# Cap and Trade with Imperfect Hedging*

Bruno Biais Johan Hombert Daniel Schmidt Pierre-Olivier Weill

February 26, 2026

**Abstract**

In a cap-and-trade system, emitters face transition risk, to the extent that emission caps and permit prices are volatile. We show, theoretically, and empirically for the EU Emissions Trading System, that (i) emitters hedge with emission permits futures bought from financials, (ii) financial constraints limit hedging, in particular by limiting and delaying emitters' purchases of permits in the spot market, implying (iii) permit prices are below the prices of replicating derivatives portfolios. Moreover, we show theoretically that constrained Pareto optima are implemented in equilibrium with cap-and-trade systems, in which the variance of emission caps is set lower than in the unconstrained case.

*We thank Richard Bateman for excellent research assistance. Support from the Investissements d'Avenir Labex (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) is gratefully acknowledged. Biais acknowledges support from the European Research Council (ERC Grant 882375, WIDE).

## 1 Introduction

In economies with no other friction than emissions externalities, Pareto optimality can be achieved by appropriately designed cap-and-trade policies (Anderson and Duanmu 2025). In practice, cap-and-trade systems play an important role, covering 23% of global greenhouse gas emissions in 2025 (International Carbon Action Partnership 2025). In these systems, the level of emissions is capped and emitters must surrender to the regulator one permit for each ton of $CO_2$ they emit. Ceteris paribus, the stricter the cap, the higher the price of permits.

Caps are set to let emitters internalize the social cost of their emissions. Due to "uncertainty about parameter values, model structure, (...) preferences" (Tol 2023) estimates of the social cost of carbon vary a lot across studies and through time.[1] Relatedly, there is uncertainty about future carbon policy, and permit prices are quite volatile. Thus, cap-and-trade systems expose emitters to transition risk.

Emitters could hedge transition risk by purchasing storable emission permits or derivatives. Thus, with frictionless markets, transition risk could be shared efficiently among emitters and non-emitters.[2] Financial constraints, however, hinder hedging. The Draghi (2024) report underscores this concern:

> "The hedging component (i.e., removing carbon price uncertainty) could also be fulfilled by sufficient ex ante purchases of ETS allowances as the later are 'bankable' (i.e., unused allowances can be saved for later use). Frontloading purchases of ETS allowances would, however, require up-front financing and may hit the financing constraints of companies."

This paper relies on theoretical and empirical analysis to study, from a positive and a normative point of view, the consequences of financial constraints on cap-and-trade systems under transition risk.

We start by documenting stylized facts about transition risk, hedging, and financial constraints in the European Union Exchange Trading System (EU ETS), which is the world's biggest cap-and-trade system in terms of total compliance costs (i.e., regulated emissions times price).

First, the price of emission permits is volatile, creating significant (transition) risk for corporate emitters. The volatility of the cost of compliance represents 40 percentage points of pre-tax

1. In his meta-analysis of the literature on the Social Cost of Carbon, Tol (2023) writes (page 533) that "the range of estimates is large." Summarizing the literature, he writes: "In the past 10 years, estimates of the Social Cost of Carbon have increased from \$9 per ton $CO_2$ to \$40 for a high discount rate and from \$122 per ton $CO_2$ to \$525 for a low discount rate."

2. Thus, without financial constraints, transition risk would not be a problem but a feature of optimal climate policies, dynamically adjusting to updates in the social cost of carbon.

profits on average.

Second, emitters hedge this risk by storing emission permits and buying futures contracts on emission permits, thereby covering two years of emissions on average. Emitters buy permit futures from financial institutions, which hedge their short futures positions by holding physical permits.

Third, hedging patterns feature symptoms of financial constraints. Less capitalized emitters have lower storage of permits, delay purchases of permits within the annual compliance cycle, and hedge relatively more using futures than permits. These correlations are consistent with permit holdings—which require upfront financing—being costly for firms with binding financial constraints.

Fourth, pricing patterns also feature symptoms of financial constraints. There is a positive basis between the discounted futures price and the spot price of permits, i.e., the law of one price does not hold. Physical permits, which require upfront financing, are cheaper than permit futures. The basis is 80 basis points annually on average.

To rationalize these facts, we consider a simple model of cap and trade with financial constraints. At time 0, emission permits are sold and emitters and financials can bid for them. Emitters and financials can also trade derivatives, e.g., futures contracts, on emission permits, and borrow or lend. At time 1, the emissions cap, which can be high or low, is drawn. Further emission permits are sold, so that the total amount of permits sold over both periods equals the cap. Uncertainty about the emissions cap reflects uncertainty about the social cost of carbon and the political process that translates this social cost into policy. Emitters purchase permits so that their total holdings of permits match their emissions. Correspondingly, the market-clearing price of permits is set and contracts are executed. For simplicity, the only source of uncertainty is the level of emissions caps (high or low). Agents are risk averse and choose their emissions, their permit holdings, as well as their positions in financial derivatives, to maximize the expected utility from their time 1 consumption.

To model financial constraints, we borrow from Bolton and Scharfstein (1990), Alvarez and Jermann (2000), DeMarzo and Fishman (2007), Rampini and Viswanathan (2010), and Biais, Hombert, and Weill (2021) the assumption that agents have limited commitments. More precisely, we assume that agents can choose to default on the contracts they wrote, absconding with a fraction $\theta$ of their assets.[3] Limited commitment implies agents can only imperfectly share transition risk.

3. Equivalently, as in Biais, Hombert, and Weill (2021), an agent can make a take-it-or-leave-it debt writedown offer, and if the agent's counterparty rejects the offer the agent's counterparty can seize only fraction $1-\theta$ of the assets.

The main features of the equilibrium arising under financial constraints match the stylized facts from the EU ETS. To hedge against transition risk, emitters take long positions in permit futures while financials correspondingly take short positions in permit futures. Because of financial constraints, less capitalized emitters have lower storage of permits at time 0, delay their purchases of permits and hedge relatively more with futures than permits. Financial constraints imply that the price of emission permits is lower than the price of the replicating portfolio of derivatives. All these predictions of the model are consistent with the data.

We complement the above positive analysis, which takes as given the cap-and-trade system, with a normative analysis, in which we derive the optimal climate policy.

First, we find that the optimal policy remains cap-and-trade in spite of the presence of financial constraints.

Second, the constrained optimal emissions cap is less responsive to updates in the social cost of carbon than in the first-best. Indeed, it is second-best optimal to distort emission caps to reduce their volatility, so as to reduce transition risk, which agents can only imperfectly hedge. Therefore, financial constraints do not uniformly raise the level of optimal emissions caps above their first best counterparts. Instead, emissions caps should be larger than their first best counterparts when the social cost of carbon is high, but lower than their first best counterparts when the social cost of carbon is low.

Third, in contrast to emission caps, permit prices in the second best are uniformly lower than in the equilibrium implementing the first best. This is because financial constraints distort the permit price not only by distorting the emissions cap, but also by limiting agents' ability to bid up for permits.

Fourth, the constrained optimal emission caps depend on the Pareto weights of the different agents in the planner's problem. Interpreting Pareto weights in terms of influence on public policy, our analysis implies that emissions increase with the political influence of emitters. This contrasts with the first best, in which Pareto weights affect transfers but not emissions and production.

Fifth, our analysis also implies that the government should not frontload the issuance of emission permits. Indeed, frontloaded issuance tightens financial constraints as purchases of permits must be financed. Instead, the government should issue for each period the number of permits that emitters (in aggregate) must surrender during this period. The EU issues permits in a way that is roughly consistent with this recommendation.

**Literature:**

Kaenzig (2023) studies the consequences of changes in carbon policy, identified using changes in carbon prices triggered by regulatory news in the EU ETS. He finds the anticipation of tighter emission caps lowers emissions and output. These findings provide motivation for our modeling of emitters' risk-exposure, but our focus on risk-sharing under financial constraints differs from Kaenzig (2023)'s focus on macroeconomic variables such as energy prices and inequality. Fuchs, Stroebel, and Terstegge (2024) emphasize uncertainty about future carbon prices, as we do, measuring it using option prices. They focus on the link between this uncertain and investment in decarbonization,[4] whereas we focus on the link between this uncertainty and carbon price insurance.

Our analysis of carbon policy under financial constraints is linked to Döttling and Rola-Janicka (2025), Heider and Inderst (2022), and Oehmke and Opp (2025),[5] but we tackle different economic issues. In these papers, carbon policy affects access to credit, in line with empirical results (e.g., Ivanov, Kruttli, and Watugala 2024). The type of inefficiency we consider differs from that analyzed by Döttling and Rola-Janicka (2025), Heider and Inderst (2022), and Oehmke and Opp (2025). While they focus on credit rationing, we focus on imperfect risk sharing. Gupta, Sockin, and Starmans (2025) study insurance against exogenous physical risk, whereas we study insurance against policy-determined transition risk.

A strand of literature studies the impact of carbon pricing on firm-level emissions (e.g., Bartram, Hou, and Kim 2022; Dechezleprêtre, Nachtigall, and Venmans 2023; Zaklan 2023; Martinsson, Sajtos, Strömberg, and Thomann 2024; Bai and Ru 2024; Colmer, Martin, Muûls, and Wagner 2025). Consistent with these studies, emitters' emissions choices decrease with the carbon price in our model.

A large literature analyzes exposure to, and pricing of, transition risk (e.g., Bolton and Kacperczyk 2021, 2023; Sautner, Van Lent, Vilkov, and Zhang 2023; Hsu, Li, and Tsou 2023; Li, Shan, Tang, and Yao 2024; Zhang 2025). In line with the literature transition risk is priced in our model, since we show that the pricing kernel overweights cash-flows received in the low-cap state relative to cash-flows received in the high cap state. Moreover, we complement the literature by analyzing how financial frictions constrain firms' ability to hedge transition risk and by characterizing the implications of financial frictions for the pricing of emission permits.

Our result on the basis between spot and futures echoes the asset pricing literature on deviation from the law of one price (e.g., Garleanu and Pedersen 2011). In contrast with Garleanu and Pedersen (2011), we show that it is always the underlying that is underpriced relative to the derivative, as in Biais, Hombert, and Weill (2021). This prediction from our theory is in line

4. See also Wang, Wurgler, and Zhang (2026) on the impact of climate policy uncertainty on green investment.
5. In fact our assumption that agents can default and abscond with a fraction of their resources is the same as in Döttling and Rola-Janicka (2025).

with our empirical finding on the basis between spot and futures permits.

Borri, Liu, Tsyvinski, and Wu (2024) study the extent to which firms trade permits, while Akey, Appel, Bellon, and Klausmann (2024) study in the cross-section of firms the link between permit trades and emissions. In contrast, we focus on the hedging motive in the EU ETS market.

In Section 2 we document stylized facts of the market for emission permits in the EU ETS. In Section 3 we present our model and positive analysis of the EU ETS and show that the implications of our theory of financially constrained risk-sharing are consistent with the stylized facts presented in Section 2. In Section 4 we turn to our normative analysis of carbon policy under financially constrained risk-sharing. We also outline the policy implications of our analysis, and in particular our result that an appropriately designed cap-and-trade system can implement the second best. Section 5 briefly concludes.

# 2 Stylized Facts From the European Union Emission Trading System

## 2.1 The European Union Emission Trading System

The Emission Trading System in the European Union (EU ETS) illustrates the importance of transition risk stemming from emissions regulation. The EU ETS was launched in 2005 and is the world's biggest cap-and-trade system. Each year, at a given date, regulated entities must surrender emission permits to cover their yearly greenhouse gas emissions.[6] Each permit is for one metric ton of $CO_2$ equivalent.

Around half the permits in the EU ETS are freely allocated, especially to sectors in which carbon leakage is an issue. The other half is sold in auctions, held on Mondays, Tuesdays and Thursdays each week.[7] In addition to this primary market, there is an active secondary market, in which participants can trade emission permits and derivatives. Derivatives include (i) one-day ahead futures, basically equivalent to spot, (ii) futures with longer maturities, the most active of which being next December, and (iii) options.[8] Compliance entities (which must surrender permits corresponding to their emissions) as well as financial firms and funds participate in the primary and secondary markets.

[Go To Figure 1]

6. Until 2025 the surrender date was April 30, but it has now been moved to September 30.
7. Auctions are uniform price auctions, organized by the European Energy Exchange (EEX).
8. Most of the secondary market trading occurs on regulated exchanges, the most important being the Intercontinental Exchange (ICE) based in Amsterdam. There is some OTC trading, but to a limited extent.

Figure 1 shows emissions for the sectors covered by the EU ETS, which, overall, account for approximately 40% of greenhouse gas emissions in the EU. During our sample period, the number of emission permits issued by the EU declined. During Phase 3 of the EU ETS (2013–2020), emission permits issues were reduced at the rate of 1.74 % per year. During Phase 4 (2021–2030), the yearly rate of reduction went up to 2.2 %. Correspondingly, yearly emissions decreased at an average annual rate of 3.5% per year from 2008 to 2023.[9]

## 2.2 Data

**Permit holdings and transactions.** Account-level data on ownership of EU ETS permits comes from the European Union Transaction Log (EUTL), run by the European Commission, which records all transactions taking place within the EU ETS. Accounts are classified into three categories: Operator Holding Accounts (OHAs) are accounts linked to an installation (for which $CO_2$ emissions are verified). OHAs receive free permits allocated to the installation and are used to surrender the permits for the installation's verified emissions. Personal Holding Accounts (PHAs) are all other non-administrative accounts. Regulated firms often hold one or more PHAs in addition to the OHAs associated to their installations, which are used to centralize the trading of permits and transfer permits to and from the firms' OHAs. Financial firms can only hold PHAs. Finally, Administrative Accounts (AAs) are the EU's and national authorities' accounts used for regulation compliance such as allocating free permits and collecting surrendered permits. In our analysis, we focus on OHAs and PHAs.

We use the EUTL data from Jan Abrell's website, which are compiled from public EUTL data of the European Commission and are explained in Abrell (2024).[10] The transactions data are released with a three-year lag at a yearly frequency. The most recent vintage of the data contains transactions up to September 2021. We only focus on transactions from 2008 onwards, which corresponds to the start of Phase 2 of the EU ETS when permits became "bankable," that is, could be stored and transferred from one phase to the next.[11] By cumulating transactions over time, we can thus reconstruct permit holdings.

We also use the EUTL to estimate firm-level positions in futures contracts, which are not publicly available. We proxy for futures positions using the settlement of futures trades, which we identify in the EUTL by transactions occurring around the settlement dates of the most

9. This reduction in emissions is calculated on a constant scope, taking into account the fact that the coverage of the EU ETS changed twice during the period: airline is added in 2012; the UK exits in 2021. By contrast, emissions plotted in Figure 1 are calculated using point-in-time (i.e., varying) scope of the EU ETS.

10. https://www.euets.info/

11. Permits from Phase 1 (2005–2007) could not be transferred to Phase 2 and are not included in our analysis. Appendix A.2 describes how we recover the transfer of permits from Phase 2 to Phase 3 in the EUTL.

traded contracts (December and March).[12]

**Financial and ownership information.** We link EUTL accounts to Bureau Van Dijk Orbis data, which we use to consolidate EUTL accounts at the corporate group level and to retrieve financial information for these corporate groups. Specifically, we rely on the publicly-available mapping provided by the JRC-EU ETS-FIRMS project, which contains approximately 11,000 links between EUTL accounts and firms in Orbis.[13] Based on information about the global ultimate owner and majority shareholders in Orbis, we then classify firms into firms belonging to a corporate group and standalone firms. For firms belonging to a corporate group, we aggregate all EUTL accounts at the corporate group level and only use consolidated financial information from Orbis. For standalone firms, we aggregate EUTL accounts at the firm level and use the (consolidated or, when the former is not available, unconsolidated) financial information for that firm.[14]

**Aggregate futures positions.** Futures position data comes from European Securities and Markets Authority (ESMA)'s Commitments of Traders Reports. The data contains the notional amount of long positions and short positions in futures contracts on EU ETS permits, aggregated across all contract maturities and by types of traders, at the weekly frequency since January 2018. We use this data to construct aggregate futures positions of emitters and of financial institutions.[15]

**Permit price.** We retrieve permit prices from Refinitiv. Specifically, we download daily spot prices from 2013 onwards (the beginning of EU ETS Phase 3) as well as daily prices for futures contracts settled in December of each year (which is the most liquid futures contract).

**Samples** Our *full sample* include all firms in the EUTL, i.e., all firms that hold permits at one point in time. We refer to "firms" as the corporate entities at the level at which we consolidate EUTL accounts and financial information, whether this corresponds to a corporate group, a standalone firm, or an account holder unmatched with Orbis.

