Climate Policy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tcpo20

Policing carbon markets

Raphael Calel, Antoine Dechezleprêtre & Frank Venmans

To cite this article: Raphael Calel, Antoine Dechezleprêtre & Frank Venmans (06 Mar 2025): Policing carbon markets, Climate Policy, DOI: 10.1080/14693062.2025.2464699

To link to this article: https://doi.org/10.1080/14693062.2025.2464699

CLIMATE POLICY
https://doi.org/10.1080/14693062.2025.2464699

Taylor & Francis
Taylor & Francis Group

Policing carbon markets

Raphael Calel[a], Antoine Dechezleprêtre[b] and Frank Venmans [c]

[a]McCourt School of Public Policy, Georgetown University; [b]Grantham Research Institute, London School of Economics and CERNA, Mines Paris - PSL; [c]Grantham Research Institute, London School of Economics

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. To shed light on the practical issues involved in policing carbon markets, this study presents the first comprehensive analysis of the EU Emissions Trading System, a single programme that was policed by up to 31 different national regulators. Since 2006, 98.8% of installations in the EU ETS have complied with the regulation according to official data. The observed non-compliance should have resulted in €13 billion in total fines, but only €2.1 billion appear to have been collected. More generally, variation in the probability and severity of fines across national jurisdictions and time explain just one-tenth of the variation in compliance rates. This pattern of *high* rates of compliance coupled with *low* rates of enforcement is known in the literature as 'Harrington's paradox'. 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. Therefore, resolving Harrington's paradox in the context of cap and trade regulation remains a fruitful area for future research.

Key policy insights

- The EU ETS has met its emissions targets in the aggregate. The compliance rate is close to 99%.
- Yet non-compliance was often not followed by enforcement action. This combination of high rates of compliance and low levels of enforcement has previously been termed 'Harrington's paradox'.
- Several monitoring and enforcement strategies are investigated – e.g. naming and shaming of non-compliant companies, having the regulator appoint third-party auditors instead of letting companies choose their own, and suspending offenders. Yet no strong causal evidence of their effectiveness was found.
- Regulators 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.
- The ETS' novelty and inexperience present real problems for enforcement, even in countries with highly developed government bureaucracies. Building on existing experienced institutions is important.

ARTICLE HISTORY
Received 18 September 2023
Accepted 4 February 2025

KEYWORDS
Pollution control; compliance; enforcement; cap-and-trade

JEL
Q50; Q52; C14

CONTACT Frank Venmans f.venmans1@lse.ac.uk Grantham Research Institute, London School of Economics, Houghton St, London WC2A 2AE

Supplemental data for this article can be accessed online at https://doi.org/10.1080/14693062.2025.2464699.

1. Introduction

The environmental integrity, political legitimacy, and economic efficiency of pollution control laws all fundamentally depend on industry's compliance. Regulators have developed many strategies to allow their limited budgets to stretch further: threatening harsh fines for non-compliers (Shimshack & Ward, 2005; Stafford, 2002), targeting past non-compliers for additional inspections (Helland, 1998), requiring self- reporting of emissions (Innes, 1999; Kaplow & Shavell, 1994) and third-party audits (Duflo et al., 2013), installing continuous pollution monitors (Giles, 2022; Henríquez, 2004), using extra-legal sanctions such as delaying permits for new projects (Decker, 2003; Fenn & Veljanovski, 1988), and naming and shaming the worst offenders (Konar & Cohen, 1997; Schlenker & Scorse, 2012).

Carbon markets have come to play a prominent role in climate change policy around the world. Their effectiveness builds on several forms of compliance: accurate emission monitoring, truthful reporting, and adequate surrender of allowances. Yet carbon markets' design undercuts many of the conventional strategies to encourage compliance. Unlike traditional command- and-control regulations, cap-and-trade programmes equalize the marginal benefit from non-compliance across plants and thus invalidate the logic of targeting past non-compliers for additional monitoring (Stranlund & Dhanda, 1999). Indeed, the marginal benefit of non-compliance for any plant is simply the market price of an allowance, which is the same for all polluters and hence independent of the plant's characteristics, including its compliance history. Then, even when a company surrenders insufficient emission allowances, the penalties for excess emissions are fixed by statute, effectively removing some of 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 non-compliers, since these programmes 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, this study conducts 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 programme has come to regulate the carbon-dioxide emissions of over 12,000 stationary sources across 31 countries (30 since Brexit), accounting for around 40% 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 programmes. 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.

The article contributes to several branches of the literature. First, it provides new empirical evidence on compliance and enforcement of pollution control regulation. Following Harrington (1988), a substantial theoretical literature has shown how regulators can use targeting and other strategies to realize high expected non-compliance costs even when the average enforcement effort is low (Bontems & Bourgeon, 2005; Fenn & Veljanovski, 1988; Harrington, 1988; Innes, 1999; Kaplow & Shavell, 1994; Kunreuther et al., 2002; Nyborg & Telle, 2004; Segerson & Tietenberg, 1992). More recent empirical studies examine the effects of these strategies in laboratory settings (Cason & Gangadharan, 2006a, 2006b; Murphy & Stranlund, 2006, 2007), in observational data (Earnhart, 2004; Helland, 1998; Keohane et al., 2009; Telle, 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 programmes may face. In addition, while earlier studies at best obtain estimates from variation in monitoring or enforcement at the polluter-level, this study exploits 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 investigated strategies have close analogies in the monitoring and enforcement of labour market regulations (Anderson, 1996; Ashenfelter & Smith, 1979; Bhorat et al., 2012; Ronconi,

2010), financial market regulations (Coates & C, 2007; Garfinkel, 1997), electricity market regulations (Wolfram, 1999), fisheries regulations (Furlong, 1991), workplace safety (Gray & Jones, 1991; Weil, 1996), food and drug safety (Law, 2006), product quality standards (Potoski & Prakash, 2005), advertising standards (Sauer & Leffler, 1990), anti-corruption measures (Cordis & Warren, 2014; Olken, 2007), antitrust (Baker, 2003; Block et al., 1981; Feinberg, 1980; Miller, 2009; Snyder, 1990), and tax law (Di Porto et al., 2013; Engel et al., 2001; Hanlon et al., 2014; Kleven et al., 2011). Our findings suggest that these strategies that have been shown to be effective in other domains may not transfer to carbon markets.

