Grantham Research Institute on Climate Change and the Environment

Centre for Climate Change Economics and Policy

Policing carbon markets

Raphael Calel, Antoine Dechêzlepretre and Frank Venmans

September 2023

Centre for Climate Change Economics and Policy Working Paper No. 425
ISSN 2515-5709 (Online)

Grantham Research Institute on Climate Change and the Environment Working Paper No. 400
ISSN 2515-5717 (Online)

Economic and Social Research Council

Grantham Foundation for the Protection of the Environment

University of Leeds

LSE The London School of Economics and Political Science

The Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council. Its third phase started in October 2018 with seven projects:

1. Low-carbon, climate-resilient cities
2. Sustainable infrastructure finance
3. Low-carbon industrial strategies in challenging contexts
4. Integrating climate and development policies for 'climate compatible development'
5. Competitiveness in the low-carbon economy
6. Incentives for behaviour change
7. Climate information for adaptation

More information about CCCEP is available at www.cccep.ac.uk

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and a number of other sources. It has 13 broad research areas:

1. Biodiversity
2. Climate change adaptation and resilience
3. Climate change governance, legislation and litigation
4. Climate, health and environment
5. Environmental behaviour
6. Environmental economic theory
7. Environmental policy evaluation
8. International climate politics
9. Science and impacts of climate change
10. Sustainable public and private finance
11. Sustainable natural resources
12. Transition to zero emissions growth
13. UK national and local climate policies

More information about the Grantham Research Institute is available at: www.lse.ac.uk/GranthamInstitute

Suggested citation:
Calel R, Dechêzlepretre A and Venmans F (2023) *Policing Carbon Markets*. Centre for Climate Change Economics and Policy Working Paper 425/Grantham Research Institute on Climate Change and the Environment Working Paper 400. London: London School of Economics and Political Science

Policing Carbon Markets

Raphael Calel[1,3], Antoine Dechêzlepretre[2,3], and Frank Venmans[3]

[1]McCourt School of Public Policy, Georgetown University, rac121@georgetown.edu

[2]Economics Department and Environment Directorate, OEC, antoine.dechezleprêtre@oecd.org

[3]Grantham Research Institute, London School of Economic, f.venmans1@lse.ac.uk

Draft: September 2023

Abstract

Carbon markets have emerged in recent decades as one of the most important tools for curbing industrial greenhouse gas emissions, but they present a number of novel enforcement challenges as compared to more conventional pollution regulations—new regulators with narrow authority, lack of legal precedent, and more. To shed light on the practical issues involved in policing carbon markets, we present the first comprehensive analysis of the EU Emissions Trading System, a single program that was policed by 31 different national regulators. We find generally *high* rates of compliance coupled with *low* rates of enforcement, a pattern that is known in the literature as 'Harrington's paradox.' Variation in the probability and severity of fines explain just one tenth of the variation in compliance rates. Meanwhile, other enforcement strategies that have been pointed to as resolutions to Harrington's paradox in other applications, such as 'naming and shaming,' appear to have had little discernible effect.

Keywords: Pollution control, Compliance, Enforcement, Cap-and-trade

JEL: Q50, Q52, C14

1 Introduction

The environmental integrity, political legitimacy, and economic efficiency of pollution control laws all fundamentally depend on industry's compliance. Regulators have consequently developed many creative strategies to allow their limited budgets to stretch further: threatening harsh fines for noncompliers (Stafford, 2002; Shimshack and Ward, 2005), targeting past noncompliers for additional inspections (Helland, 1998), requiring self-reporting of emissions (Kaplow and Shavell, 1994; Innes, 1999) and third-party audits (Duflo et al., 2013), using extra-legal sanctions such as delaying permits for new projects (Fenn and Veljanovski, 1988; Decker, 2003), and naming and shaming the worst offenders (Konar and Cohen, 1997; Schlenker and Scorse, 2012).

Carbon markets have come to play a prominent role in climate change policy around the world, but their design undercuts many of the conventional strategies to encourage compliance. Unlike traditional command-and-control regulations, cap-and-trade programs equalise the marginal benefit from non-compliance across plants and thus invalidate the logic of targeting past noncompliers for additional monitoring (Stranlund and Dhanda, 1999). The regulator's ability to verify the integrity of emissions reductions, either directly or by third-party audits, is often further circumscribed by allowing emissions permits to be traded across countries, and by allowing pollution offsets into the market from outside projects (Sigman and Chang, 2011). Then, even when a regulator does detect non-compliance, the penalties for excess emissions are commonly fixed by statute, effectively removing the regulator's discretion to apply steeper fines to repeat offenders. Neither does the regulator have the power to lean on other regulatory levers to exert pressure on noncompliers, since these programs are often administered by newly created agencies with narrow authority. One would expect all of these challenges to grow as governments ratchet down the emissions caps and the cost of compliance rises. Carbon markets therefore not only apply the tools of pollution control at a much larger scale, but also create significant new enforcement challenges.

So, have regulators been able to bring about high rates of compliance in this new regulatory context? And if so, which regulatory strategies have been most effective? To answer these questions, we conduct the first comprehensive investigation of compliance and enforcement in the European Union's Emissions Trading System (EU ETS). Launched in 2005, this cap-and-trade program has come to regulate the carbon-dioxide emissions of over 12,000 stationary sources across 31 countries, accounting for nearly half of the EU's greenhouse gas emissions. The EU ETS was the first major carbon market anywhere in the world, and continues to serve as a model for new programs. Moreover, the EU ETS offers a rare chance to study the effect of enforcement regimes on compliance—the same policy, created by an EU directive, was subsequently implemented and enforced by many different national regulators, which were empowered to use somewhat different enforcement strategies.

We find, first and foremost, that the EU ETS has met its emissions targets in the aggregate. The rates of compliance have been generally high. There are some areas of concern, though. We find suggestive evidence, using Benford's Law, that when fines increased in the second phase, misreporting increased among smaller plants, which faced less strict monitoring requirements. Moreover, in many instances of non-compliance, we have been unable to verify that any enforcement action was taken. We estimate that regulators should have collected at least €13 billion in fines for excess emissions, though public records show that only €2.1 billion have been collected. Variation in the probability and severity of fines explain just 12% of the variation in compliance rates. Even in jurisdictions that do not report taking enforcement actions against any noncompliers, the compliance rate has still been high. This combination of high rates of compliance and low levels of enforcement have previously been termed 'Harrington's paradox' (Harrington, 1988).

Prior work has pointed to regulator strategies other than fines to help resolve Harrington's paradox. We investigate several such monitoring and enforcement strategies—e.g. naming and shaming of non-compliant companies, having the regulator appoint third-party auditors instead of letting companies choose their own, and suspension of offenders. The variation in these strategies across countries and across time is weakly correlated with compliance. However, when we try to identify the causal effects of these strategies using empirical tests borrowed from prior literature, we generally do not find evidence of their effectiveness. For instance, 'naming and shaming' is expected to be more effective when companies face greater reputational risks, but we find that 'naming and shaming' is no more important for encouraging compliance from publicly traded companies, or from companies headquartered in the same country. Thus, none of these regulator strategies provide a satisfactory resolution to the paradox of high compliance and low enforcement in carbon markets.

In interviews conducted with regulators, we found they had adopted a pragmatic approach to enforcement—sometimes wary of fighting too many legal battles over fines, often choosing instead to work with polluters to bring them into compliance. This approach may have been sufficient to achieve high rates of compliance so far, but it is not clear that it will be sufficient when emissions caps tighten and compliance costs go up.

It remains an important question for future research, therefore, whether there are more robust enforcement strategies that will enable effective policing of carbon markets going forward.

The article contributes to several literatures. First, we provide new empirical evidence on compliance and enforcement of pollution control regulation. Following Harrington (1988), a substantial theoretical literature now shows how regulators can use targeting and other strategies to realise high expected non-compliance costs even when the average enforcement effort is low (Harrington, 1988; Fenn and Veljanovski, 1988; Segerson and Tietenberg, 1992; Kaplow and Shavell, 1994; Innes, 1999; Kunreuther et al., 2002; Nyborg and Telle, 2004; Bontems and Bourgeon, 2005). More recent empirical studies examine the effects of these strategies in laboratory settings (Cason and Gangadharan, 2006a,b; Murphy and Stranlund, 2006, 2007), in observational data (Helland, 1998; Earnhart, 2004; Telle, 2008; Keohane et al., 2009), and in rare field experiments (Duflo et al., 2013; Telle, 2013). We extend this literature to the study of compliance and enforcement in an active carbon market, and highlight potential constraints that such programs may face. In addition, while earlier studies at best obtain estimates from variation in monitoring or enforcement at the polluter-level, we use the rare opportunity of having many regulators trying to police the same set of rules, to look at the more policy-relevant effect of empowering or constraining regulators.

Second, our findings contribute to the broader literature on the effectiveness of these monitoring and enforcement strategies. Several of the strategies we study have close analogies in the monitoring and enforcement of labour market regulations (Ashenfelter and Smith, 1979; Anderson, 1996; Ronconi, 2010; Bhorat et al., 2012), financial market regulations (Garfinkel, 1997; Coates IV, 2007), electricity market regulations (Wolfram, 1999), fisheries regulations (Furlong, 1991), workplace safety (Gray and Jones, 1991; Weil, 1996), food and drug safety (Law, 2006), product quality standards (Potoski and Prakash, 2005), advertising standards (Sauer and Leffler, 1990), anti-corruption measures (Olken, 2007; Cordis and Warren, 2014), antitrust (Feinberg, 1980; Block et al., 1981; Snyder, 1990; Baker, 2003; Miller, 2009), and tax law (Engel et al., 2001; Kleven et al., 2011; Di Porto et al., 2013; Hanlon et al., 2014). Our findings suggest that these strategies that are well-established in other domains may not transfer to carbon markets.

Third, the EU ETS continues to serve as a model for carbon market programs around the world, and there is now a substantial literature devoted to measuring its environmental and economic effects. There are estimates of the EU ETS's effect on emissions (Ellerman and Buchner, 2008; Anderson and Di Maria, 2011; Bayer and Aklin, 2020; Dechezleprêtre et al., 2023), on employment (Anger and Oberndorfer, 2008; Chan et al., 2013), on investment (Löfgren et al., 2013), and on innovation (Calel and Dechezleprêtre, 2016;

Calel, 2020). However, a given treatment effect has a very different meaning if the underlying regulation was strictly enforced and adhered to, versus if many polluters neglected to comply without fear of punishment. Our findings therefore provide a lens through which to interpret all previous impact evaluations.

2 The life of an emissions allowance

To understand polluter non-compliance in the EU ETS, it is helpful to first situate this problem within the broader set of enforcement challenges. In a cap-and-trade program, a limited number of tradeable emissions allowances are distributed initially. Allowances can then be traded freely throughout the compliance period (or phase), but at the end of the period, polluters are required to surrender enough allowances to cover their emissions. Emissions allowances thus pass through three basic life stages—allocation, trading, surrender. In this section we review the compliance and enforcement challenges present at each stage, before proceeding to a detailed empirical analysis of the surrender-stage in later sections.

2.1 Allocation

How emissions allowances are allocated determines the quantity of scarcity rents and to whom rents will accrue. The design of allocation plans is therefore often a contentious process that requires careful balancing of local interests with overall program objectives. The US NO_X Budget Trading Program, for instance, pulled off this balancing act by having the federal government set the total number of emissions allowances for each state and then leaving each state to distribute those allowances as they saw fit. The process for allocating allowances in the US Acid Rain Program was more centralised, with the federal regulator determining both the overall cap and its distribution.

The EU ETS's allocation process has historically lacked such a clear division of responsibilities, and was instead geared toward building consensus between the central authority and local regulators. The program was created by a Directive of the European Union, adopted in October 2003. The Directive did not give the European Commission the authority to set an overall cap on allowances, but instead required each country to develop its own National Allocation Plan (NAP) stating how many allowances it intended to issue, and how it would distribute them. The NAPs had to be submitted to the European Commission 6 to 12 months before the start of each compliance phase. The first compliance phase ran from 2005-2007, the second from 2008-2012, and the third from 2013-2020. Once a NAP was submitted, the Commission had three months to review the NAPs to ensure that they (1) were consistent with internationally agreed emissions reduction obligation, (2) took due account of abatement potential, and (3) did not unduly favour or discriminate against individual sectors or companies. The national regulator would be asked to revise its plan if the Commission determined that a NAP did not meet these criteria. Although NAPs could in principle go through several iterations, the expectation was that the second version of the plan would meet approval.

The EU ETS allocation process worked mostly as intended, although the lack of central authority did create certain compliance challenges. First, very few NAPs were submitted on time for either the first and second compliance phase, and in some cases the Commission launched infringement proceedings against laggards to speed up submissions (Zapfel, 2007; Ellerman et al., 2010). Second, British and German regulators felt the Commission had not allowed them sufficient flexibility in designing their own NAPs, and both countries successfully sued the Commission in European courts (van Zeben, 2009). Third, not all countries complied when the Commission told them they should distribute fewer allowances. Ahead of the first phase, the Commission had requested that countries reduce the number of allowances from their initial NAP submissions, and was ultimately able to negotiate a 4.2% reduction. However, when the Commission requested a 10% reduction ahead of the second compliance phase, Poland and Estonia objected and filed successful law suits against the Commission.[1] The Commission, in turn, brought two successful suits against Italy and Finland for failing to fully implement the EU ETS (van Zeben, 2009).

Each country then set up its own emissions registry to distribute the allowances in accordance with the approved NAPs. First, each polluter had to register—registration required submitting a plan for monitoring and reporting emissions, and conveys a nontradeable right to conduct specific polluting activities at a given location. The regulators would then transfer the specified number of allowances into each registered polluter's account on February 28th each year. As a rule, the allowances were deposited in polluters' accounts free of charge, and distributed roughly in proportion to historical emissions.

