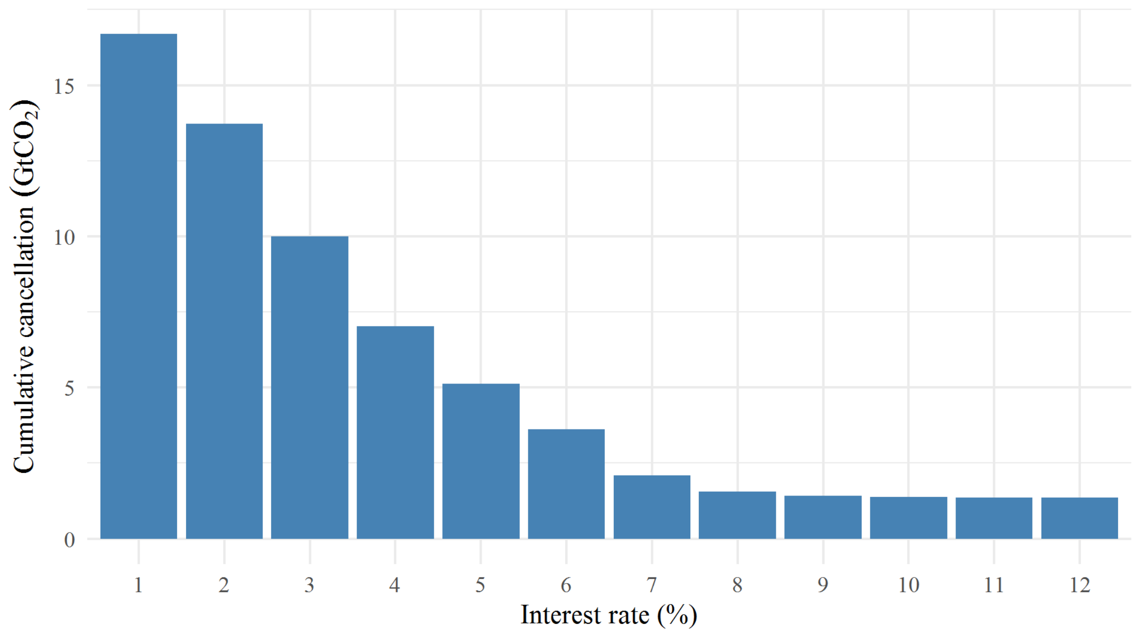


Figure C.1. Impact of interest rate on cancellations.

Another relevant assumption in the model concerns the time horizon for the MSR operation. We assume in all scenarios that certificates cannot be banked after 2057 and thus we constrain emissions from the EU ETS sector to zero after 2057. Table C.2 shows cancellations when banking is allowed until 2057 and during the entire model time horizon (2070, i.e., forever). If banking is allowed forever, total cancellations amount to 10.4 Gt. Notice that allowing banking further into the future has significant effects on MSR cancellations in this model framework because firms bank to avoid high MAC in the future. When banking is possible until 2057 there is less banking (and thus fewer cancellations) because MAC after 2057 cannot be reduced through banking. However, when the discount rate is 10%, the banking horizon has no effect on cancellations (1.4 Gt). These results point out that the differences between our estimated cancellations and those from Bruninx et al. (2019a) (13 Gt) stem from the assumptions regarding abatement costs and the banking horizon.

Table C.2. Impact of discount rate and banking horizon on cancellations (Gt).

	Discount rate (%)	
Banking horizon	**5**	**10**
Until 2057	5.1	1.4
Forever	10.4	1.4

D. Evaluation of simultaneous changes of MSR parameters and LRF

We analyse a wide range of combinations of the MSR parameters and the LRF. Among the MSR parameters, we choose the TNAC thresholds and the intake rate, as they are of greatest relevance. Figure D.1 shows a grid of 'heat maps', where colour key indicates the total cancellations. Each row of plots refer to a certain LRF and each column of plots refers to a certain intake rate. We evaluate intake rates from 12% to 100% and consider the 'critical' LRFs: a LRF of 2.2% is currently set, the LRF of 3% yields the highest cancellations (keeping the current MSR parameters, see Section 3.4.1) and the LRF of 2.4%, 2.6%, 4.1% and 5.2% were used in Section 3.4.2 to evaluate the achievement of more ambitious 2030 targets under the 'conservative' and the 'expected emissions' approaches. Due to the amount of required runs, we only evaluate lower thresholds between 0 and 1000 Mt, and upper thresholds between 200 and 1400 Mt, with a step of 200 Mt.

From Figure D.1, we observe that small increases in the LRF still have major impact on cancellations. Indeed, highest cancellations (17.5 Gt) occur when LRF equals 2.6%, intake rate 100% and lower and upper thresholds are respectively 0 and 200 Mt.

The figure also highlights that the intake rate gains in relevance when the LRF increases. Cancellations barely varies across different intake rates when the LRF is 2.2% (first row of plots), because the higher intake rate just makes the TNAC oscillates around the upper threshold implying that intake volumes increase but also decrease in certain years (see section 3.3.2). However, when the LRF is 5.2% intake rate yields significantly more cancellations when the intake rate is 100% (11.0 Gt) than when this is 12% (8.7 Gt). The larger LRF implies that the TNAC increases significantly, i.e., there is large short-term abatement in order to withhold certificates for the long-term. Correspondingly, the TNAC exceeds the upper threshold more often, and thus transfers into the MSR increase and cancellations accordingly as well.

Finally, Figure D.1 also shows that the effect of thresholds on cancellations weakens when LRF increases. As mentioned above, a very stringent cap (high LRF) leads to very high TNAC already in the short-term. The TNAC is indeed higher than the upper thresholds evaluated (up to 1400 Mt) during the period in which certificates can be transferred to the MSR, i.e., before the cap reaches 0 Mt (e.g., year 2038 when the LRF is 5.2%). As a result, cancellation does not vary across all combinations of thresholds when the LRF is 4.1% or 5.2%.

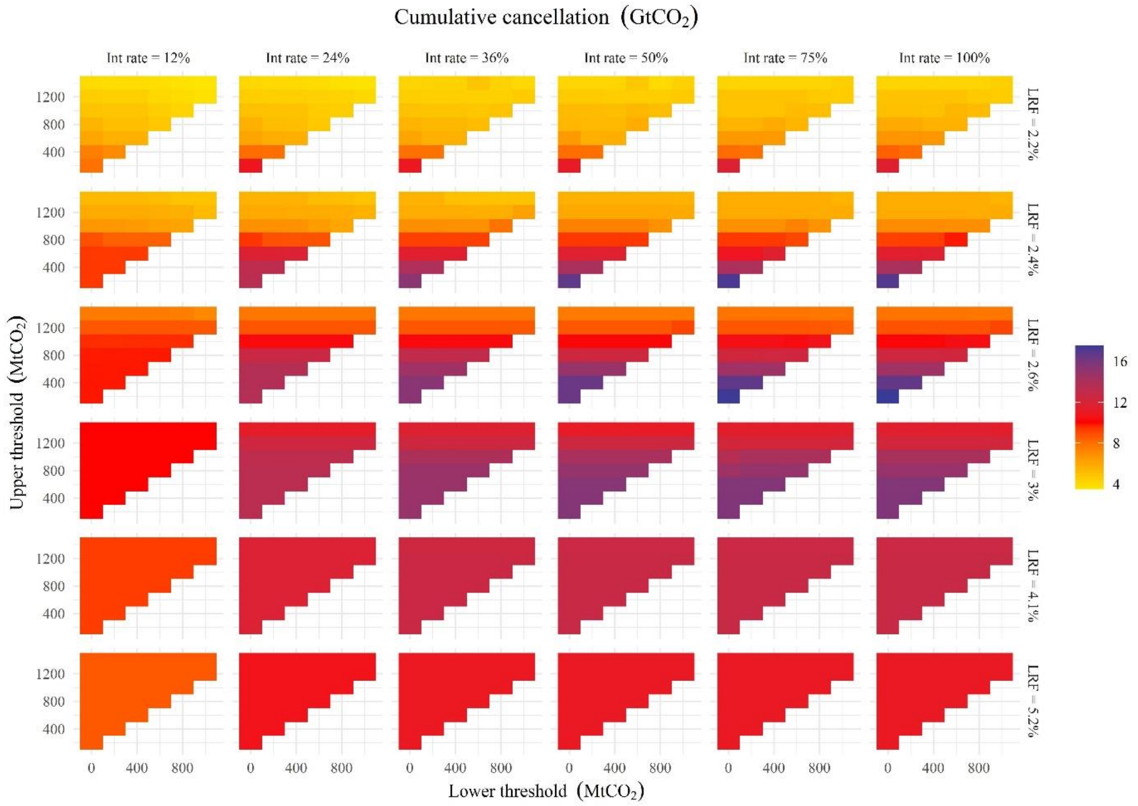


Figure D.1. Impact of simultaneous modifications of thresholds, intake rate and LRF on certificates cancellation.

Chapter 8

How to avoid history repeating itself: the case for an EU Emissions Trading System (EU ETS) price floor revisited[1]

Christian Flachsland
Michael Pahle
Dallas Burtraw
Ottmar Edenhofer
Milan Elkerbout
Carolyn Fischer
Oliver Tietjen
Lars Zetterberg

[1]Submitted to *Climate Policy*. The final version is available under https://doi.org/10.1080/14693062.2019.1682494.

How to avoid history repeating itself: the case for an EU Emissions Trading System (EU ETS) price floor revisited

Christian Flachsland[a,b,*], Michael Pahle[c], Dallas Burtraw[d], Ottmar Edenhofer[a,c,e], Milan Elkerbout[f], Carolyn Fischer[g,d], Oliver Tietjen[c], Lars Zetterberg[h]

[a] Mercator Research Institute on Global Commons and Climate Change (MCC), Berlin, Germany; [b] Hertie School of Governance, Berlin, Germany; [c] Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany; [d] Resources for the Future, Washington DC, USA; [e] Technical University Berlin, Germany; [f] Centre for European Policy Studies (CEPS), Brussels, Belgium; [g] Vrije Universiteit Amsterdam, The Netherlands; [h] Svenska Miljöinstitut IVL, Stockholm, Sweden.

