

The Role of Emission Trading in Domestic Climate Policy

Michael Hanemann*

This paper focuses on two specific issues in the design of a domestic cap and trade program for GHGs – whether the cap should be located upstream or downstream, and whether trading alone will suffice to achieve the desired reduction in GHGs or will need to be supplemented with additional regulatory measures. The paper argues for a downstream cap accompanied by measures such as a renewable portfolio standard, efficiency standards for vehicles, appliances and buildings, and a low carbon fuel standard. For this argument, it is necessary to address both the theory and the empirical evidence of emission trading. After reviewing the theory, the paper examines the actual experience in the U.S. with emission trading for SO₂, to see whether the assumptions used in the theory actually applied in practice. What actually happened deviated in several important respects from what was supposed to happen according to the conventional theorizing. The design of a cap and trade program for GHG is then discussed, first considering the similarities between the past regulation of air pollutants and the challenge posed by GHGs, and then making the case for a downstream cap and complementary policies.

1. INTRODUCTION

With the arrival of a new US administration, the debate on climate policy in Washington has moved into high gear. There is widespread agreement that we need to set a price on carbon dioxide (CO₂) and perhaps other greenhouse gasses (GHGs); the question is how best to do this. Among economists, the two obvious candidates are an emissions tax or a cap-and-trade program. In the

The Energy Journal, Volume 30 (Special Issue 2). Climate Change Policies After 2012. Copyright © 2009 by the IAEE. All rights reserved.

* University of California Berkeley, Department of Agricultural & Resource Economics, 207 Giannini Hall, Berkeley CA 94720; hanemann@are.berkeley.edu.

I am grateful to Dallas Burtraw, Frank Convery, Denny Ellerman, Meredith Fowle, Xavier Labandeira, and Juan-Pablo Montero for very helpful comments on an earlier draft. I alone am responsible for any errors.

policy arena, cap and trade has emerged as the preferred approach in the U.S. While cap and trade and a tax both generate price signals, cap and trade has the advantage with the regard to the political economy of distributional concerns and the compensation of impacted sectors.

This paper focuses on two specific issues in the design of a domestic cap and trade program for GHGs – whether the cap should be located upstream or downstream, and whether trading alone will suffice to achieve the desired reduction in GHGs or will need to be supplemented with some additional regulatory measures. The standard view – for example, Aldy and Pizer (2009) – favors an upstream cap unaccompanied by complementary measures other than an R&D policy. Instead, I argue for a downstream cap accompanied, in addition to an R&D policy, by measures such as a renewable portfolio standard, efficiency standards for vehicles, appliances and buildings, and a low carbon fuel standard.¹ The standard view is that “there is no basis in economic theory nor any empirical evidence whatsoever” for what I am proposing.² It is necessary, therefore, to address both the theory and the empirical evidence before making my case.

Section 2 reviews the theoretical analysis supporting the standard argument for emission trading. Section 3 then examines in some detail the actual experience in the U.S. with emission trading for SO₂, followed by a very brief discussion of trading for lead in gasoline, to see whether the assumptions used in the theory actually applied in practice, and whether the predictions generated by the theory were fulfilled. I will argue that what actually happened deviated in several important respects from what was supposed to happen according to the conventional theorizing, and therefore some aspects of the theory may need to be modified. Section 4 then goes on to discuss the design of a cap and trade program for GHG. Section 4.1 reviews the similarities and differences between the past regulation of air pollutants and the challenge posed by GHGs. Sections 4.2 and 4.3 make the case for a downstream cap and complementary policies. Section 5 closes with a few conclusions and caveats.

2. THEORY OF EMISSION TRADING

While the concept of emissions trading as a strategy for controlling pollution was first proposed by Crocker (1966) and Dales (1968), the first formal theoretical treatment was given by Montgomery (1972). Montgomery’s framing has remained the standard one and is commonly cited in the policy literature.³ There are one or more polluting firms which face a choice between abating their emissions or participating in an emissions trading market. The decision is based on the criterion of profit maximization or, equivalently, cost minimization. Operating

1. As explained in Farrell and Hanemann (2009) and Hanemann (2008), those are all measures being adopted in California.

2. The quotation is from Stavins (2007b, p. 9)

3. The theoretical model described below is implicitly invoked, for example, by California’s Market Advisory Committee (2007), Pizer (2007) and Stavins (2007a).

costs and capital costs are combined in an annual abatement cost function, which is known with certainty. Other input prices are known with certainty, as is output (or output price). Each firm has an initial allotment of emission permits, A . To match this to its intended emissions, e , it buys (or sells) emission permits in the amount u , such that $e = A + u$. The cost of abatement is given by $c(e)$, a decreasing and convex function. Assuming the firm is a net purchaser, the cost of obtaining emissions permits is given by $\theta(u)$, an increasing function. In the simplest version considered by Montgomery, the cost of purchasing permits takes the form $\theta(u) = \tau u$, where τ is the price per unit in the emission market, which is assumed to be taken by the firm as constant (i.e. there is no market power).⁴

Each firm determines its emissions by minimizing total cost, $c(e) + \theta(e-A)$. The first-order condition is: $-c'(e) = \theta'(e-A)$. In the simple version considered by Montgomery where $\theta'(\cdot) = \tau$, this reduces to:

$$-c'(e) = \tau. \tag{1}$$

A crucial implication, noted by Montgomery, is that with a market for emission permits the firm's emissions are *independent* of A , the initial allocation that it receives. However, as Stavins (1995), Montero (1997) and others have noted, this result follows directly from the linearity of the cost function $\theta(u)$. If there were transactions costs which varied nonlinearly with the number of emissions permits purchased (Stavins), or if there were uncertainty regarding whether the transaction will be approved which varied nonlinearly with the magnitude of the transaction (Montero), or any other nonlinearity such that $\theta''(u) \neq 0$, the optimal level of emissions, e , would *not* be independent of the permit allocation, A .⁵

Another implication of (1) is that the outcome of the emissions market is identical to the outcome with an emissions tax when the tax rate is set equal to τ . This follows because the firm's cost function with an emissions tax, $c(e) + \tau e$, differs from its cost function with an emissions market, $c(e) + \tau e - \tau A$, by a term that is independent of e .⁶ The equivalence of the emissions tax equilibrium with the emissions market equilibrium does not hold when $\theta(u)$ is nonlinear.

The foregoing is a partial equilibrium analysis focusing on the market for emissions permits. Analogous results also hold in a general equilibrium setting covering all input and output markets in the economy. The general equilibrium setup is as follows. There are N commodities, indexed by $j = 1, \dots, N$, and N producers,

4. This is also the revenue from selling permits if $u < 0$ and the firm is a net seller.

5. Montero's analysis assumes risk neutrality on the part of the firm. If there were uncertainty which varied *linearly* with the magnitude of the transaction (e.g., simple emission price uncertainty), but the firm were risk averse, this too would cause the optimal level of emissions to depend on the initial permit allocation (Baldursson and von der Fehr, 2004). Because market power leads to nonlinearity of $\theta(u)$, it also causes emissions to depend on the initial allocation of permits (Hahn, 1984).

6. I assume here that the tax revenues are disposed of through lump-sum transfers and the permits are auctioned off, so that the same amount of revenue is collected in both cases.

indexed by $k = 1, \dots, N$, who produce commodities using other commodities as inputs, like steel or electricity, and also using labor.⁷ Let $y^k = (y_1^k, \dots, y_N^k, y_{N+1}^k)$ denote the commodity vector associated with the k^{th} producer: by convention outputs are positive, ($y_j^k \geq 0$), and inputs negative, ($y_j^k \leq 0, j \neq k$). The production function for the k^{th} producer is given in implicit form by $f_k(y^k, e_k; \gamma_k) = 0$, where e_k is the amount of pollutants emitted by the k^{th} producer, treated for simplicity here as a scalar, and γ_k is an index of the technology of production for the k^{th} commodity, taken as exogenous. This representation of the production technology allows for pollution to be either an input to production or a joint product with output. On the consumption side, it is simplest to imagine a single, representative household which consumes commodities and supplies labor, with a utility function denoted $u(x, e)$, where x_j is the consumption of the j^{th} commodity, $j = 1, \dots, N$, x_{N+1} is the aggregate supply of labor, and $e = \sum_k e_k$ is the aggregate emission of pollution. The household takes e as given. Commodity balance requires $y_k^k = x_k - \sum_{j \neq k} y_j^j$. The vector of prices is given by $p = (p_1, \dots, p_N, p_{N+1})$. Taking these prices as given, producers maximize profit subject to the production function, and the household maximizes utility subject to a budget constraint.

It should be noted that this general set-up is used both in the *theoretical* analyses of pollution abatement policy and also in the *empirical* analyses with computable equilibrium (CGE) models that have been employed in the U.S. to quantify the economic impacts of climate mitigation policy.

In the case of an emissions tax, market equilibrium involves each producer choosing inputs, output, and emissions so as to maximize profit, $\sum_j p_j y_j^k - \tau e_k$, where τ is the unit emissions tax, assumed constant. With emissions trading, each producer maximizes profit, $\sum_j p_j y_j^k - \tau(e_k - A_k)$, where A_k is the initial allotment of emission permits and τ is the unit price of emission permits in the market which adjusts endogenously so that aggregate emissions just match the pre-set cap, $\sum_k e_k = \bar{e} \equiv (\sum_k A_k)$. This formulation of the emissions trading equilibrium assumes no market power, and no non-linearity in the cost of purchasing or using emission permits. Given those assumptions, it follows that the equilibrium vector of commodity prices and the wage rate, p^* , and the equilibrium price of emissions permits, τ^* , depend on the aggregate cap on emissions, \bar{e} , but *not* on the particular pattern of the individual allotments (A_1, \dots, A_N): $p_j^* = p_j^*(\bar{e}; \gamma_1, \dots, \gamma_N), j = 1, \dots, N+1$ and $\tau^* = \tau^*(\bar{e}; \gamma_1, \dots, \gamma_N)$. Moreover, the outcome with the emissions market is identical to the outcome with an emissions tax if the tax rate is set equal to τ^* . Both equilibria achieve the given overall level of emissions \bar{e} at minimum aggregate cost.

This simple framework generates the key insights regarding the design of a cap-and-trade system which are ubiquitous in the literature. First, the literature emphasizes the role of price changes as the drivers of behavioral change. Consider, for example, what happens when a price is put on CO₂ emissions from fossil fuel use, through either a tax on CO₂ or a cap on the CO₂ associated with fossil fuel production. The prices of fossil fuels rise to reflect the carbon

7. Labor will be denoted as the $N+1^{\text{st}}$ commodity. In what follows $y_{N+1}^k = 0$,

price signal. As a result, the demand for fossil fuels falls. In turn this may lower fossil fuel producers' demands for their inputs, reducing those prices and passing the cost burden backward as well as forward along the supply chain. Similarly, commodities which are fossil-fuel intensive become more expensive, leading to a reduction in their demand. At the same time, there is increased demand for inputs that might reduce the generation of emissions, including inputs to abatement and substitutes for fossil fuels such as renewables. These changes radiate throughout the economy, inducing a suite of price-driven demand and supply responses in sectors both upstream and downstream of the sectors that are taxed or capped.

Second, because equilibrium prices are the same economy-wide with an emissions market and with an emissions tax that induces the same reduction in aggregate emissions, \bar{e} , it makes no difference in theory which instrument is used – the two are economically equivalent.⁸ Third, because the economy-wide equilibrium prices are independent of the initial allocation of emission permits among individual producers, it makes no difference in theory whether there is an *upstream cap* (the point of regulation is the point of entry of fossil fuels into the economy, i.e., the point of extraction, processing or distribution) or a *downstream cap* (the point of regulation is the end user of the fossil fuels where combustion occurs, or the end user of energy derived from fossil fuels) – the two are economically equivalent.⁹

A question arises: given that an incentive-based regulation is introduced at some point in the economy – whether upstream, downstream, or somewhere in-between, and whether in the form of a cap-and-trade system or an emissions tax – where is the main response to the resulting price changes located? Obviously, responses can occur throughout the economy. But, does more of the adjustment occur *within* the particular sector that is the point of regulation (intra-sectoral adjustment),¹⁰ or does more of the adjustment occur in the *other* sectors (inter-sectoral adjustment)? Suppose, for example, that the electricity generation sector is the point of regulation. Is the main economic response a change in how electricity is produced, or is it a reduction in the downstream demand for electricity due to

8. As in footnote 3, I assume auctioning of permits and lump-sum disposition of tax revenues.

9. There are some caveats to the equivalence result (Pizer, 2007). The administrative costs of an upstream and downstream cap may be different (the former are very likely to be lower). Because of administrative costs, the coverage of the cap is likely to be different, with an upstream cap more comprehensive. Also, the equivalence may not hold if there are deviations from perfect competition in some economic sectors – for example, if the railroads which deliver low-sulfur coal have market power, or the rates charged by pipeline companies are regulated, or there are long-term fuel supply contracts that do not allow for price adjustments.