These allocations were not without controversy. The allowances were worth many billions of euros every year (IPA Energy Consulting, 2005; Sijm et al., 2006; Sandbag, 2011; Bushnell et al., 2013), even small differences in allocations across similar installations could create significant competitive advantages Ellerman et al. (2010, Ch. 4). Several suits alleging favouritism have been brought against the Commission by private parties, but to date they have been dismissed due to lack of standing (van Zeben, 2009). However,

[1] For completeness, we should note that Bulgaria, Romania, Latvia, Slovakia, and the Czech Republic filed similar suits against the Commission, but these suits were withdrawn.

some studies report that politically influential polluters did obtain more generous allocations (Anger and Oberndorfer, 2008; Hanoteau, 2014).

The free allocation of allowances also gives rise to a more indirect compliance problem. Even though allowances were free, polluters raised their prices to reflect higher opportunity costs. The resulting "windfall profits" were widely perceived as unfair, causing Ireland and Slovakia to impose new levies to claw back some of the rents from polluters. Although these were separate taxes, they effectively altered the allocation rules specified in the NAPs. The Commission sued, leading the European Court of Justice to strike down these claw-back regulations (Point Carbon, 2012a,b).

One final issue, often discussed in the context of compliance with EU ETS allocations, is the use of carbon offsets. With the start of the Kyoto Protocol compliance phase in 2008, European polluters were given the option to substitute EU emissions allowances for offset credits generated under the Clean Development Mechanism or Joint Implementation (Trotignon, 2012). This effectively delegated some of the authority to issue allowances to the United Nations, which oversaw these mechanisms. Importing these offsets can in principle improve overall compliance by reducing the marginal cost of abatement (Sigman and Chang, 2011). It soon emerged, however, that many applications to the UN Executive Board were manipulated to show emissions reductions where there were none (Consulate Mumbai, 2008; Point Carbon, 2010), and that the allure of offsets sometimes provided perverse incentives to *increase* emissions (Wara, 2007b,a; Schneider, 2011) or to break local environmental laws (Böhm and Dabhi, 2009; Bond et al., 2012). Meanwhile, several third-party auditors have been found to fail to properly vet applications for offsets (Frunza, 2013b, p. 63), and some members of the UN Executive Board had undeclared conflicts of interest (Transparency International, 2011). The European Commission has no authority to intervene in this situation, so instead moved to restrict imports of certain types of offsets in the third compliance phase.

In 2012 the system of national registries was superseded by a central EU registry, and in 2013 the NAPs were replaced by a single European allocation plan written by the Commission. This was meant to reduce the legal uncertainty in the allocation process, and moved the EU ETS one step closer to a centralised design like that of the US Acid Rain Program.

2.2 Trading

Aside from registered polluters, the EU ETS permitted financial institutions, charities, and even private persons, to open trading accounts in national registries as well. Once the emissions allowances were deposited into polluters' accounts, anyone was free to buy and sell allowances.

Overall, market fundamentals explain most of the trading behaviour (Ellerman and Trotignon, 2009; Ellerman et al., 2010; Zaklan, 2013; Betz and Schmidt, 2015). For instance, the market price of an EU allowance (EUA) rose rapidly in the program's first year, reaching a high of €30, and subsequently crashed when it was revealed that more allowances had been issued than would likely be needed for compliance (figure 1). Nearly all transactions were initially cleared bilaterally, but several brokerages and privately run exchanges were soon established across the EU and attracted a growing share of the trading volume (Ellerman et al., 2010; Kossoy and Guigon, 2012). Spot contracts quickly became secondary to futures (figure 1). Options and swaps were then introduced, allowing companies to adopt more sophisticated planning and risk management strategies (Sandor et al., 2002; Ellerman et al., 2010). By several measures, transaction costs have declined by a factor of ten or more (Jaraite et al., 2010; Jaraitė-Kažukauskė and Kažukauskas, 2014; Heindl, 2015; see also figure 1).

It is almost impossible to set up new market institutions quickly without creating some regulatory blind spots, though. For instance, because allocations are based on historical emissions in the EU ETS, allowances have been heavily concentrated among a few large companies (Trotignon and Delbosc, 2008; Ellerman et al., 2010). The effective market power of these companies is further exacerbated by the fact that many smaller companies do not appear to trade actively (Zaklan, 2013). Large polluters therefore have an opportunity to corner the allowance market and undermine market efficiency (Hahn, 1984), and the pattern of allowance holdings among large power companies does appear consistent with strategic price manipulation Hintermann (2010, 2015). No anti-trust proceedings have been brought, and it is not as yet clear which competition authority would have jurisdiction.

Other regulatory blind spots have been exploited in more obviously criminal ways. In February 2010 the German authorities discovered that cyber criminals had obtained users account data by posing as the registry. The scammers logged into these accounts and transferred the allowances to their own accounts. They then quickly sold the stolen allowances to companies in other EU countries before the intrusions were discovered, exploiting the fact that there was no central regulator that could easily trace the stolen allowances across jurisdictions. 250,000 credits were stolen, worth €3.2 million. Later in 2010 and 2011, fraudsters exploited weak internet security protocols and hacked into least 6 national registries. Over 3 million allowances were stolen, worth nearly €50 million (Frunza, 2013b, p. 75). After each intrusion was discovered, national registries shut down and trading was suspended across Europe while authorities tried

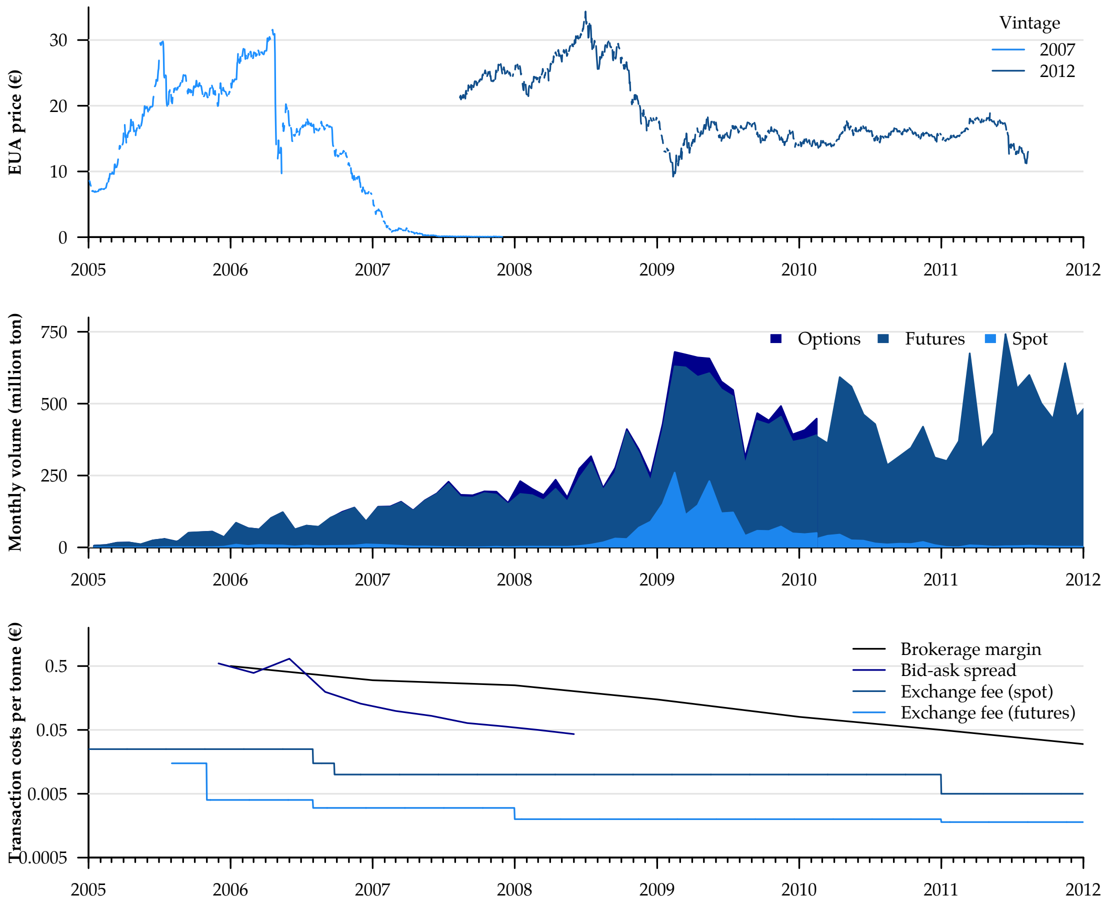

Figure 1: Market development. The **top panel** shows the evolution of the price of EUAs in phases 1 and 2 of the EU ETS. The prices of futures contracts that expire in 2007 (end of phase 1) and 2012 (end of phase 2) are taken as representative (Source: European Environment Agency). The **second panel** shows the volume of transactions for different types of contracts (Source: World Bank and Tendances Carbone). The **bottom panel** shows different measures of the variable costs of participating in the EU ETS. The gross brokerage margins are relevant for the majority of smaller polluters who lack direct access to the market (Source: Frunza, 2013b). The exchange fees are those charged by the European Energy Exchange for clearing spot and futures transactions (Source: EEX). The bid-ask spread—a common measure of total transaction costs in financial markets—is reported for futures expiring in December 2008 traded on the European Climate Exchange (Source: Frino et al., 2010).

to coordinate the search for the stolen allowances (INTERPOL Environmental Crime Programme, 2013, pp. 22-23).

Some more sophisticated criminal enterprises figured out how to exploit the regulatory blind spots in a more sustained way. Because emissions allowances were not designated as financial instruments under EU law, anyone could trade without a special license. National registries conducted their own vetting of account holders, but they were far less stringent than the Know-Your-Customer checks common in other financial markets. The weak identity checks became an entry point for several kinds of fraudulent trading.

One type of fraudulent trading was allowance "recycling." In 2010, the Hungarian government resold 2 million carbon offsets on the international market. These offsets had been previously purchased by Hungarian companies, and surrendered to the national registry in compliance with the EU ETS. However, when the Hungarian government found itself below their emissions quota, as set by the Kyoto Protocol, they could resell these offsets on the condition that these offsets would not be re-imported into the EU ETS. As a result of this restriction, the offsets were sold at a discounted price. However, the offsets were purchased by Hungarian Energy Power, which had been set up only two weeks prior. Through a quick succession of sales, these offsets were soon offered to unsuspecting buyers on the French carbon exchange BlueNext at full price, as if they were any other emissions allowances. This was illegal, but there were no procedures in place to check the serial numbers of these allowances against those sold by the Hungarian government. The scheme netted the fraudsters several million in profit while placing them out of reach of law enforcement. When the recycled allowances were discovered, BlueNext immediately suspended trading and prices crashed (INTERPOL Environmental Crime Programme, 2013, p. 15).

Another type of fraudulent trading involves VAT carousels. VAT carousels exploit the fact that cross-border transactions are exempted from VAT. VAT carousels are a well-known scheme to commit tax fraud and launder money, but it is difficult to accomplish with heavily regulated financial produces, or when you have to move physical products across borders. Emissions allowances thus presented an ideal commodity to commit these frauds—high value, easily transferable, and requiring minimal identity checks. The French carbon exchange BlueNext became a focal point for these VAT carousels when they offered immediate VAT reimbursement to traders, instead of forcing them to wait for 1-3 months like most other businesses. Trading activity on BlueNext boomed in late 2008, but estimates now suggest that 90% of trades between June 2008 and December 2009 were fraudulent, costing European taxpayers between €5 billion and €10 billion (Biegaj and Gnutek, 2010; Frunza, 2013a,b).

It is difficult to tally the costs of these regulatory blind spots—repeated market closures, stolen allowances, laundered money, and raided treasuries. It is even harder to quantify the indirect costs, although one suspects they would be substantial as well. In these tax fraud schemes, criminals have an incentive to sell below market price to increase the volume of transactions. This appears to have depressed allowances prices during 2008 and 2009 by €2-3 (Frunza, 2013a). This would have reduced the private incentive to abate, to invest, and to innovate in low-carbon technologies. These failures in enforcement also likely created distortions in linked commodity markets. Moreover, European law enforcement officials have alleged that the proceeds of these frauds have funded organized crime (Godart, 2015) and terrorism (Ferrarella and Guastella, 2014; Day and Bawden, 2014).

European regulators eventually became aware of these blind spots, and began taking steps to remedy the situation. When rumours of VAT carousels spread in 2009, the French government moved to exempt emissions allowances from VAT. This stopped the French tax authorities from paying out to fraudsters. Daily trading volumes immediately fell by over 90% on the carbon exchange BlueNext (see figure 2). The Danish registry, which had among the weakest identity checks, had accounted for approximately 45% of all new account openings and 40% of allowance transactions in this phase (Rigsrevisionen, 2012). But new account openings fell quickly once the Danish registry began requiring identification to open new accounts, and over 1,000 trading accounts were closed when the registry demanded to see identification for all existing account holders (Rigsrevisionen, 2012; see also figure 2).

Several countries kept VAT on EUAs even after this episode, though, one of which operated a spot market for EUAs: Italy. Within a few months, daily volumes increased more than a hundred-fold on the Italian exchange, GME (see figure 2). The number of participants nearly tripled to over 150, many of the new members sharing the same business address. GME quickly suspended their carbon market due to "presumed unlawful" activity. There were also some unusual spikes in trading activity on the Dutch carbon exchange, Climex, which closed a few months later.

In 2012, BlueNext settled with the French government for €32 million for its role in the systemic tax fraud, leading the exchange to close later that year (BusinessGreen, 2012; Vitelli, 2012). Several financial institutions, including the Royal Bank of Scotland and Julius Baer, were also prosecuted for their involvement in the fraud (Reuters, 2014; Carr, 2020). European law enforcement authorities also began prosecuting the individuals responsible. Europol reported that over 100 arrests had been made by December 2010 (Europol, 2010), and since then, dozens of individuals have been convicted across multiple countries, some receiving prison sentences of more than a decade (Eckert and Jones, 2011; Szabo, 2012; Hübner, 2016; Roberts, 2016; RFI, 2018).

