

The design of environmental markets: What have we learned from experience with cap and trade?

Richard Schmalensee* and Robert N. Stavins**

Abstract: This article reviews the design of environmental markets for pollution control over the past 30 years, and identifies key market-design lessons for future applications. The focus is on a subset of the cap-and-trade systems that have been implemented, planned, or proposed around the world. Three criteria led us to the selection of systems for review. First, among the broader class of tradable permit systems, our focus is exclusively on cap-and-trade mechanisms, thereby excluding emission-reduction-credit or offset programmes. Second, among cap-and-trade mechanisms, we examine only those that target pollution abatement, and so we do not include applications to natural resource management, such as individual transferable quota systems used to regulate fisheries. Third, we focus on the most prominent applications—those that are particularly important environmentally, economically, or both.

Keywords: environmental markets, cap-and-trade system, air pollution, global climate change

JEL classification: Q58, Q28, Q53, Q54

I. Introduction

Some readers may not recall that just 30 years ago the notion of a government allocating tradable rights to emit pollution was novel, indeed controversial. Most environmental advocates were hostile to the concept of trading ‘rights to pollute’, and others were sceptical about the feasibility of such market-based approaches to environmental protection. Nearly all pollution regulations were of a conventional ‘command-and-control’ variety, setting uniform emission limits or specifying the pollution-control equipment to be used.

* Sloan School of Management, Massachusetts Institute of Technology, e-mail: rschmal@mit.edu

** John F. Kennedy School of Government, Harvard University, Resources for the Future, National Bureau of Economic Research, e-mail: robert_stavins@hks.harvard.edu

Schmalensee is the Howard W. Johnson Professor of Economics and Management, Emeritus at the Massachusetts Institute of Technology. Stavins is the Albert Pratt Professor of Business and Government at the Harvard Kennedy School, a University Fellow of Resources for the Future, and a Research Associate of the National Bureau of Economic Research. This paper draws, in part, on [Schmalensee and Stavins \(2015\)](#), and has benefitted from reviews by Cameron Hepburn and Dieter Helm. The authors are responsible for any remaining errors.

doi:10.1093/oxrep/grx040

© The Authors 2017. Published by Oxford University Press.

For permissions please e-mail: journals.permissions@oup.com

Today, although arguments are still made about the ethics of tradable rights to emit pollutants (Caney and Hepburn, 2011; Sandel, 2012), throughout nearly all of academia and across much of the policy community the cap-and-trade approach has become accepted as a legitimate policy, too. It is acknowledged that inflexible command-and-control regulations ignore the fact that compliance costs typically vary greatly among sources, depending on their age, technology characteristics, and operating conditions. As a result, aggregate abatement costs under command-and-control approaches are likely to be much greater than they need to be. Instead, by establishing a price on pollutant emissions, either directly through taxes or indirectly through a market of tradable rights (called permits or allowances), market-based approaches equate marginal abatement costs rather than emissions levels, and thereby can—in principle—achieve pollution-control targets at minimum cost, that is, cost effectively.

In the early decades of the modern environmental era, which began with the first Earth Day in 1970, most experience with market-based instruments—including, most prominently, what are now called ‘cap-and-trade’ mechanisms—was in the United States, starting with the Federal government’s attention to localized air pollution, and subsequently transboundary acid rain. More recently, with increased attention to the threat of global climate change, the locus of policy action with these instruments has shifted from national to sub-national policies in the United States, and, more broadly, from the United States to overseas, in particular, Europe.

In this article, we examine the design of cap-and-trade systems over the past 30 years, and identify lessons for future applications. Drawing on three criteria, we focus on a subset of the dozens of cap-and-trade systems that have been implemented, planned, or proposed around the world. First, among the broader class of tradable permit systems, we focus exclusively on cap-and-trade mechanisms, thereby excluding emission-reduction-credit or offset systems (Stavins, 2003). Second, among cap-and-trade mechanisms, we examine only those that target pollution abatement, and so applications to natural resource management, such as individual transferable quota (ITQ) systems used to regulate fisheries, are excluded (Stavins, 2011). Third, we focus on the most prominent applications—those that are particularly important environmentally, economically, or both.

More broadly, so-called ‘market-based environmental policy instruments’ include both quantity mechanisms, such as cap-and-trade and offset systems, and pollution charges, such as emission taxes and deposit-refund systems (Stavins, 2003). It is beyond our scope to reflect on the extensive literature on pollution taxes, or to discuss the largely theoretical comparisons between such instruments and cap-and-trade systems.

II. Experience with the market design of cap-and-trade systems

A total of seven cap-and-trade systems meet our criteria: the US Environmental Protection Agency (EPA)’s leaded gasoline phasedown in the 1980s; the sulphur dioxide allowance trading programme under the Clean Air Act Amendments of 1990; nitrogen oxides (NO_x) trading in the eastern United States; the Regional Clean Air Incentives Market in southern California; the Regional Greenhouse Gas Initiative

in the north-east United States; California's AB-32 cap-and-trade system; and the European Union Emissions Trading System. For each, we describe the system's design, performance, and key lessons for market design.

(i) Leaded gasoline phasedown

Concern arose in the 1970s in many parts of the world regarding the use of lead as an additive in gasoline, the purpose of which was to boost octane ratings and preserve engine life. Although it was later documented that lead oxide emissions were a serious human health threat, the original concern was that the lead in gasoline was fouling catalytic converters, which were required in new US cars starting in 1975 to reduce emissions of carbon monoxide (CO) and unburned hydrocarbons. Because of this, in the early 1980s, the US EPA began a phasedown of lead in gasoline to 10 per cent of its original level.

A trading programme was launched in 1982 to allow gasoline refiners greater flexibility in meeting the emission standards, with EPA authorizing inter-refinery trading of lead rights to lessen the financial burden on smaller refineries, which had significantly higher compliance costs. If refiners produced gasoline with a lower lead content than was required, they earned lead credits. Unlike a textbook cap-and-trade programme, there was no explicit allocation of permits, but to the degree that firms' production levels were correlated over time, the system implicitly awarded property rights on the basis of historical levels of gasoline production (Hahn, 1989). Under banking provisions of the programme, excess reductions could be banked for later use, providing an incentive for early reductions to help meet the lower limits that existed during the later years of the phasedown.