We classify firms in the full sample into two categories: emitters and financial institutions. We classify firms as emitters if 50% or more of their ETS accounts are associated with a regulated installation (OHAs), or if their surrenders of permits are 1% or more of total trading volume. We classify other firms as financials.

12. See Appendix A.3 for a detailed description of the procedure.
13. https://data.jrc.ec.europa.eu/dataset/bdd1b71f-1bc8-4e65-8123-bbdd8981f116
14. See Appendix A.4 for details on the procedure used to consolidate corporate groups.
15. We group the regulated firms and other non-financials (categories 4 and 5) as emitters, and the funds, banks, and other financials (categories 1, 2 and 3) as financials.

Our *sample of emitters* include firm-year observations for emitters during the period 2014–2020 that satisfy the following conditions: non-missing financial information from Orbis; total assets of €1 million or more; annual sales of €1 million or more; annual emissions of 1,000 $tCO_2$ or more in at least one year over 2014-2020; median emissions over sales of at least 1 $tCO_2$ per €10,000 of sales. The final sample of emitters include 966 unique firms and 5,400 firm-year observations.[16]

## 2.3 Stylized Facts

**Fact 1.** *The price of emission permits is volatile. Emitters' exposures to permit price risk are large.*

Figure 2 illustrates Fact 1. The price of emission permits (shown in blue in the figure) has fluctuated widely since the beginning of the trading period for storable permits in 2008, reaching a low of €3 per ton of $CO_2$ in April 2013 and a peak of €100 in March 2023. Changes in permit prices can reflect demand shocks, e.g., as the Global Financial Crisis and the European Sovereign Debt Crisis reduced economic activity in the EU, the demand for emission permits was reduced and permit prices dropped. Prices also reflect the current supply of permits, and as permits are storable (since 2008) price changes can also reflect changes in expectations about the future supply of permits (see Kaenzig 2023).[17]

[Go To Figure 2]

For our sample of emitters subject to the EU ETS, the average ratio of emissions to pre-tax profits is 0.008 tons of $CO_2$ emissions per €1 of profit. For a permit price of €50 per ton, the cost of compliance represents 40% of pre-tax profits on average. Considering a price volatility of €50 per year (shown in red in Figure 2), the volatility of the cost of compliance represents 40 percentage points of pre-tax profits.

**Fact 2.** *Emitters hedge permit price risk.*

Emitters can hedge permit price risk by storing emission permits and buying futures contracts on emission permits. The left panel of Figure 3 plots emitters' holdings of permits and futures positions (using the full sample). It is consistent with emitters hedging both with storage and

16. Appendix A.1 describes the construction of variables analyzed in the next subsection and their summary statistics.

17. Low prices before 2018 were consistent with the expectation that the supply of permits would remain high, while price increases starting in 2018 are consistent with policy changes signaling tougher policies, such as the 2018 reform of the Market Stability Reserve to which backloaded and unallocated permits are transferred and the 2021 Fit for 55 climate package, whereas the decline in 2023 corresponds to political backlash against climate policy.

futures. At the end of 2020, in aggregate, emitters held around 1,500 million $tCO_2$ worth of (physical) permits and 1,000 million $tC0_2$ worth of long positions in permit futures, which together amounted to about two years of emissions. The bump in permit holdings from January to April reflects that emitters receive permits from physical settlement of the December futures contract, then receive their free permits in February, and surrender permits in April.

[Go To Figure 3]

The right panel of Figure 3 shows holdings of permits and futures positions of financial institutions. Financials are short futures for approximately the same amount as emitters' long futures positions. Therefore, emitters buy protection against permit price risk from financials. Financials hedge this risk by holding permits in approximately the same amount as their short futures positions.

**Fact 3.** *Less capitalized emitters:*

*(i) have lower storage of permits;*

*(ii) delay purchases of permits within the annual compliance cycle;*

*(iii) hedge relatively more using futures than permits.*

Figure 4 illustrates point (i) of Fact 3. It plots emitters' permit holdings as a function of emitters' book equity over total assets. Permit holdings are evaluated just after the surrender date (April 30) and thus correspond to (at least temporary) storage. The figure shows that emitters with higher equity store more permits. The interpretation is that permit holdings require upfront financing, which is costly or impractical for firms with binding financial constraints, as proxied by low equity.

[Go To Figure 4]

To complement Figure 4, Table 1 presents estimates of the regression of permit holdings normalized by emissions onto equity to total asset, permit shortfall (defined as one minus free permits over emissions), log of total assets, and tangible assets to total assets. The coefficient on the equity ratio is significantly positive, even after controlling for year fixed effects, sector shares, permits shortfall, firm size, and asset tangibility. We also find that larger firms hold more permits, as well as firms with higher asset tangibility. The latter finding is consistent with tangibility relaxing financial constraints and thus facilitating investment in permits. In the specification with all the controls (column 4), a one-standard deviation increase in the equity ratio (24 percentage points) is associated to a 29 percentage points increase in the ratio of permit holdings to annual emissions, which corresponds to a 20% increase relative to the mean.

[Go To Table 1]

Figure 5 illustrates point (ii) of Fact 3. It plots the share of permit purchases occurring in each month of the annual compliance cycle for emitters with above-median equity ratio (blue) and emitters with below-median equity ratio (red). We order months from May to April, as emission permits must be surrendered at the end of April. The purchase dates are the dates at which the transfer of permits is recorded in the transaction log. This also includes transactions corresponding to the settlement of futures trades, as permits are exchanged for cash at futures settlement.

Figure 5 shows that most of the purchases occur in December and March—the expiry of the two most traded futures contracts, and April—the month at the end of which emitters must surrender permits to cover their emissions.

Moreover, Figure 5 shows that emitters with a lower equity ratio delay their purchases of permits relative to emitters with a higher equity ratio. This is consistent with the notion that financially constrained firms delay cash outflows.

[Go To Figure 5]

Table 2 tests the relation between the timing of permit purchases and the equity ratio more formally. We define $TimeUntilBuy$ as the weighted-average number of days between the beginning of the annual compliance cycle and the date of the purchases: $\left[\sum_{d=1}^{365} d \times Purchases_{i,t}(d)\right]$ / $\left[\sum_{d=1}^{365} Purchases_{i,t}(d)\right]$, where $Purchases_{i,t}(d)$ is the number of permits purchased on day $d$ by emitter $i$ within the annual compliance cycle $t$. $d = 1$ corresponds to May $1^{\text{st}}$ and $d = 365$ to April 30 of the following year. Columns (1) and (2) show that the regression coefficient of $TimeUntilBuy$ onto the equity ratio is significantly negative. In the specification with all controls (column 2), a one-standard-deviation reduction in the equity ratio is associated with a 12 days later average time of purchases. In columns (3) and (4), we repeat the analysis using sales instead of purchases. In this case, the relationship is insignificant.

[Go To Table 2]

Figure 6 illustrates point (iii) of Fact 3. It shows that, compared with equity-rich emitters, emitters with low equity ratios hedge more with futures than with permit holdings. We define $FuturesMinusSpot$ as (proxied) positions in futures expiring during an annual compliance cycle minus permit holdings at the beginning of the cycle, as a share of annual emissions. Figure 6 plots this variable against the equity ratio, either using all emitters (blue) or restricting to emitters with nonzero futures futures positions during the sample period (red). It shows that firms with a higher equity ratio tend to rely less on futures to hedge permit price volatility.

[Go To Figure 6]

Table 3 offers a more formal test of the link between the equity ratio and the reliance on futures relative to holdings. Both for the full sample of emitters as well as for the subsample of emitters that use futures, we find that firms with a higher equity ratio hedge relatively less using futures than storage of permits, even after controlling for permit shortfalls, size, and asset tangibility.

[Go To Table 3]

**Fact 4.** *The discounted futures price minus the spot price of emission permits (also known as the basis) is on average positive.*

If the market for emissions permits was frictionless, the spot price and the futures price should be connected by the law of one price. To cover future emissions, an emitter can either buy spot today or go long in futures that settle before emission permits need to be surrendered. The difference between the futures and the spot price should thus exactly compensate emitters for the opportunity cost of committing their capital today when they buy spot. In other words, the basis, defined as the difference between the discounted futures price and the spot price, should be zero.

Figure 7 shows the evolution of the daily basis over time, where for each date the futures price is the price of the futures contract that settles in December of that year (the futures contract with the highest open interest) discounted at the Euribor rate[18] and the spot price is from the emission permits spot market on EEX.[19] As is clearly evident from the figure, the no-arbitrage restriction on futures and spot prices does not hold in the data. In violation of the law of one price, the futures price tends to be above the spot (permit) price, i.e., the basis is positive. Moreover, there is clear evidence of an annual cycle: the basis is largest at the beginning of the year, when the December futures settlement date is farther away, and it shrinks over the course of the year as the time to maturity shortens.

[Go To Figure 7]

18. Specifically, for each date, we discount the futures price by an interpolation of the two Euribor rates that straddle the time-to-maturity (e.g., when the time-to-maturity is two months, we use the average of the one-month and three-month Euribor rates). In practice, the relevant interest rate for a trader exploiting the arbitrage opportunity (leveraged long spot position and short futures) is the collateralized interest rate such as the repo rate. To the extent that the collateralized rate is below the uncollateralized rate such as Euribor, our measure of the basis is conservative.

19. Our results are robust to using auction prices or 1-day futures prices from ICE instead of the EEX spot price.

Table 4 confirms these patterns. The basis is significantly positive on average, at around €0.11 or 41 basis points when scaled by the spot price (column 3). Moreover, the basis increases with time-to-maturity (columns 2 and 4). For instance, column 4 implies that the basis is around 80 basis points ($\simeq 0.05 + 360 \times 0.0021$) at the beginning of the calendar year. The persistence of a profitable basis trade implies that arbitrage is limited, consistent with binding financial constraints.

[Go To Table 4]

## 3 Modeling Cap-and-Trade Under Financial Constraints

In this section, we present and solve a simple model of emitters and financials in a cap-and-trade system, which rationalizes the stylized facts observed in the data.

### 3.1 Setup

There are two dates $t = 0, 1$. State of nature $s = \ell$ or $s = h$ is realized at time 1 with probability $\pi_s$. There are two types of agents: emitters ($i = E$) and financials ($i = F$), each in mass one. The government manages a cap-and-trade system by issuing and collecting emission permits and verifying emissions. The government issues $\overline{n}_0 \geq 0$ emission permits at time 0 and $\overline{n}_s \geq 0$ emission permits at time 1 in state $s$. Without loss of generality we set $\overline{n}_\ell \leq \overline{n}_h$. In this section the parameters of the cap-and-trade policy $(\overline{n}_0, \overline{n}_s, s \in \{\ell, h\})$ are exogenous. The optimal policy is characterized in the next section.

**Time 0.** Agents can trade emission permits as well as Arrow securities for each time-1 state of nature. Permits trade at price $p_0$. The state $s$ Arrow security, with payoff equal to one in state $s$ and 0 in the other state, trades at price $q_s$. We denote by $n_i \geq 0$ agent $i$'s position in emission permits and by $a_{is}$ agent $i$'s position in the state $s$ Arrow security. The government rebates lump-sum the proceeds from permits issuance $p_0\overline{n}_0$ to emitters.[20] The time-0 budget constraints of emitters and financials are, respectively:

$$p_0 n_E + \sum_{s=\ell,h} q_s a_{Es} \leq p_0 \overline{n}_0, \tag{1}$$

$$p_0 n_F + \sum_{s=\ell,h} q_s a_{Fs} \leq 0. \tag{2}$$

20. We relax the assumption that the proceeds are rebated to emitters in Section 4 when we solve for the constrained optimum.

Market clearing for emission permits implies $n_E + n_F = \overline{n}_0$. Market clearing for each state-$s$ Arrow securities implies $a_{Es} + a_{Fs} = 0$.

**Time 1, state *s*.** Permits trade at price $p_s$. Emitters choose emissions $e_s$ to produce $f(e_s)$, where $f(.)$ satisfies the Inada conditions $f' > 0$, $f'' < 0$, $f(0) = 0$, $f'(0) = \infty$, and $f'(\infty) = 0$. Emitters must surrender to the government a number of permits equal to their emissions. Financials have endowment $y$, which does not depend on the state $s$, and, for simplicity, does not generate emissions. The government rebates lump-sum the proceeds from permits issuance $p_s\overline{n}_s$ to the emitters.

Agents consume $c_{is}$ and obtain utility

$$u(c_{is}) - \delta_s\overline{e}_s = \log(c_{is}) - \delta_s\overline{e}_s \tag{3}$$

where aggregate emissions $\overline{e}_s$ is a negative externality determined by the sum of individual emissions over the mass of emitters, and $\delta_s$ is the social cost of carbon. In a symmetric equilibrium, $\overline{e}_s = e_s$ since there is a mass one of identical emitters. Agents are atomistic; they do not consider the impact of their own emissions on $\overline{e}_s$. Consequently, for a given supply of permits $\overline{n}_0$ and $\overline{n}_s$, agents' behavior and equilibrium do not depend on $\delta_s$.[21] Therefore, the only relevant source of risk for agents is the supply of permits $\overline{n}_s$.

The time-1 budget constraints of emitters and financials are, respectively:

$$c_{Es} \leq f(e_s) - p_s e_s + p_s\overline{n}_s + p_s n_E + a_{Es}, \tag{4}$$

$$c_{Fs} \leq y + p_s n_F + a_{Fs}. \tag{5}$$

Emitters' time-1 budget constraint is that their consumption cannot exceed their income. Emitters income is risky for two reasons. First, emitters' output $f(e_s)$ is risky, because emissions caps vary across states. Second, the cost of compliance is risky because permit prices are volatile. These two risk exposures correspond to the notion of transition risk: more stringent environmental policies (with lower emissions caps and higher permit prices) lead to lower income for emitters.

In practice the government rebates part of the proceeds from permits issuance to emitters, and also issues free allowances. In our model, for simplicity, we assume the government fully rebates proceeds from permit issuance to emitters (as lump sums). This, however, only partly offset transition risk for emitters, because emitters are still exposed to the risk of low output,

21. $\delta_s$ will only matter when we endogenize the emissions cap in Section 4.

due to low caps.

Emitters must surrender $e_s$ emission permits. Therefore, market clearing for permits implies $e_s = \overline{n}_0 + \overline{n}_s$ in each state $s$.

**Incentive constraint.** Following Bolton and Scharfstein (1990), DeMarzo and Fishman (2007), and Biais, Hombert, and Weill (2021), we assume that at time 1 agents can abscond with fraction $\theta \in [0, 1]$ of income from production and assets held. There are several complementary motivations for this assumption. First, agents could use firms' resources to engage in activities generating private benefits for them, but depleting the amount of resources left available to pay creditors. For example, the agent could fund philanthropic activities, hire relatively inefficient employees to whom the agent is related, or tunnel resources to other firms. Second, absconding with a fraction of resources is equivalent in our model to not exerting costly unobservable effort (see DeMarzo and Fishman 2007). Third, as in Biais, Hombert, and Weill (2021), instead of assuming agents can abscond with fraction $\theta$ of resources, we could assume agents can make a take-it-or-leave-it offer to the creditors to write off part of their claim. If for creditors the cost of the write off is lower than the cost of seizing assets in a bankruptcy procedure, they accept the write off.

Under our assumption that agents can abscond with a fraction $\theta$ of resources, the incentive compatibility conditions are

$$c_{Es} \geq \theta\big[f(e_s) + p_s n_E\big], \tag{6}$$

$$c_{Fs} \geq \theta\big[y + p_s n_F\big]. \tag{7}$$

Conditions (6) and (7) apply when income from production and permits' holdings are imperfectly pledgeable, but income from Arrow securities is perfectly pledgeable. This is just for simplicity. In Appendix B.2 we show that our results are robust to relaxing this assumption and assuming that the payoff from Arrow securities is imperfectly pledgeable.

**Intertemporal budget constraint.** Agents maximize

$$E\big[u(c_{is}) - \delta_s \overline{e}_s\big] \qquad i = E, F, \tag{8}$$

subject to budget and incentive constraints. It is convenient to consolidate the budget constraints by substituting the time-1 budget constraints (4 and 5) into the time-0 budget constraints (1 and 2). We thus eliminate the terms $a_{is}$ and obtain the intertemporal budget constraint for emitters and financials, respectively:

$$\sum_{s=\ell,h} q_s c_{Es} \leq \sum_{s=\ell,h} q_s f(e_s) - \sum_{s=\ell,h} q_s p_s e_s + \Big(\sum_{s=\ell,h} q_s p_s - p_0\Big) n_E + p_0 \overline{n}_0 + \sum_{s=\ell,h} q_s p_s \overline{n}_s, \quad (9)$$

$$\sum_{s=\ell,h} q_s c_{Fs} \leq \sum_{s=\ell,h} q_s y + \Big(\sum_{s=\ell,h} q_s p_s - p_0\Big) n_F. \quad (10)$$

The time-0 prices of permits and Arrow securities are determined up to a multiplicative constant. The reason is that agents do not consume at time 0, so the risk-free interest rate $1/\sum_{s=\ell,h} q_s$ is not pinned down. Without loss of generality, we normalize the risk-free interest rate to one, i.e., $\sum_{s=\ell,h} q_s = 1$.