Spot trading resumed in 2013 on the UK-based European Climate Exchange, operating as a market for futures contracts with a 1-day maturity. This transformed them into derivatives, putting them under existing financial market regulations. In January 2017, the European Commission finally adopted a new rule explicitly designating emissions allowances as financial instruments (Cardiff et al., 2015, pp. 22-23). Regulators have also adopted rules to strengthen the Know-Your-Customer checks for new registry accounts, and closed the loop hole through which EUAs were "recycled" (Cardiff et al., 2015). This episode also contributed to the creation of Eurofisc, a network of EU tax administrators tasked with combating VAT fraud, and to increasing Europol's power to investigate financial crime.[2]

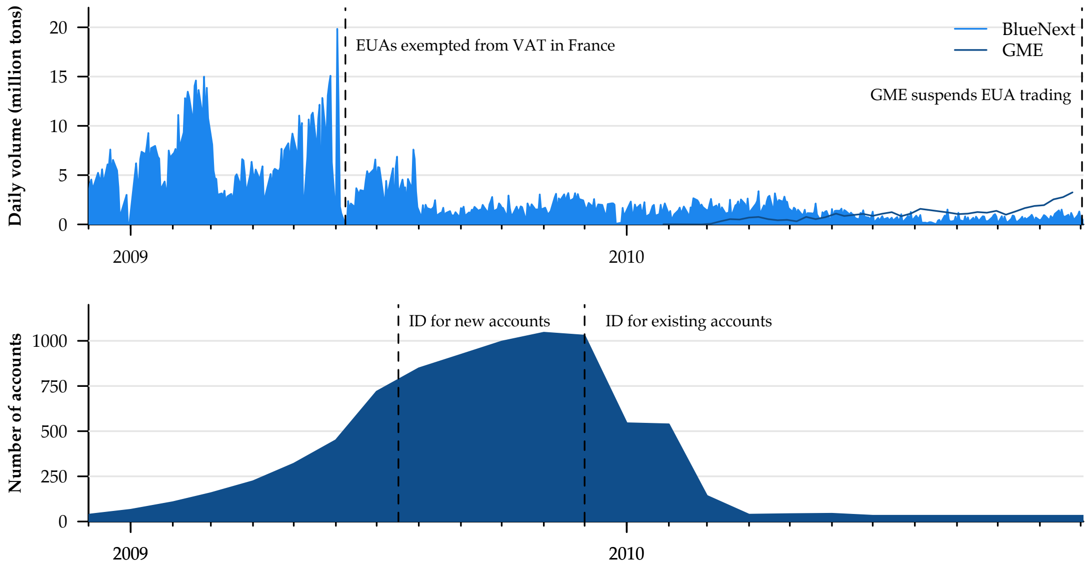

Figure 2: Carbon market fraud. The **top panel** shows the volume of EUA spot trading on the French carbon exchange BlueNext (Source: Bloomberg), and on the Italian exchange GME starting in 2010 (Source: Frunza, 2013b). The **bottom panel** shows the number of open personal holding accounts (i.e. those not associated with installations regulated under the EU ETS) in the Danish national registry (Source: Rigsrevisionen, 2012).

2.3 Surrender

The last stage of the life of an allowance—surrender—is simple in principle, but it is also where most things can go wrong. In sections 3, 4, and 5, we will conduct an in-depth quantitative analysis of these issues, but we start here by reviewing the general areas of concern, and the main strategies to encourage compliance.

The first question is how the regulator learns about each polluter's actual emissions. In the EU ETS, each polluter is required to compile a yearly emissions report, which has to be approved by an accredited third-party auditor and then submitted to the national regulator.

Third-party audits can increase reporting reliability, without the regulator having to conduct its own costly audits and inspections (Kunreuther et al., 2002). However, if the auditor is hired by the polluter, there is a fundamental conflict of interest between reporting truthfully and reporting what benefits the client. Across the EU ETS, national regulators employ a variety of strategies to limit these risks. Thirteen countries have at various times empowered the regulator to certify and decertify third-party auditors, though when

[2] It should be noted, though, that reports by both INTERPOL Environmental Crime Programme (2013) and the European Court of Auditors (Cardiff et al., 2015) have identified significant shortcomings of the new rules, and have warned of numerous ways the EU ETS can still be exploited.

the European Court of Auditors investigated three of the countries where certification was handled by an independent body, they found a lack of coordination with the regulator (Cardiff et al., 2015, pp. 36). Seven national regulators were empowered to appoint auditors to installations, rather than letting polluters select their own auditors. Another strategy that has suggested to limit conflicts of interest is to regularly rotate auditors, but in an investigation of seven countries, the European Court of Auditors found that none of the national regulators had such a policy in place (Cardiff et al., 2015, pp. 36-37).

Another strategy, which can be complementary to third-party auditing, is to require emissions reports to have higher precision. To credibly provide such precision, polluters must invest in better monitoring equipment and professional monitoring staff. This makes false reporting more difficult to maintain in the long run by reducing plausible deniability and creating more of a paper trail (Zahran et al., 2014).[3] This logic helps explain why larger polluters in the EU ETS are required to have smaller margins of error in their reports.

A more direct way to ensure the accuracy of emissions reports is to supplement third-party audits with direct inspections. Site visit from the regulator, as well as the threat of site visits, can contribute to improved self-reporting (Magat and Viscusi, 1990; Laplante and Rilstone, 1996; Telle, 2013; Lin, 2013; Lim, 2016). But inspections are costly, and regulators have historically targeted polluters with the highest marginal abatement costs—proxied by factors like plant size and complexity, past violations, etc. (Helland, 1998; Stafford, 2014).

A carbon price undermines the logic of targeted inspections, however. In equilibrium, marginal abatement costs are equal across polluters, and consequently, so is the marginal benefit of misreporting one's emissions. The regulator can do no better than to inspect at random (Stranlund and Dhanda, 1999), which reduces the value of inspections compared to the traditional command-and-control paradigm. This may explain why, even though every national regulator maintains the right to conduct inspections of EU ETS installations, public disclosures reveal that several regulators haven't actually conducted site visits.

When all else fails, harsh penalties for misreporting can encourage accurate self-reporting, even when the probability of detection is low (Lobo and Zhou, 2006; Stafford, 2007). European regulators have been given different amounts of discretion in this regard. The maximum fine for misreporting varies from €0 to €15 million, and the maximum prison sentence varies from 0 to 10 years.

[3] This strategy was used to great effect in the US Acid Rain Program, which mandated that power plants install continuous emissions monitoring systems (Henríquez, 2004). Continuous emissions monitoring has also been shown to reduce misreporting of of sulphur dioxide emissions from Chinese power plants (Xu, 2011).

A second set of challenges arise when installations surrender too few allowances to meet their obligations. European regulators have three main enforcement tools for these situations: they can impose a fine, they can "name and shame" noncompliant polluters, and in extreme cases, they can suspend installations.

Cap-and-trade programs typically mandate a fixed fine for each missing allowance (Stavins, 1998). The fine was €40 per tonne in the first compliance phase of the EU ETS, was raised to €100 in the second phase, and has been rising at the rate of inflation in the third phase. This penalty structure gives regulators little discretion to tailor fines to specific circumstances (Dion et al., 1998). On the one hand, a fixed fine removes the regulator's ability to apply escalating fines for repeat offenses (Oljaca et al., 1998; Kleit et al., 1998; Denning and Shastri, 2000; Lim, 2013), a strategy that can be used to increase compliance (Rubinstein, 1980). On the other hand, a fixed fine may encourage compliance because it reduces the ability of larger, more profitable companies to use their relationships with regulators to secure smaller fines (Gray and Deily, 1996). The net effect of removing this form of regulatory discretion is unclear.

Another way to penalize noncompliers is to "name and shame" them. This strategy has been used successfully to do everything from improving cleanliness in restaurants (Jin and Leslie, 2003; Simon et al., 2005) to cutting hospital waiting times (Besley et al., 2009). In the context of pollution control, being identified as bad actor generates negative press and hurts the stock market performance of publicly traded companies (Hamilton, 1995; Badrinath and Bolster, 1996), and ultimately encourages them to reduce their emissions (Konar and Cohen, 1997; Fung and O'rourke, 2000; Schlenker and Scorse, 2012). The EU ETS Directive has a provision for naming and shaming companies that do not surrender sufficient allowances, yet only some countries have used this tool.

In this section we have reviewed some of the most important strategies that regulators have used to encourage compliance in the EU ETS. In the remainder of this paper, we try to understand the extent to which these strategies can explain the observed patterns of compliance.

3 Trends in non-compliance

Collectively, EU ETS installations have emitted less than the emissions cap. Since the start of the program through 2020, 29.6 billion tonnes-worth of allowances have been allocated and another 1.6 billion tonnes-worth of international offsets have been imported into the EU ETS. During that time firms have emitted 29.5 billion tonnes of CO_2, corresponding to 94.6% of the total allowances available. This surplus of allowances appears to be the result of a combination of real emissions abatement, an overly generous allowance allocations, and a recession that depressed economic demand (Anderson and Di Maria, 2011; Bel and Joseph, 2015).

The annual emissions cap was not met in each individual year, however. When the first trading phase began in 2005, more allowances were issued than were needed for compliance (figure 3). Owing to difficulties setting up national registries and a general lack of familiarity with the new program, though, too few allowances were surrendered in that first year. Collectively, installations produced emissions equal to 97.6% of their total allocation of allowances in Phase 1, and the remaining surplus allowances were cancelled before the start of Phase 2.

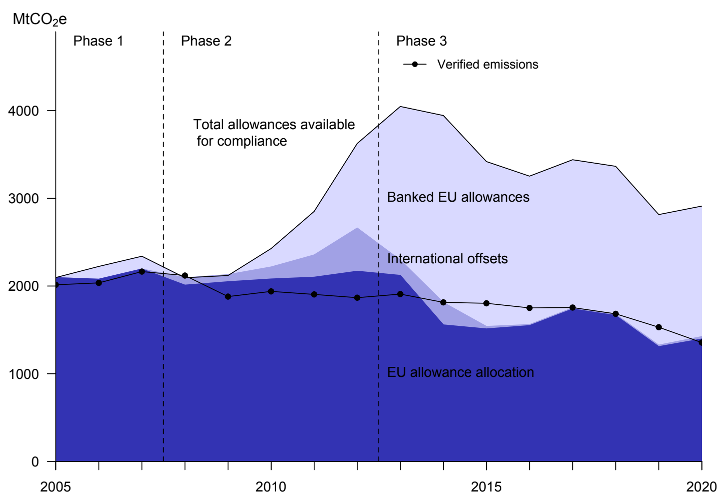

Figure 3: Trends in emissions and allowances. The dark shaded area at the bottom shows the total number of EUA and EUAA allowances allocated each year, whether freely or through auctioning. The lighter shaded area immediately above it shows the number of international offsets (CERs and ERUs) imported each year. These could be used for compliance starting in Phase 2. Together, these two areas make up the total number of new allowances available for annual compliance. The topmost shaded area shows the number of previously issued allowances that have not yet been surrendered and can be used for future compliance, which grows when there are more new allowances than are needed for compliance, and it shrinks when surrendered allowances outstrip new allowances. The vertical distance to the top black line thus shows the total number of allowances in the system that installations can in principle use for compliance. The the black dotted line shows verified emissions, which dictates how many allowances must be surrendered in total for compliance.

The second trading phase began in a similar manner as the first, with polluters not surrendering enough allowances to cover their collective emissions. The year 2008 was different from 2005, though, in that actual emissions exceeded the number of available allowances. Even so, because Phase 2 rules permitted offsetting, most of the shortfall could be made up by importing international offsets. In the following years, as emissions gradually fell and more international offsets were imported, a sizeable bank of allowances accumulated. This reserve eventually began to be drawn down in Phase 3, as the supply of new EU allowances and international offsets started to tighten significantly.

Though the overall emissions cap has been respected, compliance is determined at the installation-level. Specifically, an installation i that entered the EU ETS in year S has to answer for its cumulative emissions up to the present year T (with $T \geq S$):

$$\text{Emissions}^*_{iT} = \sum_{t=S}^{T} \text{Emissions}_{it} \tag{1}$$

This means that, if an installation fails to surrender enough allowances to cover all of its emissions in one year, the shortfall is automatically rolled over to the next year. If an installation surrenders allowances in excess of its emissions, however, it would forfeit those allowances—the regulator is not meant to serve as an allowance bank for polluters. From the regulator's point of view, the total number of allowances surrendered for compliance is therefore:

$$\text{Surrendered}^*_{iT} = \sum_{t=S}^{T} \min(\text{Surrendered}_{it}, \text{Emissions}_{it}) \tag{2}$$

Compliance status is then determined on a cumulative basis. An installation i that entered the EU ETS in year S is considered compliant in year T only if:

$$\text{Emissions}^*_{iT} \leq \text{Surrendered}^*_{iT} \tag{3}$$

Compliance has not been universal. According to official statistics, the annual rate of non-compliance fell from 22% in 2005 to 2-3% for the remainder of Phase 1, and then fell further as the EU ETS expanded to more installations in 2008 and 2013 (figure 4). Setting aside the year 2005, an average of 1.2% of installations have been officially non-compliant each year.

The official non-compliance rate does not quite match the aforementioned legal definition of non-compliance, however. Applying the legal definition to publicly available data on emissions and allowances surrendered yields an average non-compliance rate of 4%, again excluding 2005. The gap can be accounted for by two forms of discretionary behaviour on part of the regulator. First, the regulator appears to be serving as an allowance bank to the polluters. We can see in figure 4 that a large proportion of the apparent non-compliance events in the data vanish if we allow installations to put previously surrendered excess allowances toward compliance in the following year. Second, in a number of cases the regulator appears to have forgiven polluters for late surrender of allowances, retroactively modifying the official compliance status but failing to update the record in the public registry to reflect the late allowances. If we amend the definition of non-compliance to permits both allowance banking and late surrender, we get something very close to the official non-compliance rate.

The cumulative allowance shortfall for officially non-compliant installation-years was 474 $MtCO_2$ through 2020, of which nearly 80% had already accumulated by the end of 2005. At a fine of €40 per tonne in Phase 1 and €100 per tonne in Phases 2 and 3, and accounting for the fact that the same tonne may go unsurrendered for multiple years, these allowance shortfalls imply a total liability of €34.0 billion, or €13.0 billion if non-compliance in 2005 were pardoned. If the stricter definition of non-compliance were applied, there would be a more substantial shortfall. Under that definition, the cumulative shortfall was 582.3 $MtCO_2$ through 2020. The total liability would be €91.8 billion, or €76.8 billion if 2005 was pardoned.