*corresponding author

Abstract

Several years of very low allowance prices in the EU emissions trading scheme (ETS) have motivated calls to introduce a price floor to correct potential underlying distortions and design flaws, including (i) the political nature of allowance supply and related credibility issues, (ii) potential myopia of market participants and firms, and (iii) waterbed and rebound effects resulting from policy interactions. In the wake of the recent EU ETS reform, allowance prices have sharply increased. This raises the question of whether the case for introducing a price floor in the EU ETS remains valid. We argue that such a price floor, also adopted in several other greenhouse gas cap-and-trade systems worldwide, remains an important improvement in the design of the system, as long as the above-mentioned distortions and design flaws persist. An EU ETS price floor can safeguard against these issues and provides more explicit guidance on the minimum allowance price policymakers consider acceptable. Either as a complement or substitute to the current Market Stability Reserve (MSR), a price floor would thus make the EU ETS less prone to future revision in case of unexpectedly low prices. We identify and confront four prominent arguments against the introduction of an EU ETS price floor.

Key policy insights

- An EU ETS price floor would be an important institutional innovation enhancing political and economic stability, and predictability of the EUA price
- The recent Market Stability Reserve (MSR) reform has not removed the need for a carbon price floor.
- Introducing an element of price responsiveness into the so far purely quantitative design of the EU ETS would help to preserve its integrity
- In contrast to conventional wisdom, legal analysis reveals that an EU ETS price floor can be legally feasible
- Political support for a carbon price floor is gaining traction across Europe

Keywords: EU ETS, Price Floor, Market Stability Reserve

1. Introduction

The EU Emissions Trading Scheme (ETS) has for many years delivered prices below initial expectations. If low prices merely indicate low mitigation costs, they illustrate that the scheme works as theory suggests (Ellerman et al. 2016). However, the concern has been raised that low prices reflect market distortions and design flaws (Edenhofer et al. 2019a). These distortions include (i) the political nature of allowance supply and related credibility issues, (ii) myopia or inefficient discounting of market participants and firms, and (iii) waterbed and rebound effects, that is, unilateral emission reductions that are either ineffective as cumulative EU-wide emissions remain unchanged, or that even lead to an increase in cumulative emissions. Also, allowance prices in cap-and-trade are subject to large uncertainty (Borenstein et al. 2018) and market distortions may lead to inefficiency if opportunities for hedging through risk markets is limited (Tietjen et al. 2019). The potential presence of these issues risks negative long-term consequences by failing to initiate and support the technological and economic transformations necessary to decarbonize the economy (Acworth et al. 2017, Fuss et al. 2018).

To address these concerns, adding a carbon price floor to the EU ETS has been proposed, following the examples of the Regional Greenhouse Gas Initiative (RGGI), the common ETS of California and Quebec, and some Chinese pilot ETS (Knopf et al. 2014, Boehringer and Fischer 2018, ICAP 2019). However, the sharp increase of European Emission Allowances (EUA) prices in 2018 to levels around 25-30€/t (September 2019) places the EUA price at the magnitude found in economic optimization models employing a cost-effectiveness approach that takes the EU ETS cap as given (Pahle et al. 2018a). The recent EU ETS reform has increased the Linear Reduction Factor determining the annual rate of decline of the cap, and has modified the Market Stability Reserve (MSR) to increase the rate of annual allowance removals and invalidate (cancel) allowances. While it remains challenging to empirically assess the optimal level of allowance prices in cap-and-trade systems (Hintermann et al. 2016), the reform seems to have successfully addressed the broader concern over low prices. This raises the question of whether the case for a price floor in the EU ETS remains valid, which ultimately depends on the problem diagnosis.

We argue that a price floor remains an important improvement in the design of the EU ETS if there is uncertainty over the existence of the above-mentioned distortions and regulatory failures that might prevent dynamically efficient price formation. The potential existence of these issues implies the risk of a dynamically inefficient EUA price path (e.g. Salant 2016, Fuss et al. 2018). A price floor can reduce this risk, and provides more explicit guidance on the minimum EUA price policymakers consider acceptable. Either as complement or substitute to the current MSR design, a price floor would thus make the EU ETS less prone to requiring future revision in case of unexpectedly low prices.

Our contribution draws on two methods. First, we conducted a comprehensive review of the academic and policy literature on greenhouse gas (GHG) cap-and-trade system price floors and their adoption in existing systems worldwide, and on the recent EU ETS reform. Second, from 2016 to 2019, the authors organized and participated in several workshops in which academic analyses and policy questions relating to the introduction of a price floor were explicitly discussed with high-level stakeholders from academia, policy (EU and national institutions), industry, and NGOs. Building on these workshops, additional conversations with key policymakers and stakeholders, and close monitoring of the policy debate since inception of the EU ETS in 2005, as well as extensive discussions within the author team, we distilled a set of four objections to a price floor, which we set out and respond to below.

2. The recent EU ETS reform: Mainstream diagnosis and therapy

Different reasons for low EUA prices from 2012 to 2018 have been suggested, not least because understanding of the drivers of EUA prices remains poor (Hintermann et al. 2016, Friedrich et al. 2019). The mainstream view has been that the key reason for prices being lower than expected is an "imbalance" of allowance supply and demand. This is thought to result from the economic crisis in 2009, the influx of credits under the Clean Development Mechanism, and additional renewable and energy efficiency policies at the EU and member state levels that are thought to have driven emission reductions (Koch et al. 2014, Fuss et al. 2018, Ellerman et al. 2016). In this line of reasoning, allowance demand turned out to be lower than expected, leading to lower prices compared to *ex ante* expectations. This view has been guiding the recent EU ETS reform that intended to "reduce the surplus of emissions allowances […] and to improve the EU ETS's resilience to shocks" (EC 2019).

In consequence, the EU ETS reform adopted in 2018 primarily aimed at creating additional scarcity in the market. This is to be achieved, first, by strengthening the linear reduction factor, which specifies the amount that the cap will be reduced annually, from 1.74 to 2.2%. Second, the rate at which the MSR diverts allowances from auctions when the stock of allowances in circulation exceeds 833 million EUAs was doubled from 12% to 24% for the period 2019 until 2023. Third, all allowances in the MSR exceeding the level of the previous year's volume of auctioned allowances will be invalidated from 2023 on. Fourth, unilateral invalidation of allowances by member states in proportion to national regulations closing down facilities, e.g. coal plants, covered by the EU ETS is now allowed.

The reform will reduce allowance supply, thus addressing concerns about supply-demand imbalances. Estimates regarding the amount of allowances moved to the MSR and invalidation of allowances vary significantly (Beck & Kruse-Andersen 2018, Bocklet et al. 2019, Bruninx et al. 2018, Burtraw et al. 2018, Carlén et al. 2019, Danish Council on Climate Change 2017, Quemin and Trotignon 2019, Perino 2018, Perino and Willner 2017, Sandbag 2017, Silbye & Sørensen 2018, Tietjen et al. 2019). For example, Quemin and Trotignon (2019) find that up to 10 Gt will be invalidated, in contrast to 1.7 Gt in Perino and Willner (2017).

During the debate and especially after the reform was decided, the EUA price rose discernibly. According to standard economic theory, the most likely explanation is that anticipated future scarcity of allowances—reducing supply—resulted in increased current prices (Perino and Willner 2017). Other work suggests that transferring allowances into the MSR has created transitional stringency sufficient to raise prices at least in the short-term (Perino and Willner 2016; Mauer et al. 2019), especially if the hedging demand of firms for allowances is considered (Tietjen et al. 2019). A complementary interpretation is that the reform has restored market confidence in the willingness of EU policymakers to invest political capital into sustaining the ETS, triggering the comeback of allowance traders taking longer-term positions in the market (Sheppard 2018; Tagesspiegel 2018). Another interpretation is that price formation is myopically driven by short-term demand and supply, e.g. if firms have truncated planning horizons (Quemin and Trotignon 2019), and that the increased intake rates of the MSR has led to a tighter short-term market, inducing an EUA price increase.

It is uncertain, though, whether the fundamental problems of the EU ETS have been resolved for good. First, there is no solid evidence for what has driven the recent price increase. It might well be a bubble in an overconfident market (Friedrich and Pahle 2019). Second, there is a persistent risk that market confidence may be undermined again by future economic or political shocks. Given the complexity of the MSR, market actors may misjudge future effects, and unexpected outcomes may require further market interventions, possibly affecting market confidence. In fact, Phase IV of the EU

ETS envisions a formal MSR review process in 2021, and one outcome of that process may be further changes to the MSR's operation, including the intervention and invalidation parameters. Finally, the waterbed has not been effectively removed by the recent reform (see Section 4.1) and might lead to lowered allowance prices. Overall, we cannot rule out that history will repeat itself and EUA prices will drop substantially again – with potentially significant consequences for the legitimacy and political support of the policy instrument. We next argue that a price floor can help to at least partly remove related problems.

3. The case for an EU ETS price floor and implementation options

In contrast to the mainstream diagnosis of low EUA prices in recent years, another strand of literature suggests, from a theoretical perspective, that key factors depressing allowance prices were anticipated future downward price shocks or persistent doubt about the level of ambition (Salant 2016). Complementing this theoretical analysis, evidence suggests that past regulatory events, such as the backloading reform episode[1], have indeed negatively affected market credibility and were likely decisive factors in triggering the EUA price decrease (Koch et al. 2016). Inefficient discounting, for example due to myopia, regulatory risk, or incomplete risk markets, might also dampen near-term allowance prices (Kollenberg and Taschini 2019; Quemin and Trotignon 2019; Tietjen et al. 2019). In addition, reduced market confidence and low prices arguably reinforce unilateral member states' efforts to introduce additional policies to attain national climate targets, which may further drive down prices in a negative spiral due to the waterbed effect (Pahle et al. 2018a). If these were indeed the underlying problems – rather than, or in addition to, those suggested by the mainstream analysis – the recent reform then may at best just have cured the symptoms, but not their cause.