### 3.2 First Order Conditions

Maximizing expected utility (8), subject to the intertemporal budget constraints (9 and 10) and the incentive constraints (6 and 7), the first order condition for consumption in state $s$ is

$$\pi_s u'(c_{is}) + \mu_{is} = \lambda_i q_s \qquad i = E, F, \quad (11)$$

where $\mu_{is}$ is the multiplier of agent $i$'s incentive constraint in state $s$ and $\lambda_i$ is the multiplier of agent $i$'s intertemporal budget constraint constraint. Under (11), without incentive constraints, marginal utility is equal to the multiplier of the budget constraint multiplied by the state price. Incentive constraints introduce a wedge between these two terms. Agents whose incentive constraint binds in a given state have relatively lower marginal utility, reflecting that they must have sufficiently high consumption for their incentive constraint (6 or 7) to hold.

The first order condition for agent $E$'s emissions in state $s$ is

$$f'(e_s)\left(1 - \frac{\mu_{Es}\theta}{\lambda_E q_s}\right) = p_s. \quad (12)$$

Without incentive constraints, i.e., when $\mu_{Es}\theta = 0$, (12) states that the marginal productivity of emitters is equal to the price of emission permits. In contrast, with incentive constraints, there is a wedge between these two terms. When the incentive constraint of emitters is binding, the price of permits is below marginal productivity. In this context, why don't emitters buy more permits, to produce more? Raising production would tighten the emitters' binding incentive constraint. The corresponding cost, proportional to the multiplier of the emitters' incentive constraint $\mu_{Es}$, would offset the increased profit from increased production.

For agents $i$ who hold permits at time 0, the first order condition for time-0 holdings of permits is

$$p_0 = \sum_{s=\ell,h} q_s p_s \left(1 - \frac{\mu_{is}\theta}{\lambda_i q_s}\right), \tag{13}$$

while the first order condition holds with "$\geq$" instead of "$=$" for agents who do not hold permits at time 0. In a frictionless market, without incentive constraints, the permit price at time 0 would be equal to the price of the replicating portfolio of Arrow securities with the same payoff as the permit. Incentive constraints create a deviation from this no-arbitrage condition. The permit price at time 0 is below the price of the replicating portfolio, reflecting that holding permits tightens the incentive constraint, since permits can be diverted.

### 3.3 Equilibrium With Incentive Constraints

Other things equal, emitters' output, and therefore income, is lower when caps are low than when they are high. To hedge that risk, emitters should hold a portfolio of permits and Arrow securities that is worth more in the low-cap state than in the high-cap state, i.e.,

$$p_\ell n_E + a_{E\ell} > p_h n_E + a_{Eh}. \tag{14}$$

The state-$\ell$ Arrow securities purchased by emitters are sold to them by financials. This can be interpreted as insurance supply. The promised insurance payment, however, must not be so large that financials would be tempted to strategically default on it. In exchange for the state-$\ell$ Arrow securities sold to them by financials, emitters sell state-$h$ Arrow securities, but this promise must not be so large that emitters would be tempted to default on it. In line with this discussion we formulate the following conjecture:

**Conjecture 1.** *The incentive constraint of financials binds in state $\ell$, while the incentive constraint of emitters' binds in state $h$.*

In our equilibrium analysis, we will prove that this conjecture holds. Our goal is to understand how incentive constraints affect the interaction between financials and emitters. So we focus on parameter values such that incentive constraints bind, by assuming that:

$$\frac{f(e_\ell) + (1-\theta)y}{\theta y} < \frac{\pi_\ell}{\pi_h} < \frac{\theta f(e_h)}{y + (1-\theta) f(e_h)}. \tag{15}$$

(15) ensures that the state probabilities are not too unbalanced, otherwise the incentive constraint of one of the agent would always be slack. The range of values for the state probability ratio in (15) is not empty if $\theta$ is not too small. We also assume that $\overline{n}_0$ is small to ensure that emitters' divertible income (the RHS of (6)) is higher in the high-cap state than in the

low-cap state. In Section 4, we solve for the optimal emissions cap and show that it satisfies the assumption of small $\overline{n}_0$.

Under these assumptions, we obtain the following proposition:

**Proposition 1.** *If (15) holds and $\overline{n}_0$ is small, equilibrium is such that:*

*(i) Emitters' incentive constraint binds in the high-cap state and is slack in the low-cap state, while the financials' incentive constraint binds in the low-cap state and is slack in the high-cap state, confirming Conjecture 1.*

*(ii) The price of emission permits is higher in the low-cap state than in the high-cap state.*

*(iii) The price of emission permits is strictly lower than the price of the portfolio of Arrow securities replicating the payoff of the emission permits.*

*(iv) Emitters' position in Arrow securities is larger for state $\ell$ than for state $h$, and, by market clearing, financials take the exact opposite positions in Arrow securities.*

The interpretation of point (i) in Proposition 1 is the following: Incentive constraints limit the payment each agent can commit to make to the other agent at time 1, hence they limit the provision of insurance. In the low-cap state, the incentive constraint of financials (7) binds, forcing financials to consume more than absent the constraint, or equivalently, preventing financials from providing as much insurance to emitters as absent the constraint. The binding incentive constraint is reflected by the wedge $\mu_{F\ell}$ in the first-order condition for consumption (11), which implies that financials' marginal utility in this state is lower than under perfect risk sharing. Symmetrically, emitters' incentive constraint binds in state $h$, i.e., emitters consume more and have lower marginal utility than under perfect risk sharing.

The interpretation of point (ii) in Proposition 1 is the following: The first-order condition for emissions (12) implies that the permit price is below the marginal productivity of emitters when emitters are constrained (in state $h$) and equal to the marginal productivity when emitters are unconstrained (in state $\ell$). Therefore, $e_h > e_\ell$ implies:

$$p_h < f'(e_h) < f'(e_\ell) = p_\ell. \tag{16}$$

The inequality between $p_h$ and $f'(e_h)$ is strict because the incentive constraint of emitters binds in state $h$. If there were no incentive constraints, instead of a strict inequality, this would be an equality.

The interpretation of point (iii) in Proposition 1 is that the time-0 permit price is strictly less than the no-arbitrage price, as foreshadowed by our discussion of equation (13). This arises because buying permits tightens the incentive constraint, as can be seen from (6)-(7). To see why this is the case, consider an agent $i$ buying one permit financed by issuing $p_s$ units of state-$s$ Arrow securities for each $s$. This trade generates income and liabilities both equal to $p_s$ at time 1 in state $s$, but pledgeable income is only $(1-\theta)p_s$. Therefore, the incentive constraint is tightened by $\theta p_s$ in each state. Consequently, the permit price is equal to the price of the replicating portfolio of Arrow securities minus a term that reflects the shadow cost of the incentive constraint of the agent holding the permit.

The interpretation of point (iv) in Proposition 1 is the following: As mentioned above, emitters are exposed to two forms of transition risk. The first is that prices are volatile, the second is that production is low when the cap is low (in state $\ell$). Because the proceeds from permit issuance are rebated to emitters, emitters are perfectly hedged against price risk if they buy all the permits issued at time 0. This, however, does not insure emitters against the risk of low production due to low emissions caps in state $\ell$. It is to hedge that risk that emitters purchase more state-$\ell$ Arrow securities than state-$h$ Arrow securities.

While points (i) and (iii) in the proposition stem from the binding incentive constraints, points (ii) and (iv) also hold without incentive constraints.

## 3.4 Relating the Model to Stylized Facts From the EU ETS

In this subsection, we explain how the implications of our theoretical analysis fit the stylized facts from the European Union Emission Trading System documented in the previous section.

### 3.4.1 Emitters' Exposure to Transition Risk (Fact 1)

The main ingredient in our model is that emitters are exposed to transition risk. As discussed above, one component of this risk is the volatility of permit prices, which is large in practice, as documented in the previous section in Fact 1.[22] Moreover, the corresponding risk represents a large fraction of emitters' revenues.

### 3.4.2 Positions in Futures Contracts (Fact 2)

While Arrow securities can be interpreted as financial derivatives, in practice, market participants don't actually trade Arrow securities. As documented in the previous section, in the EU ETS, they trade emission permits and derivatives on emission permits. The patterns uncovered by our theoretical analysis can, however, be mapped into patterns actually observed in practice by noting that the portfolios of Arrow securities analyzed in our model can be implemented with

22. Unlike in our model, this price risk is not fully offset by rebates in the form of free permits, as only about half of permits are freely allocated while half are auctioned by the EU.

portfolios of futures on emission permits and riskfree debt. As shown below, this enables us to clarify that the implications of our theoretical model for futures holdings are in line with Fact 2 established in the previous section.

The portfolio of Arrow securities $(a_{i\ell}, a_{ih})$ can be implemented with a portfolio composed of $\psi_i$ units of risk-free asset, paying off 1 in all states, and $\phi_i$ units of a futures contract on emission permits. At time 1 in state $s$, an agent long one futures contract purchases one permit at the futures price $F_0$. Since the permit price is $p_s$, this generates payoff $p_s - F_0$ in state $s$. The portfolio composed of Arrow securities and the portfolio composed of the risk free asset and futures contracts have the same final payoff if and only if

$$\psi_i + \phi_i(p_s - F_0) = a_{is}, \forall s. \tag{17}$$

Solving the values of $\psi_i$ and $\phi_i$ such that (17) holds gives the number of units of risk-free asset

$$\psi_i = a_{i\ell}\frac{F_0 - P_h}{p_\ell - p_h} + a_{ih}\frac{P_\ell - F_0}{p_\ell - p_h}, \tag{18}$$

and the number of futures contracts

$$\phi_i = \frac{a_{i\ell} - a_{ih}}{p_\ell - p_h}, \tag{19}$$

that implement the portfolio of Arrow securities $(a_{i\ell}, a_{ih})$. To complete our implementation of the Arrow-securities equilibrium with futures contracts and the risk-free asset, we assume that at time 1, when agents must choose between fulfilling their promises and strategically defaulting, they can either fully default on all their financial contracts (futures and risk-free asset) or fulfill their obligations for all these contracts. Under that assumption the incentive constraint is the same when market participants trade Arrow securities and when they trade futures contracts and the riskfree asset.[23] In this context, market clearing, equation (19), and point (iv) of Proposition 1 yield our next proposition:

**Proposition 2.** *The equilibrium Arrow securities positions of the agents can be implemented with positions in futures and the risk-free asset. In this implementation, emitters take long positions in permits futures $\phi_E > 0$, while financials take the opposite, short, positions in permits futures: $\phi_F = -\phi_E$.*

23. The assumption rules out the case in which agents would default on one of their financial positions but not the other. For example, if the agent borrowed and bought futures and the futures position was profit making at time 1, our assumption rules out execution of the profitable futures contract simultaneously with strategic default on debt. Note that we maintain our initial assumption (giving rise to incentive constraints) that agents can default on (all) their financial contracts while absconding with a fraction $\theta$ of their non-financial assets.

Proposition 2 is in line with the stylized fact from the European Union Emission Trading System that emitters hold long positions in futures, bought from financials, who thus hold short positions in futures (see Fact 2 in the previous section). The economic interpretation of Proposition 2 is the following. Emission permits holdings, combined with rebates, enable agents to hedge price risk. To also hedge output risk (stemming from the fact that output is lower when emissions caps are tighter), agents need to also conduct financial trades, be they in Arrow securities, or, more in line with practice, in permits futures and riskfree asset.

### 3.4.3 Positions in Physical Permits (Fact 3)

While Proposition 2 characterizes futures holdings, we now study permits holdings. The analysis leading to Proposition 1 implies that in equilibrium, permits are held by the agent's type whose shadow cost of holding permits is the lowest, or by both agents' types if this shadow cost is equalized across agent types. Under condition (15) both types of agents are constrained, but which of the two agents has the lowest shadow cost of holding permits depends on parameter values. Thus, while our model delivers sharp implications on which type of agent should hold futures, it remains rather agnostic about which type of agents should hold permits. The contrast between these two implications echoes our empirical findings. As documented in the previous section, in the data, emitters (resp. financials) hold long (resp. short) positions in the futures markets. In contrast in the spot market, there is no clear difference between the aggregate position of the emitters and that of the financials. The contrast between spot and futures would not arise in a frictionless market and is due to incentive constraints. That the contrast between spot and futures suggested by our model is also present in the data is consistent with the presence of incentive constraints in the EU Emission Trading System.

While our basic model does not delivers sharp implications on the aggregate spot position of the emitters, we now consider an extension delivering implications on the distribution of permit holdings across emitters. To do so we assume emitters have heterogeneous initial endowments, resulting in different levels of capitalization, and, in turn, different (endogenous) shadow costs of holding permits. This allows us to match the empirical fact that less capitalized emitters hold fewer permits (see Fact 3).

Specifically, we now assume that there are two types of emitters $E^+$ and $E^-$, each in mass one-half. At time 0, emitters of type $E^+$ have endowment $2A$, where $A > 0$ is close to zero, while emitters of type $E^-$ have zero endowment. Both types of emitters must consume $A$ at time 0, otherwise their utility is minus infinity. It implies that, in equilibrium, emitters with low endowment borrow $A$ from emitters with high endowment. Therefore, the time-0 budget constraint of emitters of type $E^\varepsilon$, $\varepsilon \in \{+,-\}$, becomes:

$$p_0 n_{E^\varepsilon} + \sum_{s=\ell,h} q_s a_{E^\varepsilon s} \leq p_0 \overline{n}_0 + \varepsilon A. \tag{20}$$

The rest of the model is the same as in the baseline. We consider parameter values such that both types of emitters hold permits; see Appendix B.3 for a sufficient condition on parameters such that it is the case.

That emitters with high endowment hold more permits than emitters with low endowment can be established by contradiction. If both types of emitters held the same amount of permits, higher net wealth at time 0 would imply that emitters with high endowment would hold a portfolio of Arrow securities with higher value, i.e., they would have lower liabilities. Lower liabilities translate into a less binding incentive constraint in state $h$, i.e., a lower shadow cost of holding permits. Therefore, emitters with high endowment would hold more permits: a contradiction.

Thus, in equilibrium, emitters with high endowment hold more permits. This tightens their incentive constraint, up to the point where the shadow cost of holding permits is equalized for both types of emitters. Noting that the higher (resp. lower) net wealth of emitters with a positive (resp. negative) shock can be interpreted as higher (resp. lower) capitalization, our above analysis yields the next proposition:

**Proposition 3.** *Assume* (15) *holds and $\overline{n}_0$ is small. Relative to more capitalized emitters, less capitalized emitters:*

*(i) hold less permits at time 0 (i.e., $n_{E^-} < n_{E^+}$);*

*(ii) purchase more permits at time 1 (i.e., $e_{E^-s} - n_{E^-} > e_{E^+s} - n_{E^+}$);*

*(iii) hedge relatively more using futures than permits.*

Point (i) of Proposition 3 is consistent with Fact 3(i), reported in the previous section, that less capitalized emitters have lower storage of permits. Points (i) and (ii) of Proposition 3 are consistent with Fact 3(ii), that less capitalized emitters delay the purchase of permits towards the date at which they have to surrender permits. Point (iii) of Proposition 3 is consistent with Fact 3(iii), that less capitalized emitters hedge relatively more with futures than permits.

#### 3.4.4 Basis (Fact 4)

In our theoretical model, because of financial constraints, no-arbitrage pricing restrictions can be violated. In particular, point (iii) in Proposition 1 states that the permit price at time 0 should be lower than the price of the replicating portfolio of derivative contracts. This is in line with Fact 4, documented in the previous section, that there is a significantly positive basis. Thus, while inconsistent with the law of one price that should prevail in frictionless markets, our empirical findings are consistent with the predictions of our constrained equilibrium model obtaining under condition (15), which rules out cases where one agent would never be constrained (in which the basis would have been zero). In such cases, there would be no basis, in contradiction with our empirical finding.

Thus, overall, the patterns observed in the data are consistent with the implications of our positive analysis of the market equilibrium with cap-and-trade and incentive constraints. Two important questions arise: Since incentive constraints make hedging imperfect, is it still socially optimal to rely on a cap-and-trade system? And, if so, what are the characteristics of a constrained optimal cap-and-trade system? Answering these questions is the goal of the normative analysis presented next.

# 4 Constrained Optimal Emissions Cap

Our analysis above took as given the caps prevailing in the two states, $e_\ell$ and $e_h$. We now study how these caps can be set by a benevolent planner taking into account the impact of emissions on welfare.