Third, the EU ETS continues to serve as a model for carbon market programmes 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 (Anderson & Di Maria, 2011; Bayer & Aklin, 2020; Dechezleprêtre et al., 2023; Ellerman & Buchner, 2008), on employment (Anger & Oberndorfer, 2008; Chan et al., 2013), on investment (Lo¨fgren et al., 2013), and on innovation (Calel, 2020; Calel & Dechezleprêtre, 2016). 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.

The remainder of the paper is structured as follows. Section 2 describes the ETS rules and associated compliance issues reported in the literature. Section 3 describes the trends in non-compliance, Section 4 investigates to what extend fines explain compliance, whereas Section 5 investigates other explanations such as differences between countries' monitoring and enforcement strategies. Section 6 adds insights from our interviews with regulators and Section 7 concludes.

2. The life of an emission 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 programme, a limited number of tradeable emissions allowances are distributed initially, be it for free or by auction. Allowances can then be traded throughout the trading 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 (described in Appendix A), trading and surrender.

2.1. Trading

During the EU ETS' first trading period, registered polluters, but also financial institutions, charities, and even private persons, could also open trading accounts in national registries. Once the free emissions allowances were deposited into polluters' accounts, anyone was free to buy and sell allowances.

In the first and second trading period, market fundamentals explain most of the trading behaviour (Betz & Schmidt, 2015; Ellerman and Trotignon, 2009; Ellerman et al., 2010; Zaklan, 2013). For instance, the market price of an EU allowance (EUA) rose rapidly in the programme'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 6 in Appendix A). In the third period, reforms related to the cancellation of large numbers of allowances are understood to have driven prices up to €90 (Gerlagh et al., 2022).

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 (Ellerman et al., 2010; Trotignon & Delbosc, 2008). 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, 2011, 2017). However, 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. 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 such as VAT carousels, which 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 products, or when one has 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 & Gnutek, 2010; Frunza, 2013a, 2013b).

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 (Carr, 2020; Reuters, 2014). European law enforcement authorities also began prosecuting the individuals responsible. Europol reported that over 100 arrests had been made by December 2010 (Biegaj & Gnutek, 2010), and since then, dozens of individuals have been convicted across multiple countries, some receiving prison sentences of more than a decade (Eckert & Jones, 2011; Szabo, 2012; Hu¨bner, 2016; Roberts, 2016). In 2018, emission allowances were included in the Markets in Financial Instruments Directive II (MiFID II) imposing trading authorization, transparency requirements, reporting obligations, anti-money laundering rules, etc.[1]

2.2. Surrender

The third stage of the life of an allowance – surrender – is simple in principle, but it is also where most things can go wrong. This section reviews the general areas of concern, before proceeding to an in-depth quantitative analysis of these issues in sections 3 through 5.

The first set of challenges relate to learning about each polluter's emissions. In the EU ETS, each polluter has to submit a yearly emissions report approved by an accredited third-party auditor. 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 an obvious conflict of interest. Across the EU ETS, national regulators employ a variety of strategies to limit these risks. Thirteen countries have empowered the regulator to certify and decertify third-party auditors; seven national regulators were empowered to appoint auditors to installations, rather than letting polluters select their own auditors. Another strategy would be to regularly rotate auditors, but no national regulator has such a policy in place (Cardiff et al., 2015, pp. 36–37).

Another monitoring strategy 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 creating more of a paper trail and reducing plausible deniability (Zahran et al., 2014).[2] In the EU ETS, larger polluters are thus 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 (Laplante & Rilstone, 1996; Lim, 2016; Lin, 2013; Magat & Viscusi, 1990; Telle, 2013). But inspections are costly, perhaps especially as a tool for monitoring carbon markets. 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). However, since every polluter in a carbon market has marginal

[1]See appendix B for a longer description of the VAT carousel episode, as well as other criminal enforcement issues in the EU ETS.

[2]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 sulphur dioxide emissions from Chinese power plants (Xu, 2011).

abatement costs equal to the market price, the probability of non-compliance ought to be independent of plant characteristics, including its past compliance history. The regulator can therefore do no better than to inspect at random (Stranlund & Dhanda, 1999), which reduces the value of inspections compared to the traditional command-and-control paradigm. This may explain why several national regulators in the EU ETS report never having conducted site visits.

When the above strategies are insufficient, the threat of harsh penalties for misreporting can encourage accurate self-reporting, even when the probability of detection is low (Lobo & Zhou, 2006; Stafford, 2007). Member State 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]

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

The first cap-and-trade programmes mandated a fixed fine for each missing allowance (Stavins, 1998). The EU ETS adopted a similar rule: the fine was €40 per tonne in the first trading period, €100 in the second period, and has been rising at the rate of inflation in the third and fourth period.[4] On the one hand, a fixed fine limits the regulator's ability to apply escalating fines for repeat offenses (Denning & Shastri, 2000; Kleit et al., 1998; Lim, 2013; Oljaca et al., 1998), 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 companies to use their relationships with regulators to secure smaller fines (Gray & Deily, 1996).

Another way to penalize non-compliers is to 'name and shame' them. In the context of pollution control, being identified as bad actor generates negative press and hurts the stock market performance of publicly traded companies (Badrinath & Bolster, 1996; Hamilton, 1995), and ultimately encourages them to reduce their emissions (Fung & O'Rourke, 2000; Konar & Cohen, 1997; Schlenker & 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.

3. Trends in non-compliance

The European Union Transaction Log (EUTL) is the official registry of the EU ETS. It provides information on all regulated installations, including contact details, the number of allowances freely allocated each year, the verified emissions, the number of surrendered allowances, and the compliance status of each installation.[5]

EUTL data show that EU ETS installations, collectively, have emitted less than the emissions cap. Since the start of the programme through 2020, 29.6 billion tonnes-worth of allowances were allocated and another 1.6 billion tonnes-worth of international offsets were 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 allocation of allowances, and a recession that depressed economic demand (Anderson & Di Maria, 2011; Bel & Joseph, 2015; Koch et al., 2014).