10. Intra-sectoral adjustment would include input substitution and changes in the choice of technique by firms in the regulated sector. Another form of intra-sectoral adjustment arises when there is heterogeneity among establishments within the sector reflecting differences in capital vintage and/or productivity: intra-industry reallocation can then arise from exit or entry (Melitz, 2003). To the extent that empirical CGE models represent the regulated sector as a homogeneous entity characterized in terms of a single representative firm, while this may be able to capture input substitution and changes in the choice of technique, it may not account adequately for entry and exit of heterogeneous producers.

an increase in its price?¹¹ While the answer is an empirical issue, it is worth noting that price-driven inter-sectoral adjustment is strongly emphasized by writers in the policy literature arguing the case for an upstream cap on emissions.¹²

The theoretical model on which these predictions – the equivalence of an appropriately structured emissions tax and a trading system, and the equivalence of appropriately structured upstream and downstream caps – are based simplifies reality in several important respects, in addition to the linearity noted above. It is a purely static model. There is no distinction between capital cost and operating cost, no fixed capital, no investment, and no allowance for time lags associated with the turnover in capital. Also, technology (and tastes) are fixed. The question to be addressed next is whether the actual experience with emissions trading conformed to the assumptions of the model and matched its predicted outcomes.

3. THE EXPERIENCE WITH EMISSION TRADING

3.1 The Reduction in Emissions

While there have now been several emission trading systems, trading in sulfur dioxide (SO₂) emission permits under the Title IV Acid Rain program of the 1990 US Clean Air Act (CAA) Amendments is widely cited as the poster child for the cap and trade approach. Accordingly, I focus primarily on that program, with a brief note on another U.S. trading program, for lead used by refineries, at the end of this discussion.

The SO₂ component of Title IV covers emissions from fossil-fuel burning power plants located in the continental 48 states of the United States. The program was instituted in two phases. During Phase I, lasting from 1995 through 1999, electricity generating units larger than 100 MWe in generating capacity, and with an annual average emission rate greater than 2.5 pounds of SO₂ per million Btu of heat input in 1985, were required to reduce emissions to a level that would be, on average, no greater than 2.5 pounds of SO₂ per million Btu of heat input.¹³ In Phase II, beginning in 2000, the program was expanded to include fossil-fuel electricity generating units greater than 25 MWe, or virtually

11. The latter is what Goulder and Parry (2008) refer to as “the output-reduction channel.”

12. Stavins (2007a, p. 18), for example, emphasizes the role of price signals: “A cap has the same impact on the effective cost of fuel for downstream users regardless of the point of regulation. ... In either case [upstream or downstream cap] the downstream customer ultimately faces the same additional cost associated with emissions from its fuel use. ... [F]irms and consumers will undertake the same emissions reduction efforts— and thereby incur the same emissions reduction costs—in either case, because they face the same carbon price signal.”

13. There were 263 generating units required to participate during Phase I (“Table A units”). In addition, other units not originally affected until Phase II were permitted to choose to enter the program early as a substitution unit or a compensating unit as part of a multi-unit plan to fulfill the Phase I compliance obligations. Also other sources of SO₂ emissions not covered by the Acid Rain Program could voluntarily opt-in to the program. The number of participating non-Table A units varied from year to year during Phase I, reaching as many as 182 substitution/compensating units and 10 opt-ins.

all fossil-fuel power plants in the United States. The Phase II cap is equivalent to an average emission rate by 2010 of about 1.2 pounds of SO₂ per million Btu of heat input in 1985.

The total SO₂ emissions from the generating units covered by Phase I had been 9.4 million tons in 1980 and 8.7 million tons in 1990. In Phase I these units' SO₂ emissions were capped at 5.5 million tons; by 1999, their actual emissions were 3.5 million tons, a reduction of almost 60% compared to the 1990 emissions.¹⁴ With Phase II, the total SO₂ emissions from the units covered in that phase had been 17.3 million tons in 1980 and 15.7 million tons in 1990. These units' emissions were capped at a declining rate, starting at 10 million tons in 2000 and declining to 8.95 million tons in 2010, remaining fixed at that level thereafter. In 2008, these units' emissions were capped at 9.5 million tons, but their actual emissions were 7.6 million tons, a reduction of almost 52% compared to the 1990 emissions.¹⁵

How did the reductions in emissions come about? One option for the owner of a regulated unit was to install a flue gas desulfurization system (scrubber) if there was not one already. In 1991, there were 155 generating units with scrubbers in the US; by 1998, 192 units had scrubbers; and by 2007, 247 units had scrubbers.¹⁶ The number of scrubbers retrofitted for Phase I was about half the number that had been projected before emission trading began. Nevertheless, it is estimated that, by 2001, 37% of the SO₂ emission reductions was due to the installation of scrubbers.¹⁷ Thus, installation of scrubbers was an important part of the response, but not the dominant part.

Another option was to change the dispatch order, switching generation from a coal-fired unit to a nuclear unit, which generates no SO₂, or a natural gas unit, which generates only trace amounts of SO₂. Moreover, there was a shift from coal to natural gas in the construction of new generating units.¹⁸ This made a relatively limited contribution to the reduction of SO₂ emissions in Phase I, since trading was then limited to coal-fired units, but it became more important in Phase II. Ellerman and Dubroeuq (2004) estimate that the reduction in SO₂ emissions due to the introduction of efficient new combined-cycle natural gas plants accounts for about 13% of the SO₂ emission reduction occurring between 1994 and 2002.

In fact, the largest component of the reduction in SO₂ emissions was due to other factors which were essentially *innovations* in the production of low-

14. These data pertain to Table A units. Note that unused emission reductions can be banked for use in future years.

15. As of 2002, Ellerman (2003) estimates that about five sixths of the emission reductions under Phase II came from units participating in Phase I.

16. However, while there was a 59% increase in the number of units with scrubbers between 1991 and 2002, there was only a 41% increase in aggregate scrubber capacity.

17. Burtraw et al. (2005, p. 268), based on Ellerman et al. (2000, p.127).

18. The amount of natural gas generating capacity doubled from 1994 to 2005, while total US generating capacity rose by 28% over this period.

emission electricity. Ellerman and Dubroeuq (2004) estimate that about 85% of the reduction in SO₂ emissions between 1994 and 2002 was associated with a reduction of emissions at individual generating units (i.e., cleaning up old plants) as opposed to switching generation from high- to low-emitting units. As noted above, some of this reduction was due to the retrofitting of scrubbers. But, the remainder – about half or more of the total reduction in SO₂ emissions – was due to two factors other than the retrofitting of scrubbers and the substitution of gas or nuclear for coal, namely a switch from high- to low- sulfur coal, and a substantial reduction in capital and operating costs for both new and existing scrubbers.

Coal switching, in part, involved the introduction of western low-sulfur coal to power plants in the Midwest. This was a consequence of railroad deregulation in 1976 and 1980, which gave railroads greater operating freedom. The result was more efficient railroads, lower costs of operation, and more competition for freight. Busse and Keohane (2007) show that the two dominant railroads in the Powder River Basin (PRB), the source of western low-sulfur coal, engaged in price discrimination, raising the rates to transport coal to captive western power plants but lowering them for power plants in the Midwest.¹⁹ Price discrimination aside, Burtraw (1996) points out that the railroads engaged in major investment initiatives to expand the capacity for transporting PRB coal and lower the cost, including laying more track, and innovations in train technology such as increasing the size of unit trains and the number of locomotives, using aluminum cars, improving motors, and raising car dump speed. Burtraw (2000) refers to this as a “realignment of investment plans” and comments: “[t]his is not the usual form of innovation imagined as part of technical change. However, it does represent a diffusion of new technology and an organizational adaptation to market opportunity.”

In 1990, PRB coal had accounted for less than 20% of the coal burned at Table A power plants west of the Mississippi; in 1999, it accounted for 95%.²⁰ East of the Mississippi, PRB coal accounted for about 40-50% of the coal burned at Table A plants by 1999. The boilers in the Midwest units had been designed to burn bituminous Midwestern coal. “It was always recognized,” Ellerman (2003, pp 30-31) notes, “that these units could be converted to the use of sub-bituminous coals, but the higher water and ash content of the latter would lead to a significant derating, or reduction, in the generating capacity of the unit. As a result, it was expected that the predominantly high-sulfur burning units in the Midwest would either install scrubbers or switch to low-sulfur bituminous coal produced in the Appalachian region. As the effects of rail deregulation increasingly reduced the significant transportation component in the cost of *western* low sulfur, sub-bituminous coals delivered to the Midwest, power plant engineers began to experiment with blending these coals with locally produced high-sulfur

19. Railroad rates for coal shipments from the PRB to the Midwest fell by more than half between 1983 and 1993; much of this decline occurred before the enactment of Title IV (Ellerman and Montero, 1998).

20. Busse and Keohane (2007).

bituminous coals.... [T]he answer emerged that, depending on the unit and the coals being blended, mixtures of up to 60% of low-sulfur, sub-bituminous coal (and sometimes higher) could be used without significant derating in the generating capacity of the unit.” This development was “not observed before and not expected.” It was a refinement of known technology, but nevertheless an innovation.

In addition to western coal, there were what Ellerman (2003) calls “[o]ther cost-reducing changes that might be termed innovations” in the other major coal production regions. “Mid-sulfur coal mines were developed in the Midwest where none existed before. ... A similar shifting downward of the average sulfur content of coal being supplied was observed in Northern Appalachia, the other high-sulfur coal-producing region. ... The causes were new mines now made economic in local markets, changes in mining practices that reduced the sulfur content of coal being already mined, and increased sulfur removal in coal preparation plants.”

There was also innovation with regard to scrubbers. The price of new scrubbers dropped by nearly half between 1989 and 1994.²¹ The operating efficiency increased, the average removal efficiency of new scrubbers rising from about 85% before 1992 to about 95% by 1995.²² Also, there was an increase in the utilization of the retrofitted scrubbers.²³ As a result, there was a substantial reduction in the cost of using scrubbers; Ellerman (2003) estimates that scrubber O&M costs fell from an anticipated \$464 per ton of SO₂ removed prior to Phase I, to \$286 during Phase I, and \$204 during Phase II.²⁴

Another factor which Burtraw (1996) called “cost savings without emissions trading” is that the act of giving plant operators the freedom to choose between buying emission allowances, installing a scrubber, changing the dispatch order and switching fuels created an element of *competition* among the suppliers of inputs to abatement that was not formerly present. By itself, the enhanced competition among abatement alternatives caused a reduction in the cost of abatement.

In short, “[w]hat emerges from the experience with Title IV,” Ellerman (2003, p.32) notes, “is that [abatement] costs are lower for reasons beyond the ability to trade emission reductions among sources.” Carlson et al. (2000) estimate that, for a typical coal-fired generating unit, the marginal cost of abatement fell by about \$250 per ton of SO₂ between 1985 and 1995. This should be compared to a 1985 marginal cost of about \$360 per ton at the pre-Title IV level of abatement, and a 1985 estimated marginal cost of about \$540 per ton at the full Title IV level

21. GAO (1994); Taylor, Rubin and Hounshell (2003).

22. Popp (2003).

23. Ellerman and Montero (1998); Ellerman (2003) notes a 25% higher utilization of the retrofitted units. Some utilities with scrubbers actually increased their usage of higher-sulfur coal because it was cheaper and thus more profitable to use when combined with a scrubber.

24. See also Taylor (2001). Popp (2003) shows that scrubber O&M costs at covered units declined by about 22% during Phase I.

of abatement.²⁵ Carlson et al. estimate that \$200 of the cost reduction was due to the decline in the price of low-sulfur coal between 1985 and 1995, while \$50 was due to innovations in abatement technology between 1985 and 1995.