It is important to note that the mainstream diagnosis of the allowance price being lower than expected because of allowance demand turning out to be lower would not necessarily motivate intervention in the market: In absence of market or regulatory distortions, the market would simply work as it should. Lower than expected costs of meeting the cap might politically facilitate tightening of the cap to realize more environmental gains, but would not be mandated from a cost-effectiveness perspective. By contrast, any intervention aiming at supporting the EUA price without aiming at increasing the level of environmental ambition would start from the premise that some market or regulatory distortion prevents the price from achieving its cost-effective pathway. This is exactly the rationale underlying the case for a carbon price floor.

The main benefit of a price floor is that it would enhance long-term investment certainty by providing a clearer signal of regulators' commitment to achieve ambitious decarbonization targets even in the case of market and policy distortions driving the allowance price below its cost-effective pathway (Burtraw et al. 2010). Such reduced uncertainty would facilitate dynamically cost-effective allocation of investments into low-carbon technologies. This would also contribute to avoiding a situation where a lack of low-carbon investments in earlier years due to inefficiently low EUA prices might lead to significantly rising abatement costs and allowance prices in later years ("hockey stick"). Such an outcome would potentially undermine the political acceptability and environmental integrity of the cap that, as a result, might even be relaxed to avoid such high costs (Edenhofer et al. 2019a). To illustrate the relationship between downward price uncertainty and investment, when adding a price floor, firms facing investment decisions under uncertainty will, at the margin, implement 'green' investment projects that would otherwise (i.e. without a price floor) not be profitable in face of unmitigated downward EUA price risk. Conversely, firms will refrain from 'brown' high-carbon investments that are profitable only when factoring in downward EUA risk. Note that to achieve the objective of reduced uncertainty, the price floor pathway may be chosen slightly below what is anticipated to be the cost-effective trajectory – the aim is not to implement a binding price floor.

In addition, a price floor can also help avoid myopic price formation e.g. if it becomes binding and thus aligns the carbon price trajectory more closely with the dynamically efficient level and rate of increase (Fuss et al. 2018). Finally, the price floor can also reduce the waterbed effect from unilateral policies or any type of voluntary emission reductions when it turns out to be binding and is designed to induce removal of allowances from the market. This can help sustain the political acceptance of the scheme (Pahle et al. 2018a).

While it is correct that a price floor may be politically revised (downwards or upwards), and thus does not offer a perfect commitment device, we argue that policy stability and credibility is at least gradually increased, thus improving investment incentives. This important benefit comes at no cost in terms of reduced system performance. The price floor would only induce social costs if it would prevent the allowance price from dropping to an efficiently low level, e.g. in the case of significant low-carbon technology cost reductions. However, governance provisions should enable a structured review process of the price floor in any case, to enable adjustments if good reasons do emerge (note that the price floor pathway might also be adjusted upwards). Perhaps more importantly, the main cost appears to be political capital expended to initiate and implement the floor price reform in the first place.

A price floor can be implemented in the following ways: The ETS of California and Quebec, and several Chinese provincial ETS pilots, have implemented a price floor as an auction reserve price(s) below which none or only a fraction of allowances will be sold (ICAP 2019). RGGI also implements its price floor in its auction, and adds a price step known as the Emissions Containment Reserve, which provides a minimum price above the price floor that applies to 10% of the emissions cap, creating a price-responsive allowance supply (Burtraw et al. 2018). In the EU ETS, an EU-wide auction reserve price might be introduced in addition to the quantity threshold level of 833 megaton (Mt) allowances in circulation in the MSR. Unsold allowances could be moved into the MSR, where they might eventually be invalidated.

A second potential price floor implementation option is the UK carbon price floor (CPF), which requires power sector facilities covered by the EU ETS to pay a carbon price support that scales negatively with expected EUA prices to ensure that a specific domestic minimum carbon price is always achieved (Hirst 2018, Newbery et al. 2019). Currently, the support is set at £18/t (~€20) until 2021, adding to an EUA price of about €20–€25/t. To make the support rate more responsive to the actual EUA price realization than in the UK design, an ex post adjustment based on the realized EUA price is an alternative implementation option (Wood and Jotzo 2011). To avoid the waterbed effect resulting from the CPF, a proportional amount of allowances would need to be withheld from the market.

4. Debating the price floor option

In discussions with various stakeholders (see Introduction), we identified four prominent arguments against the introduction of an EU ETS price floor, which we confront below.

4.1 Objection 1: The MSR reform removes the need for a price floor

Objection: The recent price increase demonstrates that the fundamental problems of the EU ETS have been addressed. The allowance removal and invalidation features of the MSR reform eliminate the waterbed effect and reestablish fundamental allowance scarcity. The policy environment for low-carbon investments is now stable and predictable.

Response: The causality and durability of the recent EUA price increase is not yet determined. We cannot know whether the reform and economic circumstances will sustain high price levels**.** There is significant divergence in assessments of the impacts of the MSR, and market participants might misconceive the actual impact of the complex MSR invalidation mechanism. Credibility issues might easily return in case of political and economic shocks. Moreover, Perino and Willner (2016) and Kollenberg and Taschini (2019) find that the MSR (without cancellation) increases price variability, which raises the question of whether it stabilizes the market environment. A carbon price floor would more effectively constrain the downward uncertainty over future EUA prices and thus facilitate the required low-carbon investments.

The reform has at best partially addressed the waterbed effect of unilateral policies, but not eliminated it. A mechanism that perfectly accounts for unilateral policies would reduce the ETS cap by the emission abatement achieved by those policies. The amount of MSR cancellations, however, depends on the time profile and rebound effects of the emission reductions induced by unilateral policies. Several studies show that the MSR cancels less allowances in later years of its operation. Thus, the further in the future emission reductions from unilateral policies occur, the higher the waterbed effect (e.g. Beck and Kruse-Andersen 2018; Perino 2018). The reason is that the MSR absorption of allowances from the bank – the aggregate of all unused allowances held by market participants – will decline over time. In the early years of MSR operation, there will be an influx of allowances into the MSR because the allowance bank exceeds the threshold level of 833 Mt. At this time additional emission reductions increase the bank and thus the influx into the MSR. Over time, however, the MSR reduces the bank until it is lower than threshold of 833 Mt. Once the bank is low enough, unilateral policies may increase the bank but do not necessarily increase the influx into the MSR.

Moreover, Rosendahl (2019) and Pahle et al. (2019) find that unilateral policies can even lead to less MSR cancellations, implying that policies that aim to reduce emissions paradoxically increase cumulative emissions in the EU ETS. This can happen if a policy is announced today but effective in the future. Pahle et al. (2019) consider the German coal phase-out which was announced in 2019 but whose impacts would mostly unfold only from 2030 onwards. Given that market participants today anticipate lower allowance demand in the future due to the phase-out, the allowance price already drops today because of intertemporal arbitrage. This allowance price drop implies that more allowances are used early on and thus the TNAC level is lower since the coal phase-out has only minor effects at this time. In consequence, there is less short-term influx into the MSR as well, implying fewer allowances being cancelled. When the coal phase-out becomes most effective (after 2030) the decreased allowance demand has only minor effect on the influx into the MSR (see above). In the aggregate, the coal phase-out may therefore actually increase cumulative emissions in the EU ETS. In addition, policies with an immediate effect on emissions may also lead to less cancellation because of rebound or "internal carbon leakage" effects (Perino et al. 2019). This may occur if emission reductions in one country are overcompensated by expansion of emissions in other countries due to rising carbon-intensive exports (this is relevant e.g. in electricity markets, Osorio et al. 2018).

4.2 Objection 2: A price floor would transform the EU ETS from a quantitative policy instrument into a pricing instrument

Objection: Much effort has been invested in establishing the ETS as a quantitative policy instrument. This regulatory approach has ensured broad support from member states, industry, and EU institutions because it promises to achieve the climate target.

Response: A pure quantity target is not necessarily optimal; rather it is the consequence of a scientifically informed regulatory negotiation. If emissions reductions turn out to be less expensive than anticipated, then regulators would be expected to compel greater emissions reductions, and the price floor would embody such instruction to the market (Wood and Jotzo 2011). Price floors have been widely adopted in quantity-based ETSs worldwide, and for good reason (ICAP 2019). In fact, the EU ETS is increasingly becoming a special case in not featuring quantity adjustment based on rule-based price triggers. Furthermore, introducing a price floor does not imply the instrument is not based on quantity controls; if unsold allowances are invalidated, a price floor would achieve *more* ambitious environmental targets than those envisioned by the baseline cap and at prices that are below anticipated costs.[2] From economic theory, a hybrid instrument that combines elements of quantity and price regulation is likely to be superior to either approach taken alone for regulating carbon emissions under uncertainty (e.g., Weitzman 1974; Roberts and Spence 1976; Newell and Pizer 2003; Hepburn 2006; Wood and Jotzo 2011).

4.3 Objection 3: A carbon price floor is not legally feasible

Objection: A carbon tax could not be introduced in the 1990s because of the EU Council unanimity requirement of EU treaties on tax matters. This legal requirement would also make an EU ETS price floor infeasible.

Response: Fischer et al. (2019) reject the claim that introducing an auction reserve price into the EU ETS could not proceed with the ordinary legislative procedure. To trigger the special (unanimous voting) rather than ordinary (qualified majority voting) legislative procedure, a reserve price would have to be “primarily of a fiscal nature” or “significantly affect a Member State's choice between different energy sources.” The first trigger (“primarily of a fiscal nature”), although not well defined in EU law, should not apply for three reasons: First, the primary aim of the EU ETS is to reduce emissions, not to raise government revenue. Much of the allowance revenue is either freely allocated (negating the revenue motive) or earmarked for mitigation programmes (as with a fee), but not collected for general revenue (as with a tax) (Löfgren et al. 2018). Furthermore, an auction reserve price may lower or raise revenues, since the prices may rise but the number of allowances sold falls. Second, EU allowances have the status of financial instruments, and the ETS thus has already been shown not to be of fiscal nature. Third, an auction reserve price would not change the character or strictly fix the EUA price. Allowances could trade above or below that level in the secondary market, as has been the case in the California system.