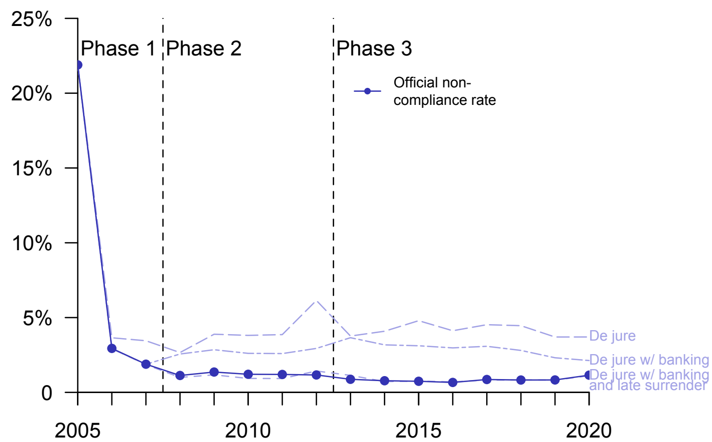

Figure 4: Rates of non-compliance. The rate of non-compliance is computed four different ways. The solid line shows the official non-compliance rate recorded in the EUTL. The top dashed line shows the non-compliance rate obtained by applying the legal definition of non-compliance to the EUTL's installation-level data on verified emissions and surrendered allowances. The middle hashed grey line show the non-compliance rate obtained by amending the legal definition to permit polluters to use previously surrendered excess allowances for compliance the following year. The bottom dashed line, which largely overlaps with solid line, shows the compliance rate if we further amend the definition to account for regulators accepting late surrender without recording these allowances in the EUTL.

These calculations presume that the underlying emissions reports are themselves trustworthy, but this need not be taken for granted. The frequency distribution of leading digits in many emissions data sets has been observed to conform to Benford's law, and deviations from this distribution have been used to identify manipulation of self-reported emissions Dumas and Devine, 2000; Marchi and Hamilton, 2006; Zahran et al., 2014; Stoerk, 2016. In our data, the distribution of leading digits substantially conforms to Benford's law (figure 5). Still, the deviations from it are telling. For instance, they are more pronounced for installations that were short of allowances ($\chi^2 = 11.4$; $MAD = 0.004$) than among installations with long positions ($\chi^2 = 3.8$; $MAD = 0.002$).[4] The degree of deviation from Benford's law may thus provide a complementary indicator, potentially allowing us to detect non-compliance along another margin.

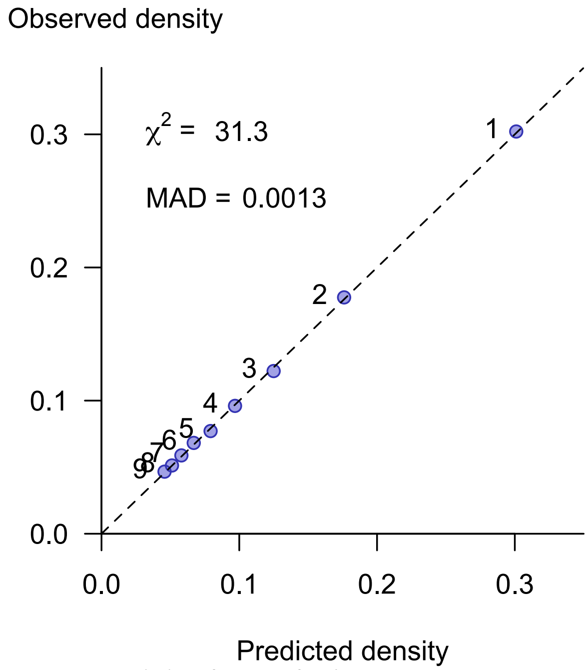

Figure 5: Misreporting as measured by deviations from Benford's law. The observed densities of first digits are here plotted against the predicted densities (e.g. 30% of values are predicted to begin with the number one, whereas 29.5% do in the data set). Dots on the 45-degree line indicate that the predicted and observed densities are equal.

To conclude, the official non-compliance rate has trended downward in the EU ETS, with an average of 1.2% since 2006. These official non-compliance codes, however, appear to incorporate banking with regulators and the fact that regulators often accepted late compliance. Benford's law does not flag major anomalies in reported emissions across the whole data set, although the companies buying allowances tend to have larger deviation from Benford's law than companies with excess allowances.

4 Do fines explain compliance?

4.1 The severity of punishment

The canonical economic model predicts that the variation in compliance should depend only on the severity and probability of punishment. Let us look at each in turn. The severity of the punishment for excess emissions was first set at €40 per tonne through 2007, and then at €100 per tonne. Other things equal,

[4]Since most data sets do not conform to Benford's law exactly, the χ^2 statistic is sensitive to sample size. Nigrini and Wells (2012) proposes that one should use the mean absolute deviation (MAD) of the densities when judging the relative conformity of samples of very different sizes.

the sudden 150% increase in the fine in 2008 would be expected to produce a decline in non-compliance, an increase in misreporting, or both. However, in 2008 the allocation of allowances also became more stringent, increasing the carbon price and hence the benefits of both non-compliance and misreporting. Taken together, the increased fine and stringency would have an ambiguous effect on overt non-compliance, but both changes would seem to encourage more covert non-compliance through misreporting.

We see almost no change in the overall non-compliance rate in 2008 (figure 4), but there is a suggestive uptick in misreporting (not shown). Figure 6 divides the sample into Small and Large installations. Small installations, with historically less than 50,000 tonnes of emissions per year, use measurement methodologies with inherently greater uncertainty and error, while larger installations are subject to much stricter reporting requirements. Smaller installations, then, have more opportunity to substitute from overt to covert non-compliance when the cost of non-compliance went up in 2008. Figure 6 shows that, although misreporting fell among larger installations, it rose slightly for small installations. This well-timed rise in misreporting among smaller installations, which face the loosest reporting requirements, is consistent with polluters substituting from overt to covert forms of non-compliance when faced with a higher fine.[5]

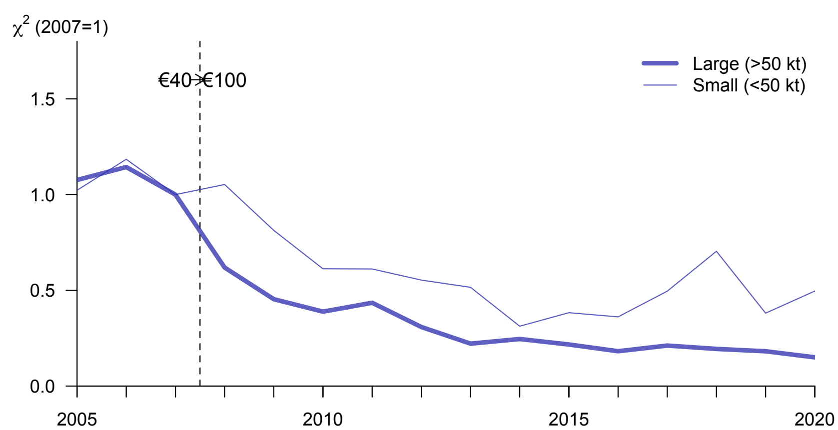

Figure 6: Misreporting by small and large installations. The two lines show the amount of misreporting among installations that were classified as Small or Large in phases 1 and 2, as measured by deviations from Benford's law. The χ^2-statistic has been normalized to 1 in 2007 to emphasize differences over time *within* each group. The raw χ^2-statistic otherwise vary significantly across groups because of differing group sizes.

[5]Benford's law allows detection of misreporting using deviations from expected patterns in the data set as a whole. This method of detection does not allow us to identify who is misreporting and how.

4.2 The probability of punishment

One other reason that we do not see a bigger change in the non-compliance rate in 2008, aside from misreporting, could be that the underlying probability of punishment was too low for the fine to make much difference. To estimate a noncomplier's probability of being punished, we divide the number of enforcement actions in each country and year by the number of non-compliant installations, and impute this average to each installation. This approximation is reasonable for a cap-and-trade program, since the marginal benefit of non-compliance should in theory be the same for all installations.

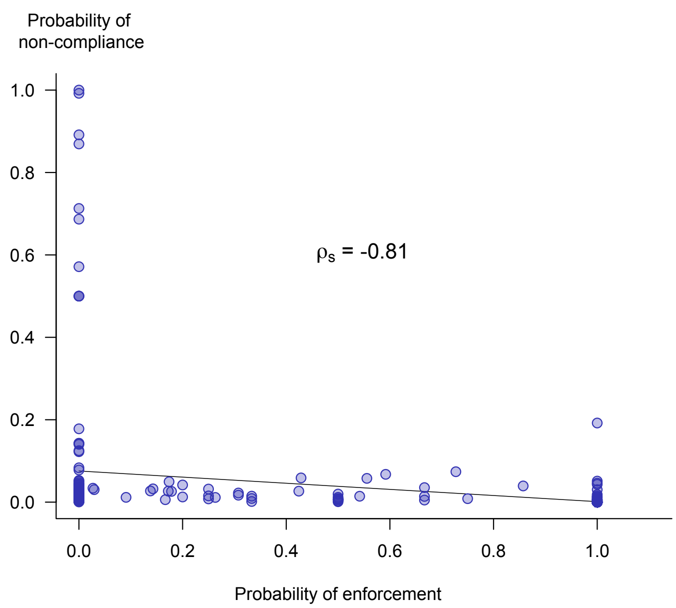

Figure 7: Association between enforcement and non-compliance. A higher probability of enforcement appears to be associated with a lower probability of non-compliance. The enforcement probabilities are estimated based on publicly available data on the number of enforcement actions. Wherever there was ambiguity about whether or not a particular enforcement action was taken, we have assumed it was. Also, whenever there is full compliance and no enforcement, we assume a unit-probability of enforcement. Our estimates are therefore an upper bound on the enforcement probabilities.

Figure 7 shows the association between the enforcement probabilities and non-compliance rates. Two features are especially noteworthy. First, there is a strong negative correlation between punishment and non-compliance (albeit highly non-linear), consistent with the economic hypothesis that a higher risk of punishment induces a lower rate of violations. This pattern is unchanged by excluding 2005 ($\rho_s = -0.81$).

Second, the risk of punishment is zero in most places most of the time. This might explain why the non-compliance rate did not change more dramatically when the fine increased from €40 to €100 in 2008. This also begs a new question: how much money was actually collected in fines for excess emissions?

Based on public reports submitted by national regulators to the European Commission, and on supplementary information solicited directly from national regulators, we have been able to verify €2.1 billion in fines. For comparison, we estimated earlier that non-compliance would have resulted in liabilities of €13 to €92 billion, depending on how strictly the standard for non-compliance is interpreted. The discrepancy represents significant foregone government revenues, similar in magnitude to the revenues from auctioning allowances (roughly €20 billion over the same phase).

These figures should be interpreted cautiously—some of the national reports mention ongoing legal proceedings, and others have not been made public. In addition, some fines go uncollected due to bankruptcy. The discrepancy between liabilities and collected fines may therefore be smaller than our numbers indicate. However, if regulators did collect fines anywhere close to the liabilities we estimate, that information is not available through public reporting or through direct requests made to national regulators.

5 Beyond the fine

The preceding section shows that, just as the canonical model predicts, the probability of non-compliance is associated with both the size of the penalty and the probability of enforcement. Yet, these factors only explain so much. If we regress the fine, the probability of enforcement, and their interaction, on the probability of non-compliance, we account for only 12% of the variation in the outcome. If we suppress everything but the interaction term, as the canonical model would have us do, this drops to just 7%. To understand whether the remainder is merely noise or due to differences in policy, we have to look at what other strategies the regulators are using to elicit compliance, beyond the fine. In this section, we look at how monitoring, enforcement, and discretionary power varies across countries and across time.

5.1 Monitoring

We observe variation in three types of monitoring strategies. In 35% of country-years, the national regulators had the power to appoint third-party auditors, rather than letting polluters select who will audit their emissions reports. These differences are not strongly associated with non-compliance rates, however (figure 8, left panel). Regulators that train and supervise third-party auditors tended to see lower non-compliance rates than those that did not (middle panel), and those that performed their own site inspections had slightly higher rates (right panel). It is worth noting that these indicators of monitoring powers are not highly correlated with each other. It is therefore unlikely that they are proxies of latent regulatory quality.[6]

It might come as a surprise that monitoring regimes are not more strongly associated with compliance (Murphy and Stranlund, 2006). However, one needs to keep in mind that stronger monitoring may, aside from reducing non-compliance behavior, also make the remaining non-compliance more visible. The net effect on the non-compliance rate could be positive or negative.

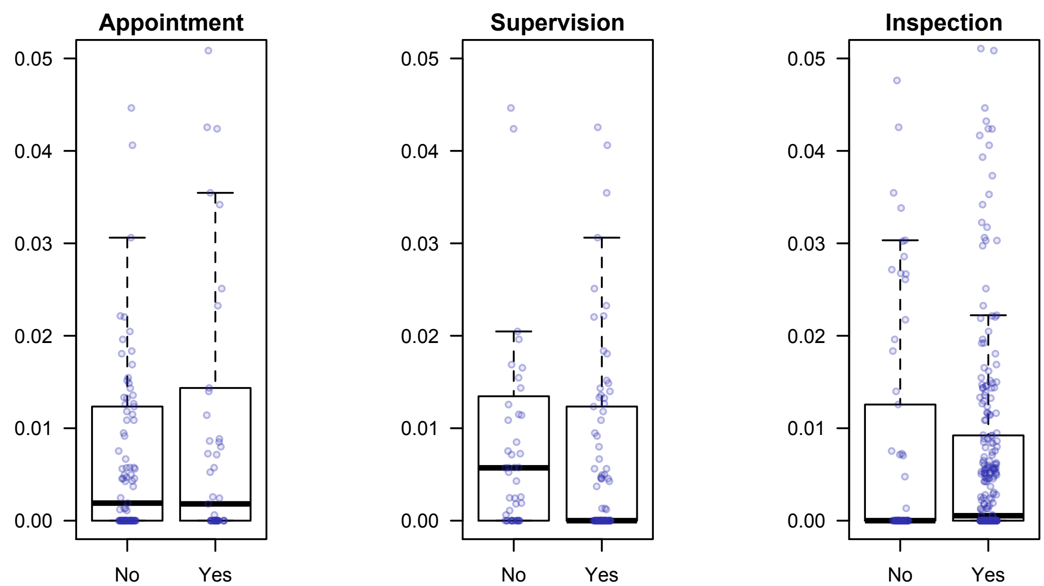

Figure 8: Associations between monitoring and non-compliance. The left panel plots the distribution of non-compliance rates for regulators that have the power to directly appoint third-party verifiers, compared to those that let polluters select who will audit their emissions reports. The middle panel compares regulators that train and supervise third-party verifiers with those who do not. The right panel compares regulators that perform their own site inspections with those that do not.