Whether or not these innovations constituted technological change is a matter of terminology. Scrubbers became cheaper, as noted above, but there does not appear to have been any fundamental change in scrubber technology. The data on patent counts relating to post-combustion SO₂ control technology actually show a decline in the number of new patents starting around 1990, and steepening after 2001.²⁶ Moreover, there was no obvious boost to other low-emission technologies for coal combustion such as Integrated Gasification Combined Cycle (IGCC) which results in lower emissions of SO₂, particulates and mercury as well as improved combustion efficiency compared to conventional pulverized coal.²⁷ Developments like the burning of low-sulfur coal in boilers designed for high-sulfur coal and the increased operating efficiency of scrubbers were primarily refinements in operating practices rather than fundamentally new technologies.

One should also note what did *not* happen. Despite increased demand, the prices of inputs used for the abatement of SO₂ in electricity production did not rise – as indicated, they fell. And there was virtually no discernible increase in the price of electricity as a result of the Title IV program. The total value of electricity sales in 2000 was \$233 billion while the total cost of compliance with the Phase II program that year is estimated to have been about \$1.5 billion (i.e., about 0.6% of sales).²⁸ Moreover, the retail price of electricity fell continuously throughout this period from (in \$2000) 9.7 cents per Kwh in 1982 to 8.05 cents in 1990, 7.66 cents in 1994, and 6.81 cents in 2000.²⁹ Thus, although SO₂ abatement may have raised the cost of electricity production by about 0.6% in 2000 compared to pre-Phase I, the retail price of electricity fell by more than 10% between 1994 and

25. The marginal cost of abatement rises with a higher level of abatement. These figures are in \$1995 and are derived from Carlson et al. (2000, figure 2).

26. By 2004, the annual number of new patents was less than half the annual average for the period 1975-1989 (Taylor, 2008). Popp (2003) shows that there is a correlation between the increase in scrubber operating efficiency after 1990 and the cumulative stock of new patents issued after 1990. But it is not clear whether the specific focus of post-1990 patents related to operating efficiency, or that the relationship found by Popp is a causal one. The cumulative stock of new patents may function in his regression as something rather similar to a time trend.

27. There are currently only two IGCC plants generating power in the U.S., which started operating in 1995 and 1996; several new IGCC plants are expected to come online in the U.S. around 2012-2020.

28. Carlson et al (2000) estimate the annual cost of Phase II by 2010 at \$1.51 billion with 1995 prices and 1995 technology, and \$1.05 billion with 1995 prices and 2010 technology. With 1995 prices and 2000 technology, the cost would have been somewhere in between, say about \$1.35 billion. Translating from \$1995 to \$2000 using the GDP implicit price deflator raises that figure to \$1.46 billion.

29. There was a spike in electricity prices in 2001, but they then fell back to the range 6.91-6.99 in 2002-2004. The data on electricity sales, prices and generation are taken from EIA (2009)

2000.³⁰ There also was no obvious reduction in the production of electricity – net generation grew from 3,247 billion Kwh in 1994 to 3,802 billion Kwh in 2000 (an increase of 17%).³¹

It also appears that the cap on SO₂ emissions had virtually no effect in promoting either energy conservation or the use of renewable sources of electricity. Net generation from renewables rose only about 6% from 337 billion Kwh in 1994 to 357 billion Kwh in 2000. One provision of Title IV created the Conservation and Renewable Energy Reserve, under which bonus allowances worth 300,000 tons of SO₂ were set aside to be allocated to utilities for energy efficiency and renewable energy development.³² However, little use was made of these allowances; as of February 2002, only about 49,000 of these allowances had been awarded (Vine, 2003).

In summary, actual SO₂ emissions were reduced from 15.7 million tons in 1990 to 11.2 million tons in 2000, and 7.6 million tons in 2008. But, the way in which the reduction occurred was not exactly what would be predicted based on the static model in section 2. Induced innovation within the electricity sector and its immediate suppliers played a key role, whether in the form of market evolution associated with increased competition, or operational innovations by coal producers, railroads, boiler operators, and scrubber manufacturers and operators. The trading program triggered a rush to find new ways of reducing SO₂ emissions at a lower cost, which built on other innovations associated with railroad deregulation that were already under way. There were not only movements *along* demand and supply curves in the various sectors of the economy but also *shifts* in those curves induced by changes in market structure and production innovations. In the terminology used above, what happened was a shift in some of the γ_j 's. Some of this may perhaps also be interpreted as a reduction in X-inefficiency within the electricity generation industry triggered by a tighter regulatory cap on pollution emissions and increased competition among abatement options.³³

30. Some of the price decline may have been a consequence of the restructuring of the electricity industry starting in the mid-1990s, but not much. Roughly 20% of the generating capacity was affected by restructuring. Bushnell and Wolfram (2005) find that restructuring had little positive and possibly negative impacts on fuel efficiency. Fabrizio, Rose and Wolfram (2007) find that, between 1990 and 1999, restructuring reduced labor and nonfuel expenses by about 3-5% in restructured plants relative to other plants, and those expenses amount to about 25% of the cost of electricity generation.

31. This is consistent with Ellerman's (2003) assumption that "no emissions reduction can be attributed to changes in demand [for electricity]." He continues: "Since the demand for electricity is price inelastic, the cost of SO₂ controls is relatively small on a kilowatt-hour basis, and the major element determining the dispatch, or utilization, of individual generating plants is the cost of fuel, the error arising from assuming no effect on demand is probably small."

32. Of this amount, 60,000 tons were set aside just for renewables. Note that participation was limited to utilities; end users and energy service companies could not claim allowances.

33. The existence of X-inefficiency in electricity generation has been suggested by Stevenson (1982) and Nelson and Primeaux (1988) among others.

3.2 Emission Trading as a Performance Standard for SO₂

These observations should not be seen as an indication that the introduction of emission trading for SO₂ was unimportant or of little economic benefit. Even setting aside the substantial reduction in SO₂ emissions mandated by Title IV of the 1990 CAA Amendments, the manner in which the emission reduction was regulated was a significant improvement compared to the prior regulatory regime established under the 1970 Clean Air Act and the 1977 CAA Amendments.

Under the prior regime, power plants in existence when the regulations implementing the 1970 CAA became effective in 1971 had to meet emission rate limits imposed by State Implementation Plans (SIPs), which the individual states were required to develop to ensure compliance with National Ambient Air Quality Standards (NAAQS) for six “criteria” air pollutants, including SO₂. New power plants constructed after that date were regulated more stringently. They were required to meet the New Source Performance Standard (NSPS) which included an emissions limit of 1.2 pounds of SO₂ per million Btu of heat input (the limit that was subsequently imposed on all power plants under Phase II of Title IV). Following the 1977 CAA Amendments, the NSPS for SO₂ was modified to require the installation of a scrubber in new power plants, ruling out any reliance on low-sulfur coal alone.

While the emissions limit established under Phase II of Title IV is the same as the prior NSPS limit, the Title IV program is vastly more flexible. Installation of a scrubber is no longer required. A power plant operator is free to use low-sulfur coal, vary the dispatch order, substitute emissions among facilities and/or purchase emission allowances in order to comply with the emissions limit. The reduction in the costs of a scrubber has already been noted. One factor that may have contributed is the flexibility provided by Title IV with regard to redundant scrubber capacity (Burtraw, 1996). Before the 1990 CAA Amendments, scrubber systems usually included a spare module to maintain low emission rates when any one module became inoperative. The alternatives afforded by Title IV lessened the need to install a spare module, thereby reducing the capital cost of a new scrubber. The key point is that, instead of mandating a particular abatement technology, Title IV established a performance standard for SO₂.

A performance standard is economically superior to a technology standard because it allows firms to attain the given emission limit at a minimum cost. Much of the existing policy literature on emission trading has focused on differences in the marginal cost of SO₂ abatement among different electricity generating units and the resulting gains from trade if emission reduction can be shifted from high- to low-cost units.³⁴ In this view, emission trading is advantageous because it reallocates emission reduction among production units in a long-run cost-minimizing manner.

34. For example Market Advisory Committee (2007), Pizer (2007), Stavins (2007a).

But, there are other forms of cost-minimization which become available through emission trading. These are associated with *short-run*, cost-minimizing operating flexibility as opposed to *long-run* cost-minimizing reallocation of emission reduction. One form of operating flexibility is the ability of a power company controlling several units to reallocate abatement among alternative units in order to cope with short-run fluctuations in operating conditions. Another source of flexibility is the ability to reallocate emission reductions over time. Companies can better control the timing of investment in new lumpy capital in an economically opportune manner. If interest rates are low, for example, a company can install new abatement equipment before it is fully needed, and either bank or sell the excess emission reduction; or, a company can postpone abatement investments until more favorable financial circumstances arise, buying emission allowances in the interim.

A third source of operating flexibility arises from the firm's ex ante uncertainty regarding what its emissions will be during the relevant regulatory period. This is only one of several sources of uncertainty that may confront participants in an emissions market. There is also likely to be uncertainty about future emission permit prices and abatement costs, and perhaps also uncertainty about prices in output markets and regulatory uncertainty regarding the administration of the emission trading system. These all can affect firms' decisions regarding investment in abatement (e.g., the installation of a scrubber) and participation in buying or selling emission permits.³⁵ But, uncertainty about *emissions* is surely ubiquitous. A power plant operator cannot know what the demand for electricity will be next week, let alone for the rest of the year, and he cannot be sure of future prices for fuels of different sulfur content.³⁶ The significance of this is that SO₂ and other emissions markets are designed so that production and compliance are *asynchronous*: emissions occur during a particular period (the calendar year, for SO₂) while the emissions permits must be turned over *at the end* of the period, or shortly thereafter.³⁷ Buying permits is a way to hedge against the uncertainty regarding emissions. Emissions permits have a *convenience value* (Burtraw, 1996) which physical abatement equipment lacks: if there is a shortfall on December 31, say, it is possible to acquire additional permits in the last minute, but it is too late to lower the year's emissions by installing a scrubber.³⁸

35. For example, Baldursson and von der Fehr (2004) discuss uncertainty regarding emission prices; Zhao (2003) deals with uncertainty regarding abatement costs; and Schennach (2000) presents a generalized treatment covering multiple sources of uncertainty.

36. Ellerman and Montero (2007, p.50).

37. SO₂ permits for a given calendar year must be turned in to the EPA by February 28 of the following year. By contrast, the model in section 2 characterizes the decision as synchronous – i.e., the allowance constraint is binding at the exact moment when the emissions decision is made. The asynchronicity is discussed by Carlson and Sholtz (1994).

38. In addition, the purchase of permits is reversible (the permits can be resold), while the installation of a scrubber is irreversible; the difference in reversibility gives rise to an option value for permits (Chao and Wilson, 1993). Rouse and Sevi (2005) analyze the implications of emission uncertainty for permit banking.

Thus, given the asynchronicity between emissions and the surrender of emissions permits, the ex ante uncertainty regarding emissions is a reason to purchase or sell permits at the end of the compliance period when actual emissions become known.³⁹ This is separate from the motivation created by the time value of money, which forms the basis for the analyses of intertemporal permit use in Rubin (1996), Kling and Rubin (1997) and Holland and Moore (2008).⁴⁰ Moreover, the asynchronicity complicates any assessment of the empirical relationship between permit allocations and emissions. The linear version of the theoretical model in section 2 implies that emissions are independent of the initial allocation of permits; but, as noted, if the firm is risk averse or there are nonlinear transaction costs or other nonlinearities, the allocation should affect the firm's choice of emissions. Even if this were true, however, with asynchronicity and uncertainty about emissions there need not be a tight link between the permit allocation and actual emissions.⁴¹

Short-run operating flexibility and long-run reallocation of emission reduction both provide motives to buy and sell emissions permits, but they are distinct phenomena and they serve different purposes. Operating flexibility, while economically valuable, does not necessarily lead to a long-run reallocation in emission reduction. For example, changing the dispatch order is a short-run fix to reduce emissions but with growth in power demand it is not necessarily a substitute for the installation of new, low-emitting generation capacity in the long-run. Thus, short-run flexibility and long-run reallocation of emission reductions are likely to lead to different patterns of participation in the emissions market, including a different timing and perhaps a different level of participation. Nevertheless, the existing literature, including the empirical literature to which I turn next, has not distinguished between these two motives for emission trading.

3.3 The Allowance Market

A striking feature of the SO₂ allowance market is that allowance prices turned out to be significantly lower than was anticipated before trading began. Burtraw and Palmer (2004) list five studies published between 1990 and 1995 which estimated the long-run (2010) annual cost of the Phase II program. The estimates clustered around about \$2.3 billion, with the lower end of one range being low as \$1.4 billion and the upper end of another range being as high as \$5.9 billion. The projected *marginal* costs were mostly about \$520-540 per ton of

39. In interviews with 114 Swedish firms during the first phase of the EU's CO₂ allowance market, Sandoff and Schaad (2009, Table 8) found that 79% of the firms may have acted like this.