Fischer et al. (2019) also reject the argument that an ETS auction reserve price might “significantly affect a Member State's choice between different energy sources.” First, an allowance price does not directly determine the energy mix of member states. Instead, its effects depend on the broader market situation (e.g., fuel prices). A legal trigger for the treaty unanimity requirement should not depend on market circumstances. Second, the EU ETS embodies an important environmental goal in justifying the competence of the European Union to introduce a cap-and-trade system establishing an EU-wide carbon price, and an incremental reform supporting the system would rely on the same competence. The European Court of Justice rejected a recent challenge by Poland to the initial version of the MSR based on this legal trigger, finding that market circumstances remain essential for the choice of energy sources and that the EU ETS constitutes a justified environmental policy.

4.4 Objection 4: Finding agreement on a common price floor will be impossible, and unilateral carbon price floors would fragment EU climate policy

Objection: Reluctant member states will strongly oppose a price floor above current or expected prices, since that effectively increases the level of ambition of the ETS. If the price floor is set at a lower level than the current price, it is irrelevant. If agreement on a common price floor cannot be achieved, a unilateral price floor implemented by one member state or a coalition of member states will reinforce political fragmentation, divergence and inefficiency in decarbonization pathways across Europe.

Response: For a few years, France was the only EU member state openly advancing the idea of a price floor (Szabo 2016). Like the UK CPF, the French initiative envisioned a price floor only for the power sector.[3] More recently, supportive signals have also come from the Netherlands (ICAP 2017), Sweden (Stam 2018), and Portugal and Spain (Brnic and Thévoz 2018). German discussions about the carbon price floor option have intensified (e.g., Hecking et al. 2017; Fernahl et al. 2017; Matthes et al. 2018; Demirdag 2018, Edenhofer et al. 2019a). Stakeholders likely to benefit from higher EUA prices (e.g., nuclear, gas and renewable power generators) can be expected to support this option.

Setting a price floor below the prevailing level of the carbon price should facilitate its adoption. The main goal of the price floor would be to provide insurance against the risk of price drops that threaten low-carbon investments. The price floor can automatically increase at a specified rate, such as the opportunity cost of capital plus inflation. For example, the California ETS price floor increases at 5% plus inflation, and the price trigger for RGGI's Emissions Containment Reserve will increase at 7%, annually independent of inflation after it is introduced in 2021 (ICAP 2019).

A harmonized EU-wide approach would be clearly preferable to avoid political fragmentation. An EU-wide approach may not be initially politically feasible though, e.g. if distributional effects of a price floor (Brink et al. 2016, Pahle et al. 2018a) cannot be addressed via well-established bargaining channels such as the reallocation of allowances (Dorsch et al. 2019). There may be reasons for a coalition to nevertheless act as a first mover, initiating a policy sequence (Pahle et al. 2018b) that would eventually lead to the remaining EU states joining. This strategy would be in line with considerations of shifting toward a Europe where "those who want more, do more" to overcome political impasse (European Commission 2017).

5. Conclusion: The way forward

An EU ETS price floor to be adopted by *all* member states could be advanced in the context of various policy processes:

- **2021**: **MSR review**. The review could be used to initiate the process for formally assessing and proposing price floor legislation, with a subsequent legislative process to be finished around 2023. For example, an EU-wide auction reserve price could be considered that would adjust the MSR such that the trigger for removal of allowances from primary auctions would be an EUA primary auction reserve price in addition to, or potentially in place of, the quantity threshold level of 833 Mt allowances in circulation.
- **2023: Paris Agreement Global Stocktake.** This international effort under the Paris Agreement could be used to initiate a process for formally assessing and proposing price floor legislation within the EU ETS, with a subsequent legislative process to be finished around 2025.

In parallel to an EU-wide price floor, a bottom-up *coalition of a few* EU countries would have more flexibility in the timing of their action to implement a unilateral price floor. The Netherlands, for example, are currently implementing legislation to implement a unilateral floor price. An agreement between Germany and France would arguably be essential to advance such a coalition. Germany has decided on a broader reform of its carbon pricing approach, including a call for adopting a price floor in the EU ETS (German Government 2019, Newbery et al. 2019, Edenhofer et al. 2019b). Over time, this coalition could grow and eventually create sufficient support for an EU-wide price floor.

To summarize, given the risk that EU ETS allowance prices might drop again due to unresolved fundamental challenges, adding a carbon floor price would go a long way towards enhancing policy credibility and thus incentivizing investments into low-carbon technologies. Implementing a floor price cooperatively at the EU level would not only make carbon pricing more dynamically efficient, but also advance and showcase the feasibility of multilateral cooperation. Ultimately, developing and testing efficient and cooperative climate policies is probably the biggest contribution the EU can make to global climate policy.

Notes

1. As a short-term measure to reduce the allowances in circulation, the EU Commission withheld allowances from auctions from 2014 to 2016 (400 million allowances in 2014, 300 million in 2015, and 200 million in 2016). Initially, it was planned to auction these in 2019 and 2020. With adoption of the MSR, these allowances were directly transferred into the MSR.

2. We do not consider the case of a price ceiling that would trigger the release of additional allowances, which might lead to non-achievement of the original quantitative target.

3. Some note that because the French electricity mix is heavily based on nuclear power, it would benefit from an increasing carbon price (Hecking et al. 2017; Pahle et al. 2018a).

Declaration of interest

No potential conflict of interest was reported by the authors.

Acknowledgements

The authors would like to particularly thank participants at the workshop “EU ETS Reform: Taking Stock and Examining Carbon Price Floor Options”, held at the Centre for European Policy Studies (CEPS) in Brussels on July 3, 2018, for a very helpful discussion of EU ETS carbon price floor options. Contributions to this paper were sponsored by the Stiftung Mercator AHEAD project, and the Mistra Carbon Exit project.

References

Acworth, W., Ackva, J., Burtraw, D., Edenhofer, O., Fuss, S., Flachsland, C., Haug, C., Koch, N., Kornek, U., Knopf, B. and Montes de Oca, M. 2017. “Emissions Trading and the Role of a Long Run Carbon

Price Signal: Achieving cost effective emission reductions under an Emissions Trading System." Berlin: ICAP.

Beck, U., Kruse-Andersen, P.. 2018. "Endogenizing the cap in a cap-and-trade system: Assessing the agreement on EU ETS phase 4." De Oekonomiske Rad Working Paper 2018:2.

Bocklet, J., Hintermayer, M., Schmidt, L., Wildgrube, T., 2019. The Reformed EU ETS - Intertemporal Emission Trading with Restricted Banking. EWI Working Paper, No 19/04.

Boehringer, C., and C. Fischer. 2018. "Emissions Floor Price Options for EU Member States." University of Oldenburg and Vrije Universiteit Amsterdam.

Borenstein, S., J. Bushnell, F. Wolak, M. Zaragoza-Watkins. 2018. "Expecting the Unexpected: Emissions Uncertainty and Environmental Market Design." Haas Working Paper.

Brink, C., H. Vollebergh, and E. van der Werf. 2016. „Carbon pricing in the EU: Evaluation of different EU ETS reform options." *Energy Policy* 97: 603-617.

Brnic, A. H., and L. F. Thévoz. 2018. "EDF hopes for European CO_2 price floor within months." *Montel*, July 31. https://www.montelnews.com/en/story/edf-hopes-for-european-co2-price-floor-within-months/922209.

Bruninx, K., Ovaere, M., & Delarue, E., 2018. A First Analysis Of the Market Stability Reserve in the European Emission Trading System. TME Working Paper.

Burtraw, D., Palmer, K., & Kahn, D. (2010). A symmetric safety valve. *Energy Policy* 38(9): 4921–4932.

Burtraw, D., A. Keyes, and L. Zetterberg. 2018. "Companion Policies under Capped Systems and Implications for Efficiency – The North American Experience and Lessons in the EU Context." RFF Working Paper. http://www.rff.org/research/publications/companion-policies-under-capped-systems-and-implications-efficiency-north.

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

Danish Council on Climate Change, 2017. The Inflated EU Emissions Trading System. Consequences of the EU ETS and Surplus of Allowances for Danish Climate Policy.

Demirdag, J. 2018. "Eleven German state ministers urge CO_2 price floor." *Montel*, July 16. https://www.montelnews.com/en/story/eleven-german-state-ministers-urge-co2-price-floor-/918136.

Dorsch, M., C. Flachsland, and U. Kornek. 2019. "Building and Enhancing Climate Policy Ambition with Transfers: Allowance Allocation and Revenue Spending in the EU ETS." *Environmental Politics,* in print.

EC – European Commission. 2019. "EU ETS Revision for Phase 4 (2021-2030)." Website information available at https://ec.europa.eu/clima/policies/ets/revision, Accessed 9 May 2019.

Edenhofer, O., C. Flachsland, L.K. Schmid (2019a): Decarbonization and EU ETS Reform: Introducing a price floor to drive low-carbon investments. In: Angrick, M., C. Kühleis, J. Landgrebe, J. Weiß (eds): 12 Years of European Emissions Trading in Germany. Stocktaking and perspectives for effective climate protection. Metropolis, p.207-232.

Edenhofer, O., C. Flachsland, M. Kalkuhl, B. Knopf, M. Pahle. 2019. „Optionen für eine CO_2-Preisreform." MCC-PIK-Expertise for the German Council of Economic Experts.

Ellerman, D., C. Marcantonini, 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. 2017. *White Paper on the Future of Europe. Reflections and Scenarios for the EU27 by 2025*. European Commission COM (2017) 2025. March 1. https://ec.europa.eu/commission/sites/beta-political/files/white_paper_on_the_future_of_europe_en.pdf.

Fernahl, A., C. Perez-Linkenheil, F. Huneke, and I. Küchle. 2017. *Wirkungsweise einer CO_2-Steuer im Stromsektor*. Berlin: Energy Brainpool.

Fischer, C., L. Reins, D. Burtraw, D. Langlet, M. Mehling, S. Weishaar, L. Zetterberg, H. van Asselt, K. Kulovesi, and A. Lofgren. 2018. „The Legal and Economic Case for an Auction Reserve Price in the EU Emissions Trading System." Working paper, May 30.