To get a better sense of whether there is a causal effect of monitoring regimes, we take inspiration from Duflo et al. (2013). When they assigned third-party auditors to Indian industrial plants, they found that pollution levels were reported more truthfully—fewer reports indicated pollution levels just below the pollution standard, compared to when the auditors were chosen by the polluters themselves. We adapt this strategy to investigate whether there is a similar effect present in the EU ETS.

Our setting is different from Duflo et al. (2013) in at least two important ways. First, monitoring regimes are not randomly assigned in the European carbon market. The fact that the three monitoring variables are so poorly correlated, however, does at least suggest their explanatory power (if any) is unlikely to derive

[6]The correlation between appointment and supervision is -0.12. The correlation between appointment and inspection is -0.14. The correlation between supervision and inspection is -0.04.

from some common unobserved factor. Second, the EU ETS is not a pollution standard. Nevertheless, polluters may wish to stay below the number of allowances they receive for free, at least on paper, since this avoids the inconvenience and cost of having to purchase extra allowances. The number of free allowances varies across polluters and across time, but it does provide a kind of focal point that parallels the pollution threshold under a standard.

To see what is happening around this focal point, then, start by subtracting the free allowances from verified emissions in each year. Polluters that emitted less than they were allocated will have a negative value, and can either bank or sell their extra allowances. Polluters that emitted more than they were allocated will have a positive value, and would have to purchase extra allowances or risk punishment. Applying the same logic as Duflo et al. (2013), when the regulator plays a stronger role in monitoring we would expect there to be a lower probability of values just below zero and a higher probability just above zero.

Following their procedure, we divide the excess emissions variable into bins and regress bin-dummies on an indicator for the monitoring regime. We include country-fixed effects, so the coefficient in each regression tells us the excess probability of polluters falling in any one bin when the monitoring regime of a country switches between a weaker and a stronger one. The results are plotted in figure 9, along with 95% confidence intervals. The standard errors are clustered at the country-year level, since regulators apply the same monitoring regime to all polluters in their jurisdiction, but do sometimes change practices over time.

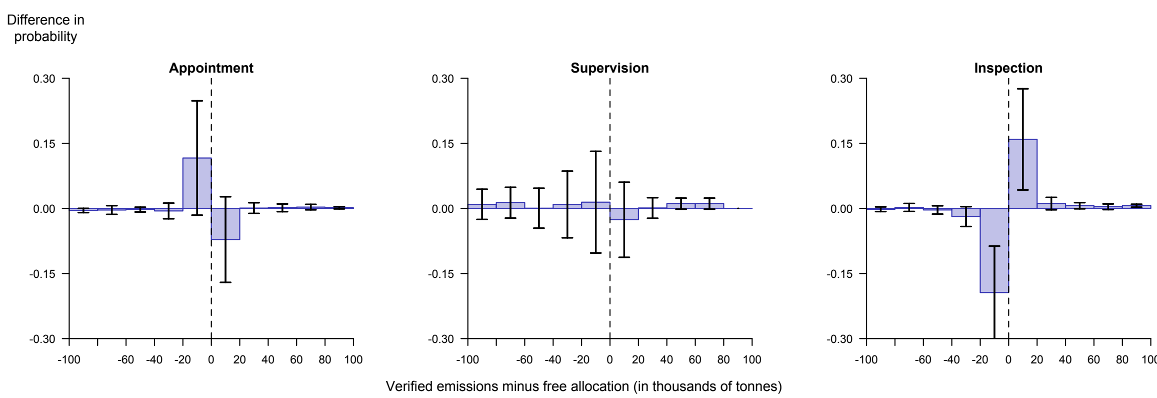

Figure 9: Verified emissions reports relative to free allowances. These three plots show how the probability of verified emissions exceeding free allowances by a certain amount differs by monitoring regime. The left panel compares regulators with and without the power to appoint auditors. The middle panel compares regulators with and without supervisory power over auditors. The right panel compares regulators with and without the power to perform their own site inspections. Standard errors are clustered at the country-year level.

Changes in monitoring regimes do not appear to be strongly associated with excess reporting above and below the focal point. Regulators with the power to appoint third-party auditors (left panel), and those with greater supervisory powers (middle panel), appear to receive *more* emissions reports just below the free allowance value, and *fewer* reports just above, but the differences are not statistically significant. Only for regulators with the power to perform site inspections (right panel) is the pattern in line with expectations, and statistically significant. Perhaps a site inspection is the only threat powerful enough to encourage more truthful reporting in the EU ETS. The implication would be that the other strategies are not adequately addressing the conflicts of interest for third-party auditors in practice. In testing three regulatory strategies, though, one should not overinterpret the one significant finding. It may therefore be most prudent to conclude that there is no clear signal that giving regulators stronger monitoring authority has increased compliance.

5.2 Enforcement

Figure 10 shows how non-compliance rates vary with regulators' enforcement powers. Regulators that "name and shame" offenders (left panel), and those that suspend offenders (middle panel), saw lower rates of non-compliance on average, as would be expected (Murphy and Stranlund, 2006). Regulators that levied greater fines for various reporting infractions saw no difference in non-compliance rates (right panel). As before, these regulatory characteristics are only weakly correlated with each other.[7]

Let us look at one of these enforcement tools in more detail: naming and shaming. Evidence from the US indicates that listed companies see their stock prices fall when they are publicly named as bad actors (Hamilton, 1995; Khanna et al., 1998), a mechanism that would seem less relevant for privately held companies. All else equal, this suggests that, when a regulator practices naming and shaming, listed companies would have lower rates of non-compliance compared to privately held companies.

This is a straightforward prediction that we can test with our data, although we should keep in mind that compliance behaviour is ostensibly public information under the EU ETS. Investors could therefore seek out this information whether or not the regulator takes the extra step of issuing a press release. Indeed, Brouwers et al. (2017) has already documented a negative relationship between allocation shortfalls in the EU ETS and firm value, but they do not test whether this is driven by jurisdictions that name and shame specifically. If we find no difference between jurisdictions, then, that would not deny the value of public disclosures, but only suggest that naming and shaming has little additional effect on listed firms in a context where investors have access to this information already.

[7] The correlation between naming and shaming and suspension is -0.18. The correlation between naming and shaming and other fines is -0.13. The correlation between suspensions and other fines is -0.04. These variables therefore do not appear to be collectively measuring the underlying strength of enforcement regime.

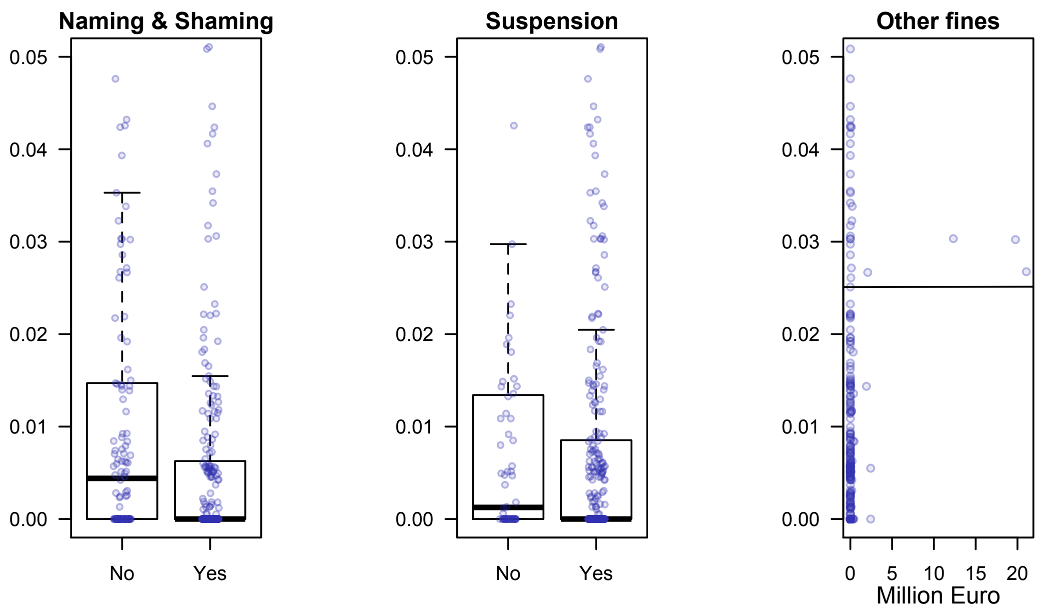

Figure 10: Associations between enforcement and non-compliance. The left panel plots the distribution of non-compliance rates for regulators that "name and shame" offenders compared to those that do not. The middle panel compares regulators that have the power to suspend offenders with those that do not. The right panel shows how the rate of non-compliance varies with the amount of fines levied for offenses other than excess emissions.

Another reason why this mechanism may be inoperative in our setting is that listed firms are more likely to be multinationals headquartered in another country. One theory about how naming and shaming works is that company leadership is afraid of bad press Campa (2018), and they may be less fearful of being named and shamed when they aren't headquartered in the same country as the offending plant. Campa (2018) observes this on a smaller scale in the US, where proximity to newspapers' headquarters appears to result in lower emissions. This gives rise to a second testable hypothesis, namely that, when a regulator practices naming and shaming, domestically owned firms should have lower non-compliance rates compared to foreign-owned firms. We should be careful to note, of course, that a null finding should not be read as evidence that naming and shaming has no effect, but merely that the effect does not vary in the way we expect.

The results are reported in figure 11. First, we regress non-compliance on dummy variables for whether a plant is operated by a listed firm, whether the regulator is practicing naming and shaming, as well as their interaction, along with country fixed effects. Standard errors are clustered at the country-year level.

Our first hypothesis implies that the coefficient on the interaction term should be negative, but our

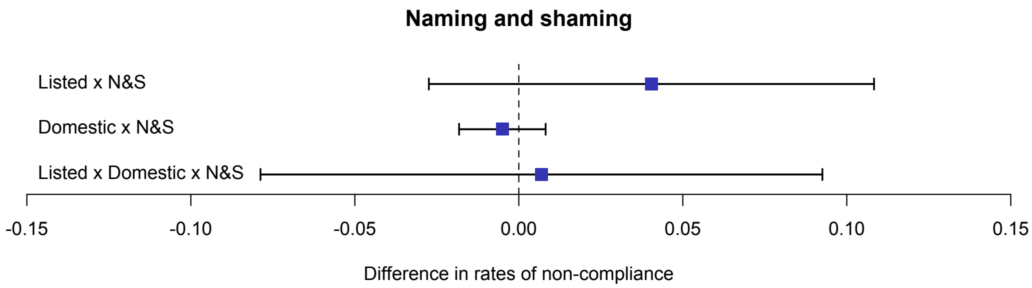

Figure 11: Testing hypotheses about naming and shaming. This plot summarizes the coefficients on the interaction term from three regressions, and their 95% confidence intervals. The hypotheses, described in the text, predict that all three coefficient would be negative. Standard errors are clustered at the country-year level.

estimate is positive and noisy. Next, we replace listing status with a dummy indicating whether or not the firm is headquartered in the same country as the plant. Our second hypothesis implies, once again, that the coefficient on the interaction term should be negative, but our estimate is zero. Finally, to account for possible correlation between listing status and domestic status, we regress compliance status on both, along with their interactions. Again, our coefficient estimate is zero. In sum, the coefficients in figure 11 show a fairly resounding null result. While this doesn't prove that naming and shaming has been ineffectual in the EU ETS, it should probably shift our beliefs in that direction.

5.3 Regulatory discretion

It has been found elsewhere that regulators sometimes treat non-compliers less favourably in other matters where the regulator has discretion, a practice that would create additional incentive to comply (Fenn and Veljanovski, 1988; Decker, 2003). One of the challenges of policing carbon markets is that the regulators have relatively few ways to exercise discretion. These newly created regulators do not oversee other areas of environmental or industrial permitting. However, one area where the national regulators may have some influence is in determining how many free allowances will be allocated to each polluter in the next trading phase.

To see whether national regulators might have been using their discretion, we look at whether non-compliance in Phase 1 is associated with lower free allocations in Phase 2. For each country, we estimate the following regression equation.

$$(\text{Phase 2 allocation}_i - \text{Phase 1 allocation}_i) = \beta_0 + \beta_1 \text{Phase 1 non-compliant}_i + u_i \tag{4}$$

where i indexes individual plants with strictly positive free allocations in both phases. The intercept measures the average change in free allocations in each country for compliant companies, while β_1 captures any remaining variation that can be explained by the plant being non-compliant in Phase 1—this is our measure of regulatory discretion. If non-compliers are systematically punished with comparatively less generous free allocations, we would expect $\beta_1 < 0$. If, on the other hand, regulators make greater concessions to non-compliers, perhaps to make it easier for them to comply in the future, then we would expect $\beta_1 > 0$.