The trading programme resulted in leaded gasoline being removed from the market faster than had been anticipated, with more than 60 per cent of the lead added to gasoline associated with traded lead credits in each year of the programme (Hahn and Hester, 1989), until the system was terminated at the end of 1987, when the lead phasedown was completed. The programme was successful in meeting its environmental targets (Newell and Rogers, 2007). The high level of trading activity and the rate at which refiners reduced their production of leaded gasoline suggest that the programme was relatively cost-effective (Kerr and Maré, 1997; Nichols, 1997). EPA estimated savings from the lead trading programme of approximately 20 per cent compared with alternative approaches that did not provide for trade, a cost savings of about \$250m per year (US EPA, Office of Policy Analysis, 1985). This was partly due to the fact that the programme provided significant incentives for cost-saving technology diffusion (Kerr and Newell, 2003).

Three lessons for market design stand out. First, EPA's leaded gasoline phasedown served as a proof of concept, validating that a cap-and-trade system could be environmentally effective and economically cost effective. Second, as in other programmes to follow, banking played a very important role, comprising a significant share of 'gains from trade'. Third, rules should be clearly defined up front, without ambiguity. For example, requiring prior government approval of individual trades may increase uncertainty and transaction costs, and thereby discourage trading (Hahn and Hester, 1989).

(ii) Sulphur dioxide allowance trading

In the late 1980s, there was growing concern in the United States and other countries that acid precipitation—due mainly to emissions of sulphur dioxide (SO₂) from coal-fired power plants reacting in the atmosphere to form sulphuric acid—was damaging forests and aquatic ecosystems. In response, the US Clean Air Act Amendments of 1990 launched an SO₂ allowance trading programme, the objective of which was to reduce total annual emissions by half relative to the 1980 level. The programme commenced in 1995 (Ellerman *et al.*, 2000).

The government gave power plants allowances denominated in tons of SO₂ emissions, based initially on actual emissions during the period 1985 to 1987, with allocations demarcated by vintage, and the total number decreasing for successive years, thereby establishing a declining cap. If annual emissions at a regulated facility exceeded the allowances allocated to that facility, the owner could buy allowances or reduce emissions. If emissions at a regulated facility were reduced below its allowance allocation, the facility owner could sell the extra allowances or bank them for future use.

Although government auctioning of allowances would have generated revenue that could have been used—in principle—to reduce distortionary taxes, thereby reducing the programme's social cost (Goulder, 1995), this efficiency argument was not advanced at the time. Because cost-of-service regulation characterized the investor-owned electric utility industry in 1990, it was assumed that the value of free allowances would be passed on to consumers and would not generate windfall profits for providers. As important, the political value of being able to allocate free allowances to address differential economic impacts across regions, states, and Congressional districts was substantial (Joskow and Schmalensee, 1998). This is key because the equilibrium allocation of pollution permits, after trading has occurred, is independent of the initial allocation (Coase, 1960; Montgomery, 1972)—at least barring particularly problematic types of transaction costs (Stavins, 1995; Hahn and Stavins, 2012). So the initial allocation of allowances could be designed to ensure the greatest political support, without fear that this would jeopardize the system's environmental performance or economic cost.¹

The programme performed well along relevant dimensions. SO₂ emissions from electric power plants decreased 36 per cent between 1990 and 2004 (US EPA, 2011), despite the fact that electricity generation from coal-fired power plants increased 25 per cent over the same period (US Energy Information Administration, 2012). Emissions reductions occurred more quickly than expected, because utilities banked allowances for future use. With a \$2,000/ton statutory penalty for emissions exceeding allowance holdings (and continuous emissions monitoring), compliance was nearly 100 per cent.

Concern existed that emissions would end up being geographically concentrated, producing 'hot spots' of high SO₂ ambient concentrations. However, the programme did not generate significant hot spots (Ellerman *et al.*, 2000; Swift, 2004). It was subsequently suggested that the use of damage-based trading ratios, where allowances would be adjusted for the marginal environmental damages of emissions from each source, would have been welfare-improving (Muller and Mendelsohn, 2009), but the practical

¹ In contrast, modifying textbook tax systems to deal with political or distributional issues almost inevitably entails efficiency losses because different actors face different effective tax rates.

challenges of setting such ratios—particularly in a political environment—would have been very great.

Abatement costs were significantly less than they would have been with a command-and-control regulatory approach, although a share of the programme's cost-effectiveness was an unanticipated consequence of the deregulation of railroad rates in the late 1970s and early 1980s (Schmalensee and Stavins, 2013). That said, costs were at least 15 per cent below and perhaps as much as 90 per cent below costs of various counterfactual policies (Carlson *et al.*, 2000; Ellerman *et al.*, 2000; Keohane, 2003). There is evidence that the programme brought down abatement costs over time by providing incentives for innovation (Ellerman *et al.*, 2000; Popp, 2003; Bellas and Lange, 2011).

Five lessons for market design emerge from this experience. First, to provide some degree of certainty to regulated entities, facilitate their planning, and limit price volatility in early years, it is valuable to put final rules in place well before the beginning of the first compliance period. As with the lead trading programme, the absence of requirements for prior approval of trades reduced uncertainty for utilities and administrative costs for government, and contributed to low transaction costs (Rico, 1995). Second, the free allocation of allowances fostered political support (Joskow and Schmalensee, 1998), an important reminder for later programmes focused on climate change that the initial allocation can be used to address issues of distributional equity.

Third, intra-sector emissions leakage can be minimized by including all sources within the sector (above some capacity or emissions level). Fourth, a robust allowance market can be fostered through a cap that is significantly below business-as-usual (BAU) emissions, combined with unrestricted trading and banking. In fact, banking was again extremely important, accounting for more than half of the programme's cost savings (Carlson *et al.*, 2000; Ellerman *et al.*, 2000). Fifth, high levels of compliance can be encouraged with the combination of effective monitoring and significant penalties for non-compliance.

(iii) NO_x trading in the eastern United States

In 1999, under US EPA guidance, 11 states and the District of Columbia developed and implemented the NO_x Budget Program, a regional NO_x cap-and-trade system. In order to reduce the adverse health effects of ground-level ozone (smog formed by the interaction of NO_x and volatile organic compounds in the presence of sunlight), the programme aimed to cut summertime ground-level ozone by more than 50 per cent relative to 1990 levels (US EPA, 2004). Upwind states were given less generous allowance allocations, but trading across zones was permitted on a one-for-one basis, and the two zones made similar reductions from baseline emissions levels (Ozone Transport Commission, 2003).