Friedrich, M., Mauer, E.M., Pahle, M., Tietjen, O., 2019. From fundamentals to financial assets: the evolution of understanding price formation in the EU ETS. Working paper.

Friedrich, M., and M. Pahle, 2019. Allowance prices in the EU ETS -- fundamental price drivers and the recent upward trend. Working paper. https://arxiv.org/abs/1906.10572

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): 220–41.

German Government (2019). „Eckpunkte für das Klimaschutzprogramm 2030". https://www.bundesregierung.de/resource/blob/997532/1673502/768b67ba939c098c994b71c0b7d6e636/2019-09-20-klimaschutzprogramm-data.pdf?download=1

Hecking, H., J. Kruse, and F. Obermüller. 2017. "Analysis of an EU-Wide Carbon Price Support: Impact on CO_2 Emissions, Cost Efficiency and Rents." EWI Energy Research & Scenarios. https://www.ewi.research-scenarios.de/cms/wp-content/uploads/2017/01/Analysis-of-an-Euwide-Carbon-Price-Support.pdf.

Hepburn, C. 2006. "Regulation by prices, quantities, or both: a review of instrument choice. *Oxford Review of Economic Policy* 22 (2): 226–247.

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): 108–28.

Hirst, D. 2018. "Carbon Price Floor (CPF) and the Price Support Mechanism." House of Commons Library Briefing Paper No. 05927. January 8. http://researchbriefings.files.parliament.uk/documents/SN05927/SN05927.pdf.

Burtraw, D., C. Holt, K. Palmer, A. Paul, and W. Shobe. 2018. "Quantities with Prices." Resources for the Future Working Paper 18-08.

ICAP (International Carbon Action Partnership). 2017. "Netherlands Proposes EUR 18 Carbon Price Floor." https://icapcarbonaction.com/fr/news-archive/491-netherlands-proposes-eur-18-carbon-price-floor.

———. 2019. "ETS Map." .

Knopf, Brigitte, et al. 2014. "The European Emissions Trading System (EU ETS): ex-post analysis, the market stability reserve and options for a comprehensive reform." FEEM Nota di Lavoro Working Paper.

Koch, Nicolas, et al. 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* 8: 121–39.

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

Löfgren, Å., D. Burtraw, M. Wråke, and A. Malinovskaya. 2018. "Distribution of Allowance Asset Values and the Use of Auction Revenues in the EU Emissions Trading System." *Review of Environmental Economics and Policy* 12(2): 284-303.

Matthes, F., et al. 2018. „Dem Ziel verpflichtet. CO_2-Mindestpreise im Instrumentenmix einer Kohle-Ausstiegsstrategie für Deutschland." https://www.oeko.de/aktuelles/2018/mindestpreis-auf-co2-fuer-mehr-klimaschutz/.

Mauer, E.M., S. Okullo, M. Pahle. 2019. „Evaluating the performance of the EU ETS MSR." Paper presented at the 2019 Mannheim Energy Conference.

Newbery, D. M., Reiner, D. M. and R. A. Ritz. 2019. "The political economy of a carbon price floor for power generation." *The Energy Journal* 40: 1-24.

Newell, R. G., and W.A. Pizer. 2003. "Regulating stock externalities under uncertainty." *Journal of Environmental Economics and Management* 45 (2):416–32.

Osorio, S., R. Pietzcker, M. Pahle., O. Edenhofer. 2018. How to Deal with the Risks of Phasing out Coal in Germany through National Carbon Pricing. CESifo Working Paper No. 7438.

Pahle, M., D. Burtraw, O. Tietjen, C. Flachsland, and O. Edenhofer. 2018a. "Preserving the Integrity of the EU ETS under Unilateral Action." Paper presented at World Congress for Environmental and Resource Economics, Gothenburg.

Pahle, M., D. Burtraw, C. Flachsland, N. Kelsey, E. Biber, J. Meckling, O. Edenhofer, and J. Zysman. 2018b. "Sequencing to ratchet up climate policy stringency." *Nature Climate Change* 8: 861–67.

Pahle, M., O. Edenhofer, R. Pietzcker, O. Tietjen, S. Osorio, C. Flachsland. 2019. „Die unterschätzten Risiken des Fahrplans für den Kohleausstieg". Energiewirtschaftliche Tagesfragen 69 (6), 31-34.

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., M. Willner. 2017. "Allowance prices, design choices and the market stability reserve in Phase IV of the EU ETS." *Climate Policy* 17 (7): 936-946.

Perino, G. 2018. "New EU ETS Phase 4 rules temporarily puncture waterbed." *Nature Climate Change* 8: 262.

Perino, G., R.A. Ritz, A. van Benthem. 2019. "Understanding Overlapping Policies: Internal Carbon Leakage and the Punctured Waterbed." NBER Working Paper No. 25643.

Quemin, S., R. Trotignon. 2019. “Emissions Trading with Rolling Horizons”. Working Paper.

Roberts, M. J., and M. Spence. 1976. “Effluent charges and licenses under uncertainty.” *Journal of Public Economics* 5 (3–4): 193–208.

Rosendahl, K.E. 2019. “EU ETS and the new green paradox.” Working Paper Series 2-2019, Norwegian University of Life Sciences, School of Economics and Business.Salant, S. W. 2016. “What ails the European Union′ s emissions trading system?” *Journal of Environmental Economics and Management* 80: 6–19.

Sandbag 2017, An agenda for strategic reform of the ETS. What’s the future for EU carbon pricing? Report.

Sheppard, D. 2018. “Hedge funds and Wall St banks cash in on carbon market’s revival.” *Financial Times*. September 6.

Silbye, F., Sørensen, P. B., 2018. Towards a more efficient European carbon market. Working Paper.

Stam, C. 2018. “Ministers highlight European divide over carbon price floor.” *Euractiv*, April 18. https://www.euractiv.com/section/emissions-trading-scheme/news/ministers-highlight-european-divide-over-carbon-price-floor/.

Szabo, M. 2016. “’We can’t wait any longer’: France floats EU ETS price support proposal.” *Carbon Pulse*, March 13. http://carbon-pulse.com/16939/.

Tagesspiegel. 2018. “CO_2-Preis sprengt alle Prognosen.” Tagesspiegel background, September 12.

Tietjen, O., Lessmann, K. and M. Pahle. 2019. "Hedging and the Temporal Permit Issuance in Cap-and-Trade Programs: The Market Stability Reserve Under Risk Aversion". https://ssrn.com/abstract=3436736

Weitzman, M. L. 1974. “Prices vs. quantities.” *Review of Economic Studies* 41 (4): 477–91.

Wood, Peter John, and Frank Jotzo. "Price floors for emissions trading." Energy Policy 39 (2011): 1746-1753.

Chapter 9

Synthesis and outlook

More and more jurisdictions around the globe are implementing emission trading systems (ETS) to reduce carbon emissions. However, real-world programs often do not work as standard theory suggests due to market and regulatory failures. For instance, the EU ETS experienced a phase of very low prices between 2011 and 2018, which raised serious doubts about its efficiency and durability.

In this thesis, I analyze underlying regulatory and market failures, their implications and possible solutions. In doing so, I aim to contribute to a better understanding and design of ETS. The focus is on the EU ETS – the longest-lived program to regulate greenhouse gas emissions in the world and the main climate policy instrument in the EU – though many results are also valid for comparable ETS like the cap-and-trade programs in the US. I rely on simple equilibrium models and the large-scale optimization model LIMES-EU to derive both tractable and quantitative results.

The central thesis is that unconstrained intertemporal allowance banking causes inefficiencies and even puts the transition to a low carbon economy at risk. The Market Stability Reserve (MSR) was implemented to improve the performance of the EU ETS, but I argue that it is not a solution to the fundamental problems. Instead, a price support mechanism (e.g., a price floor) is advisable or intertemporal trading should be constrained.

In the following section, I summarize the main findings of this thesis in more detail. Subsequently, I discuss the methods used in this thesis in Section 9.2 and the novelty and relevance of the results in a broader context in Section 9.3. Section 9.4 provides an outlook for future research and ETS policy development.

9.1 Main findings

The main findings of this thesis can be synthesized in two categories: The next section deals with the analysis of failures and their implications in the realm of the EU ETS. Thereafter, I synthesize the findings concerning potential solutions to these failures.

9.1.1 Intertemporal failures in emission trading programs and their implications

The econometric literature review in Chapter 2 reveals that typical demand-side fundamentals (e.g., gas and coal prices) are not sufficient to explain the EU ETS price. The literature alludes to regulatory and market failures, which are often linked to the discount rate applied to allowance banking. An efficient discount rate is of great importance for the functioning of the ETS with banking because it determines the growth rate of the allowance price. Therefore, the discount rate is at the center of this thesis and the following findings are structured along four failures that have an impact on the discount rate.

A first failure, considered in Chapter 3, directly follows from free intertemporal allowance trading. Unconstrained banking of allowances is inefficient because it entails that the ETS price grows at the discount rate, whereas the optimal carbon price (the social cost of carbon) grows at a lower rate. This problem is exacerbated by the inherent regulatory uncertainty about the future of the EU ETS, because it may further increase the discount rate applied to banking. The consequence of a too high rate is that the carbon damage during the transition to carbon neutrality (until the cap of the ETS is used up) is too high.

The second failure deals with capital market imperfections and the associated impact of the monetary policy of central banks on the general interest level. The general (risk-free) interest rate is part of the discount rate and as such has two important effects in the realm of emission trading: First, in Chapter 4, we show that renewable energies benefited more from the low interest rate level of past years (low cost of capital) compared to emission-intensive plants. In turn, a potentially rising interest rate level in the coming years worsens the competitiveness of renewable energies and may outweigh cost reductions due to technological progress. Hence a higher interest rate postpones investments in clean technologies. Second, the interest rate level affects the growth rate of the ETS price (see above). Due to capital market imperfections, monetary policy may distort the abatement costs and the timing of abatement.