The annual emissions cap was not met in each individual year, however. When the first trading period began in 2005, more allowances were issued than were needed for compliance (Figure 1). Owing to difficulties setting up national registries and a general lack of familiarity with the new programme, 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.

The second trading period 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,

[3]Data on countries' differences in implementation are based on regulators' reports to the European Commission (Art. 26 of Directive 2003/87/EC). We have downloaded all public reports and written to all regulators whose reports were missing or unavailable.
[4]See Directive 2003/87/EC, art. 16. As mentioned earlier, other penalties can be imposed for other offenses, such as misreporting.
[5]The EUTL data has well-known limitations. See appendix C and Pahle et al. (2023) for more details.

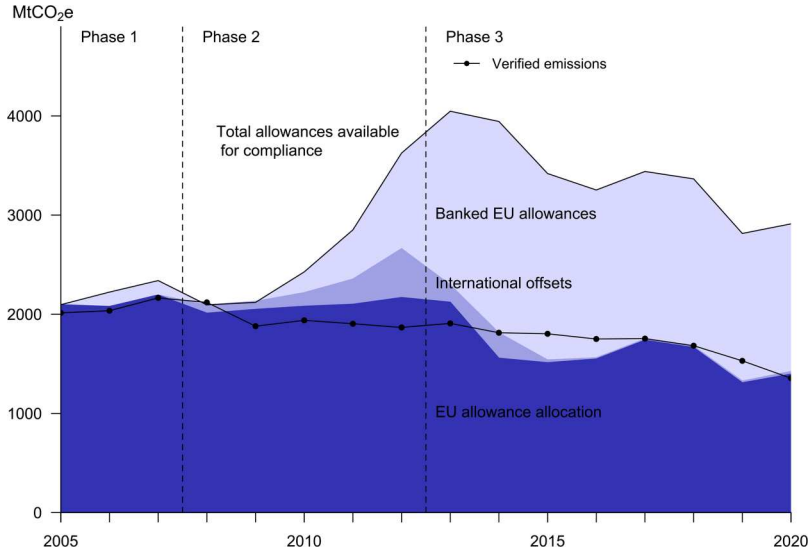

Figure 1. Trends in emissions and allowances. The dark shaded area at the bottom shows the total number of EUA and EUAA allowances allocated each year, whether freely or through auctioning. The lighter shaded area immediately above it shows the number of international offsets (CERs and ERUs) imported 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 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 black dotted line shows verified emissions, which dictates how many allowances must be surrendered in total for compliance.

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 was respected, compliance is determined at the installation-level for each year. 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:

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

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 to2–3% for the remainder of Phase 1, and then fell further as the EU ETS expanded to more

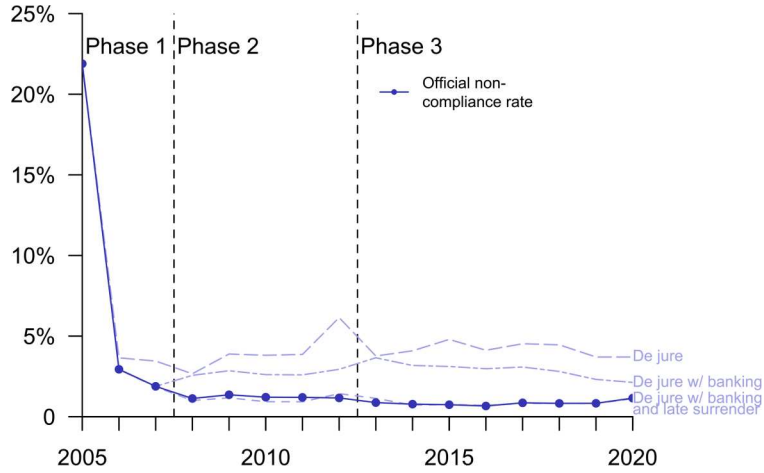

Figure 2. Rates of non-compliance. The rate of non-compliance is computed in four different ways. The solid line shows the official non-compliance rate recorded in the EUTL. The top dashed line shows the non-compliance rate obtained by applying the legal definition of non-compliance to the EUTL's installation-level data on verified emissions and surrendered allowances. The middle hashed grey line shows 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 the 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.

installations in 2008 and 2013 (Figure 2). 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. Two forms of discretionary behaviour by the regulator can account for this gap. First, the regulator appears to be serving as an allowance bank to the polluters. Figure 2 shows that a large proportion of the apparent non-compliance events in the data vanish if installations are allowed 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. When the definition of non-compliance is amended to allow for both allowance banking and late surrender, the result is 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.[6]

These calculations presume that the underlying emissions reports are themselves trustworthy, but this need not be taken for granted.

One way to test the veracity of emissions reports would be to look for inconsistencies with other data. This is challenging, since other emissions data sets are available only for small subsets of EU ETS installations. The best example of what can be achieved with this approach is perhaps Colmer et al. (2024), who look at the correlation

[6]One reviewer suggested the possibility that regulators might be crediting non-compliant installations for over-compliance at other installations operated by the same firm. This would require the regulator to draw on external data sources to identify common ownership. Using the enhanced EUTL data set from Jaraite˙ et al. (2016), this kind of firm-level compliance assessment is found to reduce the allowance shortfall by 15% at most. We thank the anonymous reviewer for this suggestion.

between emissions reported under the EU ETS and those computed on the basis of the French manufacturing energy consumption survey (EACEI). Both are self-reported, but only the EU ETS report is verified by a third party. They find a correlation of 0.96. This indicates a very high degree of correspondence, and they do not report any further analysis of the discrepant cases.[7]

In the absence of multiple comparable data sources, Benford's law provides an alternative avenue for measuring the veracity of reports. Benford's law states that the frequency distribution of leading digits in certain types of data sets will follow a particular power law (Nigrini & Wells, 2012). Many emissions data sets have been observed to conform to Benford's law, and deviations from this distribution have been used to identify manipulation of self-reported emissions (Dumas & Devine, 2000; Marchi & Hamilton, 2006; Stoerk, 2016; Zahran et al., 2014). In our data, the distribution of leading digits substantially conforms to Benford's law (Figure 3, left panel). 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 ($\chi^2 = 3.8$; $MAD = 0.002$).[8] The degree of deviation from Benford's law may thus provide a complementary indicator, potentially allowing us to detect non-compliance along another margin.