In 1998, EPA required 21 eastern states to submit plans to reduce their emissions from more than 2,500 sources. This included a model rule, which enabled states to meet their emission reduction obligations by participating in an interstate cap-and-trade programme, known as the NO_x Budget Trading Program. The trading programme went into effect in 2003, replacing the NO_x Budget Program. As in the earlier programme, states were given allowances to allocate to in-state sources.

At the outset, the NO_x Budget Program market was characterized by uncertainty because some trading rules were not yet in place, resulting in high price volatility during the programme's first year. Overall, emissions declined from about 1.9m tons in 1990 to less than 500,000 tons by 2006, with 99 per cent compliance (Butler *et al.*, 2011; Deschenes *et al.*, 2012). Abatement cost savings were estimated at 40 to 47 per cent relative to conventional regulation without trading or banking (Farrell, 2000).

Four lessons stand out from the NO_x trading programme. First, in order to avoid unnecessary price volatility—which imposes unnecessary risk on affected sources and thus raises costs—all components of an emissions trading programme (including any planned future changes in allowance allocations) should be in place well before the programme takes effect. Second, a well-designed multi-state (sub-national) process with federal guidance can be effective in coordinating what are legally state-level goals. Third, the history of NO_x trading in the eastern United States provides a precedent and model for expanding the coverage of a cap-and-trade system over time to include additional jurisdictions, such as neighbouring states. Fourth, states can be given the flexibility to allocate fixed pools of allowances among in-state sources without necessarily compromising environmental goals.

(iv) Regional clean air incentives market

The legal entity responsible for controlling pollutant emissions in southern California—the South Coast Air Quality Management District—launched the Regional Clean Air Incentives Market (RECLAIM) in 1993 to reduce NO_x emissions and in 1994 to reduce SO₂ emissions from power plants, refineries, cement plants, and other industrial sources in the Los Angeles area. Free allocations of NO_x and SO₂ RECLAIM Trading Credits (RTCs) were based on historical peak production levels, and the initial allocations proved to be well above actual emissions until the year 2000. Hence, initially there was no scarcity of allowances, which is needed to generate a positive market price of allowances. The caps declined annually until 2003, when the market reached its overall goal of a 70 per cent emissions reduction (Ellerman *et al.*, 2003; Hansjurgens, 2011). With a single-year compliance period, banking was not allowed.

An interesting aspect of the trading programme was its zonal nature, whereby trades were not permitted from downwind to upwind sources. In this way, this geographically differentiated emissions trading programme represented one small step toward an ambient trading programme (Aldy and Stavins, 2012). Under a fully-developed ambient cap-and-trade system, as posited in theory by Montgomery (1972), a matrix would be developed to characterize the impact of a unit of emissions from every source on each relevant receptor location. If the impacts could then be aggregated into a single, total measure of impact, then a vector of transfer coefficients would yield trading ratios for emission allowances from various sources.

Emissions at RECLAIM facilities were some 20 per cent lower than at facilities that were regulated with parallel, command-and-control regulations, hotspots did not appear despite concerns, and substantial cost savings were achieved (Burtraw and Szambelan, 2010). In the programme's early years, allowance prices remained in the expected range of \$500–1,000 per ton of NO_x, but California's energy crisis in 2000–1 (due to problematic restructuring of California's electricity markets) closed off some

sources of electricity, and thereby caused electricity demand and production levels at some RECLAIM generating facilities to increase dramatically, causing emissions to exceed permit allocations, thereby bringing about a dramatic spike in allowance prices to more than \$60,000/ton (Fowle *et al.*, 2012). Part of the problem was the absence of a pool of banked allowances. The programme was temporarily suspended for the affected sources, prices returned to normal levels (below \$2,000/ton) by 2002, and all sources rejoined the programme by 2007.

Three lessons for market design stand out. First, because the RECLAIM system included two zones, with trades allowed in only one direction to account for prevailing wind directions, the programme's design demonstrated the basic feasibility of an ambient as opposed to a simple emissions-based cap-and-trade system. In other words, market design can accommodate the reality of a non-uniformly mixed pollutant and attendant concerns about 'hot spots'. Second, over-allocation of allowances means there is no scarcity created and therefore no functioning allowance market. Third, provisions for emissions banking (and other cost-containment elements) are crucial in order to allow for compliance at reasonable cost in years in which unanticipated circumstances lead to greater than expected emissions.

(v) Regional greenhouse gas initiative

Nine north-eastern US states participate in the Regional Greenhouse Gas Initiative (RGGI), the first cap-and-trade system in the United States to address carbon dioxide (CO₂) emissions. Limited to the power sector, the programme took effect in 2009 with a goal of limiting emissions to then current levels in the period from 2009 to 2014. The emissions cap was set to decrease by 2.5 per cent each year, beginning in 2015, until it reached an ultimate level 10 per cent below 2009 emissions in 2019. It was originally anticipated that meeting this goal would require a reduction approximately 35 per cent below business-as-usual emissions (13 per cent below 1990 emissions levels).

However, due the combined effects of the economic recession and drastic declines in natural gas prices relative to coal prices (a result of increased supplies of unconventional natural gas, brought about by adoption of the combination of horizontal drilling and hydraulic fracturing), the programme quickly ceased to be binding. In response, the RGGI states agreed in 2014 to reduce the cap by 45 per cent in 2015, and then by 2.5 per cent per year until a 10 per cent cut would be achieved by 2019.

The programme requires participating states to auction at least 25 per cent of their allowances (and to use the proceeds for energy efficiency, renewable energy, and related improvements), but in practice, states have auctioned virtually all allowances, because of the revenues that are thereby generated for government programmes. A price ceiling is created by a cost-containment reserve, from which additional allowances are released for sale when auction prices hit specified, escalating prices (thereby expanding the cap). A price floor is created by an auction reserve price. Any unsold allowances are permanently retired after 3 years, thereby providing an automatic mechanism for tightening the cap in the face of any chronic allowance surplus. This combination provides a price collar, making the programme a hybrid—to some degree—of a cap-and-trade system and a carbon tax.