We first characterize the consumption and production plan set by the benevolent planner to maximize utilitarian welfare subject to incentive and resource constraints. Second, we show how this incentive-constrained optimal allocation can be implemented as the equilibrium of a market with a cap-and-trade system, and we spell out the implications of our analysis.

## 4.1 Second best

### 4.1.1 Planner's Problem

The social cost of emissions in state $s \in \{\ell, h\}$ is $\delta_s$. We assume $\delta_h < \delta_\ell$. The planner chooses emissions $e_s$ and consumptions $c_{is}$ in each state to maximize a weighted sum of the agents' utilities, which are given in equation (3). The weights in this objective are the Pareto weights $(\alpha_E, \alpha_F) \in [0,1]^2$ on emitters and financials, respectively, with $\alpha_E + \alpha_F = 1$. So the planner's problem is to maximize

$$\sum_{i=E,F} \alpha_i E\big[u(c_{is}) - \delta_s \bar{e}_s\big], \tag{21}$$

subject to the resource constraint at time 1 in each state

$$c_{Es} + c_{Fs} \leq f(e_s) + y, \tag{22}$$

and incentive constraints

$$c_{Es} \geq \theta f(e_s), \tag{23}$$

$$c_{Fs} \geq \theta y. \tag{24}$$

The formulation of the incentive constraints in the planner's problem differs from its counterpart in the equilibrium analyzed above. In the former, agents only hold their output $f(e_s)$ or $y$, and this is the only thing they can divert. In the latter, agents also hold permits purchased at time 0, in addition to their output. This could suggest that incentive constraints are more severe in the equilibrium than in the planner's problem. We, however, show below that the optimal incentive-constrained allocation can be implemented by a well designed cap-and-trade system. In that system, the number of permits issued at time 0 is set to 0, to relax incentive constraints. As shown below this does not constrain risk sharing, because the latter can be achieved with Arrow securities that don't tighten incentive constraints.

The first order condition for consumption of agent $i$ in state $s$ is

$$\alpha_i \pi_s u'(c_{is}) + \mu^*_{is} = \lambda^*_s, \qquad i = E, F, \tag{25}$$

where $\mu^*_{is}$ is the multiplier of agent $i$'s incentive constraint in state $s$, and $\lambda^*_s$ is the multiplier of the resource constraint in state $s$. The left-hand side of equation (25) is the benefit from increasing $c_{is}$, which reflects both the increase in agent $i$'s utility from consumption and the relaxation of this agent's incentive constraint. The right-hand side is the cost of increasing $c_{is}$, which is the multiplier of the resource constraint. Without incentive constraints, marginal utility weighted by the Pareto weight is equal to the multiplier of the resource constraint. With incentive constraints there is a wedge ($\mu^*_{is}$) between these two terms, similarly to the decentralized equilibrium.

The first order condition for emissions in state $s$ is

$$f'(e_s)\left(1 - \frac{\mu^*_{Es}\theta}{\lambda^*_s}\right) = \frac{\pi_s}{\lambda^*_s}\delta_s, \qquad s = h, \ell. \tag{26}$$

Without incentive constraints, the marginal productivity of emitters times the multiplier of the resource constraint is equal to the social cost of emissions. Again, similarly to the decentralized equilibrium, incentive constraints introduce a wedge (proportional to $\mu^*_{Es}$) between these two terms.

### 4.1.2 First-Best Emissions

To provide a benchmark, we first characterize optimal emissions cap for the case without incentives constraints, i.e., $\theta = 0$ and $\mu^*_{Es} = \mu^*_{Fs} = 0$ for all $s$. Using the first order condition for consumption (25) and the resource constraint (22), log utility implies that agents' consumption shares are equal to Pareto weights :

$$c_{is} = \alpha_i [f(e_s) + y], \qquad i = E, F. \tag{27}$$

Substituting (25) and (27) into (26), the first order condition for emissions yields :

$$\delta_s = f'(e_s)\alpha_i u'(\alpha_i(f(e_s) + y)), \qquad s = h, \ell. \tag{28}$$

Equation (28) states that the social cost of emissions is equal to marginal productivity multiplied by the Pareto weighted marginal utility of consumption. With log utility, this Pareto weighted marginal utility is equalized across agents. As shown below, this will no longer be the case with binding incentive constraints. Equation (28) implies that in the first-best emissions are decreasing in the social cost of emissions $\delta_s$. Because utility is logarithmic, equation (28) yields our next proposition:

**Proposition 4.** *In the first best, emissions ($e_s^{fb}$) are such that*

$$\delta_s = \frac{f'(e_s^{fb})}{f(e_s^{fb}) + y}, \qquad s = h, \ell, \tag{29}$$

*and thus are decreasing in the social cost of emissions and independent of Pareto weights.*

### 4.1.3 Second-Best Emissions

We now turn to the case in which incentive constraints bind, limiting the amount of resources that can be transferred between agents. To do so we assume:

$$\theta > \max(\alpha_E, \alpha_F) \tag{30}$$

$$\delta_\ell > f'\left(f^{-1}\left(\frac{(\theta - \alpha_F)y}{\alpha_F}\right)\right)\frac{\alpha_F}{\theta y} \tag{31}$$

$$\delta_h < f'\left(f^{-1}\left(\frac{\alpha_E y}{\theta - \alpha_E}\right)\right)\frac{\theta - \alpha_E}{\theta y} \tag{32}$$

These conditions imply that the variance of $\delta_s$ is large enough and pledgeability $1-\theta$ low enough that the amount of insurance agents want to trade in the first best is not incentive compatible, i.e., incentive constraints bind.[24]

In state $\ell$, the social cost of emissions is high and it is optimal to set low emissions. When emissions are low, emitters have low production, so risk sharing requires transfers from financial to emitters, tightening financials' incentive constraint. When $\delta_\ell$ is high (as stated in (31)) and $\theta$ is high, financials' incentive constraint binds, leading to

$$c_{F\ell} = \theta y. \tag{33}$$

Substituting into the resource constraint, this yields the consumption of the emitters:

$$c_{E\ell} = f(e_\ell) + (1-\theta)y. \tag{34}$$

Substituting into the first order condition for emissions, emissions are given by

$$\delta_\ell = f'(e_\ell)\alpha_E u'\big(f(e_\ell) + (1-\theta)y\big). \tag{35}$$

It is optimal to set emissions so that the social cost of emissions is equal to marginal productivity multiplied by the Pareto weighted marginal utility, of the unconstrained agents, i.e., in state $\ell$ the emitters. This reflects that in this case any marginal increase in output is consumed by the emitters. This is because, when financials are constrained, the only way to increase transfers to emitters is to allow them to produce more. As a result, emissions are higher than in the first best.

In state $h$, the social cost of emissions is low, so optimal emissions are high and so is emitters' consumption. Risk sharing thus requires transfers from emitters to financials. When $\delta_h$ is low (as stated in (32)) and $\theta$ is high, emitters' incentive constraint binds, leading to

$$c_{Eh} = \theta f(e_h), \tag{36}$$

$$c_{Fh} = y + (1-\theta)f(e_h). \tag{37}$$

Substituting into the first order condition for emissions, emissions are given by

$$\delta_h = f'(e_h)\left[(1-\theta)\alpha_F u'\big(y + (1-\theta)f(e_h)\big) + \theta\alpha_E u'\big(\theta f(e_h)\big)\right]. \tag{38}$$

24. This parameter set is non-empty because the RHS of (31) is larger than the RHS of (32). This follows from substituting $\alpha_E = 1-\alpha_F$ into the RHS of (32) and using $\theta \leq 1$ together with the fact that $f'(f^{-1}((.))$ is decreasing.

This is similar to Equation (28), in that optimal emissions equate the social cost of emissions to marginal productivity times Pareto weighted marginal utility. The difference is that, in equation (38), Pareto weighted marginal utility is averaged across emitters and financials, with weights $\theta$ and $1-\theta$ respectively. This is because, of any increase in output, emitters require fraction $\theta$ (to satisfy their incentive constraint) and only the fraction $1-\theta$ can be transferred to financials, who are unconstrained (and have higher marginal utility). The planner therefore sets an emissions cap below the first best level, thereby benefiting financials via a lower climate externality.

Building on the above analysis, the next proposition summarizes the properties of the second best.

**Proposition 5.** *In the second best:*

*(i) emissions are decreasing in the social cost of emissions ($e_\ell^{sb} < e_h^{sb}$);*

*(ii) emissions are above first-best emissions when the social cost of emissions is high ($e_\ell^{sb} > e_\ell^{fb}$) and below first-best emissions when the social cost of emissions is low ($e_h^{sb} < e_h^{fb}$), implying that emissions are less volatile than in the first best;*

*(iii) emissions are increasing in the Pareto weight of emitters.*

Point (i) of the proposition states that, as in the first best, emissions caps are decreasing with the social costs of emissions. While varying emissions in response to changes in estimated social cost of carbon creates transition risk for emitters, the planner provides some insurance against that risk by transferring resources between emitters and financials.

Incentive constraints, however, constrain transfers and therefore limit insurance. This leads to point (ii) in the proposition which arises because, to complement limited transfers, it is second best optimal to distort emissions caps, in order to provide insurance:

- When the social cost of emissions is high, emitters produce little, so the planner would like to transfer resources from financials to emitters. When these transfers are limited by incentive constraints, a second-best way to transfer resources to emitters is to allow them to increase emissions above the first-best level. Doing so transfers utility from financials, who suffer from a larger emission externality, to emitters, who enjoy higher income and consumption.

- When the social cost of emissions is low, emitters produce a lot so the planner would like to transfer resources from emitters to financials. When emitters' incentive constraint binds, only fraction $1-\theta$ of any increase in production can be transferred to financials.

> To transfer further utility from emitters to financials, it is second-best optimal to reduce emissions below their first-best level.[25]

Thus, when incentive constraints limit the ability to use transfers to provide insurance against transition risk, it is second best optimal to reduce transition risk by limiting the variance in caps.

Point (iii) of the proposition illustrates another implication of incentive constraints for the optimal policy. Without incentive constraints, the determination of optimal emissions and redistribution are separable problems. Therefore, optimal emissions do not depend on Pareto weights in the first best (see Proposition 4). Incentive constraints limits redistribution between agents. To mitigate this effect, in the second best, emissions are distorted to transfer utility to the agents who carry higher Pareto weights. When emitters carry a higher Pareto weight, emissions increase relative to the first best to transfer utility to emitters.

## 4.2 Implementation and Implications

### 4.2.1 Implementation

We now establish an implementation result.

**Proposition 6.** *A cap-and-trade equilibrium with appropriate emissions caps and no issuance of permits at time 0 implements the second best for some Pareto weights.*

Implementation of the second best rules out permits issuance at $t = 0$. Permits purchased at time 0 would tighten incentive constraints. In contrast, when there is no issuance of permits at $t = 0$ the incentives constraints in the cap-and-trade equilibrium (equations (6) and (7)) coincide with those in the second best (equations (23) and (24)). Correspondingly it is optimal to wait until time-1 to issue a number of permits equal to the desired emissions cap.

To see how the cap-and-trade equilibrium can implement the second best, compare the first order conditions in the equilibrium to their counterparts in the planner's problem. In the equilibrium the first order condition for emissions features the price of permits, while in the second best it features the social cost of emissions. Therefore, to implement the second best the permit price in the market equilibrium must depend on the social cost of emissions:

$$p_s = \frac{\pi_s}{\lambda_s^*}\delta_s. \tag{39}$$

25. There is another effect but it is dominated in the case of log utility. When financials have low consumption, their marginal utility is higher and therefore the value of transferring consumption to them is high. When RRA is large enough, this effect could dominate.

This equation also holds without incentive constraints, in which case its interpretation is that the optimal permit price is equal to the social cost of emissions divided by the marginal utility of the representative agent. By contrast, when incentive constraints bind, there is no representative agent and different agents have different marginal utilities. In this case, the optimal permit price is equal to the social cost of emissions divided by the marginal utility of the unconstrained agent.

Turning to the first order condition for consumption, there are two differences between the equilibrium and the planner's problem. First, in the planner's problem the marginal utility of the agent is multiplied by its Pareto weight. Second, in the planner's problem the multiplier of the resource constraint on the RHS is not agent specific. Therefore, the equilibrium implements the second best iff

$$\frac{\lambda_s^*}{\alpha_i} = \lambda_i q_s. \tag{40}$$

This condition is satisfied for some Pareto weights. Intuitively, the agent with a tighter budget constraint in the equilibrium (higher $\lambda_i$) has a lower Pareto weight $\alpha_i$ in the corresponding planner's problem.

The equilibrium is Pareto optimal in spite of incentive constraints because prices do not appear in the incentive constraints (when time-0 issuance of permits is zero). This implies that the behavior of one agent does not impose externalities on the incentive constraint of another agent.

While Proposition 6 states a First Welfare Theorem, a Second Welfare Theorem also holds. In Appendix B.5, we show that the second best for any Pareto weights can be decentralized as a cap-and-trade equilibrium with appropriate emissions caps, no time-0 issuance of permits, and time-0 lump-sum transfers between agents.

Note that these Welfare Theorems differ from the standard result that when externalities are the only imperfection and are perfectly observable, the first best can be implemented with emissions caps and markets for permits. Indeed, in our setting it is not the first best that can be implemented, but rather the second best, i.e., the incentive constrained Pareto optimum.

#### 4.2.2 Implications

While Proposition 6 states that the second best can be implemented as a market equilibrium with cap-and-trade, it also states that the cap-and-trade system must have certain properties to be second best optimal.

First, Proposition 6 implies that the government should not front-load the issuance of emission permits, but instead issue in each regulatory time-period the number of permits that emitters are to surrender during this period. The EU issues permits in a way that is roughly consistent with this recommendation. The EU did not issue permits worth several decades of emissions when it put in place the ETS with storable permits in 2008. Instead, each year the EU issues a number of permits roughly equal to yearly emissions.

Second, interpreting Pareto weights in terms of influence on public policy, Proposition 5 implies that emission caps should be increasing in the political influence of emitters.

Third, Propositions 5 and 6 imply that the cap-and-trade system must entail low variance in caps relative to the first best, in order to limit transition risk. That being said, financial constraints should not lead to emissions caps uniformly above their first best counterparts. Emissions caps should be larger than their first best counterparts when they are very severe, but lower than their first best counterparts when they are relatively mild.

Fourth, in contrast to emission caps, permit prices are uniformly lower in the equilibrium implementing the second best than in the first best, as stated in the next proposition:

**Proposition 7.** *The permit price in the market equilibrium implementing the second best is lower than in the market equilibrium implementing the first best.*

To understand the intuition of the proposition, consider the first order condition for emissions (12). It reflects that for emitters the marginal benefit of increasing emissions is equal to their marginal cost. The former is equal to the marginal productivity of emissions minus the shadow cost of tightening the incentive constraint. The latter is equal to the permit price.

In state $\ell$, emitters are unconstrained, so the shadow cost of the incentive constraint is equal to zero and the permit price is equal to marginal productivity. Now, since emissions are higher than in the first best, marginal productivity and hence the permit price are lower than in the first best.

In state $h$, in the second best emitters are constrained, so the shadow cost of incentives is strictly positive and the permit price is strictly lower than marginal productivity, while in the first best the permit price equal is equal to marginal productivity. As a result, the permit price is lower than in the first best.[26]

Estimates of the social cost of carbon combine analysis of climate and analysis of the economic consequences of climate (see Barrage and Nordhaus 2024). These models usually do not factor in financial constraints, so the estimates of the social cost of carbon that they generate can be interpreted as the first best permit price. In contrast, our model studies permit prices under financial constraints and shows that these prices are lower than their first best counterparts.

26. There is a countervailing effect because emissions are lower in the second best than in the first best so marginal productivity is higher. However, this countervailing effect is an effect induced by incentive problems and is always dominated by the direct effect of the shadow cost of incentives on the permit price.

This can offer a rational for the empirical observation that permit prices in practice are lower than the social cost of carbon (see Tol 2023).

Propositions 5 and 7 underscore that the implications of incentive problems are different for the optimal quantity of emissions and for the optimal price of emissions. Although the second-best price is always lower than in its first-best counterpart, emissions are not always higher in the second-best than in the first-best. Therefore, the answer to the question whether financial frictions make optimal environmental policy more or less strict depends on the metric under consideration. In terms of emissions, financial frictions make the optimal policy more strict in good states but less strict in bad states. In contrast, in terms of carbon prices, financial frictions make the price of permits lower in both states.

## 5 Conclusion

In this paper we analyze cap-and-trade systems, which subject emitters to transition risk, reflecting variability in emissions caps and permit prices. This variability, itself, reflects changes in estimates of the social cost of carbon and political preferences shocks. We examine how emitters react to this risk and how financial frictions impact this reaction.