40. Holland and Moore (2008) allow for uncertainty about abatement costs, but not uncertainty about future emissions.

41. In the EU Emission Trading System for CO₂, Reguant and Ellerman (2008) find no significant relationship between the initial allocation of permits and the daily operation of coal-fired electricity generating units in Spain. But, ex ante uncertainty regarding emissions and the flexibility to buy or sell permits after the actual emissions are known might be a factor in this finding.

SO₂.⁴² Those estimates did not anticipate the innovations that occurred.⁴³ Based on the experience with trading after 1995, Ellerman et al. (2000) subsequently estimated the long-run (2010) annual cost of the Phase II program at \$1.4 billion while Carlson et al. (2000) estimated it at \$1.1 billion; their corresponding marginal cost estimates are \$350 and \$291 per ton of SO₂ (\$1995).

In the earlier studies, it was assumed that allowance prices would generally be equal to the estimated long-run marginal cost of abatement, as implied by the first-order condition in (1). In fact, however, allowance prices with banking reflect a mix of short- and long-run factors. The prices have generally been well below long-run marginal cost -- they have mostly been under \$200 per ton of SO₂ except for a price spike between 2004 and 2008, due to a proposed tightening of emission limits under the Clean Air Interstate Rule (CAIR).⁴⁴

A second distinctive feature of the allowance market is over-compliance in emission reduction and the banking of allowances – see Figure 1. In Phase I, 30% of all allowances distributed during 1995-99 were banked; equivalently, the reduction in emissions was about twice what was required to meet the Phase I cap (Montero and Ellerman, 2007). Between 2000 and 2005 the bank was drawn down – albeit at a fairly modest rate – to cover emissions in excess of the annual allotment of new allowances. In 2007 and 2008, there has again been some over-compliance in emission reduction, and the allowance bank has started to grow. In an analysis of banking between 1995 and 2002, Ellerman and Montero (2007) concluded that, contrary to the general impression of an excessive amount of banking, the amount of banking that occurred during this period was actually

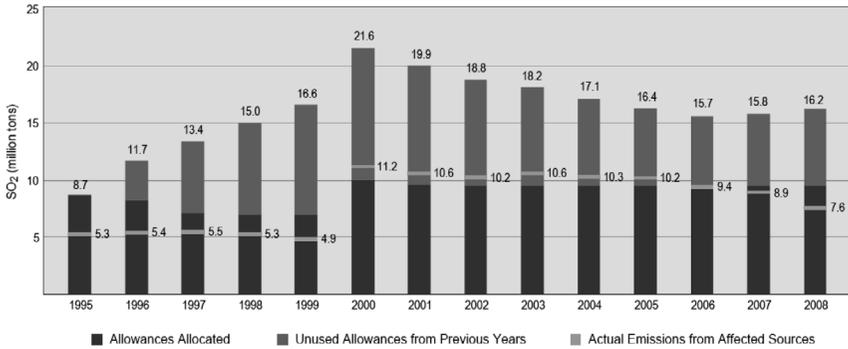
42. These costs are in \$1995.

43. Besides the failure to anticipate the use of low-sulfur coal, another reason for low allowance prices during Phase I, noted by Montero (1999), is that non-Table A firms were allowed to opt in, and firms affected by Phase I were allowed substitute alternative units. He shows that this caused adverse selection, with units being enrolled when their actual emissions were lower than the allowances that would be assigned to them, leading to an over-allocation of allowances and thus reducing the market price of allowances in Phase I.

44. The CAIR was proposed in late 2003 and finalized in March 2005. In order to ensure that 28 eastern states meet their NAAQS for ozone and fine particulates, the rule drastically reduced allowable SO₂ and NO_x emissions below the Title IV limits for power plants in those states. The effect was to devalue *future* allowances issued under Title IV in those states, and thus encourage the banking of current allowances. Pre-2010 allowances could be used on a 1:1 basis under the CAIR program, but *two* allowances issued in 2010 and subsequent years through 2014 would be required per ton of SO₂ emissions in CAIR states; starting with 2015 vintages, 2.86 allowances would be required per ton of emissions. In what may have been an “overly conservative reaction by buyers who wanted assurance they could cover current and future allowance needs” (EPA, 2006, p. 8), allowance prices rose to a peak of \$1,550 per ton at the end of 2005. A contributing factor was the spike in natural gas prices following Hurricanes Katrina and Rita in August and September 2005 (EPA, 2009b). Presumably in response to CAIR, there was a sharp increase in the number of scrubbers installed in 2006-2007. In July 2008, the U.S. Appeals Court for the D.C. Circuit vacated the rule; in December 2008, the Court rescinded the vacatur and instead remanded the rule to EPA for revision. By May 2009, the allowance price was down to \$71 per ton.

quite efficient based on economic fundamentals.⁴⁵ At the time, Ellerman and Montero projected that an efficient bank would decline to about 3 million tons in 2008. This is about half the amount actually banked now. Their analysis was conducted prior to the CAIR, which so strongly affected spot prices and could also have affected banking. Risk aversion, not included in their analysis, may also be a factor. One cannot rule out the possibility of over-compliance in emission reduction beyond that called for by pure economic efficiency.

Figure 1. SO₂ Emissions and the Allowance Bank, 1995–2008



Source: EPA (2009a, Figure 3)

A third feature of the allowance market is a significant degree of *autarky*. This is when firms do not comply with their cap by purchasing allowances in arms-length transactions; they either reduce emissions to stay within the cap or, if there are excess emissions, they draw on their own past banked allowances or on allowances available from other units that they control. One measure of this is the extent to which, when firms submit allowances to the EPA to cover their emissions, those are allowances that had originally been allocated to them rather than allowances obtained from someone else. Kreutzer (2006) examines this through 2003. Between 1997 and 1999, about 70% of the allowances retired each year were redeemed by the same unit to which they had originally been allocated, and only about 30% were originally allocated to another unit.⁴⁶ Between 2000

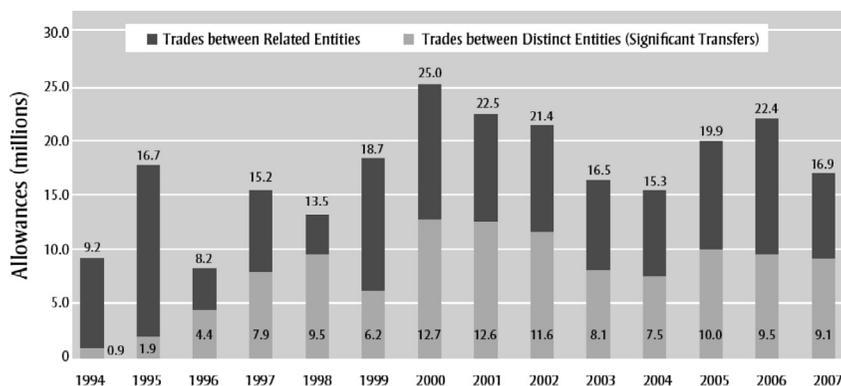
45. Efficiency aside, Burtraw and Palmer (2003) note an important political benefit of banking: “Once firms had built up a bank of unused allowances, they had a vested interest in maintaining the value of those banked credits, and thus in furthering the program itself.”

46. In a personal communication with EPA staff, I have been told that, until 2006, account IDs represented individual boilers. Therefore, “the same unit” here means the same boiler.

and 2003, the proportion of allowances redeemed by the same unit to which they had originally been allocated was lower, about 60%, but still substantial.⁴⁷

Even when allowances are transferred, this may not be an arms-length transaction since the same corporation may own multiple plants with multiple boilers, each with its own allocation of allowances. In its annual summary of allowance transaction data, the EPA distinguishes between what it calls “economically significant transactions (i.e., between economically unrelated parties),” and “transactions between related entities.”⁴⁸ In 2007, for example, the EPA deducted 8.9 million allowances from sources accounts to cover their emissions that year. In addition, “nearly 4,700 private allowance transfers moving roughly 16.9 million allowances of past, current and future vintages were recorded in the EPA allowance Tracking System. About 9.1 million (54%) were transferred in economically significant transactions.” The other 46% were transfers between related entities. As shown in Figure 2, the large proportion of transfers between related parties has been an enduring feature of the SO₂ allowance market.⁴⁹

Figure 2. SO₂ Allowances Transferred under the Acid Rain Program, 1994–2007



Source: EPA (2008, Figure 6)

47. Evidence of autarky comes from the survey of Swedish firms by Sandoff and Schaad (2009, Table 12), who found that 46% of firms thought they would handle a potential allowance deficit by reducing emissions internally rather than purchasing allowances.

48. The EPA comments that “transfers between economically unrelated parties are ‘arms length’ transactions and are considered a better indicator of an active, functioning market than transactions among the various facility and general accounts associated with a given company” (EPA, 2008, p 12).

49. Another phenomenon, commented on by Swift (2000) and Burtraw and Palmer (2003), was the relatively close physical proximity between buyers and sellers in allowance transactions, at least during Phase I. Swift showed that more than 80% of the allowances used to offset emissions during the period 1995-1998 came from the same state as the emitting source. The limited geographical extent of the market possibly implies some failure to exploit gains from trade to the fullest extent possible.

Title IV mandated a massive reduction in SO₂ emissions by 2010. As noted above, allowance trading could lower the cost of this emission reduction in two ways: by providing some short-run flexibility with respect to *when* the emission occurred, and by promoting a long-run reallocation of *where* the emission reduction occurred by reassigning abatement from relatively high-cost units to relatively low-cost one (i.e., gains from trade). To the extent that there was autarky, less of the latter would occur.

Nevertheless, significant gains from trade did occur. Carlson et al. (2000) estimate the potential cost savings attributable to formal emission trading at \$250 million per year during Phase I, compared to the counterfactual of a uniform emissions rate standard, and \$784 million per year during Phase II (\$1995); the Phase II saving is about 43% of the total compliance costs under a uniform standard. Similarly, Ellerman et al. (2000) estimate the Phase II cost savings, including savings attributable to banking, at about 55% of total compliance costs under a command-and-control approach. But, these savings do not mean that the allowance market was perfectly efficient and that all gains from trade were successfully exploited. Carlson et al (2000) find that, in the first years of Phase I, there were some differences in marginal abatement costs among facilities, and absolute compliance costs could have been reduced further with additional trades. Ellerman et al. (2000) reach a similar qualitative conclusion, although their estimate of the unexploited gains from trade is smaller. In addition, plant-level studies of production and abatement efficiency by Coggins and Swinton (1996) and Swinton (2002, 2004) indicate that some plant owners did not taken full advantage of the allowance market; they controlled emissions when it would have been cheaper to purchase allowances.⁵⁰ The conventional explanation is that, in some states, public utility regulations and other state laws deliberately discouraged power plants from purchasing allowances and, for local political reasons, pressured them instead to resort to scrubbing and/or fuel switching.⁵¹ But, a degree of preference for autarky may also have played some role.

Most of the existing empirical studies of the SO₂ allowance market use data through about 2002; more recent data have yet to be investigated. It may thus be premature to make a final assessment of the functioning of the market. Recall Ellerman and Dubroeuq's (2004) conclusion that about 85% of the reduction in SO₂ emissions between 1994 and 2002 was associated with a reduction of emissions at individual generating units as opposed to switching generation from high- to low-emitting units. If this finding continues to hold through 2009, it would suggest that allowance trading may have played a larger role in providing short-run flexibility than in promoting long-run reallocation of emission reduction.

50. Keohane (2007) reaches a similar conclusion.

51. For example, Rose et al. (1993) and Lile and Burtraw (1998) document some actions that promoted the use of locally mined coal. A dozen such studies are listed by Burtraw et al. (2005, pp. 265-6). However, Bailey (1996) concluded that local regulatory actions did *not* discourage use of the SO₂ allowance market.