Due to the non-binding nature of the cap during the programme's first compliance period and its barely binding cap since then, the direct impacts of the RGGI programme on power-sector CO₂ emissions have been trivial. But the programme's auctions have generated significant revenues for the participating states, exceeding \$1 billion. Some of this revenue has gone to financing government programmes that can reduce energy demand and hence CO₂ emissions (Hibbard *et al.*, 2011).

Monitoring costs for the programme have been low, because US power plants were already required to report their hourly CO₂ emissions to the Federal government under provisions for continuous emissions monitoring as part of the Federal SO₂ allowance trading programme. The penalty for non-compliance is that entities must submit three allowances for each allowance they are short. Due to the geographically limited scope of the system, combined with the presence of interconnected electricity markets, emissions leakage has been a significant concern (Burtraw *et al.*, 2006). One study found that if the programme were fully binding, power imports from Pennsylvania to New York State could result in emissions leakage approximating 50 per cent (Sue Wing and Kolodziej, 2008).

Four market-design lessons emerge. First, a cap-and-trade system that auctions its allowances can generate substantial revenue for government, whether or not the system has direct effects on emissions. Second, the leakage problem is potentially severe for any state or regional programme; this is particularly the case for a power-sector programme because of the inter-connected nature of electricity markets (Burtraw *et al.*, 2006). Third, a 'downstream' (CO₂ emissions) sectoral programme will inevitably be of limited scope, in comparison with an 'upstream' system that regulates the carbon content of fossil fuels as they enter the economy. Fourth and finally, a changing economy can render a cap non-binding or drive prices to excessive levels. Hence, there is an important role for price collars.

(vi) California's AB-32 cap-and-trade system

In 2006, California enacted Assembly Bill 32 (AB-32), the Global Warming Solutions Act, which is an ambitious mandate to cut the state's greenhouse gas (GHG) emissions to their 1990 level by the year 2020. This broad policy includes: energy efficiency standards for vehicles, buildings, and appliances; renewable portfolio standards for electricity producers; a low carbon fuel standard that requires refineries to reduce the carbon content of motor vehicle fuels; and a cap-and-trade system (California Environmental Protection Agency, 2014).

The cap-and-trade system began in 2013 with coverage of electricity sold in the state (whether produced in-state or imported) and large-scale manufacturing, and was expanded to include fuels in 2015, thereby covering 85 per cent of the state's emissions. The cap declines annually until 1990 emission levels are achieved in 2020. Most allowances were initially distributed via free allocation, with the programme transitioning to greater use of auctions. Auction proceeds are invested principally in other GHG reduction efforts. Banking is allowed, and regulated entities may use approved offsets.

A ceiling on allowance prices is created by an allowance price containment reserve, which releases allowances from a reserve when auction prices hit specified levels, which escalate over time. A price floor is created by an auction reservation price. This

combination produces an effective price collar. Trade-sensitive and energy-intensive manufacturers are protected from competitiveness effects through an output-based updating allocation system, whereby free allowances are conveyed in proportion to production levels in previous periods. California's system linked with a cap-and-trade system in Quebec in 2014, meaning that each system recognizes allowances from the other system for compliance purposes. In addition, joint allowance auctions are held on a quarterly basis.

Seven lessons have emerged from the California system's market design. First, California's experience demonstrates that an initial free allowance allocation that fosters political support can be transitioned over time to auctioning of allowances. But the California experience is also a reminder of the political pressures to use auction revenues for purposes other than the economist's favourite of reducing distortionary taxes. By state law, the funds 'are to be used to reduce GHG emissions and, to the extent feasible, achieve co-benefits such as job creation, air quality improvements, and public health benefits' ([California Legislative Analyst's Office, 2015](#)).

Second, California's system has demonstrated the feasibility and effectiveness of an economy-wide approach, compared with sectoral systems. Third, the risk of unanticipated allowance price changes and price volatility are reduced by employment of an effective price collar through the combination of an auction floor price and an allowance reserve. Fourth, to address concerns about competitiveness in sectors that are particularly energy-intensive and trade-sensitive, an effective mechanism is an output-based updating allocation, by which free allowances are granted to firms in specific sectors in proportion to their production levels in the previous time period. This makes the allocation of free allowances marginal, rather than infra-marginal, with allowances operating as an implicit production subsidy.

Fifth, California's expressed interest in linking its cap-and-trade system with those in other sub-national and national jurisdictions—and its implemented linkage with Quebec—highlights the importance of such linkage to reduce abatement costs, reduce price volatility, and restrain market power ([Ranson and Stavins, 2013](#)), although potentially it can also highlight some of the concerns that naturally arise with such linkage ([Bodansky et al., 2016](#)). Sixth, carbon pricing (through cap-and-trade or taxes) is necessary, but not sufficient, due to the limited sectoral scope of a carbon pricing regime, or—more broadly—due to the presence of other market failures. For example, the well-known principal-agent problem associated with energy-efficiency technology adoption decisions by landlords and renters in commercial and residential rental properties indicates that *specific* non-pricing policies *can be* complementary.

Seventh, the suite of policies within California's AB-32 also provides evidence that some 'complementary policies' conflict rather than complement. An example is provided by the state's Low Carbon Fuel Standard (LCFS), which requires that California refineries produce fuel with, on average, less than a stated amount of life-cycle carbon content. Since refineries and transportation fuels are already covered by the cap of the cap-and-trade system, the perverse result is that the refining sector therefore has less need for allowances, and so these are sold to other sectors, with the result that net CO₂ emissions are not reduced by the LCFS, but simply relocated ([Goulder and Stavins, 2011](#)), unless the allowance price floor is rendered binding. Furthermore, marginal costs are no longer equated, and so aggregate abatement costs are increased. In

addition, equilibrium allowance prices are suppressed as a result of the overall reduction in allowance demand, which causes concern with the ability of the cap-and-trade system to encourage technological change. Recent work (Borenstein *et al.*, 2016) suggests that complementary policies left so little scope for response to changes in allowance prices that those prices are very likely to be determined by an administrative price floor or ceiling.