First, we present stylized facts on the EU ETS. We find that emitters hedge transition risk by buying futures contracts on emission permits. These contracts are sold by financials, hedging their short futures positions with long positions in the spot market. We also find that more financially constrained emitters tend to delay their purchases of permits towards the surrender date, and that the basis between futures and permits is positive, i.e., spot permit prices are lower than discounted futures prices.

Second, we present and analyze a simple model of equilibrium in a cap-and-trade system. The properties of equilibrium derived in our theoretical model are in line with the above mentioned stylized facts. In particular, in our theoretical model financial constraints imply that the basis between futures and spot permits should be positive and that more financially constrained emitters should delay permits' purchases, as is the case in the data.

Third, we study cap-and-trade systems from a normative point of view. We show that, in spite of financial constraints, cap-and-trade remains optimal, in the sense that the constrained Pareto allocation can be achieved as the equilibrium of a market with cap-and-trade. We show however, that the cap-and-trade system implementing the second best, in the presence of financial frictions, differs from the cap-and-trade system implementing the first best, prevailing when there are no frictions. In particular, because financial frictions hinder emitters' hedging, to reduce emitters' risk bearing, in the second best optimal cap-and-trade system there is less variability in emission caps than in the first best. That is, optimal caps are less reactive to changes in the estimate of the social cost of carbon in the second best than in the first best.

## References


Abrell, Jan. 2024. "Database for the European Union Transaction Log." *Working paper.*

Akey, Pat, Ian Appel, Aymeric Bellon, and Johannes Klausmann. 2024. "Do carbon markets undermine private climate initiatives?" *Available at SSRN.*

Alvarez, Fernando, and Urban J Jermann. 2000. "Efficiency, equilibrium, and asset pricing with risk of default." *Econometrica* 68 (4): 775–797.

Anderson, Robert, and Haosui Duanmu. 2025. "Cap-and-trade and carbon tax meet Arrow–Debreu." *Econometrica* 93 (2): 357–393.

Bai, Jennie, and Hong Ru. 2024. "Carbon emissions trading and environmental protection: International evidence." *Management Science* 70 (7): 4593–4603.

Barrage, Lint, and William Nordhaus. 2024. "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model." *Proceedings of the National Academy of Sciences* 121 (13): e2312030121.

Bartram, Söhnke M, Kewei Hou, and Sehoon Kim. 2022. "Real effects of climate policy: Financial constraints and spillovers." *Journal of Financial Economics* 143 (2): 668–696.

Biais, Bruno, Johan Hombert, and Pierre-Olivier Weill. 2021. "Incentive Constrained Risk Sharing, Segmentation, and Asset Pricing." *American Economic Review* 111:3575–3610.

Bolton, Patrick, and Marcin Kacperczyk. 2021. "Do investors care about carbon risk?" *Journal of financial economics* 142 (2): 517–549.

———. 2023. "Global pricing of carbon-transition risk." *The Journal of Finance* 78 (6): 3677–3754.

Bolton, Patrick, and David S Scharfstein. 1990. "A theory of predation based on agency problems in financial contracting." *American economic review,* 93–106.

Borri, Nicola, Yukun Liu, Aleh Tsyvinski, and Xi Wu. 2024. "Inefficiencies of Carbon Trading Markets." *arXiv preprint arXiv:2408.06497.*

Colmer, Jonathan, Ralf Martin, Mirabelle Muûls, and Ulrich J Wagner. 2025. "Does pricing carbon mitigate climate change? firm-level evidence from the european union emissions trading system." *Review of Economic Studies* 92 (3): 1625–1660.

Dechezleprêtre, Antoine, Daniel Nachtigall, and Frank Venmans. 2023. "The joint impact of the European Union emissions trading system on carbon emissions and economic performance." *Journal of Environmental Economics and Management* 118:102758.

DeMarzo, Peter M, and Michael J Fishman. 2007. "Optimal long-term financial contracting." *The Review of Financial Studies* 20 (6): 2079–2128.

Döttling, Robin, and Magdalena Rola-Janicka. 2025. "Too levered for Pigou: carbon pricing, financial constraints, and leverage regulation." *Journal of Financial Economics* 172:104105.

Draghi, Mario. 2024. *The Future of European Competitiveness.* Technical report. Presented to the European Commission on 9 September 2024. European Commission, September.

ESMA, Markets Authority. 2022. *Emission Allowances and Associated Derivatives.*

Fuchs, Maximilian, Johannes Stroebel, and Julian Terstegge. 2024. *Carbon vix: Carbon price uncertainty and decarbonization investments.* Technical report. National Bureau of Economic Research.

Garleanu, Nicolae, and Lasse Heje Pedersen. 2011. "Margin-based asset pricing and deviations from the law of one price." *The Review of Financial Studies* 24 (6): 1980–2022.

Gupta, Deeksha, Michael Sockin, and Jan Starmans. 2025. "Insuring Climate Risks in Integrated Markets." *Available at SSRN 5430974.*

Heider, Florian, and Roman Inderst. 2022. "A Corporate Finance Perspective on Environmental Policy." *CEPR Discussion Paper.*

Hsu, Po-Hsuan, Kai Li, and Chi-Yang Tsou. 2023. "The pollution premium." *The Journal of Finance* 78 (3): 1343–1392.

International Carbon Action Partnership. 2025. *Emissions Trading Worldwide: ICAP Status Report 2025.* Technical report. Accessed: 2025-06-06. Berlin: International Carbon Action Partnership, April.

Ivanov, Ivan T, Mathias S Kruttli, and Sumudu W Watugala. 2024. “Banking on carbon: Corporate lending and cap-and-trade policy.” *The Review of Financial Studies* 37 (5): 1640–1684.

Kaenzig, Diego. 2023. “The Unequal Economic Consequences of Carbon Pricing.” *Working paper.*

Li, Qing, Hongyu Shan, Yuehua Tang, and Vincent Yao. 2024. “Corporate climate risk: Measurements and responses.” *The Review of Financial Studies* 37 (6): 1778–1830.

Martinsson, Gustav, László Sajtos, Per Strömberg, and Christian Thomann. 2024. “The effect of carbon pricing on firm emissions: Evidence from the Swedish CO2 tax.” *The Review of Financial Studies* 37 (6): 1848–1886.

Oehmke, Martin, and Marcus M Opp. 2025. “A theory of socially responsible investment.” *Review of Economic Studies* 92 (2): 1193–1225.

Rampini, Adriano A, and S Viswanathan. 2010. “Collateral, risk management, and the distribution of debt capacity.” *The Journal of Finance* 65 (6): 2293–2322.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang. 2023. “Firm-level climate change exposure.” *The Journal of Finance* 78 (3): 1449–1498.

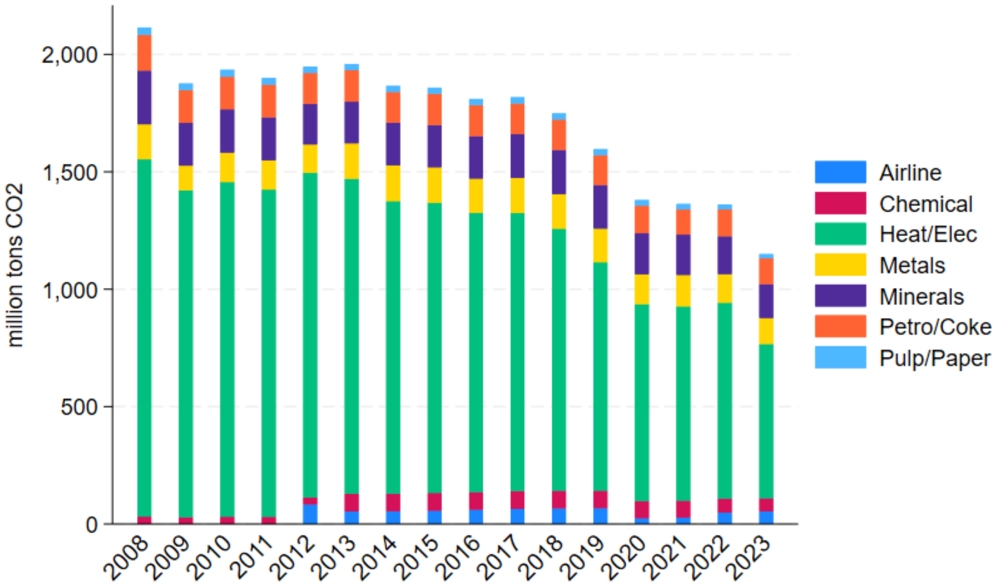
Tol, Richard. 2023. “Social cost of carbon estmiates have increased over time.” *Nature Climate Change* 13 (6): 532–536.

Wang, Mengyu, Jeffrey Wurgler, and Hong Zhang. 2026. “Policy uncertainty reduces green innovation.” *Journal of Financial Economics* 175:104189.

Zaklan, Aleksandar. 2023. “Coase and cap-and-trade: Evidence on the independence property from the European carbon market.” *American Economic Journal: Economic Policy* 15 (2): 526–558.

Zhang, Shaojun. 2025. “Carbon returns across the globe.” *The Journal of Finance* 80 (1): 615–645.

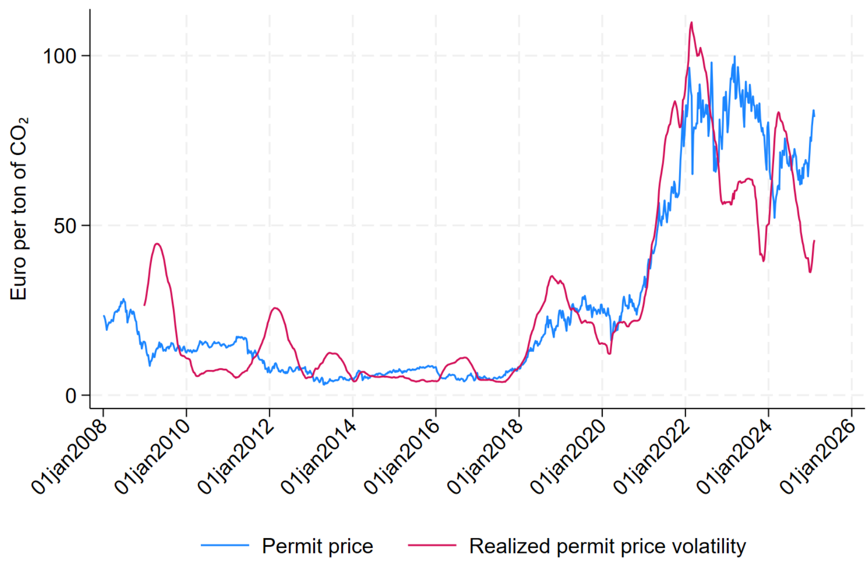
Figure 1: Emissions Covered by the EU ETS




Total emissions covered by the EU ETS (in million metric tons of $CO_2$ equivalent) as reported in the EUTL data (see Section 2.2). The scope of the EU ETS changed twice during this period: airline is added in 2012; and the UK exits in 2021, which explains why the post-Covid rebound in emissions in 2021 is not visible in the figure. Source: EUTL.

[Back to Section 2.1]

Figure 2: Emission Permit Price (Fact 1)

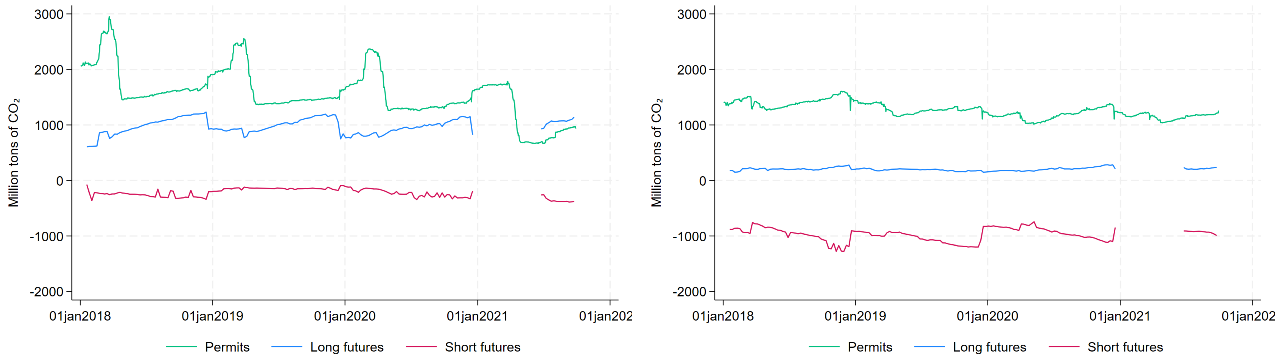



Blue: Price of emission permits (in euro per metric ton of $CO_2$ equivalent). Red: 52-week rolling-window realized volatility of weekly prices (in euro per ton of $CO_2$ equivalent). Source: Refinitiv.

[Back to Fact 1]

Figure 3: Hedging With Physical Permits and Permit Futures (Fact 2)

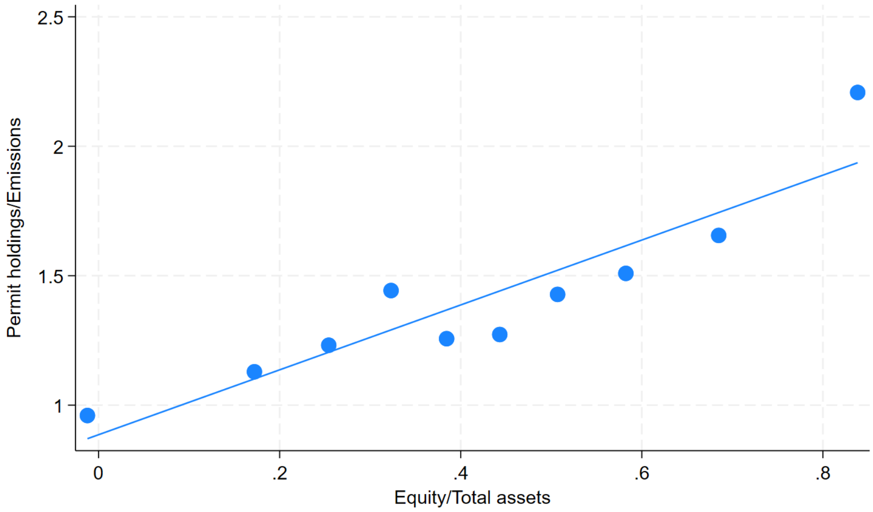
A. Emitters B. Financials




Green: Aggregate permit holdings from the full EUTL sample. Blue (red): Long (short) futures positions from ESMA Commitment of Traders. Left panel: Emitters. Right panel: Financial institutions.

[Back to Fact 2]

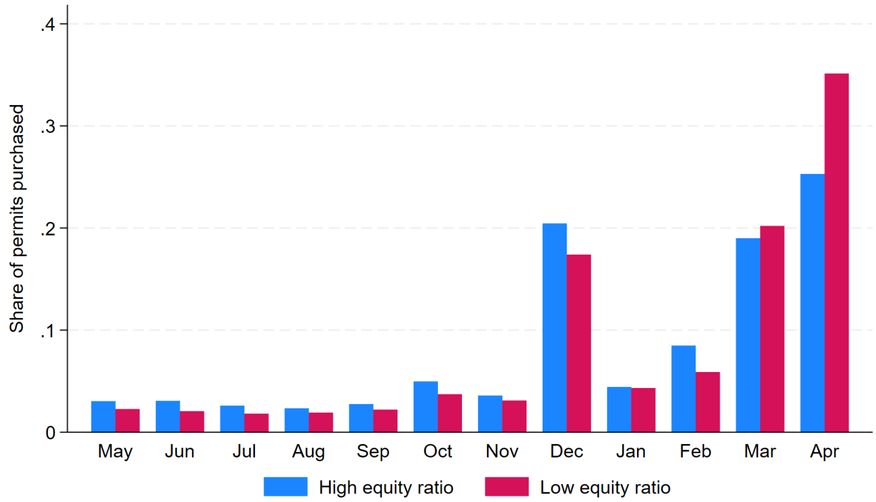
Figure 4: Emitters' Permit Holdings Are Increasing in the Equity Ratio (Fact 3i)




Sample of emitters, 2014–2020. The figure plots permit holdings at the end of the annual compliance cycle (April 30) as a share of annual emissions, as a function of lagged equity over total asset. We group emitters into ten deciles of equity ratio using year-specific breakpoints. Each dot corresponds to the average calculated over all the firm-year observations in one of the ten deciles. The straight line shows the linear fit.