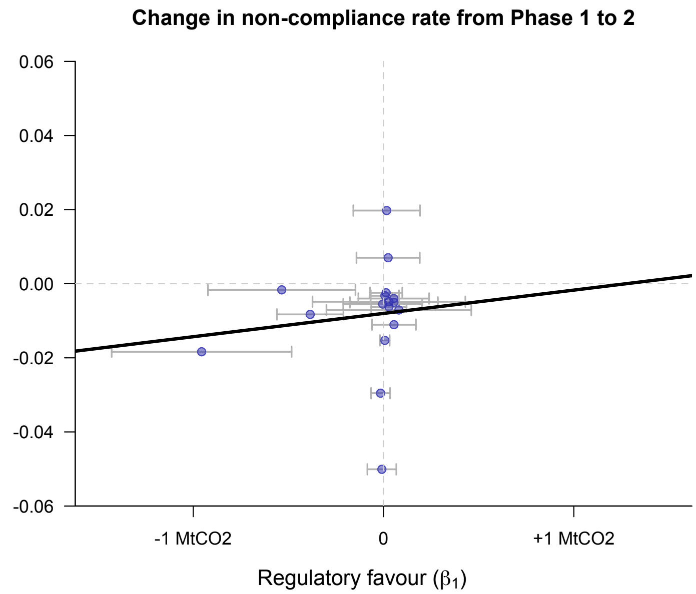

Figure 12: Association between regulatory favour for non-compliers and rates of non-compliance. The horizontal axis measures regulatory favour bestowed on non-compliers (β_1), with 95% confidence intervals plotted in grey. The vertical axis measures the change in the annual average rate of non-compliance between Phases 1 and 2. The year 2005 is excluded from Phase 1 here due to the special circumstances of that year, as discussed in section 3. Each dot represents a country that participated in both phases of the EU ETS, and line of best fit is drawn in black.

Figure 12 plots our estimated β_1 for each country, along with the change in non-compliance rates from Phase 1 to 2. The first thing to notice is most β_1s are statistically and substantively indistinguishable from from zero, which suggests that carbon market regulators did not generally exercise this kind of discretionary power. A small number of national regulators stand out, however, and it is interesting to observe that these regulators saw somewhat greater reductions in non-compliance on average. We are talking about only a handful of regulators, of course, and the coefficients measured with error at that, so the association is unsurprisingly weak. As such, we can neither confirm or reject the mechanisms studied by Fenn and Veljanovski (1988) and Decker (2003), but only observe that they are consistent with the pattern in our data.[8]

[8]We also estimated equation 4 in logarithmic changes instead of level-changes. The levels-specification is more sensitive

6 Conclusion and discussion

Carbon markets have emerged as one of the main policy instruments for curbing industrial greenhouse gas emissions. As they begin to ratchet down their emissions caps, the rising cost of compliance is going to contribute to making it more challenging to police these markets effectively. Where previous work has theorised a number of important enforcement challenges that arise specifically in the context of carbon markets (Stranlund and Dhanda, 1999), and has studied them in lab settings (Murphy and Stranlund, 2007), we have investigated how regulators are tackling these challenges in practice. To this end, we have conducted the first comprehensive analysis of compliance and enforcement behavior in the EU ETS. Although it is often impossible for an individual regulator to assess the effectiveness of particular enforcement strategies, we are able to gain additional insights by comparing the different ways that 31 countries have tried to enforce the same rules.

Our main finding is that there have been high rates of compliance in the EU ETS, and this despite relatively low levels of enforcement. In prior literature, this pattern has been called Harrington's paradox. The emissions caps have been respected in the aggregate, and the rate of non-compliance has been around 1-3%, depending on how it is measured. Meanwhile, our best estimate is that total fines did not exceed €2.1 billion, compared to our lower bound estimate of €13 billion in liabilities for excess emissions.

We have also investigated a number of potential resolutions to Harrington's paradox, but can only explain a small fraction of the variation in compliance behaviour. Expected fines, which play a central role in the canonical economic model, explain only about a tenth of the variation in non-compliance rates. Beyond fining, we do not find compelling evidence of the effectiveness of other regulatory tools, such as having more oversight over third-party auditors, naming and shaming offenders, or using regulatory discretion to reduce the free allowances of offenders. Across all of the monitoring and enforcement strategies we have examined, we have only found suggestive evidence that, when regulators have the power to perform site inspections, emissions are reported more accurately. The combination of widespread compliance and weak enforcement in the EU ETS thus remains an empirical puzzle.

As much as we can learn from a broad quantitative assessment, then, we find ourselves in need of new

to cases where regulators were applying substantial penalties to a small number of outlier cases as a way to set an example, whereas the logarithmic specification would pick up regulators that were applying a consistent proportional penalty. The logarithmic specification yields a similar line of best fit, but none of the coefficients are statistically significant. This suggests that none of the regulators are applying proportional penalties in a consistent fashion, which we also confirmed by examining the sensitivity to outliers in each country. The overall pattern, then, seems to be driven more by a small number of regulators that, on rare occasions, made examples of large non-compliers.

hypotheses and empirical strategies. To this end, we conclude here by summarizing qualitative findings from interviews with 11 national regulators conducted in the course of writing this paper, as well as one interview with the EU Transactions Log in Brussels. By triangulating between the regulator's first-person experiences and our quantitative findings, we hope future researchers will have a head start in generating new hypotheses.

When speaking to regulators about this period, one is reminded just how novel this policy was at first. Several regulators reported to us that, in 2005, when the EU ETS launched, there were firms that did not know whether or not they were required to comply with the new policy. Several national regulators reported that they tried to contact all the firms that should have been included in the EU ETS, but it was sometimes challenging to make these determinations on account of not having adequate data on either historical emissions or on plant-level production capacity. In some cases, even years later, regulators would call up firms that had not surrendered allowances or verification reports in the days before the deadline.

Meanwhile, the national regulators were in many cases newly minted entities themselves. While educating polluters about their obligations, they had to stand up entirely new bureaucracies. Unsurprisingly, they ran into many unforeseen problems. The Belgian registry, for example, required users to submit electronic identification with an identity card in order to gain access, but had not accounted for the fact that US citizens did not have such identity cards. The Czech Republic had similar problems with identification via text message in 2012. This locked some polluters out of their accounts just as they were expected to surrender emissions allowances. The French registry experienced a software error during the first trading phase, which deleted imported allowances from the database, turning many compliers into non-compliers.

In addition to these growing pains, the regulators sometimes faced obstacles put into place by other parts of the government. In France, national law requires the regulator to send a warning when a polluter is out of compliance, and there will be no penalty as long as the firm comes into compliance within one month. In one instance, the national EU ETS regulator actually tried to levy a fine, but the regional judicial authority refused to enforce the penalty, so the company was not fined. This means, in effect, that the real deadline for French polluters is one month later than it is meant to be under EU ETS rules.

Too often, polluters found whatever opportunity they could to make the regulators' job even harder. Fully aware of the legal novelty of these new emissions allowances, firms often chose to dispute the fines in court. Firms in financial trouble could also exploit the situation by selling their remaining allowances before declaring bankruptcy. One person told us,

> A lot of operators are closed because of the [2008 financial] crisis. The site is closed, there are no more people. In that case we [could] go ahead [with an enforcement action], but we understand that we will not get the money. We can try to get the allocation that is missing, but that is difficult. In practice, we block the account.

Some struggling firms, we were told, would even wait to collect their allotment of free allowances for the following year and sell those off too, just prior to declaring bankruptcy, leading to two years of non-compliance. Some regulators have tried to collect what's owed from the liquidator, but closing the account is often the only thing they can do.

Against this backdrop, it is easy to understand why many of the national regulators chose to exercise discretion in enforcement. One regulator told us,

> If companies surrendered units the following year because there was a disagreement over reporting, we cannot hold them responsible for that, and we consider them as compliant.

Another national regulator confirmed to us that they do in fact let polluters bank their permits with the registry for future compliance, contrary to how the European registry defines compliance.

> The EUTL automatically generates these codes regarding the compliance status based on whether there is a difference between verified emission data and the surrendered units... A company may be compliant according to the cumulative data but the code is generated every year and accordingly it can happen that it surrendered more units earlier than it should have done based on the verified emission report. Consequently the following year it shall surrender less units and it can result in code B [which means non-compliant] for that specific year.

Rather than punishing firms for failing to surrender allowances in a particular year, the regulators have focused their efforts on working with the firms to get those allowances somehow, whether by counting previously surrendered allowances or by letting them make up shortfalls even after the deadline has passed. The result is that firms are mostly choosing to comply, even while regulators are letting off non-compliers for billions. This pattern may look a bit less paradoxical once we have the benefit of a first-person perspective. Our quantitative analysis should therefore not be read as an indictment of the regulators, but more as a testament to how challenging a job they were asked to do.

Another finding in our interviews was on the persistence of database problems. When faced with our detailed questions about the compliance status of individual plants, one person told us "none of us has time to go through the data as you have." In some instances, we were able to identify discrepancies between the information held by national regulators and the public database compiled and published by the EU Transactions Log, discrepancies that remained years after the fact. We have tried to correct for these problems whenever possible (see appendix), and it seems relatively unlikely that any arbitrary database errors that remain could account for the correlations we have found. Nevertheless, this adds another obvious layer of doubt to the interpretation of our results.

Put together, our quantitative evidence put together with the first-person perspective speaks to the challenges of setting up and policing carbon markets. Even in countries with motivated politicians and highly developed government bureaucracies, novelty and inexperience present real problems. As such, many of the institutional challenges we have discussed may provide a preview of what could happen if and when policy makers try to expand carbon markets to smaller polluters and to non-industrial sectors, or when trying to link multiple emissions trading systems. More broadly, it seems prudent for any government that is implementing a carbon market to avoid creating everything from scratch, and instead try to make use of pre-existing regulatory infrastructure put in place by tax collectors and financial regulators whenever possible. Relying on more familiar and tested reporting systems and legal mechanisms may help to avert problems that already have solutions, and to avoid creating new blind spots along the ill-defined borders between old and new regulatory bodies.

Despite running into many of the problems of novelty, European regulators have been mostly successful in eliciting a high degree of compliance so far. This is a welcome finding, but is not a reason for complacency. As carbon markets grow in scope and in ambition, we should be prepared for the possibility that the challenges involved in policing carbon markets will grow as well. If more stringent emissions caps raise the cost of compliance substantially, it may force regulators to rely more heavily on their more formal powers to enforce the rules. We hope our evaluation of these regulatory tools can help policy makers and regulators identify what strategies are more or less promising, and can serve as a basis for further experimentation and reform.

After all, effective policing of carbon markets is embedded in a complex context of institutional rules, traditions, and perceptions.

Acknowledgements The authors gratefully acknowledge feedback from the AERE conference and the McCourt Faculty Seminar at Georgetown University. Frank Venmans acknowledges financial support from the

Grantham Research Institute on Climate Change and the Environment, at the London School of Economics and the ESRC Centre for Climate Change Economics and Policy (CCCEP) (ref. ES/R009708/1).
Declaration of interest: none.

References

Anderson, B. and Di Maria, C. (2011). Abatement and Allocation in the Pilot Phase of the EU ETS. *Environmental and Resource Economics*, pages 1–21.

Anderson, B. E. (1996). The Ebb and Flow of Enforcing Executive Order 11246. *The American Economic Review*, 86(2):298–301.

Anger, N. and Oberndorfer, U. (2008). Firm performance and employment in the eu emissions trading scheme: An empirical assessment for germany. *Energy Policy*, 36(1):12–22.

Ashenfelter, O. and Smith, R. S. (1979). Compliance with the Minimum Wage Law. *Journal of Political Economy*, 87(2):333–350.

Badrinath, S. G. and Bolster, P. J. (1996). The role of market forces in EPA enforcement activity. *Journal of Regulatory Economics*, 10(2):165–181.

Baker, J. B. (2003). The Case for Antitrust Enforcement. *The Journal of Economic Perspectives*, 17(4):27–50.

Bayer, P. and Aklin, M. (2020). The European Union Emissions Trading System reduced CO2 emissions despite low prices. *Proceedings of the National Academy of Sciences*, 117(16):8804–8812.

Bel, G. and Joseph, S. (2015). Emission abatement: Untangling the impacts of the EU ETS and the economic crisis. *Energy Economics*, 49:531–539.

Besley, T. J., Bevan, G., and Burchardi, K. B. (2009). Naming & Shaming: The impacts of different regimes on hospital waiting times in England and Wales. CEPR Discussion Paper 7306, C.E.P.R. Discussion Papers.

Betz, R. A. and Schmidt, T. S. (2015). Transfer patterns in Phase I of the EU Emissions Trading System: a first reality check based on cluster analysis. *Climate Policy*, 0(0):1–22.

Bhorat, H., Kanbur, R., and Mayet, N. (2012). Estimating the Causal Effect of Enforcement on Minimum Wage Compliance: The Case of South Africa. *Review of Development Economics*, 16(4):608–623.

Biegaj, A. and Gnutek, A. (2010). Europol review: General report on europol Activities. Technical report, European Police Office.

Block, M. K., Nold, F. C., and Sidak, J. G. (1981). The Deterrent Effect of Antitrust Enforcement. *Journal of Political Economy*, 89(3):429–445.

Böhm, S. and Dabhi, S. (2009). *Upsetting the Offset: The Political Economy of Carbon Markets*. MayFly-Books.

Bond, P., Sharife, K., Allen, F., Amisi, B., Brunner, K., Castel-Branco, R., Dorsey, D., Gambirazzio, G., and Hathaway, T. (2012). The CDM cannot deliver the money to Africa. Technical Report 2, EJOLT.

Bontems, P. and Bourgeon, J.-M. (2005). Optimal environmental taxation and enforcement policy. *European Economic Review*, 49(2):409–435.

Brouwers, R., Schoubben, F., and Hulle, C. V. (2017). The eu ets and corporate environmental abatement. *Applied Economics Letters*, 24(5):334–337.

Bushnell, J. B., Chong, H., and Mansur, E. T. (2013). Profiting from Regulation: Evidence from the European Carbon Market. *American Economic Journal: Economic Policy*, 5(4):78–106.

BusinessGreen (2012). BlueNext agrees €32m carbon fraud settlement. *BusinessGreen*.

Calel, R. (2020). Adopt or Innovate: Understanding technological choices under cap-and-trade. *American Economic Journal: Economic Policy*, 12(3):170–201.

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

Campa, P. (2018). Press and leaks: Do newspapers reduce toxic emissions? *Journal of Environmental Economics and Management*, 91:184–202.

Cardiff, K., Kerrigan, M., Macys, G., Friel, C., Fésűs, E., den Engelsen, S., Dentamaro, M., Dürrwanger, A., Gómez-Valcárcel, M. L., Gores, L., Kokot, J., Puisais-Jauvin, L., Soblet, F., and Zalega, A. (2015). The integrity and implementation of the EU ETS. Special report 6/2015, European Court of Auditors.