(vii) European Union Emissions Trading System

The European Union Emission Trading System (EU ETS), a CO₂ cap-and-trade system, is the world's largest carbon pricing regime (European Commission, 2012). Adopted in 2003 with a pilot phase that became active in 2005, the EU ETS covers about half of EU CO₂ emissions in 30 countries, including all 27 member states plus Iceland, Liechtenstein, and Norway (Ellerman and Buchner, 2007). A total of 11,500 regulated emitters include large sources such as oil refineries, electricity generators, coke ovens, cement factories, ferrous metal production, glass and ceramics production, and pulp and paper production. Most sources in the transportation, commercial, and residential sectors are not included.

The EU ETS was implemented in phases: a pilot or learning phase from 2005 to 2007, a Kyoto phase from 2008 to 2012, and a series of subsequent phases. Penalties for violations increased from €40 per ton of CO₂ in the first phase to €100 in the second phase. The process for setting caps and allowances in member states was initially decentralized (Kruger *et al.*, 2007), with each member state responsible for proposing its own national carbon cap, subject to review by the European Commission. This created incentives for individual countries to be generous with their allowances to protect their economic competitiveness (Convery and Redmond, 2007). Not surprisingly, the result was an aggregate cap that exceeded business-as-usual emissions, which led to a dramatic fall in allowance prices (Convery and Redmond, 2007). The 'over-allocation' was concentrated in a few countries, particularly in Eastern Europe, and in the non-power sectors (Ellerman and Buchner, 2007, 2008).

The first and second phases of the EU ETS required member states to distribute almost all of the emissions allowances freely to regulated sources, but beginning in 2013, member states were allowed to auction larger shares of their allowances. The system's cap was tightened for Phase II (2008–12). Allowance prices increased to over €20/tCO₂ in 2008, then fell when economic recession brought decreased demand for allowances due to reduced output in energy-intensive sectors and lower electricity consumption. Prices were down to €10/tCO₂ by the fall of 2011, and have remained in the range of €5/tCO₂ to €10/tCO₂ since then.

The programme has been extended through its Phase III, 2013–20, with a centralized cap becoming increasingly stringent (20 per cent below 1990 emissions), a larger share of the allowances subject to auctioning, tighter limits on the use of offsets, and unlimited banking of allowances between Phases II and III.

Concern continues in the EU regarding relatively low allowances prices. These low prices largely reflect the slow pace of European economic recovery, as well as the fact that other EU climate policies, such as renewable portfolio standards and energy efficiency standards, reduce emissions and thereby reduce demand for allowances

(Fankhauser *et al.*, 2010). The perverse interactions present in the design of California's AB-32 suite of policies are also found in the EU portfolio.

Five market-design lessons stand out. First, to avoid unnecessary price fluctuations in the early years of a trading programme, sufficient time is needed between the adoption of the programme and the beginning of implementation to develop necessary rules. Second, good data are potentially important for sound allowance allocation decisions to avoid the type of over-allocation that occurred in the EU ETS's pre-Kyoto phase. Third, to avoid price volatility (in the form of a price collapse) at the end of a compliance period, it is necessary to allow for banking from one period to the next. The European system did not do this from its pilot phase (I) to its Kyoto phase (II), and the unsurprising result was that first-period (2005–7) allowance prices fell to zero as that period came to a close.

Fourth, more broadly, as with the California system, the EU ETS illustrates the perverse outcomes that are fostered when so-called 'complementary policies' are put in place under the cap of a cap-and-trade system. In the European context, these are largely renewable portfolio standards, other renewable policies, and some energy efficiency initiatives. Unless those policies address sources outside of the cap or other market failures, the result is to relocate emissions, drive up aggregate abatement costs, and suppress allowance prices.

Fifth and finally, the EU attempts to deal with competitiveness concerns through free (inframarginal) allowance allocations. The result is a wealth transfer to companies that receive free allowances, but without an effect on competitiveness or emissions leakage. The free allocation of allowances benefits the recipient firms, because the allowances are worth thousands of dollars on the market. But because the free allocation is inframarginal, it has no effect on competitiveness. A firm's marginal cost of production is not affected.

(viii) Other cap-and-trade systems

Additional cap-and-trade systems are in various stages of development, including trading of rights for ozone depleting substances (ODS) in several countries during the ODS phasedown from 1991 to 2000 under the 1987 Montreal Protocol (Klaassen, 1999; Stavins, 2003; US EPA, 2014), as well as CO₂ cap-and-trade systems in New Zealand, Japan, South Korea, Kazakhstan, and the provinces of Quebec and Ontario in Canada (Ranson and Stavins, 2013; Sopher and Mansell, 2014; Kossoy *et al.*, 2014). Also, an international CO₂ cap-and-trade system has operated since 2008 under Article 17 of the Kyoto Protocol. However, because the trading agents are nations, rather than firms, there has been little activity, an outcome that was anticipated (Hahn and Stavins, 1999).

Cap-and-trade systems have been planned or proposed in other countries for the purpose of reducing CO₂ emissions. Most important, in 2015, China announced that it would launch a national carbon market in 2017, which is intended to become the world's largest environmental market. Most design elements have not been revealed, including how the national system will incorporate the existing pilot systems (Jing, 2015). Cap-and-trade systems have also been proposed in other countries at levels of governance ranging from sub-municipal to national, including in: Brazil, as well as in São Paulo and Rio de Janeiro; Manitoba, Canada; Chile; Costa Rica; Japan; Mexico;

Taiwan; Thailand; Turkey; Ukraine; the United States; and Vietnam (Kossoy *et al.*, 2014; OECD and World Bank Group, 2015).

III. Conclusions

While there has been a significant amount of experience over the past 30 years with the use of cap-and-trade instruments for environmental protection in the United States and Europe, market-based instruments have not replaced nor come close to replacing conventional approaches. When and where cap-and-trade has been used, often with considerable success, applications have usually not been designed or implemented perfectly. Hence, it is important to ask what both positive and negative experiences suggest about good market design.

First, in terms of the basics, cap-and-trade has long since proved to be environmentally effective and economically cost-effective (lead phasedown, SO₂ allowance trading). Economy-wide systems have been shown to be feasible (AB-32), although downstream, sectoral programmes have been commonly employed (RGGI, EU ETS). Transaction costs have turned out to be low to trivial, particularly when compliance entities have been homogeneous (lead phasedown, SO₂ allowance trading). In the context of climate policy, CO₂ emissions trading programmes have inevitably been limited in scope of coverage, in contrast with textbook, upstream trading of rights associated with the carbon content of fossil fuels (World Bank, 2016).