[Back to Fact 3]

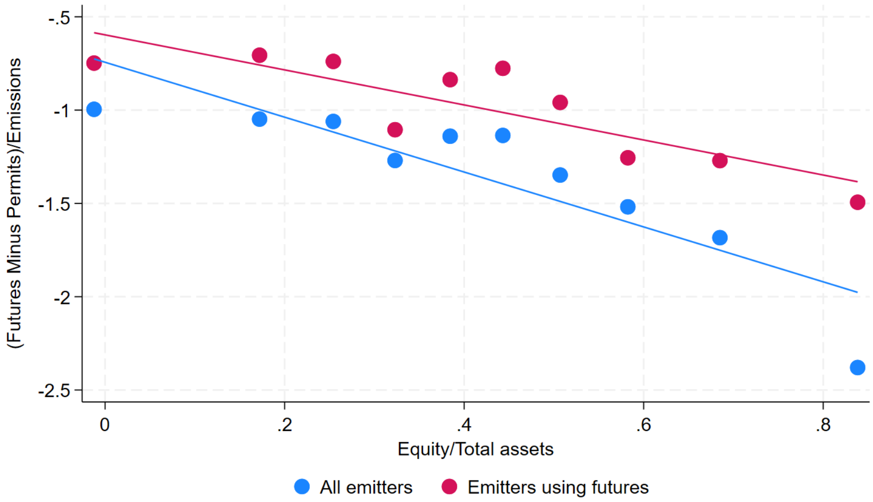
Figure 5: Timing of Permit Purchases (Fact 3ii)




We group emitters into above-median and below-median equity to total assets using year-specific breakpoints. For each group, we calculate the share of permits purchased in each month of the year.

[Back to Fact 3]

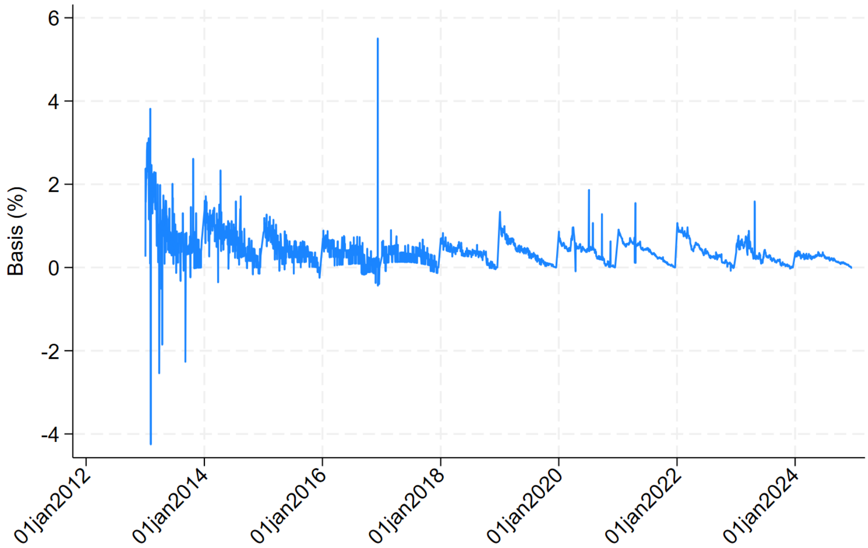
Figure 6: Emitters' Use of Permit Futures Relative to Physical Permits Is Decreasing in Equity Ratio (Fact 3iii)




Sample of emitters, 2014–2020. The figure plots permits futures positions minus permit holdings as a share of annual emissions, as a function of lagged leverage. We group emitters into ten deciles of leverage using year-specific breakpoints, either using all emitters (blue) or restricting to emitters with nonzero futures positions during the sample period (red).

[Back to Fact 3]

Figure 7: Emission Permits Basis is Positive Most of the Time (Fact 4)




The figure plots the basis over time, where the basis is defined as the price of the next December futures contract discounted by the average maturity-matched Euribor rate minus the spot price. The basis is scaled by the spot price.

[Back to Fact 4]

Table 1: Emitters' Permit Holdings Are Increasing in the Equity Ratio (Fact 3i)

| | Holdings of permits/Emissions | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Equity/Total assets | 1.1*** | 1.1*** | 1.2*** | 1.2*** |
| | (.23) | (.21) | (.21) | (.21) |
| Permit shortfall/Emissions | | -1.9*** | -1.9*** | -1.9*** |
| | | (.15) | (.15) | (.15) |
| Log(Total assets) | | | .057*** | .053*** |
| | | | (.016) | (.016) |
| Tangible assets/Total assets | | | | .54** |
| | | | | (.22) |
| Observations | 5,400 | 5,400 | 5,400 | 5,400 |
| Year FE | y | y | y | y |
| Sector shares | y | y | y | y |

Sample of emitters, 2014–2020. OLS regressions at the firm-year level, 2014–2020. The dependent variable is holdings of permits at the end of April divided by annual emissions. Equity is scaled by total assets. All specifications include year fixed effects and the share of emissions in each of the 37 activity sectors covered by the EU ETS. Standard errors clustered by firm are in parenthesis.

[Back to Fact 3]

Table 2: Timing of Permit Purchases (Fact 3ii)

| | Time Until Buy | | Time Until Sell | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Equity/Total assets | -45*** | -49*** | 4.1 | 3.8 |
| | (8.4) | (8) | (11) | (11) |
| Permit shortfall/Emissions | | -5.1 | | 22*** |
| | | (4.4) | | (5.3) |
| Log(Total assets) | | -4.4*** | | 1.4 |
| | | (.74) | | (1.1) |
| Tangible assets/Total assets | | -5.7 | | 14 |
| | | (8.6) | | (12) |
| Observations | 3,482 | 3,482 | 2,533 | 2,533 |
| Year FE | y | y | y | y |
| Sector shares | y | y | y | y |

Sample of emitters, 2014–2020. OLS regressions at the firm-year level, 2014–2020. The dependent variable in columns 1 and 2, Time Until Buy is the weighted-average number of business days between the beginning of the annual compliance cycle and the date of the purchases: $\left[\sum_{d=1}^{365} d \times Purchases_{i,t}(d)\right] \; / \; \left[\sum_{d=1}^{365} Purchases_{i,t}(d)\right]$, where $Purchases_{i,t}(d)$ is the number of permits purchased on day $d$ by emitter $i$ within the annual compliance cycle $t$. $d = 1$ corresponds to May $1^{\text{st}}$ and $d = 365$ to April 30 of the following year. The dependent variable in columns 3 and 4, Time Until Sell, is defined similarly based on the date of sales. Equity is scaled by total assets. All specifications include year fixed effects and the share of emissions in each of the 37 activity sectors covered by the EU ETS. Standard errors clustered by firm are in parenthesis.

[Back to Fact 3]

Table 3: Emitters' Use of Permit Futures Relative to Physical Permits Is Decreasing in Equity Ratio (Fact 3iii)

| | (Futures Minus Permits)/Emissions | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Equity/Total assets | -1.3*** | -1.2*** | -.66** | -.86*** |
| | (.31) | (.27) | (.32) | (.28) |
| Permit shortfall/Emissions | | 2.7*** | | 2.6*** |
| | | (.22) | | (.3) |
| Log(Total assets) | | -.018 | | -.047** |
| | | (.022) | | (.023) |
| Tangible assets/Total assets | | -.32 | | -.043 |
| | | (.29) | | (.32) |
| Observations | 4,369 | 4,369 | 2,876 | 2,876 |
| Year FE | y | y | y | y |
| Sector shares | y | y | y | y |
| Sample: | All emitters | | Emitters using futures | |

OLS regressions at the firm-year level, 2014–2020. The dependent variable is (proxied) positions in futures expiring during an annual compliance cycle minus permit holdings at the beginning of the cycle over emissions. Equity is scaled by total assets. All specifications include year fixed effects and the share of emissions in each of the 37 activity sectors covered by the EU ETS. Standard errors clustered by firm are in parenthesis.

[Back to Fact 3]

Table 4: Average Basis and Correlation with Time to Maturity (Fact 4)

| | Basis (×0.01 €) | | Basis (%) | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Constant | 11*** | .99 | .41*** | .05*** |
| | (.6) | (.7) | (.014) | (.018) |
| Time to maturity | | .056*** | | .0021*** |
| | | (.0062) | | (.00014) |
| Observations | 2,947 | 2,947 | 2,947 | 2,947 |

Sample of emitters, 2014–2020. OLS regressions, daily frequency, 2013–2024. The dependent variable is the basis, which is defined as the price of the next December futures contract discounted by the average maturity-matched Euribor rate minus the spot price. In columns 1 and 2, the basis is in cents of €. In columns 3 and 4, the basis is in percentage points of the spot price. Time to Maturity is the number of days until expiry of the next December futures contract. Newey-West adjusted standard errors (with 7 lags) are reported in parenthesis.

[Back to Fact 4]

# Appendix

## A Data

### A.1 Variables

- *Permit holdings/Emissions:* Average is permit holdings at the end of April of year $t+1$, divided by emissions in year $t$. We winsorize the right tail at 10 (approximately 98th percentile).

- *Time until buy:* Weighted-average number of days between the beginning of the annual compliance cycle and the date of the purchases: $\left[\sum_{d=1}^{365} d \times Purchases_{i,t}(d)\right] / \left[\sum_{d=1}^{365} Purchases_{i,t}(d)\right]$, where $Purchases_{i,t}(d)$ is the number of permits purchased on day $d$ by emitter $i$ within the annual compliance cycle $t$. $d = 1$ corresponds to May $1^{\text{st}}$ of year $t$ and $d = 365$ to April 30 of year $t+1$. It is defined if the firm has at least one buy transaction during the yearly compliance cycle.

- *Time until sell:* Defined similarly for sales.

- *(Futures Minus Permits)/Emissions:* Estimated futures positions in the December-year $t$ and March-year $t+1$ contracts minus permit holdings at the end of April-year $t+1$, divided by emissions in year $t$. The estimation of futures positions from settlement transactions is described in Appendix A.3.

- *Permit shortfall/Emissions* is free allowances received at the beginning of year $t$ minus emissions in year $t$, divided by emissions in year $t$. *Equity/Total assets.*

- *Equity/Total assets:* Equity divided by total assets, at the end of year $t$.

- *Log(Total assets):* Log of total assets at the end of year $t$.

- *Tangible assets/Total assets:* Tangible fixed assets divided by total assets, at the end of year $t$.

Table A.1 reports summary statistics for these variables for the period 2014–2020.

Table A.1: Summary Statistics

| | N | Mean | SD | P25 | P50 | P75 |
|---|---|---|---|---|---|---|
| Permit holdings/Emissions | 5,400 | 1.4 | 1.8 | .26 | .88 | 1.8 |
| Time Until Buy | 3,482 | 269 | 73 | 229 | 279 | 330 |
| Time Until Sell | 2,533 | 266 | 85 | 223 | 291 | 334 |
| (Futures Minus Permits)/Emissions | 4,369 | -1.4 | 2.2 | -1.9 | -.93 | -.19 |
| Permit shortfall/Emissions | 5,400 | .23 | .47 | -.025 | .25 | .56 |
| Equity/Total assets | 5,400 | .42 | .24 | .26 | .41 | .58 |
| Log(Total assets) | 5,400 | 19 | 2.4 | 18 | 19 | 21 |
| Tangible assets/Total assets | 5,400 | .48 | .22 | .33 | .46 | .64 |

## A.2 Link Between Phase 2 Accounts and Phase 3 Accounts

At the beginning of Phase 3, the EU exchanged Phase 2 permits held in each account for Phase 3 permits in a so-called banking transaction. However, the recording of the transfer in the EUTL involve several inconsistencies, which we correct through the following procedure.

First, the EUTL only shows one leg of the banking transaction (transaction type 10-33 or 10-34), in which permits are transferred out of accounts, while the other leg in which the same number of permits are transferred in is missing. We therefore add the second leg of the transfer.

Second, it is sometimes the case that accounts appear to transfer out more permits than have been transferred in, resulting in negative balance. This occurs following the migration from Phase 2 accounts (account type *Former Operator Holding Account* – hereafter "old accounts") to Phase 3 accounts (account type *Operator Holding Account* – hereafter "new accounts). Such apparent negative balance appears because we rarely observe the transfer of permits from the old accounts to the new accounts. This is a significant issue as old accounts cease trading with a total closing balance of approximately 3 billion permits.

To link old accounts to new accounts, we rely on account names, account holders, and opening dates, the latter because new accounts carry forward the opening date of the account which they replace. We proceed in two steps.

1. Taking the old and new account types (120-0, 121-0, 100-7, 100-8, 100-12) with the same account holder, we look for two accounts with the same name and opening date. If we find only two accounts with the same name and opening date and one is an old type while the other is a new type, we link these accounts. We leave the accounts unlinked in the few cases where we find more than two accounts with the same name and opening date.

2. As account holder information can occasionally be different across old and new accounts, we link unmatched accounts with different account holders. We match accounts by, in this order, account name, first term in account name, last term in account name. Of the potential matches we retain those that have the same opening date and account holder address information (main address, postcode). If we find only two accounts with either the same name, first term, or last term, as well as the same opening date and the same account holder address information, and one account is an old type while the other is a new type, we then link these two accounts. We leave the accounts unlinked in the few cases where we find more than two accounts with this procedure.

Next, we assign all trades carried out by the old account to its linked new account. This assigns 89% of the closing balances in all old accounts to a new account.

329 million permits remain in unlinked old accounts, and consequently some unlinked new accounts continue to exhibit negative balance. We rebase the balance of unlinked account with negative balance by providing them with an opening balance equal to their minimum negative balance during Phase 3, which ensures that the balance remains positive at all time. The total number of permits we add to accounts in this way equals 323 million permits, which is only 2% less than the number of permits remaining in unlinked old accounts. Thus our procedure only slightly underestimates the number of credits carried into Phase 3 for these firms.

Overall, through the linking of old accounts to new accounts and the rebasing of unlinked new accounts, we link 99.8% of the balance of permits in Phase 2 accounts to Phase 3.

## A.3 Estimation of Futures Positions from the EUTL

Two clearing houses facilitate the trading of permit futures contracts, ICE and EEX. Trading in both contracts ceases on the same day. Both contracts call for the physical delivery of permits. ICE contracts have a delivery period of three business days following the last trading day, while EEX delivers the permits on the second business day after the last trading day.

We rely on the timing of transactions reported in the EUTL to identify the transactions related to the settlement of futures contracts. We focus on the December contract, which is the most traded contract, and the March contract, which is the second most traded contract (ESMA 2022).

Figure A.1 shows the daily total transaction volume in the EUTL (blue) and the daily volume involving the clearing houses (red) around each settlement period (delineated by the vertical dashed lines). It shows that the transaction volume spikes during the settlement period: average daily volume over 2014-2021 is 17 million permits outside of settlement periods, and 286 million permits during settlement periods.

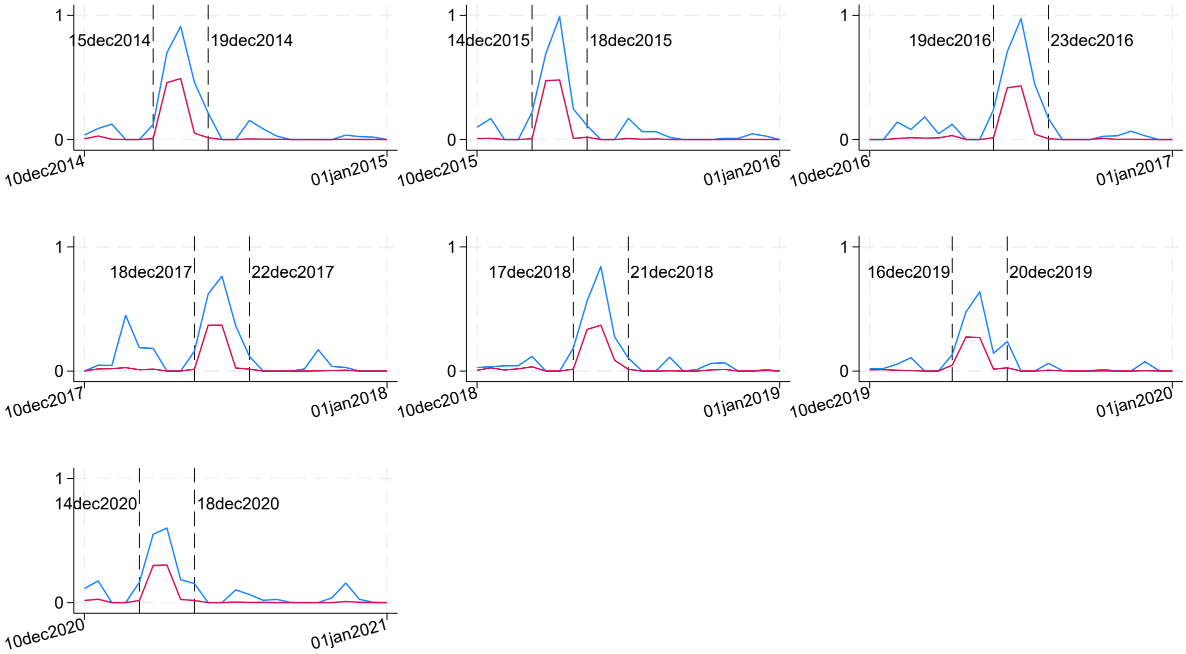
60% of the transactions during settlement periods do not involve a clearing house. This is because the accounts which receive permits from the clearing houses during the settlement period are often financial intermediaries which pass on the permits through another transfer in the days following their transfer from the clearing house. This is also the case for the accounts sending permits to the clearing houses on the day before the contracts expire. This implies that there can be several transfers related to the expiration of each futures contract. For example, the permit may pass from a company with a short position in the futures contract to an intermediary, then from the intermediary to the clearing house, then from the clearing house to a second intermediary, and finally from the second intermediary to the company which was long the futures contract.