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 (Becker, 1968; Ehrlich, 1996). Let us look at each in turn. The severity of the punishment for excess emissions was first set at €40 per tonne, and then at €100 per tonne from 2008 onwards. Yet, we see almost no change in the overall overt non-compliance rate in 2008 (Figure 2). Note however that at the same time in 2008, the carbon price increased. The higher fine *decreases* the incentive for overt non-compliance, while the higher carbon price *increases* the incentive to do so. The sum of both effects is ambiguous.

By contrast, both the higher fine and higher carbon price increase the incentive for misreporting. A non-compliant company has two options, surrender insufficient allowances and hope that the fine is not applied (overt non-compliance), or underreport its emissions (covert non-compliance). When the fine increases, misreporting becomes more attractive, especially for companies with less stringent reporting requirements.

Figure 3 (right panel) 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 non-compliance to misreporting when the cost of overt non-compliance went up in 2008. Figure 3 shows that, although misreporting fell among larger installations (the expected effect of experience), it rose slightly for small installations, which face the loosest reporting requirements. This well-timed rise in misreporting among smaller installations is consistent with a substitution from overt to covert forms of non-compliance when faced with a higher fine.[9]

[7]Note that, due to the sensitivity of these survey data, the French government only permits publication of summary statistics for sufficiently large samples, which would likely prevent analysis of the small number of discrepant cases.

[8]Since most data sets do not conform to Benford's law exactly, the $\chi 2$ 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.

[9]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.

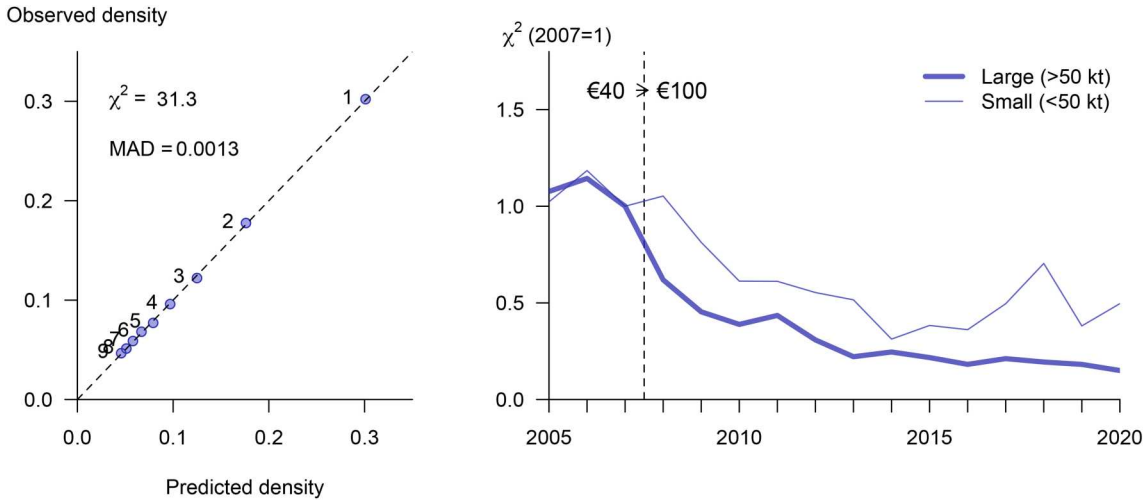

Figure 3. Misreporting as measured by deviations from Benford's law. Left panel: The observed densities of first digits are here plotted against the predicted densities (e.g. 30% of values are predicted to begin with the number one, whereas 29.5% do in the data set). Dots on the 45-degree line indicate that the predicted and observed densities are equal. Right panel: 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.

4.2. The probability of punishment

Another reason for the small change in the non-compliance rate in 2008, aside from mis-reporting, 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 programme, since the marginal benefit of non-compliance should in theory be the same for all installations.[10]

Figure 4 shows the association between the enforcement probabilities and non-compliance rates. Two features are especially noteworthy. First, there is a strong negative correlation between the probability of 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 failure to surrender sufficient allowances in a timely manner. For comparison, according to our estimates above, 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 period).

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.

[10]What matters for firms' decisions to comply is the *perceived* or *expected* probability of being punished. The standard model of 'rational expectations' in economics posits that firms will have enough sophistication to make unbiased predictions, so that their average expectations coincides with the average realized outcome in large samples. As a consequence, an unbiased estimate of what installation operators believe is the realized probability after the fact.

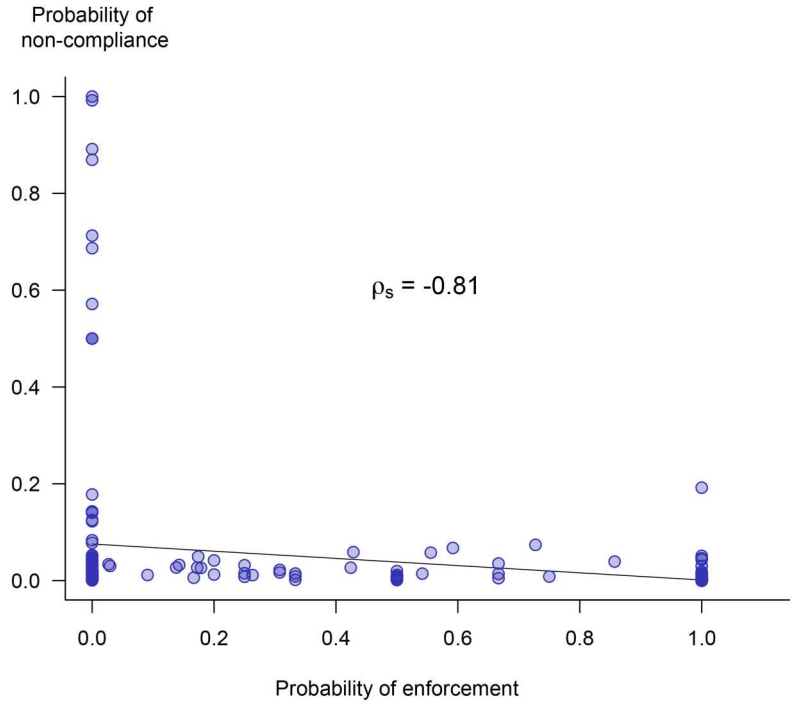

Figure 4. Association between enforcement and non-compliance. A higher probability of enforcement appears to be associated with a lower probability of non-compliance. The enforcement probabilities are estimated based on publicly available data on the number of enforcement actions. Enforcement was assumed wherever there was ambiguity. Also, unit-probability of enforcement was assumed whenever there was full compliance and no enforcement. Our estimates are therefore an upper bound on the enforcement probabilities.