It would be useful to have a more extensive analysis of which firms did not trade, which did, and the pattern of their trading. If allowance trading functions as an instrument for the long-run reallocation of emission reduction one would expect to find that, since 1995 or 2000, certain power plants have consistently purchased allowances and emitted beyond their original allocation, while others have consistently sold allowances and emitted below their allocation, with the former having high abatement costs and the latter low abatement costs. If certain plants consistently purchased allowances for a number of years, but then installed a scrubber and switched to consistently selling allowances, this could be an indication of the use of the allowance market to obtain flexibility in the timing of long-run abatement. Alternatively, if low-abatement-cost plants commonly buy allowances, and high-abatement-cost plants sell allowances, or if plants randomly switch from buying allowances in one year to selling allowances in another year, this could be an indication of the use of permits to smooth abatement in the face of uncertainty regarding annual emissions. Only by investigating the pattern of abatement and trading at the individual firm level can one pinpoint the particular contribution of allowance trading to the very successful reduction of SO₂ emissions since 1994.⁵²

3.4 SO₂ Trading Behavior

As noted above, a definitive assessment of the SO₂ allowance market over the period 1995-2009 remains to be undertaken. Here we offer some preliminary, and still speculative, conclusions. When viewed through the lens of the simple economic model outlined in section 2, there appear to have been some anomalies. Not all market transactions that ought to have occurred from the perspective of a cost-minimizing allocation of abatement appear to have occurred – some firms which ought to have bought allowances, abated their SO₂ emissions instead. And some firms which ought not to have purchased allowances, or ought to have sold allowances, accumulated allowances and banked them. There are at least overtones of two types of behavior not accounted for in the simple model of section 2. One is the use of the allowance market to provide short-run operational flexibility of various forms. The other is occasional indications that managers' behavior might sometimes be influenced by the allotted limit on their emissions; this could lead some firms to stay out of the market.⁵³

There are several possible reasons for this behavior, including both transaction costs and managerial perceptions and attitudes. There could be a pure preference for autarky – a wish not to bother with the complexity of the market.

52. For example, Keohane (2007) analyzes firms' decisions to install a scrubber and how this was affected by their being regulated under Phase I of the trading program; but there is no analysis of whether a firm then sold allowances if it installed a scrubber, or purchased allowances if it did not.

53. In a different context, a recent experimental study of allowance trading by Goeree et al. (2009) finds some indication of a predilection to autarky when allowances are grandfathered, although not when they are auctioned.

There could be a preference to comply with the law by keeping emissions within the limit set by the EPA.⁵⁴ There could be risk aversion in the face of uncertainty about future emissions, about future allowance prices, about future program regulations. Another possible factor is loss aversion, which Kreutzer (2006) suggests might have come into play and promoted autarky. A firm that does not sell allowances when it should do so foregoes a gain; a firm that sells allowances when it should not do so, and then has to buy them back at a higher price, suffers a loss which may receive a heavier weight than a gain of the same magnitude.⁵⁵

While these are different stories, they could all end up generating a profit maximization decision which takes the form

$$\begin{aligned} & \sum_j p_j y_j^k - (\tau + \delta)(e_k - A_k) \quad \text{if } e_k < A_k \\ \text{maximize} & & & (2) \\ & \sum_j p_j y_j^k - (e_k - A_k) \quad \text{if } e_k \leq A_k \end{aligned}$$

where $\delta > 0$ is a preference weight in favor of autarky.

While there may have been some bias against participation in the allowance market, the opening of the allowance market ignited an unexpected wave of innovation leading to a significant reduction in the cost of abating SO₂ emissions, innovation which “went beyond the ability to trade” in the allowance market. It could be that the cap on emissions, A_k , made abatement more salient for senior managers and attracted their attention more strongly than would have occurred with a price signal, τ , alone.

An intriguing thought-experiment is: what would have happened if there had been a tax on coal based on its sulfur content rather than the cap and trade program for SO₂? This is equivalent to an upstream cap on sulfur. Would it have led to the same outcomes – the same reduction in aggregate SO₂ emissions at about the same economic cost? Such a tax/upstream cap would have fostered the use of low sulfur coal, including the development of mid-sulfur mines in the Midwest and Northern Appalachia. But, to the extent that the SO₂ cap functioned as a performance standard on emissions, a sulfur tax/upstream cap alone might not have triggered the same wave of innovation in the operation of scrubbers, and the same reduction in the cost of using them. Also, to the extent that managers’ behavior conformed to the structure in (2) with $\delta > 0$, there might not have been the same degree of degree of over-compliance in SO₂ abatement, and the same reduction in aggregate SO₂ emissions.

54. Tyler (1990).

55. Greene et al. (2009) argue for a similar loss aversion on the part of a consumers contemplating the purchase of a new, fuel-efficient vehicle. Risk aversion and loss aversion both create a convexity in the marginal valuation function. Loss aversion is similar to the “excess caution” which Borenstein et al. (2007) detected in natural gas distribution companies, leading them to forego surplus-increasing gas trades.

3.5 Comparison with Lead Trading

Before the SO₂ allowance trading program, there was the lead trading program which operated in the U.S. from 1982 to 1987. This was a success and is widely cited by proponents of emission trading. I consider it briefly here because it sheds light on some of the issues raised above in connection with the SO₂ program.

The phase-out of lead in gasoline in the U.S. began in July 1974 when the EPA both required the use of unleaded gasoline in new cars equipped with catalytic converters, and ordered gasoline retailers to offer unleaded gasoline for sale. In 1979, the EPA issued individual facility performance standards for refineries. These varied according to the size of the refinery. Large refineries were required to produce gasoline (leaded and unleaded combined) averaging no more than 0.5 gram of lead per gallon (gpg) per calendar quarter. Small refineries faced less stringent limits, starting out at 2.65 gpg to 0.8 gpg depending on the refinery size, because it was believed to be more difficult for them to achieve high octane levels without lead. However, small refineries were scheduled to reach 0.5 gpg by October 1982. In early 1982, the Reagan administration proposed to defer the 0.5 gpg limit for small refineries and also to consider relaxing the limit on large refineries. Following strong opposition, the administration announced that it was no longer planning to relax the standard and it would significantly narrow the definition of “small” refineries that were regulated less tightly. The final rule, issued in October 1982, changed the basis of lead regulation and introduced emission trading. The new rule set a standard that limited the allowable content of lead in leaded gasoline specifically, set at a quarterly average of 1.1 grams per leaded gallon (gplg).⁵⁶ The standard applied to large refineries immediately, and to small refineries starting July 1, 1983. In addition, the rule allowed *averaging* of lead content across refineries. The inter-refinery averaging was in effect a trading program. Each refinery could use each quarter an amount of lead (in grams) equal to 1.1 times the number of leaded gallons of gasoline it produced, with any shortfall in the amount of lead used being available for transfer to another refinery.

Further analysis by the EPA in 1984 showed that the lead standard could and should be tightened further. To provide refineries with some flexibility and time for adjustment, the standard was reduced to 0.5 gplg starting July 1, 1985, and 0.1 gplg starting January 1, 1986. Also, beginning in January 1985, refineries were permitted to bank excess lead rights for use in a future quarter. Trading and banking of lead rights were terminated at the end of 1987.⁵⁷

56. This was, in effect, slightly tighter than the current standard for large refineries (Nichols, 1997). See also Newell and Rogers (2003).

57. Starting in January 1988 a performance standard of 0.1 gplg was imposed on each refinery individually, with no option for trading. Lead was eventually banned as a fuel additive in the U.S. in 1996.

Thus, the lead trading program was used to bring the overall lead content of leaded gasoline sold in the U.S. down from a little over 1.1 gplg in 1982 to 0.1 gplg by the end of 1987.⁵⁸

Four points about the lead program should be noted. First, in the terminology introduced above, lead trading served as a mechanism for short-run flexibility, as opposed to a mechanism for the long-run reallocation of abatement. It smoothed the transition to a low lead standard. By January 1988, all refineries that produced leaded gasoline had found a way to reduce their individual lead usage to 0.1 gplg. The notion that lead trading was a mechanism for short-run flexibility is supported by the observation of Kerr and Mare (1998, p. 15) of a “lack of consistency” in trading behavior. They found in their data that two thirds of both consistent traders and occasional traders swapped sides of the market, switching between buying permits in some quarters and selling in others.

Second, Kerr and Mare found a significant amount of non-arms length trading, in the form of internal transfers among different refineries owned by the same company. In their data, 67% of the quantity of lead bought was bought within the same company, and 70% of the lead sold was sold within the same company.⁵⁹ This is consistent with a possible preference for autarky.

Third, although unleaded gasoline was more costly to produce than leaded gasoline, the cost increase was small: the wholesale price of unleaded gasoline was only 3-4% higher than that of leaded gasoline between January 1983 and January 1987; by January 1988, the wholesale price differential was less than 1%.⁶⁰ Thus, while lead trading may be regarded as an example of an upstream cap and trade program (Stavins, 2007), there was not much of a price signal downstream.

Fourth, there does not appear to have been a wave of innovation comparable to that which occurred with SO₂. Kerr and Newell (2003) focus on a technology known as isomerization used by refineries to replace octane when lead is restricted. The technology was introduced in the late 1960s. By 1980, the cost of isomerization had fallen by about 40%; during the 1980s, the cost fell hardly at all. More than half the adoption of isomerization occurred after 1986, when the phase-out of lead trading had been set.⁶¹

58. During this period, the share of leaded gasoline in total U.S. gasoline sales by volume fell from about 50% in January 1983 to 23% in January 1988 (Borenstein, 1991, Table 2).

59. Kerr and Mare (1998, p. 8).

60. Borenstein (1991, Table 2). The retail price of unleaded gasoline was 6-8% higher than that of leaded gasoline, but Borenstein finds the higher retail margin on unleaded gasoline to have been the result of local market power and price discrimination by gas stations. By contrast, the wholesale prices of gasoline seemed to be competitively determined.

61. Kerr and Newell (2003, Figures 2, 4). The short-run flexibility afforded by lead trading may have permitted some refineries to postpone the installation of isomerization, say, from 1985 to 1987.

4. EMISSION TRADING FOR GREENHOUSE GASSES

4.1 SO₂ versus GHGs: Similar or Different?

The successful use of emission trading in the U.S. for lead and SO₂ is commonly cited in support of emission trading for GHGs. In this section I consider the similarities between the past regulation of pollutants and the challenge posed by the regulation of GHGs.

GHGs might seem to be a better candidate for emission trading than SO₂ because there can be no “hot spots” for them: the environmental consequences of their emission in terms of climate change depend just on the aggregate volume of emissions, regardless of the location where the emissions occur.⁶² However, there are other physical and economic differences between SO₂ and GHGs which confound the comparison.

The emission of GHGs, especially CO₂, is profoundly dispersed throughout the economy, unlike SO₂. About 70% of SO₂ emissions in the U.S. in 1995 came from electricity generation, but only 34% of the GHG emissions in the U.S. today come from electricity generation.⁶³ Nearly as many GHG emissions (28%) come from transportation.⁶⁴ The reduction of SO₂ emissions engaged a relatively small number of decision makers – fewer than 300 power companies accounted for the vast majority of SO₂ emissions. Narrowing the focus of GHG reduction to electricity generation alone would leave the majority of U.S. emissions untouched. To accomplish a more substantial reduction in GHG emissions requires engaging with millions of actors in other sectors, including large industrial facilities, automobiles and other modes of transportation, agriculture, commercial buildings, businesses and households.

With SO₂, a 50% reduction in emissions was accomplished virtually overnight without any major technological breakthrough. The emission reduction was accomplished through technologies that had long been familiar. By 1995, scrubbers were a mature technology and, as noted above, if anything there appears to have been a decline in new patents relating to scrubber technology after emission trading was initiated. Burning low-sulfur coal in boilers designed for high-sulfur coal was a challenge for boiler operators and represented a significant operational advance, but was hardly a major technology breakthrough.

62. The impact depends on the *composition* of GHGs emitted (with methane, for example, being more potent in the short run than CO₂), but not the location. The consensus in the literature is that hot spots have not so far been a major problem with SO₂ allowance trading (Swift, 2000; Burtraw et al, 2005).

63. Electricity’s share of CO₂ emissions alone is 39% (EPA 2009c). CO₂ accounts for about 85% of the 7.15 billion metric tons of CO₂-equivalent GHG emissions in the United States in 2007; but methane, nitrous oxides, and other gases also contributed. About 94% of the CO₂ comes from the combustion of fossil fuels, with the rest from changes in land use (deforestation, etc). The methane comes mainly from landfills and cows.

64. In California, by contrast, electricity (including electricity imported from out of state) accounts for 22% of the CO₂ emissions, and transportation accounts for 40% of the emissions.