It is clear from basic economic theory and is now validated by experience that a robust market requires a cap that is significantly below BAU emissions (SO₂ allowance trading, RECLAIM). Likewise, high levels of compliance require monitoring, reporting, and verification combined with significant penalties for non-compliance (SO₂ allowance trading). Also, it has been shown to be important for final rules to be established well before commencement of a system's first compliance period to avoid unnecessary price volatility (SO₂ allowance trading, NO_x Budget, EU ETS).

Turning to specific elements of design, experience argues for systems to be designed to allow for a broad set of compliance alternatives, in terms of both timing and technological options. For example, allowing flexible timing and intertemporal trading of allowances—that is, banking for future use—played a very important role in the SO₂ allowance trading programme's performance, much as it did in the US lead rights trading programme a decade earlier. One of the most significant benefits of using market-based instruments may simply be that technology and uniform performance standards are thereby avoided. Less flexible systems would not have led to the technological change that appears to have been induced by market-based instruments (Bohi and Burtraw, 1992; Ellerman and Montero, 1998; Keohane, 2003; Schmalensee and Stavins, 2013), or the induced process innovations that have resulted (Doucet and Strauss, 1994).

Provisions for banking of allowances have proven to be very important. Such intertemporal trading has represented a large share of the realized gains from trade (lead phasedown, SO₂ allowance trading). In contrast, the absence of banking provisions can lead to price spikes (RECLAIM) and price collapses (EU ETS). More broadly, a changing economy can render a cap non-binding (RGGI, EU ETS) or drive prices to excessive levels (RECLAIM). Hence, there is a distinct role in cap-and-trade systems for price

collars, which reduce the risk of unanticipated allowance price changes and price volatility by combining an auction price floor with an allowance reserve (RGGI, AB-32).² On the other hand, excessive constraints on offset use can lead to a thin market that fails to be effective for cost-containment purposes (RGGI, AB-32).

Simplicity is important, and transparent formulae—including for allowance allocation—are difficult to contest or manipulate. Such rules should be clearly defined up front, without ambiguity. By avoiding any requirements for prior government approval of individual trades, uncertainty and transaction costs are decreased (lead phasedown, SO₂ allowance trading).

The allocation of allowances is inevitably a major political issue, because of the large distributional impacts that can be involved. A striking and important experience has been that free allowance allocation has proven to foster political support, although it means that the opportunity is forgone to cut the programme's overall social cost by refunding revenues from auctioning allowances through cuts of distortionary taxes (SO₂ allowance trading, AB-32). However, empirical experience has revealed that political pressures exist to use auction revenue not for the economist's favourite of cutting distortionary taxes, but to fund new or existing government programmes or relieve deficits (AB-32, RGGI). Indeed, cap-and-trade allowance auctions can and have generated very significant revenue for governments (RGGI, AB-32).

Another prominent political concern when cap-and-trade systems have been designed has been the possibility of emissions and economic leakage and related competitiveness impacts. This issue is not specific to cap-and-trade or to market-based instruments. Virtually any meaningful environmental policy will increase production costs and thereby could raise this concern, but attention given to this issue has been greatest when market-based instruments are being considered. In practice, leakage from cap-and-trade systems can range from non-existent (lead phasedown) to potentially quite serious (RGGI). It is most likely to be significant for programmes of limited geographic scope, particularly in the power sector because of interconnected electricity markets (RGGI, AB32). Attempts to reduce leakage and competitiveness threats through free allocation of allowances *per se* do not address the problem (EU ETS), but an output-based updating allocation—in principle—can do so (AB-32).

No government policy exists in a vacuum, and this includes cap-and-trade mechanisms. Carbon pricing (through cap-and-trade or taxes) may be said to be necessary to address climate change, but it is surely not sufficient, due to the limited sectoral scope of some carbon pricing regimes, and—more broadly—due to the presence of other market failures that inhibit the perfect functioning of markets, as in the case of principal-agent problems associated with energy-efficiency decisions in rental properties. Hence, there can be an appropriate role for complementary policies. But actual suites of so-called 'complementary policies' have frequently conflicted rather than complemented by addressing emissions under the cap, thereby relocating rather than reducing emissions, driving up abatement costs, and suppressing allowance prices. Sadly, this perverse

² A laboratory experiment by [Holt and Shobe \(2016\)](#), which compared price collars, as used in RGGI and AB-32, with quantity collars (modelled after the EU ETS Market Stability Reserve), found that the price collar was superior to the quantity collar in terms of reducing allowance price volatility and increasing efficiency.

situation has characterized two of the most prominent applications of cap-and-trade in the climate policy context (AB-32, EU ETS).

Taken together, these many lessons from experience suggest that careful design of cap-and-trade merits serious attention when regions, nations, or sub-national jurisdictions seek to develop policies to reduce GHG emissions. But political responses to possible market-based approaches to climate policy in most countries will likely be—at least in part—a function of issues and structural factors that transcend the scope of environmental and climate policy. Because a meaningful climate policy will have significant impacts on economic activity in a wide variety of sectors and in every region of a country, it is not surprising that proposals for such policies have sometimes brought forth significant opposition.

At the same time, political revealed preference has emerged around the world for employing cap-and-trade systems to address GHG emissions. This includes the regional system in Europe, national systems in New Zealand and South Korea, and sub-national systems in Canada, China, Japan, and the United States. Additional cap-and-trade systems are in various stages of planning (or at least proposing) in Brazil, Canada, Chile, China, Costa Rica, Japan, Mexico, Taiwan, Thailand, Turkey, Ukraine, the United States, and Viet Nam. In international climate negotiations, leading up to the 2015 Paris Agreement under the United Nations Framework Convention on Climate Change, many parties endorsed key roles for regional, national, and sub-national carbon markets, and broad recognition emerged of the importance of international linkage among these systems.