However, if regulators did collect fines anywhere close to our estimated liabilities, 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. When regressing the fine, the probability of enforcement, and their interaction, on the probability of non-compliance, only 12% of the variation in the outcome is explained. When suppressing everything but the interaction term, as the canonical model would suggest, this drops to just 7%. To understand whether the remainder is merely noise or due to differences in policy, other compliance strategies, beyond the fine, need to be considered. Therefore, this section investigates how monitoring, enforcement, and discretionary power varies across countries and across time.

5.1. Monitoring

We observe variation in three types of monitoring strategies (Table 1 and Figure 11 in Appendix E). 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. Regulators that train and supervise third-party auditors tended to see higher compliance rates than those that did not, while those that performed their own site inspections had slightly lower compliance rates than those that did not. It is worth noting that these indicators of monitoring

Table 1. Mean levels of observed non-compliance for different monitoring and enforcement designs.

Monitoring and enforcement strategies	Yes	No	Difference in non-compliance
Regulators can directly appoint third-party verifiers	0.2%	0.2%	0.0%
Regulators can train and supervise third-party verifiers	0.0%	0.6%	−0.6%
Regulators can perform their own site inspections	0.1%	0.0%	0.1%
Regulators can suspend offenders	0.0%	0.2%	−0.2%
Regulators can apply naming & shaming	0.0%	0.4%	−0.4%
Naming & shaming: listed companies vs non-listed			4.0%
Naming & shaming: domestically owned vs foreign owned			−0.5%
Naming & shaming: both listed and domestically owned vs others			0.7%

Note: Negative numbers in the last column implies lower non-compliance. The first 5 lines are descriptive statistics, whereas the last 3 lines are outcomes of 3 regressions with a dummy for listed companies, domestically owned companies and their interaction respectively. These regressions include country fixed effects. The outcomes are not statistically significant.

powers are not highly correlated with each other. It is therefore unlikely that they are proxies of latent regulatory quality.[11]

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

Establishing the causal effect of monitoring regimes is challenging. The study by Duflo et al. (2013) assigned third-party auditors to Indian industrial plants, and 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. A similar effect may be present in the EU ETS.

Our setting is different from Duflo et al. (2013) in at least two important ways. First, monitoring regimes in the European carbon market cannot be assigned randomly, so there may be common drivers that affect both monitoring regimes and compliance. However, the fact that the three monitoring variables are so poorly correlated does at least suggest that their explanatory power (if any) is unlikely to derive 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, a lower probability of values just below zero and a higher probability just above zero is expected.[12]

The results are plotted in Figure 5, along with 95% confidence intervals. 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

[11]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.

[12]Following their procedure, we divide the excess emissions variable into bins and regress bin-dummies on an indicator for the monitoring regime. The effects are identified only through changes over time that occur within a country, by including country-fixed effects. The coefficient in each regression therefore 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 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.

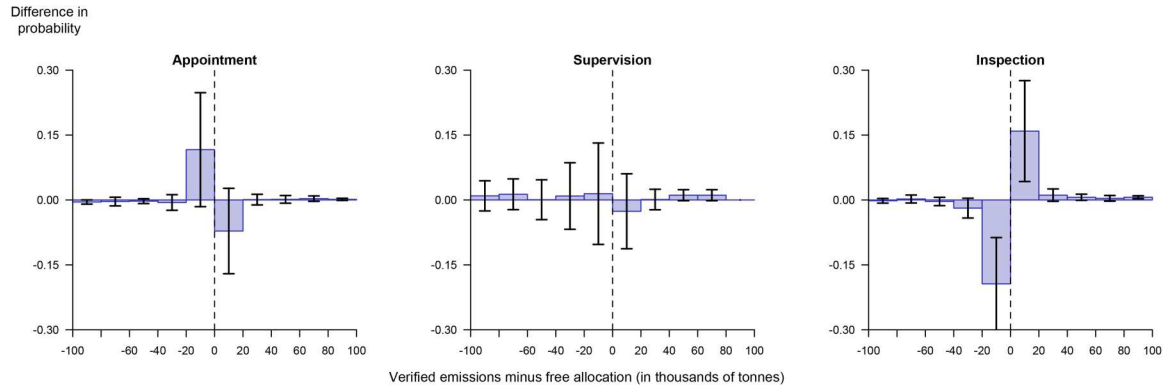

Figure 5. Verified emissions reports relative to free allowances. These three plots show how the probability of verified emissions exceeding free allowances by a certain amount differs by monitoring regime. The left panel compares regulators with and without the power to appoint auditors. The middle 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. This right panel shows that if regulators have the power to perform their own site inspections, the likelihood of companies reporting emissions just above their free allocation is 16% higher whereas the likelihood of reporting emissions just below their free allocation is 20% lower. Whiskers indicate the 95% confidence intervals using standard errors clustered at the country-year level.

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

Table 1 (line 4 onwards) shows how non-compliance rates vary with regulators' enforcement powers (see also Figure 12 in Appendix E). Regulators that 'name and shame' offenders, and those that suspend offenders, saw higher rates of compliance on average, as would be expected (Murphy & Stranlund, 2006). Regulators that levied greater fines for various reporting infractions (i.e. fines for things other than excess emissions) saw no difference in compliance rates (Figure 12 right panel). As before, these regulatory characteristics are only weakly correlated with each other.[13]

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 can be tested with our data, although one 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 there is 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.

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

[13]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.

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. Note, however, 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 expected way.

To test our first hypothesis, non-compliance is regressed 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 (see Table 1 line 6 and Figure 13 in Appendix E). If naming and shaming is particularly effective for listed firms, the coefficient on the interaction term should be negative, but our estimate is positive and noisy. Next, the listing status is replaced with a dummy indicating whether or not the firm is headquartered in the same country as the plant, to test if naming and shaming is particularly effective for these firms. Again, 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, compliance status is regressed on both, along with their interactions. Again, our coefficient estimate is zero.