The third failure analyzed in this thesis is the potentially softening of the ETS cap as a result of political lobbying (Chapter 5). Specifically, this failure builds upon the previous failures: Relying on the large-scale model LIMES-EU, we show in a scenario for the EU ETS that a higher discount rate (due to a tighter monetary policy and more market risks) leads to a relatively low ETS price early on and higher renewable energy costs. This slows down the deployment of clean technologies, implying relatively weak political power of the green (emission-free) lobby coalition, while for the brown (emission-intensive) coalition the opposite is true. As the brown market share remains high for a considerable time, it needs to decline fast towards the end of the transition to carbon neutrality if a given cap shall be reached. However, the combination of strong lobby power and high adjustment costs (due to fast abatement) for the brown coalition raises the probability that the cap is softened such that emission targets may not be reached.

The fourth failure is related to incomplete risk markets. In Chapter 6, we consider the case in which firms want to hedge their profits via ETS allowances or futures contracts. However, financial traders as trading counterparts face liquidity constraints such that hedging becomes costly for the firms. The costs for hedging are reflected by a risk premium as part of the discount rate. The time-dependent risk premium depends on the hedging demand of firms and the size of the allowance bank. In a numerical application to the EU ETS, we show that the risk premium can be negative and thus, reduces the discount rate applied to allowance banking.

9.1.2 Enhancing emission trading programs

Given the failures described in the previous section, how should an efficient policy be designed? Ideally, each failure will be corrected by an additional measure so that the first best outcome can be reached. However, the considered market failures can hardly be directly addressed (capital market imperfections, risk allocations) or the origin of the failure lies in regulation (cap softening, regulatory risk). Therefore, the following measures under discussion are not necessarily theoretical ideal solutions but rather practical design improvements.

Since all considered failures are related to intertemporal trading, a first straightforward solution would be to restrict banking. If banking would be banned, the regulator could, in theory, directly set the optimal cap for each period (e.g., year) and thus, the ETS price would rise at the optimal rate. However, besides the problem of determining the optimal periodic cap, banking also has welfare advantages as it reduces the impact of uncertainty because shocks spread to more periods (Chapter 3). A related, second, option would be to implement intertemporal trading ratios, which adjust the discount rate applied to banking. In Chapter 3, I argue, however, that trading ratios have practical implementation problems in actual climate policy and are therefore an unlikely option.

As an alternative, I suggest a third option in Chapter 3: a subsidy for clean production to correct the growth rate of the carbon price in ETS (with free intertemporal trading). The optimally set subsidy path raises the too-low carbon price early on, reduces the too-high carbon price later on and also lowers the ETS price volatility. However, unless the regulator is able to commit to the optimal subsidy path, the subsidy partly crowds out the ETS price due to the waterbed effect. Moreover, the subsidy induces some inefficiencies due to overconsumption, and if the subsidy is set to generous, it may undermine the credibility of the ETS program. Hence, a subsidy can improve the outcome of ETS but it is not an ideal solution because it entails considerable risks.

A fourth measure to enhance emission trading is the MSR implemented in the EU ETS. We show in Chapter 6 that the MSR indeed has a positive effect on the ETS price by reducing the allowance bank because a smaller bank increases the hedging value of allowances. However, a smaller bank also reduces the discount rate applied to banking. As a result, the price hardly increases or even decreases over time in our numerical simulation of the EU ETS. Since it is hardly possible to calibrate the MSR to a targeted discount rate, it is not a suitable solution to correct inefficient growth rates. A potentially positive long-term effect of the MSR is gained by its cancellation mechanisms, which reduces the overall cap. Here we find a higher number of cancellations compared to many other studies because we consider that the MSR reduces the discount rate applied to allowance banking which, in turn, raises cancellations.

Moreover, using the detailed model LIMES-EU, we show in Chapter 7 the significance of several ETS parameters for the number of MSR cancellations. Of particular importance is the linear reduction factor (LRF). The LRF is the main parameter to adjust the overall cap in the EU ETS, and it is bound to be adjusted in the coming years to reach more ambitious EU climate targets. We find that a higher LRF (lower cap) also increases MSR cancellations; therefore, only a moderately higher LRF is required to achieve significantly more ambitious emission targets. Moreover, adjusting certain parameters as MSR intake thresholds can have a strong effect on cancellations. The discount rate also has a strong impact such that cancellations can vary greatly and achieving specific emission targets becomes more complicated with the MSR. Based on such findings and other results from the literature, we argue in Chapter 8 that the MSR may even increase regulatory risk and price variability. Moreover, the MSR does not solve the waterbed problem: overlapping policies (e.g., coal phase-out policies) still reduce the ETS price and may even raise cumulative emissions (less MSR cancellations).

As an alternative and final option to enhance emission trading programs, we suggest implementing a price floor, or more generally, a price-responsive allowance supply in Chapter 8. Such a measure tackles the problems of the EU ETS more directly: Constraining the price (from below) reduces the risk and improves planning certainty such that investments in clean technologies are facilitated. If policy makers are able to implement a reasonable price floor (path), they would signal their commitment to the ETS and the emission targets so that the (perceived) regulatory risk would be lower. Also, price reductions through the waterbed effect of additional climate policies are mitigated because fewer allowances are auctioned if the price declines. As a result, a price floor reduces dynamic inefficiencies, essentially by avoiding too-low prices early on and signaling long-term ambition.

9.2 Discussion of the applied economic models

Besides literature reviews (Chapters 2 and 8) and conceptual approaches (Chapter 5), the results of this thesis rely on economic models as the central method. Specifically, a simple equilibrium model in two different versions was developed by the author of this thesis and used in Chapters 3 and 6. In addition, the large-scale numerical model LIMES-EU was extended by an explicit representation of the EU ETS including the MSR and used in Chapters 5 and 7. In this section, I discuss these two model approaches.

In general, both model types have a partial equilibrium interpretation under perfect competition: In the small model of Chapters 3 and 6, two representative firm types maximize their profits, where electricity and ETS prices are considered as given. Although LIMES-EU is a cost-minimization model without explicit firms, it induces the same outcome as perfectly competitive firms. Perfect competition is often seen as a critical assumption since larger firms might be able to apply market power in reality. Considering market power is certainly an interesting aspect, especially in the context of the MSR (see also discussion in Chapter 7), but given the relatively low market concentration in the EU ETS, perfect competition seems to be a reasonable assumption for the research questions of this thesis.

A further critical assumption of both model types is the rationality and farsightedness of market agents. In this regard, the econometric literature review of Chapter 2 points to behavioral aspects as potentially important ETS price drivers, which, however, is out of scope of the model analyses of this thesis. Moreover, in emission trading programs with banking, such as the EU ETS, shortsightedness is sometimes considered an important market failure (Fuss et al. 2018; Quemin and Trotignon 2019; Perino and Willner 2019; Willner 2018). Although such myopia, is not explicitly modeled, variations of the discount rate are included as sensitivity analyses in Chapters 6 and 7, where a high rate can be interpreted as a shortcut to implement myopia. As both the effect of hedging in Chapter 6 and the number of MSR cancellations in Chapter 7 decline with higher discount rates, shortsightedness has a potentially strong impact on the results.

Regarding the MSR analyses in Chapters 6 and 7, a potential caveat is the possibility of multiple equilibria. The reason is that the MSR induces a discontinuous allowance supply in the ETS (Gerlagh et al. 2019). In order to find different equilibria, the numerical models of Chapters 6 and 7 are solved with different starting values. Yet, differences between model runs were rare and negligible in size, such that multiple equilibria seem to be only a minor issue.

A striking difference between the two model approaches applied in this thesis is the level of detail: LIMES-EU explicitly includes 29 model regions (largely EU member states), grid connections between them, more than 20 electricity generation technologies and different time scales (short-term time slices to reflect weather and demand variations, representative days and years). In contrast, the small model only includes one region and two different firm types, each using a different technology reflected by quadratic cost functions. In Chapter 6, plant capacities are also explicitly considered. The limited complexity of the small model implies that its numerical results should only be understood as illustrations and rough estimates. In turn, the advantage of detailed LIMES-EU is more precise numerical results, including interaction effects between different variables, such that more policy-relevant results can be generated. This is of particular importance in Chapter 7, where we analyze different ETS and MSR parameters in light of the upcoming ETS reform.

At the same time, small size is also an advantage, because it allows full tracking of model outcomes, whereas the large size of LIMES-EU impairs tractability. Therefore, smaller models are particularly useful when economic mechanisms are examined. Especially if the model can be solved analytically, as partly in Chapters 3 and 6, mechanisms can be easily tracked. A further advantage of the smaller model is that it can be extended to more complex market conditions. This includes policy interactions in Chapter 3, market frictions (hedging, incomplete markets) in Chapter 6 and stochasticity in Chapters 3 and 6.

9.3 Relevance and novelty of this work

On a general level, the relevance of this thesis lies in the analysis of (1) failures of the EU ETS and cap-and-trade programs in general, and (2) potential solutions to these issues. The overarching distinctiveness of this work is the focus on intertemporal issues of ETS as an instrument to drive emissions down to zero in the long term. Moreover, the papers of this thesis are among the first to analyze the MSR as a new policy instrument, and thus, they contribute to a better understanding of its effects and to the improvement of the EU ETS.

A first, more specific addition to the literature is the structured review of the empirical literature on the EU ETS in Chapter 2. Compared to other recent reviews (e.g., Hintermann et al. 2016), we emphasize financial market issues and their effects on the discount rate applied to allowance banking. For the plausible case that the discount rate applied to allowance banking is too large (see Section 9.1.1), I provide a new perspective on the instrument choice to regulate stock pollutants (Fell et al. 2012; Newell and Pizer 2003) in Chapter 3. While others show that the discount rate can be adjusted by intertemporal trading ratios (Kling and Rubin 1997; Pizer and Prest 2020), I examine the case where subsidies for clean energy are added to an ETS as a second best alternative. In doing so and in contrast to previous work (Böhringer and Rosendahl 2010; Fankhauser et al. 2010), I show that the waterbed effect of such overlapping policies can be welfare-enhancing.