In order to identify the end-buyers and end-sellers in futures transactions, we estimate settlement of futures positions for each EUTL account as its net trading with non-administrative accounts over a period of five business days, starting on the last day of trading for the futures, and finishing on the day after the last day for settlement.

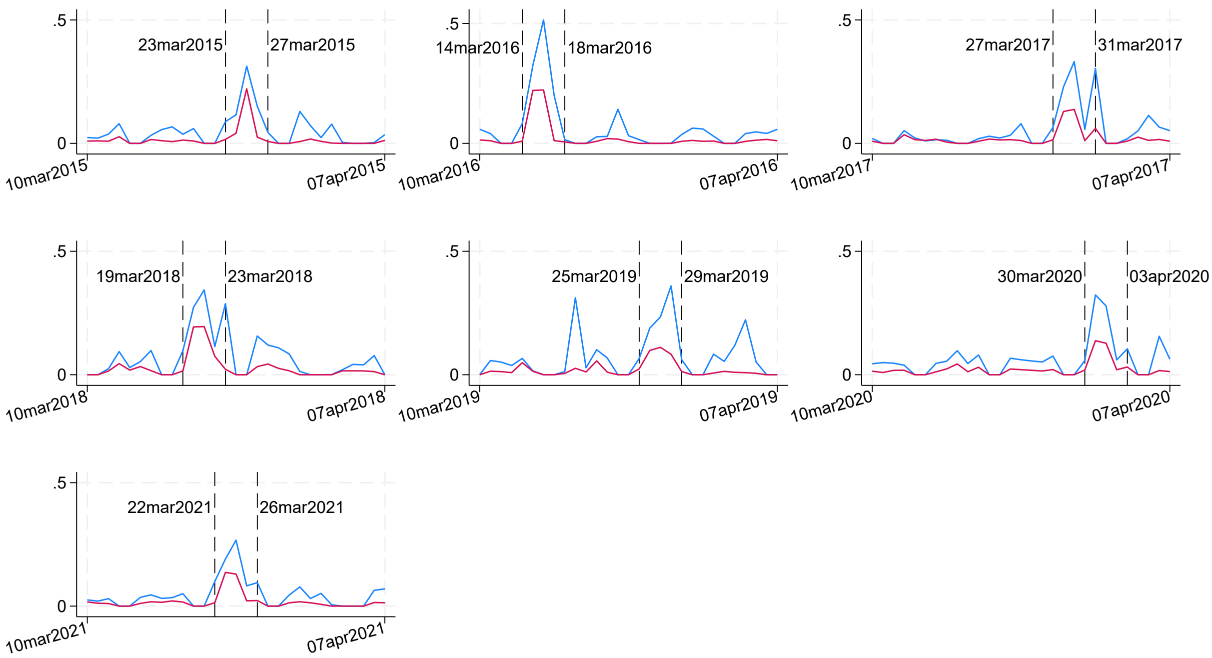
Settlements observed in the EUTL under-estimate futures positions because futures positions are rolled over before expiry do lead to settlement. We correct this mismeasurement using information on aggregate futures positions of emitters, which is disclosed by ESMA in the Commitments of Traders data. Specifically, we calculate the ratio of emitters' aggregate futures positions divided by our estimate of settlements in the EUTL. This ratio is equal to 2.7 on average over our sample period. We proxy for firm-level futures positions as observed settlements

Figure A.1: Trading Volume during the Settlement Period ($\times 1$ billion permits)

A. December Contract




B. March Contract




Blue: Daily total transaction volume of permits. Red: Daily transaction volume of permits in which one of the party is a clearing house (ICE or EEX). Daily volume on the vertical axis is in billion of permits.

multiplied by this ratio (equal to 2.7).

### A.4 Consolidation of Corporate Groups

We link EUTL accounts to Bureau Van Dijk ORBIS data, which we use to consolidate EUTL accounts at the corporate group level and to retrieve financial information for these corporate groups. We proceed in four steps.

#### Step 1: Link EUTL Accounts to ORBIS

We match EUTL accounts to ORBIS using the mapping provided by the JRC-EU ETS-FIRMS project available at https://data.jrc.ec.europa.eu/dataset/bdd1b71f-1bc8-4e65-8123-bbdd8981f116.

#### Step 2: Link Firms to their Corporate Group

We classify each firm in ORBIS matched to an EUTL account as either standalone or part of a group and, if the firm is part of a group, we identify the head of the group. A corporate group is defined as a set of firms with a common Global Ultimate Owner (GUO). GUO information is often removed from ORBIS when a firm becomes inactive. Therefore, we carry out extra steps to find the head of the group of inactive firms with no GUO. We treat each firm differently depending on whether it is linked to a GUO and whether it is currently active.

- If the firm is linked to a GUO and there is more than one firm in the corporate group, we classify the firm as part of a group and assign the GUO as the head of the group.
- If the firm is linked to a GUO, but the GUO provided is the firm itself and there are no other subsidiaries in the group, we classify the firm as standalone.
- If the firm is not linked to a GUO and is active, we classify the firm as standalone.
- If the firm is not linked to a GUO and is inactive, we download the complete shareholder history from ORBIS. We define the most recent majority shareholder as a corporate entity with a shareholding of strictly more than 50%, if any.

  If there is no such majority shareholder, we classify the firm as standalone.

  If there is such a majority shareholder (referred to as "S1" thereafter), we classify the firm as being part of a corporate group. We assign the head of the group as follows. If S1 is linked to a GUO, we assign that GUO as the head of the group. If S1 is active and not linked to a GUO, we assign S1 as the head of the group . If S1 is inactive and not linked to a GUO, we carry out the same step again, searching in S1's ownership information to find its most recent majority shareholder (S2), if any. If there is no such S2, we assign S1 as the head of the group despite of being inactive. If there is such as an S2 and S2 is linked to a GUO, we assign that GUO as the head of the group. If there is such an S2 and S2 is active but not linked to a GUO, we assign S2 as the head of the group. If S2 is inactive and is not linked to a GUO, it means that the firm is part of a corporate group but we do not manage to link the firm to its corporate group. In this case, we drop the firm so that the associated EUTL account is not matched with ORBIS.

#### Step 3: Reconstruct Corporate Groups

When the head of a corporate group is a government owning several firms, decision making is likely made at the firm level rather than at the government level. Therefore, in those cases, we step down the aggregation level below the government level, as described below in detail.

When a firm is part of a group owned by an individual or a family, ORBIS does not provide financial statements for the individual. In those cases again, we step down the aggregation level.

When the head of the group is a corporate entity but ORBIS does not provide consolidated financial statements at this level, we also step down the aggregation level.

To summarize, we apply the following rule. If the head of the group is a corporate entity with consolidated financials, we do nothing. Otherwise, we trace the ownership chain from the firm up to its head of the group and find the highest corporate entity with consolidated financials. We redefine the head of the group as this corporate entity. If there is no corporate entity with consolidated accounts in the ownership chain, we drop the firm.

#### Step 4: Consolidate Financial Information

We use consolidated accounts for firms which are part of a corporate group. For standalone firms, ORBIS may provide unconsolidated accounts, consolidated accounts, or both. If both are provided, we use consolidated accounts. Otherwise, we use whichever is provided.

# B Proofs

## B.1 Proof of Proposition 1

*Proof of point (i).* We conjecture that emitters' IC binds in state $h$ and is slack in state $\ell$, while financials' IC binds in state $\ell$ and is slack in state $h$, and derive sufficient conditions on parameters such that the conjecture holds.

In state $h$, emitters' IC binds, so consumption is given by:

$$c_{Eh} = \theta f(e_h) + \theta p_h n_E, \tag{B.1}$$

$$c_{Fh} = y + (1-\theta) f(e_h) - \theta p_h n_E. \tag{B.2}$$

In state $\ell$, financials' IC binds, so consumption is given by:

$$c_{F\ell} = \theta y + \theta p_\ell n_F, \tag{B.3}$$

$$c_{E\ell} = f(e_\ell) + (1-\theta) y - \theta p_\ell n_F. \tag{B.4}$$

First, we verify that financials' IC is slack in state $h$ and emitters' IC is slack in state $\ell$ if

$$\overline{n}_0 < \frac{1-\theta}{\theta} \frac{y + f(e_\ell)}{f'(e_\ell)}. \tag{B.5}$$

Equation (B.5) ensures that aggregate pledgeable income is positive in each state. Financials' IC is slack in state $h$ if $c_{Fh} \geq \theta y + \theta p_h n_F$. Substituting $c_{Fh}$ using (B.2), and using $n_E + n_F = \overline{n}_0$ and $p_h \leq f'(e_h)$, we obtain that financials' IC is slack in state $h$ if

$$\overline{n}_0 \leq \frac{1-\theta}{\theta} \frac{y + f(e_h)}{f'(e_h)}. \tag{B.6}$$

(B.5) and $e_1 > e_2$ imply that (B.6) holds, i.e., financials' IC is slack in state $h$. Following similar steps, (B.5) ensures that emitters' IC is slack in state $\ell$.

Second, we check that emitters' IC in state $h$ and financials' IC in state $\ell$ bind. We need to show that $\mu_{Eh} > 0$ and $\mu_{F\ell} > 0$. The first order condition for consumption of emitters in each state implies:

$$\text{state } h: \quad \frac{\mu_{Eh}}{\lambda_E q_h} = 1 - \frac{\pi_h u'(c_{Eh})}{\lambda_E q_h} \tag{B.7}$$

$$\text{state } \ell: \quad \pi_\ell u'(c_{E\ell}) = \lambda_E q_\ell \tag{B.8}$$

Combining both equations and denoting $q \equiv q_\ell/q_h$ and $\pi \equiv \pi_\ell/\pi_h$, we obtain:

$$\omega_h \equiv \frac{\mu_{Eh}}{\lambda_E q_h} = 1 - \frac{q}{\pi}\frac{u'(c_{Eh})}{u'(c_{E\ell})} = 1 - \frac{q}{\pi}\frac{f(e_\ell) + (1-\theta)y - \theta p_\ell n_F}{\theta f(e_h) + \theta p_h n_E} \tag{B.9}$$

Similarly for financials:

$$\text{state } h \qquad \pi_h u'(c_{Fh}) = \lambda_F q_h \tag{B.10}$$

$$\text{state } \ell: \qquad \frac{\mu_{F\ell}}{\lambda_F q_\ell} = 1 - \frac{\pi_\ell u'(c_{F\ell})}{\lambda_F q_\ell} \tag{B.11}$$

Combining both:

$$\omega_\ell \equiv \frac{\mu_{F\ell}}{\lambda_F q_h} = q - \pi\frac{u'(c_{F\ell})}{u'(c_{Fh})} = q - \pi\frac{y + (1-\theta)f(e_h) - \theta p_h n_E}{\theta y + \theta p_\ell n_F} \tag{B.12}$$

Emitters' intertemporal budget constraint is:

$$\sum_{s=\ell,h} q_s c_{Es} = \sum_{s=\ell,h} q_s f(e_s) - \left(\sum_{s=\ell,h} q_s p_s - p_0\right)(\overline{n}_0 - n_E)$$

Substituting $c_{Es}$ on the LHS and rearranging, we obtain:

$$q = \frac{f(e_h)}{y} + \frac{-bn_F + q\theta p_\ell n_F - \theta p_h n_E}{(1-\theta)y} \tag{B.13}$$

where $b \equiv (\sum_s q_s p_s - p_0)/q_h = \min(\theta\omega_h p_h, \theta\omega_\ell p_\ell) \leq \min(\theta p_h, \theta q p_\ell)$. To ensure that $q > 0$, we assume that

$$n_0 < \frac{1-\theta}{\theta}\frac{f(e_h)}{f'(e_h)}. \tag{B.14}$$

Using $b \leq \theta q p_\ell$, $n_E \leq \overline{n}_0$, and (B.14) in (B.13), we obtain $q > 0$.

We are looking for conditions such that $\omega_h > 0$ and $\omega_\ell > 0$. When $\overline{n}_0 \to 0$, we have:

$$q \to \frac{f(e_h)}{y} \equiv \hat{q} \tag{B.15}$$

$$\omega_h \to 1 - \frac{1}{\pi}\frac{f(e_\ell) + (1-\theta)y}{\theta y} \equiv \hat{\omega}_h \tag{B.16}$$

$$\omega_\ell \to \frac{f(e_h)}{y} - \pi\frac{y + (1-\theta)f(e_h)}{\theta y} \equiv \hat{\omega}_\ell \tag{B.17}$$

$\hat{\omega}_h > 0$ and $\hat{\omega}_\ell > 0$ if and only if

$$\underline{\pi} \equiv \frac{f(e_\ell) + (1-\theta)y}{\theta y} < \pi < \frac{\theta f(e_h)}{y + (1-\theta)f(e_h)} \equiv \overline{\pi} \tag{B.18}$$

which corresponds to assumption (15) in the main text.

When $\theta > \overline{\theta}$ and $\pi \in (\underline{\pi}, \overline{\pi})$, using a continuity argument, we have $\omega_h > 0$ and $\omega_\ell > 0$ when $\overline{n}_0$ is close to zero. In addition, conditions (B.5) and (B.14) are satisfied when $\overline{n}_0$ is close to zero. To summarize, sufficient conditions for emitters' IC binds in state $h$ and is slack in state $\ell$ and financials' IC binds in state $\ell$ and is slack in state $h$, are (B.18) and $\overline{n}_0$ close to zero.

*Proof of point (ii).* This follows from equation (16).

*Proof of point (iii).* The first order condition for time 0-permits holdings (Equation 13) holds with an equality for agents holding the asset. By market clearing, some agents must hold permits, either emitters or financials or both. Consider an agent $i$ holding permits. From point (i) of the proposition, we know that each type of agent is constrained in one state. Denote $s_i$ the state in which $i$'s IC binds. Equation 13 implies

$$p_0 = \sum_{s=\ell,h} q_s p_s - \frac{\mu_{is_i}\theta}{\lambda_i q_{s_i}} < \sum_{s=\ell,h} q_s p_s. \tag{B.19}$$

*Proof of point iv).* Substituting $c_{Es}$ in emitters' time-1 budget constraint for each state:

$$a_{Eh} = -(1-\theta) f(e_h) + \theta p_h n_E + p_h(\overline{n}_0 - n_E) \tag{B.20}$$

$$a_{E\ell} = (1-\theta) y - \theta p_\ell(\overline{n}_0 - n_E) + p_\ell(\overline{n}_0 - n_E) \tag{B.21}$$

Therefore:

$$\begin{aligned} a_{E\ell} - a_{Eh} &= (1-\theta)(f(e_h) + y) + (1-\theta)(p_\ell - p_h)(\overline{n}_0 - n_E) - \theta p_h \overline{n}_0 && \text{(B.22)} \\ &\geq (1-\theta)(f(e_h) + y) - \theta p_h \overline{n}_0 && \text{(B.23)} \\ &> 0 && \text{(B.24)} \end{aligned}$$

where the first inequality follows from $p_\ell > p_h$ and $n_E \in [0, \overline{n}_0]$, and the second inequality follows the fact the financials' incentive constraint is slack in state $h$.

### B.2 When Arrow Securities Are Imperfectly Pledgeable

In this appendix, we extend our analysis to the case in which Arrow securities are imperfectly pledgeable, like the other resources of the firm. So, at time 1, agents can abscond with a fraction $\theta$ of long positions in Arrow securities. In this case, we obtain the following proposition:

**Proposition 8.** *Under condition*

$$\overline{n}_0 < \min\left\{(1-\theta)\frac{f(e_\ell)}{f'(e_\ell)}, \frac{1-\theta}{\theta(2-\theta)}\frac{f(e_\ell)+(1-\theta)y}{f'(e_\ell)}\right\}, \tag{B.25}$$

*the equilibrium allocations and prices, as well as the states and types of agents for which incentive constraints bind, are the same when Arrow securities' payoffs are perfectly pledgeable and when they are imperfectly pledgeable.*

While Proposition 8 points to the robustness of our results, we acknowledge it comes at the cost of a parameter restriction. Indeed, Assumption (B.25) is a tighter upper bound on $\bar{n}_0$ than Assumption (B.5).