Lastly, it has been found elsewhere that regulators sometimes treat non-compliers less favourably in other domains where the regulator has discretion, a practice that would create additional incentive to comply (Decker, 2003; Fenn & Veljanovski, 1988). This may be difficult for carbon market regulators, whose authority pertains only to one domain. Nevertheless, national regulators might be trying to apply this kind of pressure by awarding lower free allocations in the second trading period for installations that failed to comply in the first. Appendix E shows that this kind of mechanism could be operating in a small number of countries, though it is not present in most countries.

In sum, we find 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.

6. Insights from interviews with regulators

To complement the quantitative assessment, interviews with 11 national regulators and with the EU Transactions Log in Brussels were conducted.[14]

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, many firms did not know whether or not they were required to comply with the new policy. Regulators 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 period, 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

[14]The interviews were semi-structured and took place in 2014 and 2015. Three interviews were in person, the others took place over the phone. Duration between 30 min and 1h30. Interviews included questions about specific particularities and anomalies observable in the national registries.

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 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 appear less paradoxical with 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, in a context where fines associated with non-compliance could have threatened the viability of large industrial facilities upon which the local economy may crucially depend.

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. These problems were corrected whenever possible (see appendix C), 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 uncertainty to the interpretation of our results.

7. Conclusion

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 theorized a number of important enforcement challenges that arise specifically in the context of carbon markets (Stranlund & Dhanda, 1999), and has studied them in lab settings (Murphy & Stranlund, 2007), this study investigates how regulators are tackling these challenges in practice and presents the first comprehensive analysis of compliance

and enforcement behaviour in the EU ETS. While it is often impossible for an individual regulator to assess the effectiveness of particular enforcement strategies, additional insights are obtained by comparing the different ways that up to 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, despite relatively low levels of enforcement: 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. In prior literature, this pattern has been called Harrington's paradox.

This study investigates a number of potential resolutions to Harrington's paradox, but only a small fraction of the variation in compliance behaviour is explained. 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, no compelling evidence is found regarding 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 investigated monitoring and enforcement strategies, 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.

Our quantitative evidence put together with the qualitative evidence from selected interviews 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 revealed institutional challenges 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, when trying to link multiple emissions trading systems or when setting up border carbon adjustment mechanisms, such as the EU CBAM which requires importers to calculate and declare emissions embedded in certain imported energy-intensive goods.

More broadly, it seems prudent for any government that is implementing a carbon market to avoid creating new institutions from scratch whenever possible. The EU ETS was layered on top of the Integrated Pollution Prevention and Control Directive (now the Industrial Emissions Directive), integrating it with some pre-existing frameworks for pollution control. Similar foresight on integration with tax collection and financial regulation would likely have avoided many later compliance and enforcement issues. 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, the challenges involved in policing carbon markets may 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.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Frank Venmans acknowledges financial support from the Grantham Research Institute on Climate Change and the Environment, at the London School of Economics, the ESRC Centre for Climate Change Economics and Policy (CCCEP) (ref. ES/R009708/1) and the BIOADD NERC grant (ref. NE/X002292/1).

ORCID

Frank Venmans http://orcid.org/0000-0002-4264-6606

References

Anderson, B. E. (1996). The Ebb and flow of enforcing executive order 11246. *The American Economic Review*, *86*(2), 298–301.

Anderson, B., & Di Maria, C. (2011). Abatement and allocation in the pilot phase of the EU ETS. *Environmental and Resource Economics*, *48*, 83–103. doi:10.1007/s10640-010-9399-9

Anger, N., & Oberndorfer, U. (2008). Firm performance and employment in the EU emissions trading scheme: An empirical assessment for Germany. *Energy Policy*, *36*(1), 12–22. doi:10.1016/j.enpol.2007.09.007

Ashenfelter, O., & Smith, R. S. (1979). Compliance with the minimum wage Law. *Journal of Political Economy*, *87*(2), 333–350. doi:10.1086/260759

Badrinath, S. G., & Bolster, P. J. (1996). The role of market forces in EPA enforcement activity. *Journal of Regulatory Economics*, *10*(2), 165–181. doi:10.1007/BF00133530

Baker, J. B. (2003). The case for antitrust enforcement. *The Journal of Economic Perspectives*, *17*(4), 27–50. doi:10.1257/089533003772034880

Bayer, P., & 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. doi:10.1073/pnas.1918128117

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, *76*(2), 169–217. doi:10.1086/259394

Bel, G., & Joseph, S. (2015). Emission abatement: Untangling the impacts of the EU ETS and the economic crisis. *Energy Economics*, *49*, 531–539. doi:10.1016/j.eneco.2015.03.014

Betz, R. A., & 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*, *16*(4), 474–495. doi:10.1080/14693062.2015.1028319

Bhorat, H., Kanbur, R., & 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. doi:10.1111/rode.12007

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

Block, M. K., Nold, F. C., & Sidak, J. G. (1981). The deterrent effect of antitrust enforcement. *Journal of Political Economy*, *89*(3), 429–445. doi:10.1086/260979

Bontems, P., & Bourgeon, J.-M. (2005). Optimal environmental taxation and enforcement policy. *European Economic Review*, *49*(2), 409–435. doi:10.1016/S0014-2921(03)00036-9

Brouwers, R., Schoubben, F., & Hulle, C. V. (2017). The EU ETS and corporate environmental abatement. *Applied Economics Letters*, *24*(5), 334–337. doi:10.1080/13504851.2016.1189504

Business Green. (2012). *Bluenext agrees C32 m 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. doi:10.1257/pol.20180135

Calel, R., & 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. doi:10.1162/REST_a_00470

Campa, P. (2018). Press and leaks: Do newspapers reduce toxic emissions? *Journal of Environmental Economics and Management*, *91*, 184–202. doi:10.1016/j.jeem.2018.07.007