References

- Aldy, J. E., and Stavins, R. N. (2012), 'The Promise and Problems of Pricing Carbon: Theory and Experience', *Journal of Environment and Development*, **21**(2), 152–80.
- Bellas, A. S., and Lange, I. (2011), 'Evidence of Innovation and Diffusion under Tradable Permit Programs', *International Review of Environmental and Resource Economics*, **5**(1), 1–22.
- Bodansky, D., Hoedl, S., Metcalf, G., and Stavins, R. N. (2016), 'Facilitating Linkage of Climate Policies through the Paris Outcome', *Climate Policy*, **16**(8), 956–72.
- Bohi, D. R., and Burtraw, D. (1992), 'Utility Investment Behavior and the Emission Trading Market', *Resources and Energy*, **14**(1–2), 129–53.
- Borenstein, S., Bushnell, J., Wolak, F. A., and Zaragoza-Watkins, M. (2016), 'Expecting the Unexpected: Emissions Uncertainty and Environmental Market Design', Energy Institute at Haas, Working Paper WP-274, University of California, Berkeley, August.
- Butler, T. J., Vermeylen, F. M., Rury, M., Likens, G. E., Lee, B., Bowker, G. E., and McCluney, L. (2011), 'Response of Ozone and Nitrate to Stationary Source Nox Emission Reductions in the Eastern USA', *Atmospheric Environment*, **45**(5), 1084–94.
- Burtraw, D., and Szambelan, S. J. (2010), 'US Emissions Trading Markets for SO₂ and NO_x', in B. Hansjürgens (ed.), *Permit Trading in Different Applications*, New York, Routledge.
- Kahn, D., and Palmer, K. (2006), 'CO₂ Allowance Allocation in the Regional Greenhouse Gas Initiative and the Effect on Electricity Investors', *The Electricity Journal*, **19**(2), 79–90.
- California Environmental Protection Agency (2014), *Assembly Bill 32 Overview*, Air Resources Board.
- California Legislative Analyst's Office (2015), *Governor's May Revision: 2015–16 Cap-and-Trade Expenditure Plan*, Sacramento, 17 May.
- Caney, S., and Hepburn, C. (2011), 'Carbon Trading: Unethical, Unjust and Ineffective?', *Royal Institute of Philosophy Supplement*, **69**, 201–34.

- Carlson, C., Burtraw, D., Cropper, M., and Palmer, K. (2000), 'SO₂ Control by Electric Utilities: What Are the Gains from Trade?', *Journal of Political Economy*, **108**(6), 1292–326.
- Coase, R. H. (1960), 'The Problem of Social Cost', *Journal of Law and Economics*, **3**, 1–44.
- Convery, F., and Redmond, L. (2007), 'Market and Price Developments in the European Union Emissions Trading Scheme', *Review of Environmental Economics and Policy*, **1**, 66–87.
- Deschenes, O., Greenstone, M., and Shapiro, J. S. (2012), 'Defensive Investments and the Demand for Air Quality: Evidence from the NO_x Budget Program and Ozone Reductions', Working Paper w18267, Cambridge, MA, National Bureau of Economic Research.
- Doucet, J. A., and Strauss, T. (1994), 'On the Bundling of Coal and Sulphur Dioxide Emissions Allowances', *Energy Policy*, **22**(9), 764–70.
- Ellerman, A. D., and Buchner, B. K. (2007), 'The European Union Emissions Trading Scheme: Origins, Allocation, and Early Results', *Review of Environmental Economics and Policy*, **1**(1), 66–87.
- (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), 267–87.
- Montero, J. P. (1998), 'The Declining Trend in Sulfur Dioxide Emissions: Implications for Allowance Prices', *Journal of Environmental Economics and Management*, **36**(1), 26–45.
- Joskow, P. L., and Harrison, D. (2003), 'Emissions Trading in the US: Experience, Lessons and Considerations for Greenhouse Gases', prepared for the Pew Center on Global Climate Change.
- Schmalensee, R., Montero, J. P., and Bailey, E. M. (2000), *Markets for Clean Air: The US Acid Rain Program*, Cambridge, Cambridge University Press.
- European Commission (2012), 'Emissions Trading System', http://ec.europa.eu/clima/policies/ets/index_en.htm.
- Fankhauser, S., Hepburn, C., and Park, J. (2010), 'Combining Multiple Climate Policy Instruments: How Not To Do It', *Climate Change Economics*, **1**(3), 209–25.
- Farrell, A. (2000), 'The NO_x Budget: A Look at the First Year', *The Electricity Journal*, **13**(2), 83–93.
- Fowle, M., Holland, S. P., and Mansur, E. T. (2012), 'What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NO_x Trading Program', *American Economic Review*, **102**(2), 965–93.
- Goulder, L. H. (1995), 'Environmental Taxation and the "Double Dividend": A Reader's Guide', *International Tax and Public Finance*, **2**(2), 157–83.
- Stavins, R. N. (2011), 'Challenges from State–Federal Interactions in US Climate Change Policy', *American Economic Review Papers and Proceedings*, **101**(3), 253–7.
- Hahn, R. W. (1989), 'Economic Prescriptions for Environmental Problems: How the Patient Followed the Doctor's Orders', *Journal of Economic Perspectives*, **3**(2), 95–114.
- Hester, G. L. (1989), 'Marketable Permits: Lessons for Theory and Practice', *Ecology Law Quarterly*, **16**, 361–406.
- Stavins, R. N. (1999), *What Has the Kyoto Protocol Wrought? The Real Architecture of International Tradeable Permit Markets*, Washington, DC, The AEI Press.
- (2012), 'The Effect of Allowance Allocations on Cap-and-Trade System Performance', *The Journal of Law and Economics*, **54**(4), S267–S294.
- Hansjürgens, B. (2011), 'Markets for SO₂ and NO_x—What Can We Learn for Carbon Trading?', *Wiley Interdisciplinary Reviews: Climate Change*, **24**, 635–46.
- Hibbard, P. J., Tierney, S. F., Okie, A. M., and Darling, P. A. (2011), *The Economic Impacts of the Regional Greenhouse Gas Initiative on Ten Northeast and Mid-Atlantic States*, Boston, MA, Analysis Group.
- Holt, C. A., and Shobe, W. M. (2016), 'Price and Quantity Collars for Stabilizing Emission Allowance Prices: Laboratory Experiments on the EU ETS Market Stability Reserve', *Journal of Environmental Economics and Management*, **76**, 32–50.
- Jing, L. (2015), 'China Upbeat on Prospects for National Carbon Market', *South China Morning Post*, 25 June.
- Joskow, P. L., and Schmalensee, R. (1998), 'The Political Economy of Market-based Environmental Policy: The US Acid Rain Program', *Journal of Law and Economics*, **41**(1), 37–83.