By contrast, reducing CO₂ emissions is more difficult. There is no low-CO₂ coal. CO₂ is the primary product of breaking down hydrocarbon chains; there is no end-of-pipe treatment per se. There is the possibility of carbon capture and sequestration, but this has hardly yet been practiced at the scale of a coal-fired generating plant. When this technology does become operational, it is unlikely to be applicable to existing power plants because retrofitting is estimated to be as expensive as building a brand new plant.⁶⁵

Working with existing power plants, one can reduce CO₂ emissions by changing the dispatch order – electricity generated from nuclear power and renewables has no CO₂ emissions, electricity from natural gas has about half the CO₂ emissions of coal, and electricity from oil is about half way between natural gas and coal. CO₂ can also be reduced by co-firing coal with biomass which, since it is renewable, has effectively zero carbon intensity. Until recently it was thought this could be done on a very limited scale (under 5% of biomass). It is now believed this can be done with up to 15% of biomass.⁶⁶ This may improve further with experience, as happened with the use of low-sulfur coal. But still, it seems unlikely that co-firing with biomass will be as much of a boon as low-sulfur coal was for SO₂ reduction.

Co-firing with biomass aside, the only way to significantly reduce CO₂ emissions from *existing* fossil-fuel power plants is to operate them less. For the electricity sector, there needs to be a sharper focus on influencing the design of *new* fossil-fuel power plants. In particular, it will be necessary to promote greater use of renewables and perhaps nuclear power. And, if new fossil-fuel plants are built, it is important that they attain higher thermal efficiencies than what is conventional now, using state-of-the-art technologies such as supercritical combustion or IGCC, and that they are designed from the beginning to accommodate carbon capture and sequestration once that technology has been refined and becomes available. Apart from this, the electricity sector today needs to focus on promoting conservation and energy efficiency. It should be noted that the SO₂ trading program had no measurable effect in promoting either the use of renewable energy or energy conservation.

There are three fundamental differences between GHGs and other the other air pollutants regulated by the Clean Air Act that affect the potential for cap-and-trade as a tool for climate policy. (1) We cannot rely on existing technologies to achieve the requisite emission reduction. It will be necessary to develop new technologies for the generation of electricity with low emissions of GHGs, for CO₂ capture, storage and sequestration, and for reducing the reliance on fossil fuels (e.g., electric or fuel cell vehicles). (2) There will have to be a very significant turnover of capital stocks in order to deploy the new technologies. The problem is that today we have the wrong types of power plants, the wrong types of motor vehicles (gas guzzlers rather than fuel-efficient, smaller vehicles), and

65. Massachusetts Institute of Technology (2007).

66. I owe this observation to Dallas Burtraw.

perhaps the wrong types of urban development with urban sprawl and limited public transit. With the other air pollutants, emission reduction was accomplished for the most part with existing capital assets: it was possible to retrofit existing production facilities with end-of-pipe treatment and/or to modify the production process. There was not a wholesale retirement of capital assets well before the end of their normal lives. It is unlikely that the major reductions in GHG emissions required by mid-century can be accomplished with existing capital stocks left mostly unscathed. (3) In the interim, while new technologies are being developed and capital is being replaced, it will be necessary to promote greater energy efficiency in major energy-using sectors of the economy for example by promoting a degree of behavioral change among energy users

This is a more daunting challenge than was faced when dealing with SO₂. With regard to the distinctive differences of GHGs just noted, Montgomery and Smith (2007) comment: “These features limit the need for and value of immediate emission reductions in GHGs, which is the primary and well-proven advantage of the cap-and-trade approach. At the same time, these features demand a policy prescription that is effective in stimulating future technology development without imposing inefficiently high costs in the near-term stages of the policy. The ability of the cap-and-trade approach to perform in the latter manner has not been demonstrated by any of its previous applications, nor has it been explored adequately in modeling or in theory.”⁶⁷

Besides these differences between GHGs and SO₂, the political climate in the U.S. today for reducing GHGs is more clouded than it was in 1990 for SO₂. During the 1980s, over 70 pieces of legislation had been assayed to deal with SO₂ emissions and related environmental problems. One major proposal in 1983 would have reduced SO₂ emissions by 10 million tons from pre-1980 levels (Burtraw and Palmer, 2004), which is about the same amount of reduction as was eventually required under the 1990 CAA Amendments.⁶⁸ By 1990, scrubbing was well accepted and there was finally a readiness to move forward using a cap-and-trade approach. But there was still disagreement about the size of the overall emission reduction: the electric utility industry wanted a total reduction of only 8 million tons of SO₂ by 2010, while environmentalists wanted a reduction of 12

67. They continue: “Thus, the main problem that cap-and-trade resolves so handily is how to find and induce use of the best controls from a large array of options. In practice it has almost entirely focused on existing technology, and making reductions that are economically feasible to achieve right now. In doing so, cap-and-trade applications have also demonstrated an ability to motivate innovation that further reduces costs of meeting caps beyond what would have been the least cost in the absence of such innovation. However, this incentive for innovation has only been demonstrated in situations where innovation could have immediate implications for cost savings (because the policies in question imposed the full force of the intended emissions reductions within only five years or less from program initiation), and where the ultimate cap would be achievable at a socially-acceptable cost using existing technologies, even if the innovations were not to have materialized.” In their view, that situation does not apply to GHGs.

68. 10 million tons includes mandated *non-utility* emission reductions required by the 1990 Amendments.

million tons. The final figure of 10 million tons was a compromise.⁶⁹ With CO₂ there exists less agreement, with the GHG reduction proposals in the U.S. ranging from about 5 to 15% relative to 2005 emission levels, and with some uncertainty as to whether there yet exists a sufficient majority of votes in the U.S. Senate to pass meaningful CO₂ emission legislation this year.

4.2 Upstream Cap or Downstream Cap?

A topic of active debate in policy circles is where emissions should be capped in the U.S. economy – upstream or downstream. Several economists have advocated an upstream approach, including Stavins (2007a), and Aldy and Pizer (2009). For example, the latter argue that the dispersed nature of CO₂ emissions throughout the economy makes regulation at the point of production/import both most complete and least cumbersome administratively. They emphasize that downstream approaches entail some incompleteness of coverage, which is economically inefficient. The most cost-effective emission reduction exploits abatement opportunities among as many sources as possible, while downstream approaches facilitate leakage. However, other policy analysts have advocated a downstream approach, arguing that firms are more likely to change their behavior in response to direct regulation, and will not react equally to changes in the prices of fuel inputs.⁷⁰ Aldy and Pizer (2009) characterize this as an “anecdotal” concern. Stavins (2007b, p. 9) rejects it, saying “There is no basis in economic theory nor any empirical evidence whatsoever for this proposition. On the contrary, the strong financial incentives for emissions reductions are identical with an upstream system, as powerful price signals are sent throughout the economy.”

As earlier noted, the equivalence between upstream and downstream caps arises because of certain assumptions built into the theoretical model, including the assumption of linearity (no risk aversion, no loss aversion, etc), and the static model formulation (no distinction between operating and capital costs, no investment decisions per se, etc). It is unclear whether the assumptions really hold. If firm behavior corresponded instead to the model in (2), whether because of risk aversion, a desire to comply with the law, a preference for autarky, or something else, this could justify a downstream approach: the cap, A_k , would influence a firm’s behavior, as well as the allowance price, τ .

Moreover, with SO₂ and lead trading, essentially no price signals were sent downstream from the capped sector. All the adjustment took place within the capped sector itself. Thus, the existing experience with emission trading provides no evidence one way or the other with regard to the strength of the price signals

69. This was an economically favorable deal for electric utilities: they accepted 2 million more tons of emission reduction in exchange for a trading system where they could make money by selling unused allowances.

70. This view was held by some members of California’s Market Advisory Committee (2008) who were “convinced that incentives for reducing emissions are strongest when downstream entities must submit allowances.”

radiating throughout the economy from the capped sector, or the magnitude of the response to those signals in other sectors of the economy.

Stavins (2007b) refers to the “the dramatic change that was anticipated and has been observed in the SO₂ allowance market” but, according to the major observers of the market, innovation occurred for reasons “beyond the ability to trade;” the innovation was characterized as “cost savings without emission trading.” While there was substantial trading activity, the full economic significance of the trading remains to be determined – it may have served to provide short-term flexibility rather than long-run, cost-minimizing abatement reallocation. While most of the firms in the regulated sector reduced their emissions, it is not yet clear whether this was in response to the caps that they faced, as opposed to the allowance price signal. At any rate, some firms clearly exhibited a degree of autarkic behavior and there was a significant amount of banking.

Aldy and Pizer are correct in saying that one ideally wants to equalize the marginal cost of abatement across all emission sources in the economy. But, this is not necessarily accomplished through the price signal from an upstream cap once one allows for behavioral considerations and/or recognizes the dynamic context in which abatement may occur. In the real world, short- and long-run elasticities are often different, in ways that vary across sectors, and equalizing the marginal cost of abatement across sectors in the long-run is not necessarily best accomplished by equalizing it in the short-run.

To illustrate the power of a price signal from an upstream cap, Aldy and Pizer (2009, p. 187) cite the recent run-up in energy prices. Consider the consumer response to an increase in retail gasoline prices. In 2005, the average retail price for regular gasoline in the U.S. was about \$2.24 per gallon. Imagine a 10% increase in that price, an increase of 22 cents. This increase is equivalent to the imposition of a carbon price of about \$25 per ton of CO₂. How might this affect a household’s vehicular CO₂ emissions? The Congressional Budget Office (2008) uses a short run price elasticity of demand for gasoline of -0.06. Thus, there might be an 0.6 per cent reduction in household gasoline usage in response to a carbon price of \$25/CO₂ ton.⁷¹ Arguably, a downstream cap on automobile manufacturers based on the emissions of their annual fleet of new vehicles, with the same carbon price of \$25/CO₂ ton, would do more to induce them to offer a line up of fuel-efficient vehicles.

Decision makers do respond to prices. But, they respond to prices *that are visible and salient to them*. Not all price changes are equally visible or salient to all decision makers. Moreover, decision makers respond to *other* factors as well, besides price.

71. To be sure, in the summer of 2008 at the height of the run-up in gas prices, there was a larger consumer response. Retail gas prices in the U.S. rose from an average of about \$2.92/gallon in June – August 2007 to an average of \$3.91 in June – August 2008, an increase equivalent to a carbon price of about \$113/CO₂ ton. While gas prices rose by 34%, vehicle miles travelled in the U.S. fell during this period by about 4.6%. The implicit short-run elasticity was therefore about 0.14. This higher elasticity reflects both the very high price and the large price increase. But still, the elasticity is not large.

When price changes are small, or the background level of prices is subject to significant variation, a given price change may not be very visible to decision makers. In 2005, for example, before the major run-up in gas prices, the average weekly retail price of regular gasoline in the U.S. varied from as low as \$1.75/gallon (first week of January) to as high as \$3.04 (first week of September); the average for the year was \$2.24. In that context, an increase of 22 cents per gallon due to a carbon price of \$25/CO₂ ton is not likely to be visible to most motorists; hence the limited predicted response. In addition to visibility, there is a growing literature in behavioral economics on the importance of salience. For example, Chetty et al. (2007) analyze the effect of salience on consumer's responses to prices in a supermarket. They find that when commodity taxes are included in the *posted* price that consumers see when shopping (and thus are more salient) they have a much larger effect on consumer demand. Similarly, Gallagher and Muehlegger (2008) show that more salient sales tax waivers given at the time of purchase have seven times as large an effect on hybrid vehicle purchases as income tax credits of an equivalent amount.⁷² In the terminology of Gabaix and Laibson (2006), such studies show that prices may become "shrouded attributes."

Furthermore, even if a price is visible and salient, it may not always influence behavior as theory supposes. In the world of economic theory, firms are considered to act as a single, unitary decision maker, with a single objective, namely profit maximization. In real life, there are multiple decision makers and, quite possibly, multiple objectives. They have different responsibilities, they face different incentives, which are subject to the principal-agent problem.⁷³ Thus, the decision-maker when a firm faces a cap on its emissions is not necessarily the same as when it faces an increase in its fuel price, and the response may be different. An emissions cap may turn out to be a more visible and more salient trigger than an increase in fuel price.

4.3 Why Complementary Measures?

With SO₂, not only was a 50% reduction in emissions accomplished virtually overnight, but the only policy action required of the federal government was to establish the allowance trading program under Title IV. Will the same hold true for CO₂ emissions – will allowance trading alone secure whatever reduction in CO₂ emissions is eventually mandated? Here, too, opinion is divided.