Cardiff, K., Kerrigan, M., Macys, G., Friel, C., Fe´su¨s, E., den Engelsen, S., Dentamaro, M., Du¨rrwanger, A., Go´mez-Valca´rcel, M. L., Gores, L., Kokot, J., Puisais-Jauvin, L., Soblet, F., & 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., & Gangadharan, L. (2006a). Emissions variability in tradable permit markets with imperfect enforcement and banking. *Journal of Economic Behavior & Organization*, *61*(2), 199–216. doi:10.1016/j.jebo.2005.02.007

Cason, T. N., & Gangadharan, L. (2006b). An experimental study of compliance and leverage in auditing and regulatory enforcement. *Economic Inquiry*, *44*(2), 352–366. doi:10.1093/ei/cbj019

Chan, H. S. R., Li, S., & Zhang, F. (2013). Firm competitiveness and the European union emissions trading scheme. *Energy Policy*, *63*, 1056–1064. doi:10.1016/j.enpol.2013.09.032

Coates, I. V., & C, J. (2007). The goals and promise of the Sarbanes-Oxley act. *The Journal of Economic Perspectives*, *21*(1), 91–116. doi:10.1257/jep.21.1.91

Colmer, J., Martin, R., Muûls, M., & Wagner, U. J. (2024). Does pricing carbon mitigate climate change? Firm-level evidence from the European union emissions trading scheme. *The Review of Economic Studies*, rdae055. doi:10.1093/restud/rdae055

Cordis, A. S., & 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. doi:10.1016/j.jpubeco.2014.03.010

Dechezleprêtre, A., Nachtigall, D., & 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. doi:10.1016/j.jeem.2022.102758

Decker, C. S. (2003). Corporate environmentalism and environmental statutory permitting. *Journal of Law and Economics*, *46*(1), 103–129. doi:10.1086/345586

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

Di Porto, E., Persico, N., & Sahuguet, N. (2013). Decentralized deterrence, with an application to labor tax auditing. *American Economic Journal: Microeconomics*, *5*(1), 35–62. doi:10.1257/mic.5.1.35

Duflo, E., Greenstone, M., Pande, R., & 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. doi:10.1093/qje/qjt024

Dumas, C. F., & 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. doi:10.1162/003465304323023895

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

Ehrlich, I. (1996). Crime, punishment, and the market for offenses. *Journal of Economic Perspectives*, *10*(1), 43–67. doi:10.1257/jep.10.1.43

Ellerman, A., & 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. doi:10.1007/s10640-008-9191-2

Ellerman, A. D., Convery, F. J., & de Perthuis, C. (2010). *Pricing carbon: The European union emissions trading scheme*. Cambridge University Press.

Ellerman, A. D., & Trotignon, R. (2009). Cross border trading and borrowing in the EU ETS. *The Energy Journal*, *30*(2_suppl), 53–78. doi:10.5547/ISSN0195-6574-EJ-Vol30-NoSI2-4

Engel, E. M. R. A., Galetovic, A., & Raddatz, C. E. (2001). A note on enforcement spending and VAT revenues. *Review of Economics and Statistics*, *83*(2), 384–387. doi:10.1162/rest.2001.83.2.384

Feinberg, R. M. (1980). Antitrust enforcement and subsequent price behavior. *The Review of Economics and Statistics*, *62*(4), 609–612. doi:10.2307/1924786

Fenn, P., & Veljanovski, C. G. (1988). A positive economic theory of regulatory enforcement. *The Economic Journal*, *98*(393), 1055–1070. doi:10.2307/2233719

Frunza, M.-C. (2013a). Aftermath of the VAT fraud on carbon emissions markets. *Journal of Financial Crime*, *20*(2), 222–236. doi:10.1108/13590791311322382

Frunza, M.-C. (2013b). *Fraud and carbon markets: The carbon connection*. Routledge.

Fung, A., & 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. doi:10.1007/s002679910009

Furlong, W. J. (1991). The deterrent effect of regulatory enforcement in the fishery. *Land Economics*, *67*(1), 116–129. doi:10.2307/3146490

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. doi:10.1016/S0929-1199(96)00009-0

Gerlagh, R., Heijmans, R. J., & Rosendahl, K. E. (2022). Shifting concerns for the EU ETS: Are carbon prices becoming too high? *Environmental Research Letters*, *17*(5), 1–6. doi:10.1088/1748-9326/ac63d6

Giles, C. (2022). *Next generation compliance, environmental regulation for the modern Era*. Oxford University Press.

Gray, W. B., & 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. doi:10.1006/jeem.1996.0034

Gray, W. B., & 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. doi:10.2307/2109575

Hahn, R. (1984). Market power and transferable property rights. *The Quarterly Journal of Economics*, *99*(4), 753–765. doi:10.2307/1883124

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. doi:10.1006/jeem.1995.1007

Hanlon, M., Hoopes, J. L., & 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. doi:10.2308/atax-50820

Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics*, *37*(1), 29–53. doi:10.1016/0047-2727(88)90003-5

Helland, E. (1998). The enforcement of pollution control laws: Inspections, violations, and self- reporting. *Review of Economics and Statistics*, *80*(1), 141–153. doi:10.1162/003465398557249

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

Hintermann, B. (2011). Market power, permit allocation and efficiency in emission permit markets. *Environmental and Resource Economics*, *49*(3), 327–349. doi:10.1007/s10640-010-9435-9

Hintermann, B. (2017). Market power in emission permit markets: Theory and evidence from the EU ETS. *Environmental and Resource Economics*, *66*, 89–112. doi:10.1007/s10640-015-9939-4

Hu¨bner, 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. doi:10.1016/S0047-2727(98)00101-7

Jaraite˙, J., Jong, T., Kažukauskas, A., Zaklan, A., & Zeitlberger, A. (2016). Matching EU ETS accounts to historical parent companies a technical note.