Carr, M. (2020). Royal Bank of Scotland Ignored Carbon Trading Fraud, Court Rules. *Bloomberg*.

Cason, T. N. and Gangadharan, L. (2006a). Emissions variability in tradable permit markets with imperfect enforcement and banking. *Journal of Economic Behavior & Organization*, 61(2):199–216.

Cason, T. N. and Gangadharan, L. (2006b). An Experimental Study of Compliance and Leverage in Auditing and Regulatory Enforcement. *Economic Inquiry*, 44(2):352–366.

Chan, H. S. R., Li, S., and Zhang, F. (2013). Firm competitiveness and the European Union emissions trading scheme. *Energy Policy*, 63:1056–1064.

Coates IV, J. C. (2007). The Goals and Promise of the Sarbanes-Oxley Act. *The Journal of Economic Perspectives*, 21(1):91–116.

Consulate Mumbai (2008). Carbon credits sufficient but not necessary for sustaining clean energy projects of major Indian business groups. Wikileaks reference ID: 08MUMBAI340.

Cordis, A. S. and Warren, P. L. (2014). Sunshine as disinfectant: The effect of state Freedom of Information Act laws on public corruption. *Journal of Public Economics*, 115:18–36.

Day, M. and Bawden, T. (2014). Briton 'used carbon trading to fund terror'. *Indepedent*.

Dechezleprêtre, A., Nachtigall, D., and Venmans, F. (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.

Decker, C. S. (2003). Corporate Environmentalism and Environmental Statutory Permitting. *Journal of Law and Economics*, 46(1):103–129.

Denning, K. and Shastri, K. (2000). Environmental performance and corporate behavior. *Journal of Economic and Social Research*, 2(1):13–38.

Di Porto, E., Persico, N., and Sahuguet, N. (2013). Decentralized Deterrence, with an Application to Labor Tax Auditing. *American Economic Journal: Microeconomics*, 5(1):35–62.

Dion, C., Lanoie, P., and Laplante, B. (1998). Monitoring of Pollution Regulation: Do Local Conditions Matter? *Journal of Regulatory Economics*, 13(1):5–18.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2013). Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*. *The Quarterly Journal of Economics*, 128(4):1499–1545.

Dumas, C. F. and Devine, J. H. (2000). Detecting Evidence Of Non-Compliance In Self-Reported Pollution Emissions Data: An Application Of Benford's Law. 2000 Annual meeting, July 30-August 2, Tampa, FL 21740, American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association).

Earnhart, D. (2004). Panel Data Analysis of Regulatory Factors Shaping Environmental Performance. *Review of Economics and Statistics*, 86(1):391–401.

Eckert, V. and Jones, K. (2011). German court finds six guilty in CO2 fraud trial. *Reuters*.

Ellerman, A. and Buchner, B. (2008). Over-allocation or abatement? A preliminary analysis of the EU ETS based on the 2005–06 emissions data. *Environmental and Resource Economics*, 41(2):267–287.

Ellerman, A. D., Convery, F. J., and de Perthuis, C. (2010). *Pricing Carbon: The European Union Emissions Trading Scheme*. Cambridge University Press.

Ellerman, A. D. and Trotignon, R. (2009). Cross Border Trading and Borrowing in the EU ETS. *The Energy Journal*, 30:53–77.

Engel, E. M. R. A., Galetovic, A., and Raddatz, C. E. (2001). A Note on Enforcement Spending and VAT Revenues. *Review of Economics and Statistics*, 83(2):384–387.

Europol (2010). Further investigations into VAT fraud linked to the Carbon Emissions Trading System. Press release.

Feinberg, R. M. (1980). Antitrust Enforcement and Subsequent Price Behavior. *The Review of Economics and Statistics*, 62(4):609–612.

Fenn, P. and Veljanovski, C. G. (1988). A Positive Economic Theory of Regulatory Enforcement. *The Economic Journal*, 98(393):1055–1070.

Ferrarella, L. and Guastella, G. (2014). La grande truffa dell'Iva in Italia per finanziare i gruppi islamici. *Corriere della Sera*.

Frino, A., Kruk, J., and Lepone, A. (2010). Liquidity and transaction costs in the European carbon futures market. *J Deriv Hedge Funds*, 16(2):100–115.

Frunza, M. (2013a). Aftermath of the VAT fraud on carbon emissions markets. *Journal of Financial Crime*, 20(2):222–236.

Frunza, M.-C. (2013b). *Fraud and Carbon Markets: The Carbon Connection*. Routledge.

Fung, A. and O'rourke, D. (2000). Reinventing Environmental Regulation from the Grassroots Up: Explaining and Expanding the Success of the Toxics Release Inventory. *Environmental Management*, 25(2):115–127.

Furlong, W. J. (1991). The Deterrent Effect of Regulatory Enforcement in the Fishery. *Land Economics*, 67(1):116–129.

Garfinkel, J. A. (1997). New evidence on the effects of federal regulations on insider trading: The Insider Trading and Securities Fraud Enforcement Act (ITSFEA). *Journal of Corporate Finance*, 3(2):89–111.

Godart, B. (2015). 2nd interpol environmental compliance and enforcement events. http://www.interpol.int/Crime-areas/Environmental-crime/Events/Meetings/2nd-INTERPOL-Environmental-Compliance-and-Enforcement-Events.

Gray, W. B. and Deily, M. E. (1996). Compliance and Enforcement: Air Pollution Regulation in the U.S. Steel Industry. *Journal of Environmental Economics and Management*, 31(1):96–111.

Gray, W. B. and Jones, C. A. (1991). Are Osha Health Inspections Effective? A Longitudinal Study in the Manufacturing Sector. *The Review of Economics and Statistics*, 73(3):504–508.

Hahn, R. (1984). Market power and transferable property rights. *The Quarterly Journal of Economics*, 99(4):753–765.

Hamilton, J. T. (1995). Pollution as News: Media and Stock Market Reactions to the Toxics Release Inventory Data. *Journal of Environmental Economics and Management*, 28(1):98–113.

Hanlon, M., Hoopes, J. L., and Shroff, N. (2014). The Effect of Tax Authority Monitoring and Enforcement on Financial Reporting Quality. *The Journal of the American Taxation Association*, 36(2):137–170.

Hanoteau, J. (2014). Lobbying for carbon permits in Europe. *Recherches économiques de Louvain*, 80(1):61–87.

Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics*, 37(1):29–53.

Heindl, P. (2015). The impact of administrative transaction costs in the EU emissions trading system. *Climate Policy*, 0(0):1–16.

Helland, E. (1998). The Enforcement of Pollution Control Laws: Inspections, Violations, and Self-Reporting. *Review of Economics and Statistics*, 80(1):141–153.

Henríquez, B. P. (2004). Information Technology: The unsung hero of market-based environmental policies. *Resources*, pages 9–12.

Hintermann, B. (2010). Market Power, Permit Allocation and Efficiency in Emission Permit Markets. *Environmental and Resource Economics*, 49(3):327–349.

Hintermann, B. (2015). Market Power in Emission Permit Markets: Theory and Evidence from the EU ETS. *Environmental and Resource Economics*, pages 1–24.

Hübner, A. (2016). Former Deutsche banker jailed for carbon trading fraud. *Reuters*.

Innes, R. (1999). Remediation and self-reporting in optimal law enforcement. *Journal of Public Economics*, 72(3):379–393.

INTERPOL Environmental Crime Programme (2013). Guide to Carbon Trading Crime. Technical report, INTERPOL.

IPA Energy Consulting (2005). Implications of the EU Emissions Trading Scheme for the UK Power Generation Sector. *Report to the UK Department of Trade and Industry*.

Jaraite, J., Convery, F., and Maria, C. D. (2010). Transaction costs for firms in the EU ETS: lessons from Ireland. *Climate Policy*, 10(2):190–215.

Jaraitė-Kažukauskė, J. and Kažukauskas, A. (2014). Do Transaction Costs Influence Firm Trading Behaviour in the European Emissions Trading System? *Environmental and Resource Economics*, pages 1–31.

Jin, G. Z. and Leslie, P. (2003). The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. *The Quarterly Journal of Economics*, 118(2):409–451.

Kaplow, L. and Shavell, S. (1994). Optimal Law Enforcement with Self-Reporting of Behavior. *Journal of Political Economy*, 102(3):583–606.

Keohane, N. O., Mansur, E. T., and Voynov, A. (2009). Averting Regulatory Enforcement: Evidence from New Source Review. *Journal of Economics & Management Strategy*, 18(1):75–104.

Khanna, M., Quimio, W. R. H., and Bojilova, D. (1998). Toxics Release Information: A Policy Tool for Environmental Protection. *Journal of Environmental Economics and Management*, 36(3):243–266.

Kleit, A. N., Pierce, M. A., and Hill, R. C. (1998). Environmental Protection, Agency Motivations, and Rent Extraction: The Regulation of Water Pollution in Louisiana. *Journal of Regulatory Economics*, 13(2):121–137.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, 79(3):651–692.

Konar, S. and Cohen, M. A. (1997). Information As Regulation: The Effect of Community Right to Know Laws on Toxic Emissions. *Journal of Environmental Economics and Management*, 32(1):109–124.

Kossoy, A. and Guigon, P. (2012). State and trends of the carbon market 2012. Annual report, World Bank.

Kunreuther, H. C., McNulty, P. J., and Kang, Y. (2002). Third-Party Inspection as an Alternative to Command and Control Regulation. *Risk Analysis*, 22(2):309–318.

Laplante, B. and Rilstone, P. (1996). Environmental Inspections and Emissions of the Pulp and Paper Industry in Quebec. *Journal of Environmental Economics and Management*, 31(1):19–36.

Law, M. T. (2006). How do Regulators Regulate? Enforcement of the Pure Food and Drugs Act, 1907–38. *Journal of Law, Economics, and Organization*, 22(2):459–489.

Lim, J. (2013). *Evaluating the impact of monitoring and enforcement in the Clean Air Act*. PhD thesis, Vanderbilt University.

Lim, J. (2016). The impact of monitoring and enforcement on air pollutant emissions. *Journal of Regulatory Economics*, pages 1–20.

Lin, L. (2013). Enforcement of pollution levies in China. *Journal of Public Economics*, 98:32–43.

Lobo, G. J. and Zhou, J. (2006). Did Conservatism in Financial Reporting Increase after the Sarbanes-Oxley Act? Initial Evidence. *Accounting Horizons*, 20(1):57–73.

Löfgren, Å., Wråke, M., Hagberg, T., and Roth, S. (2013). The Effect of EU-ETS on Swedish Industry's Investment in Carbon Mitigating Technologies. *University of Gothenburg Working Papers in Economics*, (565).

Magat, W. A. and Viscusi, W. K. (1990). Effectiveness of the EPA's Regulatory Enforcement: The Case of Industrial Effluent Standards. *Journal of Law and Economics*, 33(2):331–360.

Marchi, S. d. and Hamilton, J. T. (2006). Assessing the Accuracy of Self-Reported Data: an Evaluation of the Toxics Release Inventory. *Journal of Risk and Uncertainty*, 32(1):57–76.

Miller, N. H. (2009). Strategic Leniency and Cartel Enforcement. *American Economic Review*, 99(3):750–68.

Murphy, J. J. and Stranlund, J. K. (2006). Direct and market effects of enforcing emissions trading programs: An experimental analysis. *Journal of Economic Behavior & Organization*, 61(2):217–233.

Murphy, J. J. and Stranlund, J. K. (2007). A laboratory investigation of compliance behavior under tradable emissions rights: Implications for targeted enforcement. *Journal of Environmental Economics and Management*, 53(2):196–212.

Nigrini, M. and Wells, J. (2012). *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. Wiley Corporate F&A. Wiley.

Nyborg, K. and Telle, K. (2004). The role of warnings in regulation: keeping control with less punishment. *Journal of Public Economics*, 88(12):2801–2816.

Oljaca, N., Keeler, A. G., and Dorfman, J. (1998). Penalty Functions for Environmental Violations: Evidence from Water Quality Enforcement. *Journal of Regulatory Economics*, 14(3):255–264.

Olken, B. A. (2007). Monitoring Corruption: Evidence from a Field Experiment in Indonesia. *Journal of Political Economy*, 115(2):200–249.

Point Carbon (2010). *Carbon 2010 - Return of the sovereign*.

Point Carbon (2012a). Ireland abandons CO2 tax after losing court battle. *Point Carbon News*.

Point Carbon (2012b). Slovakian court suspends CO2 permit tax. *Point Carbon News*.

Potoski, M. and Prakash, A. (2005). Green Clubs and Voluntary Governance: ISO 14001 and Firms' Regulatory Compliance. *American Journal of Political Science*, 49(2):235–248.

Reuters (2014). Julius Baer asked for 4.7*mlnguaranteeinFrenchcarbonmarketprobe.Reuters*.

RFI (2018). 36 jailed for huge French carbon tax fraud.

Rigsrevisionen (2012). Report on the administration of the Danish Emission Trading Registry - Rigsrevisionen. Technical report, Public Accounts Committee of Denmark.

Roberts, A. (2016). Multi-billion EU carbon market fraud operated from Poland. *Euroactiv*.

Ronconi, L. (2010). Enforcement and Compliance with Labor Regulations in Argentina. *Industrial & Labor Relations Review*, 63(4):719–736.

Rubinstein, A. (1980). On an anomaly of the deterrent effect of punishment. *Economics Letters*, 6(1):89–94.

Sandbag (2011). Carbon Fat Cats 2011: The Companies profiting from the EU Emissions Trading Scheme.

Sandor, R., Bettelheim, E., and Swingland, I. (2002). An overview of a free-market approach to climate change and conservation. *Philosophical Transactions: MathematicalPhysical and Engineering Sciences*, 360(1797):1607–1620.

Sauer, R. D. and Leffler, K. B. (1990). Did the Federal Trade Commission's Advertising Substantiation Program Promote More Credible Advertising? *The American Economic Review*, 80(1):191–203.

Schlenker, W. and Scorse, J. (2012). Does Being a "Top 10" Worst Polluter Affect Environmental Releases? Evidence from the U.S. Toxic Release Inventory. *Working paper*.