- Keohane, N. O. (2003), 'What Did the Market Buy? Cost Savings under the US Tradeable Permits Program for Sulfur Dioxide', Working Paper YCELP-01-11-2003, Yale Center for Environmental Law and Policy.
- Kerr, S., and Maré, D. (1997), 'Efficient Regulation through Tradeable Permit Markets: The United States Lead Phasedown', Department of Agricultural and Resource Economics, University of Maryland, College Park, Working Paper 96-06 (January).
- Newell, R. (2003), 'Policy-induced Technology Adoption: Evidence from the US Lead Phasedown', *Journal of Industrial Economics*, **51**(3), 317–43.
- Klaassen, G. (1999), 'Emissions Trading in the European Union: Practice and Prospects', in S. Sorrell and J. Skea (eds), *Pollution for Sale: Emissions Trading and Joint Implementation*, 83–100, Cheltenham, Edward Elgar.
- Kossoy, A., Opperman, K., Platanova-Oquab, A., Suphachalasai, S., Höhne, N., Klein, N., Gilbert, A., and Wong, L. (2014), *State and Trends of Carbon Pricing 2014*. Washington, DC, World Bank.
- Kruger, J., Oates, W. E., and Pizer, W. A. (2007), 'Decentralization in the EU Emissions Trading Scheme and Lessons for Global Policy', *Review of Environmental Economics and Policy*, **1**, 112–33.
- Montgomery, D. W. (1972), 'Markets in Licenses and Efficient Pollution Control Programs', *Journal of Economic Theory*, **5**(3), 395–418.
- Muller, N. Z., and Mendelsohn, R. (2009), 'Efficient Pollution Regulation: Getting the Prices Right', *American Economic Review*, **99**(5), 1714–39.
- Newell, R. G., and Rogers, K. (2007), 'The Market-based Lead Phasedown', in J. Freeman and C. Kolstad (eds), *Moving to Markets in Environmental Regulation: Lessons from Twenty Years of Experience*, Oxford, Oxford University Press, 173–93.
- Nichols, A. L. (1997), 'Lead in Gasoline', in R. D. Morgenstern (ed.), *Economic Analyses at EPA: Assessing Regulatory Impact*, Washington, DC, Resources for the Future, 49–86.
- OECD and World Bank Group (2015), *The FASTER Principles for Successful Carbon Pricing: An Approach Based on Initial Experience*, Washington, DC, Organization for Economic Cooperation and Development and World Bank Group.
- Ozone Transport Commission (2003), *NO_x Budget Program: 1999–2002 Progress Report*, <http://www2.epa.gov/sites/production/files/2015-08/documents/otcreport.pdf>
- Popp, D. (2003), 'Pollution Control Innovations and the Clean Air Act of 1990', *Journal of Policy Analysis and Management*, **22**(4), 641–60.
- Ranson, M., and Stavins, R. N. (2013), 'Post-Durban Climate Policy Architecture Based on Linkage of Cap-and-trade Systems', *Chicago Journal of International Law*, **13**, 403–38.
- Rico, R. (1995), 'The US Allowance Trading System for Sulfur Dioxide: An Update of Market Experience', *Environmental and Resource Economics*, **5**(2), 115–29.
- Sandel, M. J. (2012), *What Money Can't Buy: The Moral Limits of Markets*, New York, Farrar, Straus and Giroux.
- Schmalensee, R., and Stavins, R. N. (2013), 'The SO₂ Allowance Trading System: The Ironic History of a Grand Policy Experiment', *Journal of Economic Perspectives*, **27**(1), 103–22.
- — (2015), 'Lessons Learned from Three Decades of Experience with Cap-and-trade', Working Paper 21742, Cambridge, National Bureau of Economic Research.
- Sopher, P., and Mansell, A. (2014), *Japan. The World's Carbon Markets: A Case Study Guide to Emissions Trading*, International Emissions Trading Association.
- Stavins, R. N. (1995), 'Transaction Costs and Tradeable Permits', *Journal of Environmental Economics and Management*, **29**(2), 133–48.
- (2003), 'Experience with Market-based Environmental Policy Instruments', *Handbook of Environmental Economics*, **1**, 355–435.
- (2011), 'The Problem of the Commons: Still Unsettled After 100 Years', *American Economic Review*, **101**(1), 81–108.
- Sue Wing, I., and Kolodziej, M. (2008), 'The Regional Greenhouse Gas Initiative: Emission Leakage and the Effectiveness of Interstate Border Adjustments', Regulatory Policy Program Working Paper RPP-2008-03, Cambridge, MA, Mossavar-Rahmani Center for Business and Government, John F. Kennedy School of Government, Harvard University.

- Swift, B. (2004), 'Emissions Trading and Hot Spots: A Review of the Major Programs', *Environment Reporter*, 35(19), 1–16.
- US Energy Information Administration (2012), 'Electricity Net Generation: Total (All Sectors), 1949–2011', Table 8.2a, *Annual Energy Review 2011*.
- US EPA (1985), *Costs and Benefits of Reducing Lead in Gasoline, Final Regulatory Impact Analysis*, US Environmental Protection Agency, Office of Policy Analysis, Washington, DC.
- (2004), *Evaluating Ozone Control Programs in the Eastern United States: Focus on the NO_x Budget Trading Program*, US Environmental Protection Agency, <http://www.epa.gov/airtrends/2005/ozon-enbp.pdf#page=1>
- (2011), *National Emissions Inventory (NEI) Air Pollutant Emissions Trends Data: 1970–2011*, US Environmental Protection Agency, <http://www3.epa.gov/ttnchie1/trends/>
- (2014), *Chlorofluorocarbon Taxes and Allowance Trading*, US Environmental Protection Agency, <http://yosemite1.epa.gov/EE/epa/eed.nsf/dcee735e22c76aef85257662005f4116/86a2750ed83fe1d88525774200597f3a!OpenDocument>
- World Bank (2016), *Emissions Trading in Practice: Handbook on Design and Implementation*, Partnership for Market Readiness and International Carbon Action Partnership, Washington, DC, World Bank.