Kaplow, L., & Shavell, S. (1994). Optimal law enforcement with self-reporting of behavior. *Journal of Political Economy*, *102*(3), 583–606. doi:10.1086/261947

Keohane, N. O., Mansur, E. T., & 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., & Bojilova, D. (1998). Toxics release information: A policy tool for environmental protection. *Journal of Environmental Economics and Management*, *36*(3), 243–266. doi:10.1006/jeem.1998.1048
Kleit, A. N., Pierce, M. A., & 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. doi:10.1023/A:1008056615258
Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, *79*(3), 651–692. doi:10.3982/ECTA9113
Koch, N., Fuss, S., Grosjean, G., & Edenhofer, O. (2014). Causes of the EU ETS price drop: Recession, cdm, renewable policies or a bit of everything?-new evidence. *Energy Policy*, *73*, 676–685. doi:10.1016/j.enpol.2014.06.024
Konar, S., & 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. doi:10.1006/jeem.1996.0955
Kunreuther, H. C., McNulty, P. J., & Kang, Y. (2002). Third-Party inspection as an alternative to command and control regulation. *Risk Analysis*, *22*(2), 309–318. doi:10.1111/0272-4332.00029
Laplante, B., & 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. doi:10.1006/jeem.1996.0029
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. doi:10.1093/jleo/ewj014
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*, *49*, 203–222. doi:10.1007/s11149-016-9293-7
Lin, L. (2013). Enforcement of pollution levies in China. *Journal of Public Economics*, *98*, 32–43. doi:10.1016/j.jpubeco.2012.11.004
Lobo, G. J., & Zhou, J. (2006). Did conservatism in financial reporting increase after the Sarbanes-Oxley act? Initial evidence. *Accounting Horizons*, *20*(1), 57–73. doi:10.2308/acch.2006.20.1.57
Lo¨fgren, Å, Wråke, M., Hagberg, T., & 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., & Viscusi, W. K. (1990). Effectiveness of the EPA's regulatory enforcement: The case of industrial effluent standards. *The Journal of Law and Economics*, *33*(2), 331–360. doi:10.1086/467208
Marchi, S. d., & 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. doi:10.1007/s10797-006-6666-3
Miller, N. H. (2009). Strategic leniency and cartel enforcement. *American Economic Review*, *99*(3), 750–768. doi:10.1257/aer.99.3.750
Murphy, J. J., & 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. doi:10.1016/j.jebo.2005.03.013
Murphy, J. J., & 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. doi:10.1016/j.jeem.2006.08.003
Nigrini, M., & Wells, J. (2012). *Benford's Law: Applications for forensic accounting, auditing, and fraud detection*. Wiley Corporate F&A. Wiley.
Nyborg, K., & Telle, K. (2004). The role of warnings in regulation: Keeping control with less punishment. *Journal of Public Economics*, *88*(12), 2801–2816. doi:10.1016/j.jpubeco.2004.02.004
Oljaca, N., Keeler, A. G., & Dorfman, J. (1998). Penalty functions for environmental violations: Evidence from water quality enforcement. *Journal of Regulatory Economics*, *14*(3), 255–264. doi:10.1023/A:1008031307103
Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, *115*(2), 200–249. doi:10.1086/517935
Pahle, M., Quemin, S., Abrell, J., Millischer, L., Taschini, L., Baudry, M., Raude, M., Jo, A., Cludius, J., Zaklan, A., Cantillon, E., & Slechten, A. (2023). *Feedback from members of the scientific research community*. European Commission.
Potoski, M., & Prakash, A. (2005). Green clubs and voluntary governance: ISO 14001 and firms' regulatory compliance. *American Journal of Political Science*, *49*(2), 235–248. doi:10.1111/j.0092-5853.2005.00120.x
Reuters. (2014). Julius Baer asked for 4.7*mlnguaranteeinFrenchcarbonmarketprobe. Reuters.* RFI (2018). 36 jailed for huge French carbon tax fraud.
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. doi:10.1177/001979391006300409
Rubinstein, A. (1980). On an anomaly of the deterrent effect of punishment. *Economics Letters*, *6*(1), 89–94. doi:10.1016/0165-1765(80)90062-2
Sauer, R. D., & 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., & Scorse, J. (2012). Does being a "Top 10" worst polluter affect environmental releases? Evidence from the U.S. *Toxic Release Inventory. Working paper*.
Segerson, K., & Tietenberg, T. (1992). The structure of penalties in environmental enforcement: An economic analysis. *Journal of Environmental Economics and Management*, *23*(2), 179–200. doi:10.1016/0095-0696(92)90028-U

Shimshack, J., & Ward, M. (2005). Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*, *50*(3), 519–540.
Snyder, E. A. (1990). The effect of higher criminal penalties on antitrust enforcement. *The Journal of Law and Economics*, *33*(2), 439–462. doi:10.1086/467213
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. doi:10.1006/jeem.2001.1204
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. doi:10.1002/pam.20249
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. doi:10.1016/j.irle.2013.07.009
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. doi:10.1257/jep.12.3.69
Stoerk, T. (2016). Statistical corruption in Beijing's air quality data has likely ended in 2012. *Atmospheric Environment*, *127*, 365–371. doi:10.1016/j.atmosenv.2015.12.055
Stranlund, J. K., & Dhanda, K. K. (1999). Endogenous monitoring and enforcement of a transferable emissions permit system. *Journal of Environmental Economics and Management*, *38*(3), 267–282. doi:10.1006/jeem.1999.1092
Szabo, M. (2012). UK CO2 fraudsters trade Rolls Royces for jail time. *Reuters*.
Telle, K. (2009). The threat of regulatory environmental inspection: Impact on plant performance. *Journal of Regulatory Economics*, *35*(2), 154–178. doi:10.1007/s11149-008-9074-z
Telle, K. (2013). Monitoring and enforcement of environmental regulations: Lessons from a natural field experiment in Norway. *Journal of Public Economics*, *99*, 24–34. doi:10.1016/j.jpubeco.2013.01.001
Trotignon, R., & Delbosc, A. (2008). *Allowance trading patterns during the EU ETS trial period: What does the CITL Reveal? Climate Report 13*. CDC Climat.
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), 595–596.
Weil, D. (1996). If OSHA is so bad, why is compliance so good? *The RAND Journal of Economics*, *27*(3), 618–640. doi:10.2307/2555847
Wolfram, C. D. (1999). Measuring duopoly power in the British electricity spot market. *The American Economic Review*, *89*(4), 805–826. doi:10.1257/aer.89.4.805
Xu, Y. (2011). Improvements in the operation of SO2 scrubbers in China's coal power plants. *Environmental Science & Technology*, *45*(2), 380–385. doi:10.1021/es1025678
Zahran, S., Iverson, T., Weiler, S., & 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. doi:10.1007/s11166-014-9204-1
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.