To establish Proposition 8 we proceed in three steps.

- In the first step we consider the case in which Arrow securities are perfectly pledgeable. In that case, we show that, under (B.25), $a_{Eh}$ and $a_{F\ell}$ are negative.[27] The proof of this result is given below.

- In the second step we consider the case in which Arrow securities are imperfectly pledgeable, and the agent can abscond with a fraction $\theta$ of their payoff when that payoff is positive. In that case, the incentive constraints become

$$c_{Es} \geq \theta\big[f(e_s) + p_s n_E + \max(a_{Es}, 0)\big], \tag{B.26}$$

$$c_{Fs} \geq \theta\big[y + p_s n_F + \max(a_{Fs}, 0)\big]. \tag{B.27}$$

  Apart from these two constraints, all the equilibrium conditions remain the same as when Arrow securities are perfectly pledgeable.

- In the third step, we note that (for the parameter values we consider) when Arrow securities are perfectly pledgeable only two incentive constraints are binding: the constraint of emitters in state $h$ and the constraint of financials in state $\ell$. This is where we invoke the result from our first step, that $a_{Eh}$ and $a_{F\ell}$ are negative. This result implies that the incentive constraints which are binding with perfectly pledgeable Arrow securities are unchanged and remain binding when Arrow securities are imperfectly pledgeable. Moreover, as shown below, the incentive constraints which are slack with perfectly pledgeable remain slack with imperfectly pledgeable Arrow securities.

27. The intuition is that financials provide insurance against state $\ell$ to emitters, who reciprocate by providing insurance against state $h$ to financials.

The three steps above imply our conclusion that equilibrium allocation and prices, as well as the states and the type of agent for which incentive constraints bind, are the same when Arrow securities are imperfectly pledgeable and when they are perfectly pledgeable.

To finalize the proof of Proposition 8 it only remains to finalize the proof of the claims made in the first and third steps. The proof is the following:

Binding the incentive constraint of the emitters (6) for state $h$, we obtain an expression for $c_{Eh}$, which we equate with the expression for $c_{Eh}$ from the binding budget constraint (4). This yields

$$a_{Eh} = -(1-\theta)f(e_h) + p_h\overline{n}_0 - (1-\theta)p_h n_E, \tag{B.28}$$

where we have used $\overline{n}_0 = e_h - \overline{n}_h$. Using $n_E \geq 0$ and $p_h \leq f'(e_h)$, (B.28) implies that $a_{Eh} \leq 0$ if $\overline{n}_0 \leq (1-\theta)\frac{f(e_h)}{f'(e_h)}$. $e_\ell < e_h$, $f' > 0$ and $f'' < 0$ imply $\frac{f(e_h)}{f'(e_h)} > \frac{f(e_\ell)}{f'(e_\ell)}$. Hence, the first upper bound in (B.25) implies $a_{Eh} \leq 0$ in the equilibrium prevailing when Arrow securities are perfectly pledgeable.

Similarly, using the binding incentive constraint (7) to substitute $c_{Fh}$ into the budget constraint (5), we obtain

$$a_{F\ell} = -(1-\theta)y - (1-\theta)p_\ell n_F. \tag{B.29}$$

$n_F \geq 0$, $p_\ell = f'(e_\ell)$, the first upper bound in (B.25), and (B.29) imply $a_{F\ell} \leq 0$.

It remains to show that emitters' incentive constraint in state $\ell$ and financials' incentive constraint in state $h$ remain slack when Arrow securities are imperfectly pledgeable. The long Arrow positions now enter these incentive constraints.

Using $a_{F\ell} \leq 0$, financials' binding incentive constraint in state $\ell$ together with the budget constraint implies

$$a_{F\ell} = -(1-\theta)y - (1-\theta)p_\ell n_F, \tag{B.30}$$

which we proved is negative. By market clearing,

$$a_{E\ell} = -a_{F\ell} = (1-\theta)y + (1-\theta)p_\ell n_F \geq 0. \tag{B.31}$$

Using the budget constraint (4) to substitute $c_{E\ell}$ and (B.31) to substitute $a_{E\ell}$ into emitters' incentive constraint (B.26), $p_\ell = f'(e_\ell)$, and $n_E + n_F = \overline{n}_0$ we obtain that emitters' incentive constraint in state $\ell$ is slack if the second upper bound in (B.25) holds.

Finally, using $a_{Fh} \geq 0$ and $c_{Fh}$ from the budget constraint (5), financials' incentive constraint (B.27) in state $h$ is slack.

## B.3 Proof of Proposition 3

*Sufficient condition for both types of emitters holding permits at time 0.* First, we look for values of parameters other than $A$ such that, in the baseline model with a single type of emitters for these parameter values, emitters hold permits and financials don't hold permits. This happens when emitters' shadow cost of holding permits is less than that of financials: $\theta p_h \omega_h < \theta p_\ell \omega_\ell$. Since $p_h < f'(e_h)$ and $p_\ell = f'(e_\ell)$, a sufficient condition is $f'(e_h)\omega_h < f'(e_\ell)\omega_\ell$. When $\overline{n}_0$ is close to zero, a sufficient condition is $f'(e_h)\hat{\omega}_h < f'(e_\ell)\hat{\omega}_\ell$. Since $\hat{\omega}_h$ is increasing in $\pi$ by (B.16) and equal to zero for $\pi = \underline{\pi}$, while $\hat{\omega}_\ell$ is decreasing in $\pi$ by (B.17) and equal to zero for $\pi = \overline{\pi} > \underline{\pi}$, the sufficient condition is $\pi \in (\underline{\pi}, \tilde{\pi})$ where $\tilde{\pi} \in (\underline{\pi}, \overline{\pi})$.

Second, consider the case with two types of emitters, $\pi \in (\underline{\pi}, \tilde{\pi})$ and $A$ close to zero. By continuity, the equilibrium prices, consumption and holdings of agents are close to those in the baseline. In particular, both types of emitters hold permits.

*Proof of point (i) in the proposition.* Since both types of emitters hold permits, the first order condition for time-0 permit holdings implies

$$\frac{\mu_{E^+h}}{\lambda_{E^+}} = \frac{\mu_{E^-h}}{\lambda_{E^-}}. \tag{B.32}$$

The first order condition for emissions in state $h$ implies

$$e_{E^+h} = e_{E^-h} \tag{B.33}$$

The first order condition for consumption of emitters in state $h$ implies:

$$\frac{\mu_{E^\varepsilon h}}{\lambda_{E^\varepsilon}} = q_h - \frac{\pi_h}{\lambda_{E^\varepsilon}\theta\big[f(e_{E^\varepsilon h}) + p_s n_{E^\varepsilon}\big]}, \qquad \varepsilon \in \{+,-\}, \tag{B.34}$$

where we have substituted $c_{E^\varepsilon h}$ using emitters' binding incentive constraint in state $h$.

Suppose, by contradiction, that $n_{E^+} \leq n_{E^-}$. Using (B.32) on the LHS of (B.34) and (B.33) on the RHS of (B.34), we obtain $\lambda_{E^+} \geq \lambda_{E^i}$. The first order condition for consumption of emitters in state $\ell$ implies $c_{E^+\ell} < c_{E^-\ell}$. Emitters' binding incentive constraint in state $h$ implies $c_{E^+h} < c_{E^-h}$. Therefore, emitters with high endowment have lower expected utility than emitters with low endowment, which is a contradiction.

*Proof of point (ii).* First, we note that both types of emitters have the same level of emissions in each state. The first order condition for emissions in state $\ell$ implies $e_{E^+\ell} = e_{E^-\ell}$. We showed in the proof of point (i) that $e_{E^+h} = e_{E^-h}$.

Purchase of permits at time 1 is equal to emissions minus purchase of permits at time 0. Using point (i), it implies that emitters with high endowment purchase fewer permits at time 1 emitters low endowment.

*Proof of point (iii).* Combining emitters' the first order condition for consumption in state $\ell$ with that in state $h$, we obtain

$$\frac{c_{E^{\epsilon}h}}{c_{E^{\epsilon}\ell}} = \frac{\pi_h}{\pi_\ell} \frac{q_\ell}{q_h - \frac{\mu_{E^{\epsilon}h}}{\lambda_{E^{\epsilon}}}} \tag{B.35}$$

(B.32) implies that the RHS of (B.35) is equalized across both types of emitters. Therefore

$$\frac{c_{E^+h}}{c_{E^+\ell}} = \frac{c_{E^-h}}{c_{E^-\ell}} \tag{B.36}$$

(B.36) means that emitter type $E^+$'s consumption is a constant proportion of $E^-$'s consumption across states. Since $E^+$ has higher initial wealth than $E^-$, it implies $E^+$ has higher expected utility, therefore higher consumption in each state. (B.36) implies

$$c_{E^+h} - c_{E^+\ell} = c_{E^+\ell}\frac{c_{E^-h}}{c_{E^-\ell}} - c_{E^+\ell} = \frac{c_{E^+\ell}}{c_{E^-\ell}}(c_{E^-h} - c_{E^-\ell}) > c_{E^-h} - c_{E^-\ell} \tag{B.37}$$

where the inequality follows from $c_{E^+\ell} > c_{E^-\ell}$. Using the time 1-budget constraint to substitute consumptions on each side of (B.37), and using the fact that emissions are the same for both types of emitters, we obtain

$$a_{E^+\ell} - a_{E^+h} + (p_\ell - p_h)n_{E^+} < a_{E^-\ell} - a_{E^-h} + (p_\ell - p_h)n_{E^-} \tag{B.38}$$

Using $\phi_{E^\epsilon} = (a_{E^\epsilon \ell} - a_{E^\epsilon h})/(p_\ell - p_h)$, we can rewrite (B.38)

$$\phi_{E^+} + n_{E^+} < \phi_{E^-} + n_{E^-} \tag{B.39}$$

Finally, $n_{E^+} > n_{E^-}$ implies

$$\phi_{E^+} - n_{E^+} < \phi_{E^-} - n_{E^-} \tag{B.40}$$

## B.4 Proof of Proposition 5

First, we check that financials' incentive constraint binds in state $\ell$. We need to show that $\mu^*_{F\ell} > 0$. Using (33) and (34) to substitute consumption in (25), we obtain

$$\mu^*_{F\ell} = \frac{\alpha_E \pi_\ell}{f(e_\ell) + (1-\theta)y} - \frac{\alpha_F \pi_\ell}{\theta y}. \tag{B.41}$$

Replacing $\alpha_E = 1 - \alpha_F$ and using $\theta - \alpha_F > 0$, we obtain that $\mu^*_{F\ell} > 0$ holds if and only if $f(e_\ell) < (\theta - \alpha_F)y/\alpha_F$, i.e.,

$$e_\ell < f^{-1}\left(\frac{(\theta - \alpha_F)y}{\alpha_F}\right) \equiv \underline{e}_\ell. \tag{B.42}$$

The RHS of (35) is decreasing in $e_\ell$. Therefore, $\mu^*_{F\ell} > 0$ holds if and only if $\delta_\ell$ is greater than the RHS of (35) for $e_\ell = \underline{e}_\ell$, which after some algebra is equal to the assumed lower bound for $\delta_\ell$ in condition (31).

Similarly, we check that emitters' incentive constraint binds in state $h$. We need to show that $\mu^*_{Eh} > 0$. Using (36) and (37) to substitute consumption in (25), we obtain

$$\mu^*_{Eh} = \frac{\alpha_F \pi_h}{y + (1-\theta)f(e_h)} - \frac{\alpha_E \pi_h}{\theta f(e_h)}. \tag{B.43}$$

Replacing $\alpha_F = 1 - \alpha_E$ and using $\theta - \alpha_E > 0$, we obtain that $\mu^*_{Eh} > 0$ holds if and only if $f(e_h) > \alpha_E y/(\theta - \alpha_E)$, i.e.,

$$e_h > f^{-1}\left(\frac{\alpha_E y}{\theta - \alpha_E}\right) \equiv \overline{e}_h. \tag{B.44}$$

The RHS of (38) is decreasing in $e_h$. Therefore, $\mu^*_{Eh} > 0$ holds if and only if $\delta_h$ is lower than the RHS of (38) for $e_h = \overline{e}_h$, which after some algebra is equal to the assumed upper bound for $\delta_h$ in condition (32).

*Proof of point (i).* Using (B.42) and (B.4) and the fact that $f^{-1}(.)$ is increasing, a sufficient condition for $e_\ell < e_h$ is

$$\frac{(\theta - \alpha_F)y}{\alpha_F} \leq \frac{\alpha_E y}{\theta - \alpha_E}. \tag{B.45}$$

This condition holds because $\theta \leq 1$ and $\alpha_F = 1 - \alpha_E$.

*Proof of point (ii).* To show that $e_\ell > e_\ell^{fb}$, note that the RHS of (35) is decreasing in $e_\ell$ and increasing in $\theta$, therefore $e_\ell$ is increasing in $\theta$. The first best corresponds to a decrease in $\theta$ up to the point where incentive constraints do not bind. Hence $e_s$ is higher than in first best. Similarly, to show that $e_h < e_h^{fb}$, note that the RHS of (38) is decreasing in $e_h$ and in $\theta$, therefore $e_h$ is decreasing in $\theta$ and is lower than in the first best.

*Proof of point (iii).* The RHS of (35) is decreasing in $e_\ell$ and increasing in $\alpha_E$, therefore $e_\ell$ is increasing in $\alpha_E$. The RHS of (38) is decreasing in $e_h$ and its derivative with respect to $\alpha_E$, using that $\alpha_F = 1 - \alpha_F$, is equal to

$$f'(e_h)\left[-\frac{1-\theta}{y + (1-\theta)f(e_h)} + \frac{1}{f(e_h)}\right] > 0. \tag{B.46}$$

Therefore $e_h$ is increasing in $\alpha_E$.

## B.5 Proof of Proposition 6

*We start by proving the converse of the proposition (Second Welfare Theorem): The second best for any Pareto weights can be implemented as a cap-and-trade equilibrium with time-0 transfers.*

First, we augment the equilibrium outlined in Section 3 with time-0 transfers $t_i$ from the government to each agent $i \in \{E, F\}$. The government's budget constraint implies $t_E + t_F = 0$. The time-0 budget constraints are modified as follows:

$$c_{Es} \leq f(e_s) - p_s e_s + p_s \overline{n}_s + p_s n_E + a_{Es} + t_E, \tag{B.47}$$

$$c_{Fs} \leq y + p_s n_F + a_{Fs} + t_F. \tag{B.48}$$

The other equations remain unchanged.

Second, we prove by construction that the constrained optimal allocation $(c, e)$ for any Pareto weights can be decentralized as an equilibrium with time-0 transfers. To do so, we show that there exist emissions caps and transfers $(\overline{n}, t)$, prices $(p, q)$, permit holdings $n$, and multipliers $(\lambda_i, \mu_{is})$ such that agents' first order conditions, budget constraints, incentive constraints, and market clearing are satisfied for $(c, e)$.

We set time-0 permit issuance to zero, i.e., $\overline{n}_0 = n_E = n_F = 0$, which implies $\overline{n}_s = e_s$. Markets clear by feasibility of the planner's allocation. The incentive constraints are satisfied by construction for $n_i = 0$. The budget constraints are satisfied by choosing $t_i$ appropriately. The first order conditions (11) and (12) are satisfied with the following prices and multipliers

$$q_s = q\lambda_s^*, \tag{B.49}$$

$$p_s = \frac{\pi_s}{\lambda_s^*}\delta_s, \tag{B.50}$$

$$\lambda_i = \frac{1}{q\alpha_i}, \tag{B.51}$$

$$\mu_{is} = \frac{\mu_{is}^*}{\alpha_i}, \tag{B.52}$$

for some $q > 0$.

*Proof of the proposition (First Welfare Theorem).* We need to show that there exist Pareto weights such that the equilibrium with appropriate emissions cap and zero time-0 transfers implements the second best with these Pareto weights.

The time-0 transfer required to implement the second best is equal to $t_E = \sum_s q_s(c_{Es} - f(e_s))$. In the second best for $\alpha_E = 0$, it is optimal to bind emitters' incentive constraint in every state, so $t_E = \sum_s q_s(\theta f(e_s) - f(e_s)) < 0$. In the second best for $\alpha_E = 1$, it is optimal to bind financials' incentive constraint in every state, so $t_E = \sum_s q_s(f(e_s) + (1-\theta)y - f(e_s)) > 0$. The solution to the planner's second best problem is continuous in $\alpha_E$. Therefore, by the intermediate value theorem, there exists $\alpha_E \in (0, 1)$ such that $t_E = 0$.