72. Chetty (2009) provides a number of additional examples. He points out that, when salience is a factor, "contrary to conventional wisdom ... the economic incidence of a tax depends on its statutory incidence." The same can apply in the production context with a tax versus a cap. To the extent that visibility and salience mute the response to price, they sever the tight link between demand elasticity and marginal cost of abatement. Inferring the marginal cost of abatement from the response to a price change may require some additional manipulation analogous to that applied by Chetty for consumer behavior.

73. As Borenstein et al. (2007) demonstrate, this can induce excessively cautious behavior, leading the agent to forego surplus maximizing market transactions.

There is general agreement of the need for some form of R&D policy for GHG mitigation because, as noted earlier, low-carbon and zero-carbon technologies are still in their infancy. Private incentives to invest in R&D are weakened by the public good nature of knowledge which reduces the ability to appropriate all the rents from an innovation. Thus, Aldy and Pizer argue for a government-sponsored climate-oriented R&D policy to complement mitigation policy; they recommend this be targeted especially at basic research. They also point out that policies which impose a price on GHG emissions, such as cap and trade or a carbon tax, would also induce technological change.

As earlier noted, the actual experience with emission trading for SO₂ and lead was mixed. Emission trading promoted the adoption of technologies which were already available. For SO₂, at least, the evidence is that the rate of patenting declined when emission trading began. Schumpeter famously identified three stages in the process of technological change: invention, innovation and diffusion. Invention is the first development of a scientifically or technically new product or process, which may involve both basic and applied research. Innovation is accomplished when the new product or process is commercialized, i.e., made available on the market. Diffusion is when the product or process comes to be widely used through adoption by many firms or individuals. In the case of climate change, invention and innovation are the core issues – the development and commercialization of technologies that do not exist yet or, at best, are still highly experimental (e.g., carbon capture and sequestration). With emission trading for SO₂ and lead, in Schumpeter's terminology, while there was diffusion, there was no invention or innovation.⁷⁴

R&D policy aside, some believe that a cap and trade program should suffice for a GHG mitigation policy while others believe that complementary measures will be required. The Regional Greenhouse Gas Initiative (RGGI) set up by ten Northeastern and Mid-Atlantic states relies solely on cap-and-trade.⁷⁵ By contrast, the seven states which constitute the U.S. members of the Western Climate Initiative (WCI) have determined that GHG emission reduction requires a suite of other policies working in concert with cap and trade. These complementary policies include efficiency standards for motor vehicles (a GHG emission limit per vehicle mile), energy efficiency standards for appliances and buildings, a renewable energy standard, a low-carbon fuels standard, and government procurement policies. At the national level, the legislation currently being considered by Congress, HR 2454 (Waxman-Markey), calls for a renewable portfolio standard and energy efficiency standards for buildings, appliances and vehicles.

Stavins (2007b) and Aldy and Pizer (2009) regard such complementary measures as unwise. They fear that these measures may only raise the cost of meeting a GHG reduction target because their narrow focus relative to that of

74. Except for the burning of low-sulfur coal in boilers originally designed for high-sulfur coal.

75. At least 25% of the allowances have to be auctioned, with auction revenues directed to strategic energy investments.

an (upstream) cap and trade program may edge out other, cheaper abatement alternatives.

The force of this argument depends on one's assessment of the potency of a price signal from an (upstream) cap and trade program. As noted earlier, some of the benefits of the existing trading programs for SO₂ and lead appear to have been due to the fact that the cap functioned as a performance standard, rather than to the reallocation of abatement driven by a price signal. To the extent that the complementary measures are crafted as performance standards they may be similarly efficacious.

Much of the existing theoretical literature on R&D and pollution abatement assumes that the firm which does the invention is *the same* as the firm which causes pollution and invests in abatement. In fact, however, this is generally false: the vast majority of the inventors are not the polluters but rather machinery suppliers and other outside sources.⁷⁶ The fact that different parties engage in invention, innovation and pollution abatement creates the possibility of what is known as a *coordination problem* (Rodrik, 1996); the coordination problem arises because each of the various actors does not know the others' expectations and intentions. For an innovation to be successful, these expectations need to be coordinated, especially when the innovation is costly and capital intensive. The polluting firm must be reasonably confident it can recoup the investment if it invests in an expensive new production process that lowers GHG emissions. The equipment supplier, likewise, must be reasonably confident that there will be an industrial market before he invests in commercializing this product. And the venture capitalist must be reasonably confident that the product will be marketable if it is successfully commercialized before he finances the R&D. Given the time lags and uncertainties, the coordination problem is likely to be especially severe for many carbon control technologies. How likely is it that a carbon price signal from an (upstream) cap and trade program will generate sufficient coordination? May additional policy instruments not be required?

This point of view underpins the recommendation for complementary measures along with allowance trading. The complementary measures are targeted at particular groups of decision makers who play a strategic role in the reduction of GHG emissions. In addition to the carbon price signal generated by allowance trading in capped sectors, the automotive GHG emission standards target the automobile manufacturers; the low carbon fuel standard targets the oil companies and may discourage them from committing investment funds to the development of highly CO₂-intensive oil from Canadian tar sands; the building efficiency standards target developers and owners; the appliance efficiency standards target appliance manufacturers; and the renewable portfolio standard targets the retail

76. Lanjouw and Mody (1996) estimate that machinery suppliers were the source of about 80% of the patents for the control of industrial air pollution, water pollution, oil spills, and the exploitation of non-fossil fuel energy sources. Taylor (2007) shows that electric utilities and oil companies accounted for only about 18% of the patents for SO₂ control, while 82% of the patents are held by research institutions and, especially, other entities.

sellers of electricity. Taken together, and in combination with a cap and trade program, they are likely to induce more decisive actions by a broader group of decision makers than the price signal from allowance trading alone.

It is sometimes said that complementary measures are not needed because the cap in a cap and trade system will by itself ensure the desired reduction in emissions. The market *has to* clear, albeit at some price. But this argument misses the point. If the market price is excessively high, it will become politically unsustainable. The market will be suspended, either because a price cap mechanism kicks in, or because politicians intervene in an ad hoc way and suspend or bypass the market. If the price cap is triggered too frequently, or trading is suspended too frequently, this could destroy the credibility of the market. Thus, there is a public policy interest in not having the market clear at an excessively high price. This is where complementary measures can play a useful role. They have the potential to change the demand and supply curves for emission reductions and make them more price elastic. If used successfully, they reduce the possibility that the market clearing price reaches an excessively high and politically unsustainable level.

Aldy and Pizer (2009, p. 202) see the issue differently. Because complementary measures accelerate the deployment of existing technologies, they argue, “this will tend to lower the allowance price in the emissions market, as the new regulation takes some of the pressure off the cap and trade program to reduce emissions. This lower allowance price lowers the incentive for innovation and technological adoption in those industries that are *not* targeted by the performance standard.” Which of us is right depends in part on the stringency of the cap. If the cap is quite tight, high allowance prices may become a problem. If the cap is lax, low allowance prices may become a problem; but, if so, this does not necessarily produce a strong case for relying on emission trading alone.

While I think that well-crafted complementary measures may do more than the price signal from an allowance market alone to ignite a period of innovation in emission reduction, there are two caveats. First, at some point there needs to be flexibility to trade off between the measures. Second, to the extent that the allowance market becomes more liquid over time, and also that there is learning by doing and/or induced technical change in abatement, there is less need for the supplementary measures and they can eventually be phased out. I see them primarily as mechanisms to smooth the transition to GHG reduction.

5. CONCLUSIONS

The purpose of this paper is not to attack cap and trade. I do not doubt that emission trading for SO₂ or lead was economically very beneficial, and I agree that emission trading should be one of the policy instruments used now for GHGs. The article’s purpose is to argue the need for a more nuanced understanding of how emission trading has functioned in the past, and how it may function for GHGs. Against the conventional wisdom, I have argued for a

downstream system, accompanied by complementary measures. I have suggested that some of the benefits attributed to allowance trading were due to the use of downstream caps that functioned as a performance standard and provided short-run flexibility, rather than a long-run reallocation of abatement effort. Allowance trading triggered a wave of adoption (in Schumpeter's sense) and operational experimentation. Whether this was due just to the price signal or also the cap is the question that needs to be answered.

While I have stressed the behavioral elements of firm and household decision making regarding GHG reduction, I recognize that behavior evolves as preferences change and new information is acquired. People may come to see things differently: risk-aversion, loss aversion, and a preference for autarky are attitudes that can change. If they do, this will increase the potency of market price signals.

Also, while I have stressed the features of GHGs which may make them a more daunting challenge than SO₂ or lead, I recognize that "history is always a surprise and innovation takes many forms."⁷⁷ A cap and trade system creates incentives; not all the responses to those incentives can be predicted ahead of time and there may be some surprises, as was the case with SO₂. But, what causes the surprises – is it just the price signal or also the cap?

REFERENCES

- Aldy, Joseph E., and William A. Pizer (2009). "Issues in Designing U.S. Climate Change Policy." *The Energy Journal* 30(3): 179-209.
- Bailey, Elizabeth M. (1996). *Allowance Trading and State Regulatory Rulings: Evidence from the U.S. Acid Rain Program*, MIT-CEEPR Working Paper 96-002. Massachusetts Institute of Technology.
- Baldursson, Fridrik M., and Nils-Henrik M. von der Fehr (2004). "Price Volatility and Risk Exposure: On Market-based Environmental Policy Instruments," *Journal of Environmental Economics and Management* 48: 682-704.
- Ben-David, Shaul, David Brookshire, Stuart Burness, Michael McKee, and Christian Schmidt (2005). "Attitudes Toward Risk and Compliance in Emission permit Markets," *Land Economics* 76(4): 590-600.
- Borenstein, Severin (1991). "Selling Costs and Switching Costs: Explaining Retail Gasoline Margins." *RAND Journal of Economics* 22(3): 354-369.
- Borenstein, Severin, Meghan Busse and Ryan Kellogg (2007). "Principal-Agent Incentives, Excess Caution, and Market Inefficiency: Evidence from Utility Regulation." University of California Energy Institute Working Paper CSEM WP 174. Berkeley CA.
- Burtraw, Dallas (2000). "Innovation Under the Tradable Sulfur Dioxide Emissions Permits Program in the US Electricity Sector." Resources for the Future Discussion Paper 00-38.
- Burtraw, Dallas (1996). "The SO₂ Emissions Trading Program: Cost Savings without Allowance Trades," *Contemporary Economic Policy* 14 (April): 79-94.
- Burtraw D, and K. Palmer (2004). "The SO₂ Cap-and-Trade Program in the United States: a "Living Legend" of Market Effectiveness." In Winston Harrington, Richard D. Morgenstern and Thomas Sterner, eds., *Choosing Environmental Policy: Comparing Instruments and Outcomes in the United States and Europe*. Resources for the Future: Washington D.C.