In Chapter 4, we show the importance of monetary policy and the interest rate level for the costs of renewable energies. Although the cost-increasing effect of higher interest rates on renewable energies is well-known (Schmidt 2014; Hirth and Steckel 2016), it has not been linked to monetary policy. That is, in contradiction to the views of many observers, who expect ever-falling renewable costs (Clark 2017; Obama 2017), we show that a tighter monetary policy may even lead to higher costs in the coming years, which could slow down the renewable energy deployment. In addition, a higher interest rate also raises the growth rate of the EU ETS price, and as a result, investments in renewable energies may be further postponed. It has been argued before that a fast growing or "too high" ETS price may lead to retroactive changes in the cap (Burtraw et al. 2010; Edenhofer et al. 2017), but we go a step further in Chapter 5 and conceptualize how different ETS price paths lead to different feedback effects on the ETS policy. Put differently, we link ETS price paths to the stickiness of the cap, which is of high significance for the credibility of cap-and-trade programs.

Furthermore, we contribute to the literature by analyzing the effect of firm hedging on the ETS price in an intertemporal setting, especially regarding the MSR. Hedging is considered to be an important price driver in commodity markets in general (Acharya et al. 2013; Kang et al. 2020) and in the EU ETS in particular (Chevallier 2013; Trück and Weron 2016). Parts of the MSR design are also motivated by the hedging demand of electricity producers covered by the EU ETS (Schopp et al. 2015). Nonetheless, the MSR has not been analyzed based on the classical hedging pressure theory (Keynes 1930; Hicks 1939; Acharya et al. 2013; Ekeland et al. 2019). We fill this literature gap in Chapter 6.

In Chapter 7, we also add to the MSR literature by examining the MSR with a highly detailed model, LIMES-EU, and by showing the impact of several ETS and MSR parameters on the amount of canceled allowances. This is highly policy relevant for upcoming MSR reviews and cap updates to reach more ambitious emission targets. Of similar policy relevance is our broader MSR assessment in Chapter 8. We are among the first to evaluate the MSR in light of the EU ETS problems and compare it to a price floor.

9.4 Outlook and directions for future research

To reach the Paris target of limiting the increase of the global temperature to "well below 2 řC" or even 1.5 řC (UNFCCC 2015), more and essentially more ambitious climate instruments are required around the globe. The increasing number of ETS programs worldwide, including the new Chinese program (ICAP 2020), is therefore a promising development. The EU considers itself as a forerunner in climate policy and aspires to climate neutrality by 2050 (European Commission 2018; European Parliament 2019; European Council 2019). For this purpose, it is expected that the EU will increase its current reduction target for 2030 of 40% to 55% relative to 1990. To reach this target, a strengthening of the EU ETS is needed because it is the major EU climate policy instrument. Moreover, the scope of the EU ETS might be extended so that it covers more sectors (European Commission 2019; European Parliament 2020). This is a desirable development, as exposing more sectors to carbon pricing makes climate policy more efficient and effective, and the EU ETS is, in principle, an adequate instrument for this endeavor.

However, this thesis points to several problems of the EU ETS that hamper its efficiency and effectiveness, and thus, questions whether the EU is indeed a role model for climate policy. Most problems discussed in this work are related to intertemporal trading, a feature that is present in many cap-and-trade programs around the world (ICAP 2020). I argue that the discount rate applied to allowance banking is unlikely to be optimal for several reasons, which leads to welfare losses and, even worse, threatens the durability of the EU ETS. By analyzing the MSR and discussing price floors, I also consider potential solutions to these issues. However, this thesis cannot cover all relevant issues of emission trading, nor provide irrefutable results for the analyzed issues. Therefore, I present some further important avenues for future research in the following.

A research direction deserving more attention in the realm of ETS programs is behavioral finance, especially related to expectation and belief formation. While in the models used in this work market agents are assumed to be rational and sufficiently forward-looking, the econometric literature review in Chapter 2 suggests that agents may behave differently. For example, several empirical papers find herding behavior, under- and overreaction to new information and different trader types, all of which are ignored in typical ETS studies as well as in this thesis. Such issues are not new to the finance literature, but were rarely transferred to cap-and-trade markets. Especially because ETS markets are politically created and the allowance price reflects expectations about the mid- to long-term, belief formation and information processing seems to be of even greater relevance for ETS compared to other commodity markets.

On a broader level, the interaction of ETS and financial or capital markets is an interesting research field. For instance, financial market frictions may affect the cost of capital for abatement technologies and distort their deployment. In Chapter 6, we consider liquidity constraints for financial traders as friction and, as a result, the hedging demand of firms induces risk premiums distorting the discount rate applied to allowance banking. While many empirical studies analyze the effect of hedging pressure on risk premiums in other commodity markets (Acharya et al. 2013; Kang et al. 2020), comparable work is, to the best of my knowledge, missing for ETS markets.[1]

A related question is how far do market agents look into the future? To date, there exists no definite empirical answer to the question of whether myopia plays a role in ETS markets (Fuss et al. 2018). Nonetheless, some theoretical ETS papers have taken shortsightedness into account (Quemin and Trotignon 2019; Perino and Willner 2019; Willner 2018), but its implications and measures against it are not yet fully understood. Therefore, both more theoretical and empirical papers in this direction would be of great value.

In this thesis, the MSR gets relatively bad marks for several reasons (see Chapter 8). Yet at least regarding its ability to smooth out shocks and thus, to improve the stability of the EU ETS, the jury is still out. A first incidence to show its usefulness is certainly the economic recession caused by the COVID-19 pandemic. In early March 2020, when the economic consequences of the pandemic became clearer, the ETS price plummeted from about 25 ď/t to below 17 ď/t – as one would expect without a stabilization mechanism. However, by the end of June, the price had already recovered, suggesting that the market believes that the MSR absorbs additional freed allowances due to the economic crisis, which is in line with the theoretical analysis by Gerlagh and Heijmans (2018). On the other hand, Quemin (2020) finds that stabilization through the MSR is limited, and Bruninx et al. (2019) show that the MSR cancels more when abatement costs are higher in the future and thus may amplify the impact of shocks. These results indicate that the kind of shock, especially concerning its timing and duration (correlation over time), is crucial for the stabilization function of the MSR, which, however, requires more research (see Gerlagh et al. 2020 for a first analysis of the COVID-19 shock). A related aspect is the influence of market power when the allowance bank level is close

[1] Existing empirical papers on risk premiums in the EU ETS do not include an independent variable for hedging pressure in their analysis, but rather interpret deviations from the frictionless cost-of-carry model as risk premiums (e.g., Trück and Weron 2016).

to the thresholds levels of the MSR. Future work could examine the ability of firms to manipulate the market.

Research into such directions helps to improve cap-and-trade programs in general, and the EU ETS in particular. However, to put improvements into practice and raise the stringency of climate policy, barriers such as distributional issues need to be considered as well. This possibly requires a sequence of policies because the ideal instrument might be politically not feasible without intermediate steps (Pahle et al. 2018). An important step could be the introduction of a price floor in the EU ETS, which is now considered by several states including France and Germany (Pahle and Tietjen 2019). Other promising ideas to enhance the EU ETS include a ban of intertemporal trading and the introduction of a price-responsive allowance supply (Traeger et al. 2020). The willingness of the EU and a growing number of its member states to improve the EU ETS allows us to be cautiously optimistic. The European Green Deal and the upcoming MSR review are of great importance to put the EU on the way to climate neutrality.

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.

Böhringer, C. and K. E. Rosendahl (2010). Green promotes the dirtiest: On the interaction between black and green quotas in energy markets. *Journal of Regulatory Economics 37*(3), 316–325.

Bruninx, K., M. Ovaere, K. Gillingham, and E. Delarue (2019). The unintended consequences of the EU ETS cancellation policy. *MPRA Paper No. 96437*.

Burtraw, D., K. Palmer, and D. Kahn (2010, sep). A symmetric safety valve. *Energy Policy 38*(9), 4921–4932.

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

Clark, P. (2017). The Big Green Bang: how renewable energy became unstoppable.

Edenhofer, O., C. Flachsland, C. Wolff, L. Schmid, A. Leipprand, N. Koch, U. Kornek, and M. Pahle (2017). Decarbonization and EU ETS Reform: Introducing a price floor to drive low-carbon investments. Technical report.

Ekeland, I., D. Lautier, and B. Villeneuve (2019). Hedging pressure and speculation in commodity markets. *Economic Theory 68*(1), 83–123.

European Commission (2018). A Clean Planet for all. A European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Soc. *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0773*.

European Commission (2019). The European Green Deal. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN*.

European Council (2019). European Council conclusions, 12 December 2019. *https://www.consilium.europa.eu/media/41768/12-euco-final-conclusions-en.pdf*.

European Parliament (2019). Climate change. European Parliament resolution of 14 March 2019 on climate change âĂŞ a European strategic long-term vision for a prosperous, modern, competitive and climate neutral economy in accordance with the Paris Agreement. *https://www.europarl.europa.eu/doceo/document/TA-8-2019-0217_EN.html*.

European Parliament (2020). The European Green Deal. European Parliament resolution of 15 January 2020 on the European Green Deal. *https://www.europarl.europa.eu/doceo/document/TA-9-2020-0005_EN.html*.

Fankhauser, S., C. Hepburn, and J. Park (2010). Combining Multiple Climate Policy Instruments: How not to do it. *Climate Change Economics 1*(3), 209–225.

Fell, H., I. A. MacKenzie, and W. A. Pizer (2012). Prices versus quantities versus bankable quantities. *Resource and Energy Economics 34*(4), 607–623.

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.

Gerlagh, R., R. Heijmans, and K. E. Rosendahl (2019). Endogenous Emission Caps Always Induce a Green Paradox. *CESifo Working Paper No. 7862*.

Gerlagh, R., R. J. Heijmans, and K. E. Rosendahl (2020). COVID-19 Tests the Market Stability Reserve. *Environmental and Resource Economics 76*(4), 855–865.

Gerlagh, R. and R. J. R. K. Heijmans (2018). Regulating Stock Externalities. *CESifo Working Paper No. 7383*.

Hicks, J. (1939). *Value and Capital.* Cambridge: Oxford University Press.