Schneider, L. R. (2011). Perverse incentives under the CDM: an evaluation of HFC-23 destruction projects. *Climate Policy*, 11(2):851–864.

Segerson, K. and Tietenberg, T. (1992). The structure of penalties in environmental enforcement: An economic analysis. *Journal of Environmental Economics and Management*, 23(2):179–200.

Shimshack, J. and Ward, M. (2005). Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*.

Sigman, H. and Chang, H. F. (2011). The Effect of Allowing Pollution Offsets with Imperfect Enforcement. *The American Economic Review*, 101(3):268–272.

Sijm, J., Neuhoff, K., and Chen, Y. (2006). CO_2 cost pass-through and windfall profits in the power sector. *Climate Policy*, 6(1):49–72.

Simon, P. A., Leslie, P., Run, G., Jin, G. Z., and al, e. (2005). Impact of Restaurant Hygiene Grade Cards on Foodborne-Disease Hospitalizations in Los Angeles County. *Journal of Environmental Health*, 67(7):32–6, 56; quiz 59–60.

Snyder, E. A. (1990). The Effect of Higher Criminal Penalties on Antitrust Enforcement. *Journal of Law and Economics*, 33(2):439–462.

Stafford, S. L. (2002). The Effect of Punishment on Firm Compliance with Hazardous Waste Regulations. *Journal of Environmental Economics and Management*, 44(2):290–308.

Stafford, S. L. (2007). Should you turn yourself in? The consequences of environmental self-policing. *Journal of Policy Analysis and Management*, 26(2):305–326.

Stafford, S. L. (2014). Will additional federal enforcement improve the performance of pipelines in the U.S.? *International Review of Law and Economics*, 37:137–146.

Stavins, R. N. (1998). What Can We Learn from the Grand Policy Experiment? Lessons from SO2 Allowance Trading. *The Journal of Economic Perspectives*, 12(3):69–88.

Stoerk, T. (2016). Statistical corruption in Beijing's air quality data has likely ended in 2012. *Atmospheric Environment*, 127:365–371.

Stranlund, J. K. and Dhanda, K. K. (1999). Endogenous Monitoring and Enforcement of a Transferable Emissions Permit System. *Journal of Environmental Economics and Management*, 38(3):267–282.

Szabo, M. (2012). UK CO2 fraudsters trade Rolls Royces for jail time. *Reuters*.

Telle, K. (2008). The threat of regulatory environmental inspection: impact on plant performance. *Journal of Regulatory Economics*, 35(2):154–178.

Telle, K. (2013). Monitoring and enforcement of environmental regulations: Lessons from a natural field experiment in Norway. *Journal of Public Economics*, 99:24–34.

Transparency International (2011). *Global Corruption Report: Climate Change*. Earthscan, London.

Trotignon, R. (2012). Combining cap-and-trade with offsets: lessons from the EU-ETS. *Climate Policy*, 12(3):273–287.

Trotignon, R. and Delbosc, A. (2008). Allowance Trading Patterns During the EU ETS Trial Period: What does the CITL Reveal? Climate Report 13, CDC Climat.

van Zeben, J. A. (2009). The European Emissions Trading Scheme Case Law. *Review of European Community & International Environmental Law*, 18(2):119–128.

Vitelli, A. (2012). NYSE Euronext to Shut Carbon Market After Failed Auction Bid. *Bloomberg*.

Wara, M. (2007a). Is the global carbon market working? *Nature*, 445(8):595–569.

Wara, M. (2007b). Measuring the clean development mechanism's performance and potential. *UCLA L. Rev.*

Weil, D. (1996). If OSHA is So Bad, Why is Compliance So Good? *The RAND Journal of Economics*, 27(3):618–640.

Wolfram, C. D. (1999). Measuring Duopoly Power in the British Electricity Spot Market. *The American Economic Review*, 89(4):805–826.

Xu, Y. (2011). Improvements in the Operation of SO2 Scrubbers in China's Coal Power Plants. *Environmental Science & Technology*, 45(2):380–385.

Zahran, S., Iverson, T., Weiler, S., and Underwood, A. (2014). Evidence that the accuracy of self-reported lead emissions data improved: A puzzle and discussion. *Journal of Risk and Uncertainty*, 49(3):235–257.

Zaklan, A. (2013). Why Do Emitters Trade Carbon Permits? Firm-Level Evidence from the European Emission Trading Scheme. SSRN Scholarly Paper ID 2239656, Social Science Research Network, Rochester, NY.

Zapfel, P. (2007). A brief but lively chapter in EU climate policy: the Commission's perspective. In Ellerman, A. D., Buchner, B. K., and Carraro, C., editors, *Allocation in the European Emissions Trading Scheme: Rights, Rents and Fairness*. Cambridge University Press.

Compliance code	Explanation
A	The number of allowances and ERUs/CERs surrendered by 30 April is greater than or equal to verified emissions.
B	The number of allowances and ERUs/CERs surrendered by 30 April is lower than verified emissions.
C	Verified emissions were not entered until 30 April.
D	Verified emissions were corrected by competent authority after 30 April of year X. The competent authority of the Member State decided that the installation is not in compliance for year X-1 (pre-2013).
E	Verified emissions were corrected by competent authority after 30 April of year X. The competent authority of the Member State decided that the installation is in compliance for year X-1 (pre-2013).
–	No Compliance Obligations.

Table 1: Compliance codes

A Measuring non-compliance

Our measures of non-compliance are based on the publicly available EU Transactions Log (EUTL) database.[9] This database records both the quantity of verified emissions and the number of allowances surrendered for each installation and year. In addition, it assigns a compliance code to each installation to indicate whether or not each installation was in compliance as of May 1 of each year. The allowance and emissions data are submitted by national authorities to the European registry, while the compliance codes are assigned by the EUTL based on the submitted data.

Table 1 provides a list of definitions for the different compliance codes. A is compliant, B is non-compliant, and C indicates that data was not entered in time. Prior to 2013, compliance codes D and E were used to designate cases where the data had been corrected after May 1, but starting in 2013 the EUTL switched to adding a star to the previously assigned compliance code to indicate that the data had been changed in some way. For example A^* means that verified emissions did not exceed surrendered allowances on May 1, but that the data base has since been updated.

Between 2005 and 2016 there were 3,826 instances where installations were judged to have surrendered an insufficient number of allowances in time (code B/B^* or $D/D*$), though more than half of these non-compliance events were recorded in just the first year of the program (see table 2). There were another 1,330

[9] http://ec.europa.eu/environment/ets/

Code	2005	2006	2007	2008	2007	2010	2011	2012	2013	2014	2015	2016
A/A^*	8,038	10,367	10,895	10,733	10,721	10,807	10,488	10,697	10,912	10,877	10,670	10,528
B/B^*	2,277	313	218	98	127	116	118	133	97	95	80	71
C/C^*	7	25	44	94	167	196	195	153	136	116	109	88
D/D^*	0	2	4	30	18	17	12	0	0	0	0	0
E/E^*	1	0	5	198	196	52	42	0	0	0	0	0
–	0	0	0	0	0	0	0	0	0	0	1	1
Missing	243	26	84	451	330	173	409	227	130	268	565	849

Table 2: Caption.

instances where a verified emissions report had not yet been entered in time (code C/C^*), and 3,755 instances where no compliance code was assigned. We have deferred to the registry' word on which installations are compliant and non-compliant when calculating non-compliance rates in this paper, which means we treat an installation as non-compliant only if it has been assigned a compliance code of B/B^* or D/D^*, and we include it in the denominator only if it has been assigned a compliance code. One could reasonably argue that the definition of non-compliance should be expanded to include failures to submit a verified emissions report (C/C^*). As figure 13 shows, this results in about a doubling of the non-compliance rate in most years.

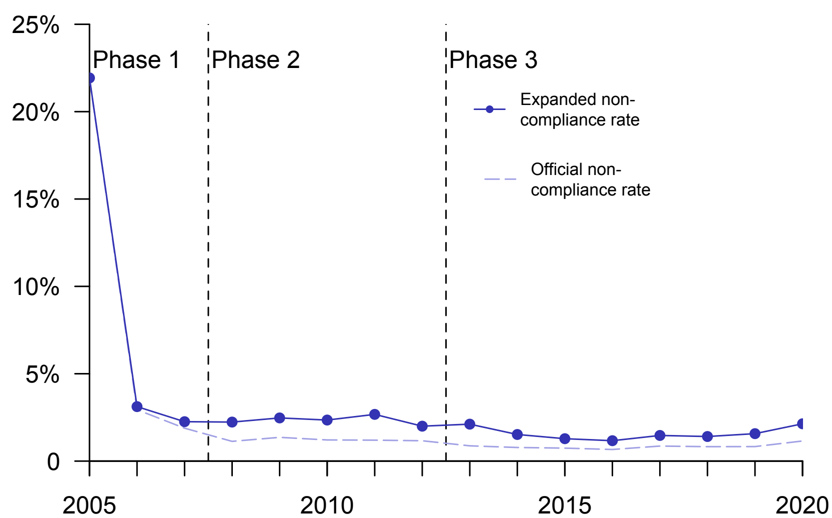

Figure 13: Alternative definitions of non-compliance rates.

Alternatively, we can ourselves compare the data on verified emissions and surrendered allowances to determine whether or not an installation is in compliance. This exercise is not entirely straightforward, however, on account of missing data. There are 671 instances where either verified emissions or the number of surrendered allowances are not available. And when either of these values are missing in one year, it also means we cannot determine whether an installation is compliant in any subsequent year. Recall, being in compliance requires not only that you surrender sufficient allowances in any given year, but also that you have surrendered enough allowances to cover your cumulative emissions. These 671 instances of missing data therefore create a total of 5,725 installation-years with indeterminate compliance status. In other words, for every instance of missing data we lose the ability to determine compliance status for an average of 8 installation-years.

Figure 14 shows the upper and lower bounds on the non-compliance rate, where we assume that all indeterminate cases are non-compliant to get an upper bound, or that they are all compliant to get a lower bound. Notice first how wide this range is. The observed data, then, is consistent with a non-compliance rate as high as 10% or as low as 5% for a single year. To narrow this down we need to make some assumption about the what data is missing. For instance, the line through the middle, which actually hews closer to the lower bound, shows the non-compliance rate that we get if were to assume that the data were missing at random.

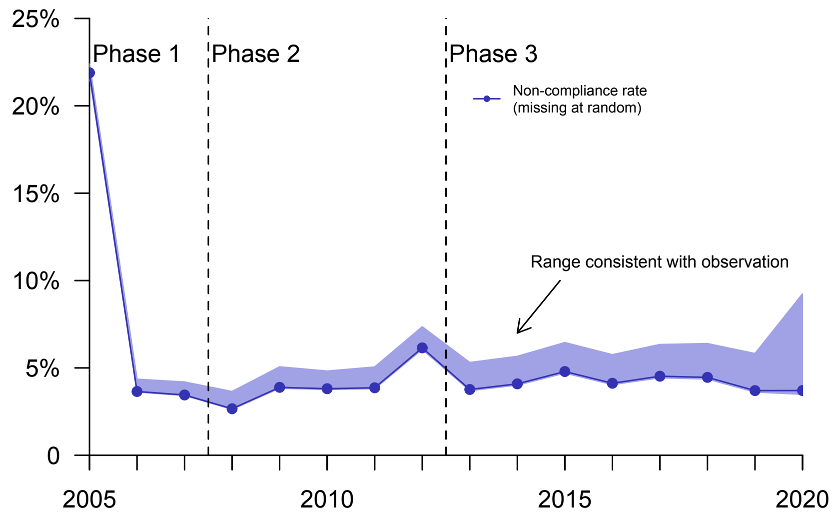

Figure 14: Bounds on non-compliance rates calculated from verified emissions reports.

B Patterns of non-compliance

For completeness, figure 15 presents the rates of non-compliance by country and by economic sector.

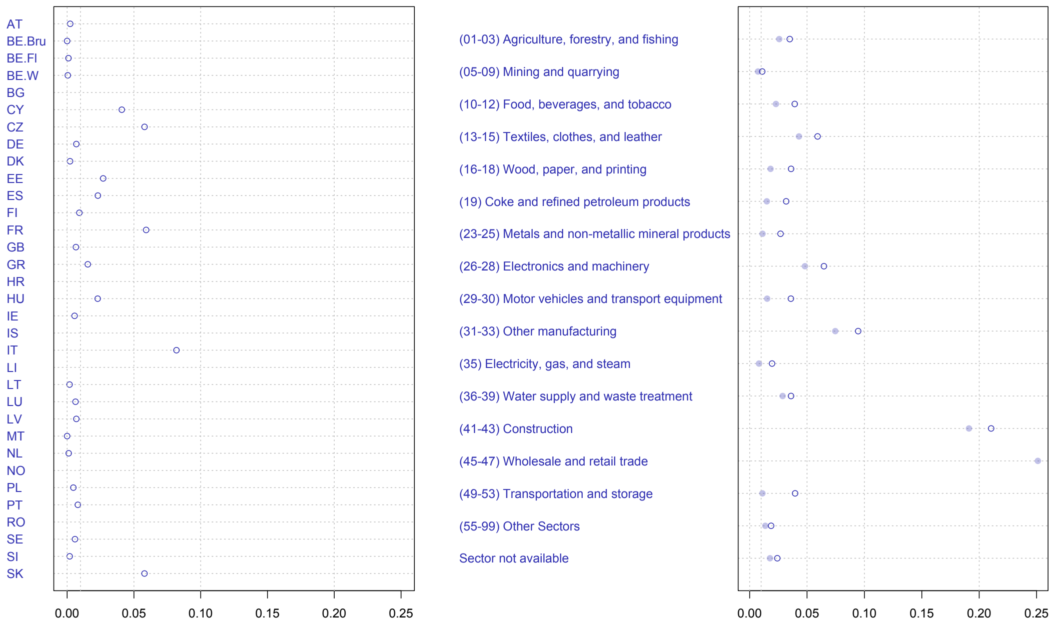

Figure 15: Patterns of non-compliance. The left panel plots the annual average rates of non-compliance by country, either including or excluding the first year of operation, represented respectively by the hollow circles and the filled in circles. The right panel provides the same information broken down by economic sector.