77. This line is taken from Frank Convery.

- Burtraw, Dallas, David A. Evans, Alan Krupnick, Karen Palmer and Russell Toth (2005). "Economics of Pollution Trading for SO₂ and NOx." *Annual Review of Environment and Resources* 30: 253-289.
- Bushnell, James B., and Catherine Wolfram (2005). "Ownership Change, Incentives and Plant Efficiency: The Divestiture of U.S. Electric Generation Plants." *Center for the Study of Energy Markets Working Paper CSEMWP-140*. University of California, Berkeley CA.
- Busse, Meghan R. and Nathaniel O. Keohane (2007). "Market Effects of Environmental Regulation: Coal, Railroads, and the 1990 Clean Air Act." *RAND Journal of Economics* 38(4): 1159-1179.
- Carlson, Dale A. and Anne M. Sholtz (1994). "Designing Pollution Market Instruments: Cases of Uncertainty" *Contemporary Economic Policy* XII: 114-125.
- Carlson, Curtis P., Dallas Burtraw, Maureen Cropper, and Karen Palmer (2000). "SO₂ Control by Electric Utilities: What are the Gains from Trade?" *Journal of Political Economy* 108(6): 1292-1326.
- Chao, Hung-Po and Robert Wilson (1993). "Option Value of Emission Allowances." *Journal of Regulatory Economics* 5:233-249.
- Chetty, Raj (2009). "The Simple Economics of Salience and Taxation." NBER Working Paper 15246. Cambridge, MA.
- Chetty, Raj, Adam Looney and Kory Kroft (2007). "Salience and Taxation: Theory and Evidence." NBER Working Paper 13330. Cambridge, MA.
- Coggin, J. S. and J. R. and Swinton (1996). "The Price of Pollution: a Dual Approach to Valuing SO₂ Allowances." *Journal of Environmental Economics and Management* 30(1):58-72
- Crocker, Thomas D. (1966). "The Structuring of Atmospheric Pollution Control Systems," In H. Wolozin, ed., *The Economics of Air Pollution* New York: Norton. 61-86.
- Dales, J. H. (1968). *Pollution, Property and Prices* University of Toronto Press.
- Ellerman, A. D. (2003). "Ex Post Evaluation of Tradable Permits: the US SO₂ Cap-and-Trade Program." Working Paper CEEPR 03-003, Massachusetts Institute of Technology.
- Ellerman, A. Denny, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth Bailey (2000). *Markets for Clean Air: The U.S. Acid Rain Program*. Cambridge University Press
- Ellerman A. Denny and Florence Dubroeuq (2004). "The Sources of Emission Reductions: Evidence from U.S. SO₂ Emissions from 1985 through 2002" Working Paper CEEPR 04-001, Massachusetts Institute of Technology.
- Ellerman A. Denny and Juan-Pablo Montero (2007). "The Efficiency and Robustness of Allowance Banking in the U.S. Acid rain Program." *The Energy Journal* 28(4): 47-71.
- Ellerman, A. Denny and Juan-Pablo Montero (1998). "The Declining Trend in Sulfur Dioxide Emissions: Implications for Allowance Prices." *Journal of Environmental Economics and Management* 36(1): 26-45
- Fabrizio, Kira A., Nancy L. Rose and Catherine D. Wolfram (2007). "Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency." *American Economic Review* 97(4) 1250-1277.
- Farrell, Alexander E. and W. Michael Hanemann (2009). "Field Notes on the Political Economy of California Climate Policy." In Henrik Selin and Stacy Van Deveer, eds., *Changing Climates in North America: Institutions, Policymaking and Multilevel Governance* MIT Press, Cambridge, MA. 87-109.
- Gabaix, Xavier and David Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121(2): 505-540.
- Gallagher, Kelly Sims and Erich Muehlegger (2008). "Giving Green to Get Green: Incentives and Consumer Adoption of Hybrid Vehicle Technology." John F. Kennedy School of Government Faculty Research Working Paper RWP08-009. Cambridge, MA: Harvard University.
- Goeree, Jacob K., Charles A. Holt, Karen Palmer, William Shobe and Dallas Burtraw (2009). "An Experimental Study of Auctions versus Grandfathering to Assign Pollution Permits." Institute for Empirical research in Economics Working Paper No. 429. University of Zurich.
- Goulder, Lawrence H. and Ian W. H. Parry (2008). "Instrument Choice in Environmental Policy." *Review of Environmental Economics and Policy* 2: 152-174.

- Greene, D.L., J. German and M.A. Delucchi (2009). "Fuel Economy: The Case for Market Failure," in D. Sperling and J.S. Cannon, eds., *Reducing Climate Impacts in the Transportation Sector*, Springer Science.
- Hahn, Robert W. (1984). "Market Power and Transferable Property Rights." *Quarterly Journal of Economics* 99(4): 753-765.
- Hanemann, Michael (2008). "California's New Greenhouse Gas Laws" *Review of Environmental Economics and Policy* 2(1): 114-129.
- Hennessy David A., and Jutta Roosen (1999). "Stochastic Pollution, Permits, and Merger Incentives." *Journal of Environmental Economics and Management* 37: 211-232
- Holland, Stephen P. and Michael Moore (2008). "When to Pollute, When to Abate? Intertemporal Permit Use in the Los Angeles NOx Market." NBER Working Paper 14254 Cambridge, MA..
- Innes, Robert (2003). "Stochastic Pollution, Costly Sanctions, and Optimality of Emission Permit Banking," *Journal of Environmental Economics and Management* 45: 546-568.
- Keohane, Nathaniel O. (2007). "Cost Savings from Allowance Trading in the 1990 Clean Air Act: Estimates from a Choice-Based Model." In Jody Freeman and Charles D. Kolstad, eds., *Moving To Markets in Environmental Regulation*. New York: Oxford University Press. 194-229.
- Kerr, Suzi and David Mare (1998). "Transaction Costs and Tradable Permit Markets: The United States Lead Phasedown." Working paper, Motu Economic and Public Policy Research Trust, New Zealand. Available at SSRN: <http://ssrn.com/abstract=1082596>.
- Kerr, Suzi and Richard G. Newell (2003). "Policy-Induced Technology Adoption: Evidence from the U.S. Lead Phasedown." *The Journal of Industrial Economics* LI(3) 317-343.
- Kling, Catherine and Jonathan Rubin (1997). "Bankable Permits for the Control of Environmental Pollution." *Journal of Public Economics* 64: 101-115.
- Kreutzer, Jacob (2006). "Cap and Trade: A Behavioral Analysis of the Sulfur Dioxide Emissions Market." *New York University Annual Survey of American Law* 62: 125-162.
- Lanjouw, Jenny Olson and Ashoka Mody (1996). "Innovation and the International Diffusion of Environmentally Responsive Technology." *Research Policy* 25: 549-571.
- Lile, R. and D. Burtraw (1998). "State-level Policies and Regulatory Guidance for Compliance in the Early Years of the SO₂ Emission Allowance Trading Program." Resources for the Future Discussion Paper 98-35. Washington, DC
- Market Advisory Committee (2008). "Recommendations for Designing a Greenhouse Gas Cap-and-Trade System for California: Recommendations to the California Air Resources Board." June 30. Sacramento, CA.
- Massachusetts Institute of Technology (2007). *The Future of Coal: Options for a Carbon-Constrained World*, Massachusetts Institute of Technology, Cambridge, MA,
- Melitz, Marc J. (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71(6): 1695-1725.
- Montero, Juan-Pablo (1997). "Marketable Pollution Permits with Uncertainty and Transaction Costs." *Resource & Energy Economics* 20: 27-50.
- Montgomery, W. David (1972). "Markets in Licenses and Efficient Pollution Control Programs" *Journal of Economic Theory* 5: 395-418.
- Montgomery, W David and Anne E. Smith (2007). "Price, Quantity, and Technology Strategies for Climate Change Policy." In Michael E. Schlesinger, Haroon S. Kheshgi, Joel Smith, Francisco C. de la Chesnaye, John M. Reilly, Tom Wilson and Charles Kolstad, eds., *Human-Induced Climate Change: An Interdisciplinary Assessment* Cambridge: Cambridge university Press. 328- 342.
- Nelson, Randy A., and Walter J. Primeaux (1988). "The Effects of Competition on Transmission and Distribution Costs in the Municipal Electric Industry." *Land Economics* 64(4): 338-346.
- Newell, Richard G. and Kristian Rogers (2003). "The Market-based Lead Phasedown." Resources for the Future Working Paper 03-37. Washington, DC.
- Nichols, Albert L. (1997). "Lead in Gasoline." In Richard D. Morgenstern, ed., *Economic Analyses at EPA: Assessing Regulatory Impact* Washington, DC: Resources for the Future, 49-86.

- Pizer, William A. (2007). "Scope and Point of Regulation for Pricing Policies to Reduce Fossil-Fuel CO₂ Emissions." In Raymond J. Kopp and William A. Pizer, eds., *Assessing U.S. Climate Policy Options*. Washington, DC: Resources for the Future. 69-78.
- Popp, David (2003). "Pollution Control Innovations and the Clean Air Act of 1990" *Journal of Policy Analysis and Management* 22(4):641-60
- Reguant, Mar and A. Denny Ellerman (2008). "Grandfathering and the Endowment Effect." MIT Center for Energy and Environmental Policy Research, Working Paper 08-018.
- Rose K., A. S. Taylor, and M. Harunuzzaman (1993). Regulatory Treatment of Electric Utility Compliance Strategies, Costs and Emission Allowances, The National Regulatory Research Institute, Ohio State University.
- Rodrik, Dani (1996). "Coordination Failure and Government Policy: a Model with Applications to East Asia and Eastern Europe." *Journal of International Economics* 40 (1-2): 1-22.
- Rousse, Olivier and Benoit Sevi (2005). "Behavioral Heterogeneity in the US Sulfur Dioxide Emissions Allowance Trading Program" Paper presented at the 45th Congress of the European Regional Science Association, Vrije University, Amsterdam
- Rubin, Jonathan D. (1996). "A Model of Intertemporal Emission Trading, Banking and Borrowing." *Journal of Environmental Economics and Management* 31: 269-286.
- Sandoff, Anders and Gabriela Schaad (2009). "Does EU ETS Lead to Emission Reductions Through Trade? The Case of the Swedish Emissions Trading Sector Participants." *Energy Policy* (37): 3967-3977.
- Schennach, Susanne M. (2000). "The Economics of Pollution Permit Banking in the Context of Title IV of the 1990 Clean Air Act Amendments," *Journal of Environmental Economics and Management*, 40: 189-210
- Stavins, Robert N. (2007a). *A US Cap-and-Trade System to Address Global Climate Change* The Hamilton Project, Discussion Paper 2007-13. Washington, DC: The Brookings Institution.
- Stavins, Robert N. (2007b). "Comments on the Recommendations of the Market Advisory Committee to the California Air Resources Board, 'Recommendations for Designing a Greenhouse Gas Cap-and-Trade System for California.'" Letter to Winston Hickox, Chair, Market Advisory Committee, June 15.
- Stavins, Robert N. (1995). "Transaction Costs and Tradeable Permits" *Journal of Environmental Economics and Management* 29: 133-148.
- Stevenson, Rodney (1982). "X-Inefficiency and Interfirm Rivalry: Evidence from the Electric Utility Industry." *Land Economics* 58(1): 52-66
- Swift, Byron (2000). "Allowance Trading and SO₂ Hot Spots: Good News from the Acid Rain Program," *Environment Reporter*, 31:19 (May 12): 954-59. Washington, DC: BNA, Inc.
- Swinton, J. R. (2004). "Phase I Completed: an Empirical Assessment of the 1990 CAAA." *Environmental and Resource Economics* 27: 227-46.
- Swinton, J. R. (2002). "The Potential for Cost Savings in the Sulfur Dioxide Allowance Market: Empirical Evidence from Florida." *Land Economics* 78(3):390-404.
- Tyler, Tom R.(1990). *Why People Obey the Law*. New Haven: Yale University Press.
- Taylor, Margaret R. (2008). "Induced Technical Innovation under Cap-and-Trade Abatement Programs." Working Paper, Goldman School of Public Policy, University of California, Berkeley.
- Taylor, Margaret (2001). *The Influence of Government Actions on Innovative Activities in the Development of Environmental Technologies to Control Sulfur Dioxide Emissions from Stationary Sources*. PhD thesis. Carnegie Mellon University, Pittsburgh, PA
- Taylor, M.R., E. S. Rubin, and D. A. Hounshell (2003). "The Effect of Government Actions on Technological Innovation for SO₂ Control." *Environmental Science and Technology* 37 (20): 4527-4534.
- U.S. Energy Information Administration (EIA) (2009). *Electric Power Annual 2007*. Report DOE/EIA-0348(2007). Washington, D.C.
- U.S. Environmental Protection Agency (2009a). *Acid Rain and Related Programs: 2008 Emission, Compliance and Market Data*. Washington, D.C.

- U.S. Environmental Protection Agency (2009b). "Allowance Markets Assessment: A Closer Look at the Two Biggest Price Changes in the Federal SO₂ and NO_x Allowance Markets." White Paper, accessed at <http://www.epa.gov/airmarkt/resource/docs/marketassessmnt.pdf>
- U.S. Environmental Protection Agency (2009c). *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2007*. EPA 430-R-09-004. Washington, D.C.
- U.S. Environmental Protection Agency (2008). *Acid Rain and Related Programs 2007 Progress Report*. Washington, D.C.
- U.S. Environmental Protection Agency (2006). *Acid Rain Program 2005 Progress Report* EPA-430-R-06-015. Washington, D.C..
- U.S. General Accounting Office (1994). *Allowance Trading Offers an Opportunity to Reduce Emissions at Less Cost*. Report RCED-95-30. Washington, D.C.
- Vine, E. (2003). "Opportunities for Promoting Energy Efficiency in Buildings as an Air Quality Compliance Approach." *Energy* 28: 319-341.
- Wirl, Franz and Juerhen Noll (2008). "Abatement and Permits when Pollution is Uncertain and Violations are Fined" *Environmental & Resource Economics*, 40: 299-312.
- Zhao, Jinhua (2003). "Irreversible Abatement Investment under Cost Uncertainties: Tradable Emission Permits and Emissions Charges" *Journal of Public Economics* 87: 2765-2789.