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.

Hirth, L. and J. C. Steckel (2016). The role of capital costs in decarbonizing the electricity sector. *Environmental Research Letters 11*(11), 114010.

ICAP (2020). *Emissions Trading Worldwide: Status Report 2020.* Berlin, Germany: International Carbon Action Partnership.

Kang, W., K. G. Rouwenhorst, and K. Tang (2020). A Tale of Two Premiums: The Role of Hedgers and Speculators in Commodity Futures Markets. *The Journal of Finance 75*(1), 377–417.

Keynes, J. (1930). *A treatise on money, Vol 2.* London: Macmillan.

Kling, C. and J. Rubin (1997). Bankable permits for the control of environmental pollution. *Journal of Public Economics 64*(1), 101–115.

Newell, R. G. and W. A. Pizer (2003). Regulating stock externalities under uncertainty. *Journal of Environmental Economics and Management 45*(2, Supplement), 416–432.

Obama, B. (2017). The irreversible momentum of clean energy. *Science.*

Pahle, M., D. Burtraw, C. Flachsland, N. Kelsey, E. Biber, J. Meckling, O. Edenhofer, and J. Zysman (2018, oct). Sequencing to ratchet up climate policy stringency. *Nature Climate Change 8*(10), 861–867.

Pahle, M. and O. Tietjen (2019). The coalition for an EU-ETS carbon price floor is reaching critical mass. *energypost.eu, 4 Dec. 2019, https://energypost.eu/the-coalition-for-an-eu-ets-carbon-price-floor-is-reaching-critical-mass/.*

Perino, G. and M. Willner (2019). Rushing the Impatient: Allowance Reserves and the Time Profile of Low-Carbon Investments. *Environmental and Resource Economics 74*(2), 845–863.

Pizer, W. A. and B. C. Prest (2020). Prices versus Quantities with Policy Updating. *Journal of the Association of Environmental and Resource Economists 7*(3), 483–518.

Quemin, S. (2020). Using Supply-Side Policies to Raise Ambition: The Case of the EU ETS and the 2021 Review. *GRI Working Paper 366.*

Quemin, S. and R. Trotignon (2019). Emissions Trading with Rolling Horizons. *Working paper*.

Schmidt, T. S. (2014). Low-carbon investment risks and de-risking. *Nature Climate Change 4*(4), 237–239.

Schopp, A., W. Acworth, D. Huppmann, and K. Neuhoff (2015). Modelling a market stability reserve in carbon markets. *DIW Discussion Papers 1483*.

Traeger, C., G. Perino, K. Pittel, T. Requate, and A. Schmitt (2020). The Flexcap. An Innovative CO2 Pricing Method for Germany. *ifo DICE Report 18*(1), 42–48.

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.

UNFCCC (2015). *Adoption of the Paris Agreement, Proposal by the President, Draft decision.*

Willner, M. (2018). Consulting the chrystal ball: Firm's foresight and a cap-and-trade scheme with endogenous supply adjustments. *WiSo-HH Working Paper Series 46*.

Statement of contribution

Chapters 2 to 8 are based on individual research papers. Apart from Chapter 3, these papers are the result of collaborations between the author of the thesis and colleagues. In the following, it is explained how the author of this thesis and the other authors have contributed to each paper.

Chapter 2 *From fundamentals to financial assets: the evolution of understanding price formation in the EU ETS.*

Marina Friedrich and Michael Pahle proposed the idea for this paper. The research design and the structure of the paper was developed by all authors. The review of the empirical literature in Section 2 ("Demand-side fundamental price drivers) and Section 3 ("Political and regulatory changes") was written by Marina Friedrich. The review of the empirical literature in Section 4.1 ("Financial market frictions") was written by the author of this thesis and in Section 4.2 ("Behavioral aspects") by Eva-Maria Mauer. The text on the theoretical model was written by Michael Pahle and the author of this thesis.

Chapter 3 *Reducing the cost of delay: on the interaction of cap-and-trade and subsidies for clean energy.*

The author of this thesis carried out the entire work for this chapter.

Chapter 4 *Adverse effects of rising interest rates on sustainable energy transitions.* Published in Nature Sustainability (2019), Vol. 2, pages 879-885.

Tobias S. Schmidt, Bjarne Steffen, Florian Egli, Michael Pahle, Ottmar Edenhofer and the author of this thesis developed the research idea and design of the study. The numerical model was developed by Tobias S. Schmidt, Bjarne Steffen and Florian Egli and all authors interpreted the results. Tobias S. Schmidt, together with the other authors, wrote the paper.

Chapter 5 *The risk of softening the cap in emissions trading systems.*

Michael Pahle, Sebastian Osorio, Florian Egli, Bjarne Steffen, Tobias S. Schmidt, Ottmar Edenhofer and the author of this thesis developed the research idea and design of the study. The text was written by Michael Pahle in cooperation with the author of this thesis and all other authors. The author of this thesis developed the code for technology-specific risk premiums in the LIMES-EU model. The LIMES-EU model runs were conducted by Sebastian Osorio and all authors interpreted the results.

Chapter 6 *Hedging and temporal permit issuances in cap-and-trade programs: the Market Stability Reserve under risk aversion.*

The author of this thesis developed the research idea and design of this chapter after joint discussions with Kai Lessmann and Michael Pahle. The analytical and numerical models were developed by the author of this thesis and he also carried out the numerical simulations and interpreted the results. The text was written by the author of this thesis where Kai Lessmann and Michael Pahle provided comments and suggested refinements.

Chapter 7 *Reviewing the Market Stability Reserve in light of more ambitious EU ETS emission targets.*

The research idea was developed in joint discussions by Sebastian Osorio, Michael Pahle and the author of this thesis. Sebastian Osorio and the author of this thesis wrote the paper where the other authors provided comments and refinements. Sebastian Osorio included the Market Stability Reserve in the LIMES-EU model with the support by the author of this thesis. Sebastian Osorio also conducted the model runs with LIMES-EU and all authors interpreted the results.

Chapter 8 *How to avoid history repeating itself: the case for an EU Emissions Trading System (EU ETS) price floor revisited.* Published in Climate Policy (2020), Vol. 20 (1), pages 133-142.

This chapter is based on extensive discussions among the author team and stakeholders from academia, policy, industry, and NGOs held in several workshops about price floors in the EU ETS. Christian Flachsland, Michael Pahle and Dallas Burtraw proposed to distill the main objections to price floors from these discussions and respond to them in this paper. In addition, a comprehensive literature review on emission trading and price floors was conducted. The text was mainly written by Christian Flachsland where all other authors provided ideas and text for several sections of the article. The author of this thesis reviewed the literature on the Market Stability Reserve and the waterbed effect and he wrote the respective text in Section 4.1 ("Objection 1: the MSR reform removes the need for a price floor Objection").

Tools and resources

Some parts of this thesis are based on numerical modeling. This section lists the software used to run the models and to analyze their output.

In Chapters 5 and 7, the **L**ong-term **I**nvestment **M**odel for the **E**lectricity Sector of **EU**rope (LIMES-EU) is applied to obtain the quantitative results. LIMES-EU is implemented in the General Algebraic Modeling System (GAMS) with the solver CPLEX. The simulations in Chapters 3 and 6 are also implemented in GAMS with the solvers JAMS, PATH and CPLEX. The model output was analyzed using R and Microsoft Office. For more information, refer to the following websites (all accessed on 30/07/2020):

- LIMES-EU:

  `https://www.pik-potsdam.de/research/transformation-pathways/models/limes/limes`

- GAMS: `https://www.gams.com`
- CPLEX: `https://www.gams.com/latest/docs/S_CPLEX.html`
- JAMS: `https://www.gams.com/latest/docs/S_JAMS.html`
- PATH: `https://www.gams.com/latest/docs/S_PATH.html`
- R: `https://www.r-project.org/`
- Microsoft Office: `https://www.office.com/`

Acknowledgments

This thesis would not have been possible without the support of my colleagues, my friends and my family. Thank you very much to all of you.

I would especially like to mention my day-to-day supervisor Michael Pahle, whose ideas and critical comments improved this thesis greatly. Thank you for letting me be part of several interesting projects, but also for the necessary freedom to pursue my own ideas. I enjoyed the discussions we had about our research and about (climate) economics and politics in general.

I am very grateful to my supervisor Ottmar Edenhofer for giving me the opportunity to conduct research in an exciting environment like PIK. His economic and political insights were inspiring and a valuable input to this thesis. I also want to thank Grischa Perino for being supervisor of this thesis and for providing helpful comments.

Many thanks to my other research collaborators Dallas Burtraw, Florian Egli, Milan Elkerbout, Carolyn Fischer, Christian Flachsland, Marina Friedrich, Sabine Fuss, Brigitte Knopf, Kai Lessmann, Eva-Maria Mauer, Sebastian Osorio, Simon Quemin, Robert Pietzcker, Wolf-Peter Schill, Tobias S. Schmidt, Bjarne Steffen and Lars Zetterberg for all the interesting joint projects, workshops and discussions.

Special thanks to my colleagues at the ESED/CAEP group at PIK, Yiyi Bai, Laura Delsa, Anouk Faure, Marina Friedrich, Adeline Gueret, Fabian Hein, Emma Krause, Eva-Maria Mauer, Paul Nahmmacher, Samuel Okullo, Sebastian Osorio, Michael Pahle, Eva Schmid, as well as to Max Franks, Beatriz Gaitan, Christoph Gornott, Kai Lessmann, Christina Roolfs, Anselm Schultes and Björn Sörgel - I enjoyed the lunches, coffees and walks through the Einstein-Wissenschaftspark and the TU Berlin campus with all of you. I also thank Nicole Field, Dorothe Ilskens, Simon Kleineweber, Ramona Lorenz, Kristyana Neumann and Nicole Reinhardt for their support. I was very fortunate to be Christian Gambardella' s office mate. I thank him for his friendship, for sharing his views on the world and challenging mine, and last but least, for all his contributions to my research.

I have to thank Nika for the encouragement, and in particular for being so patient. Thank you, Renate and Jürgen, for the support in many different ways.