CEEPR Center for Energy and Environmental Policy Research

Considerations for Designing a Tradeable Permit System to Control SO_2 Emissions in China (Including Chinese translation courtesy of Hongjun Kan, RAND Graduate School of Policy Studies)

by

A. Denny Ellerman

01-009 WP **October 2001**

A Joint Center of the Department of Economics, Laboratory for Energy and the Environment, and Sloan School of Management

CONSIDERATIONS FOR DESIGNING A TRADABLE PERMIT SYSTEM TO CONTROL SO_2 EMISSIONS IN CHINA

A. Denny Ellerman[1]

1 INTRODUCTION

As China's economy has grown, atmospheric pollution has become a greater problem and a matter of increasing concern to policymakers at all levels of government. One of the principal pollutants has been sulfur dioxide (SO_2), which is emitted in varying intensity when coal, China's most abundant fossil energy resource, is burned. Excessive SO_2 emissions can cause serious health problems locally from high ambient concentrations, as well as non-health-related damages that can occur from acidification at some distance from the source of emissions.

In recent years, Chinese environmental authorities have expressed interest in the use of tradable permits as a regulatory instrument to control SO_2 emissions. This interest arises from the greater priority now placed on controlling pollutants in China, the increasing attention given throughout the world to the use of market-based instruments as a means of achieving environmental objectives, and the successful use of tradable permits to reduce and to limit SO_2 emissions in the United States, another large country in which coal use is significant. Moreover, the least-cost property of market-based instruments is particularly appealing in China, where the competition to meet social needs is great and the income level is relatively low.

[1] Senior lecturer, Sloan School of Management, and Executive Director, Center for Energy and Environmental Policy Research at the Massachusetts Institute of Technology. Comments from Stephanie Benkovic, Noreen Clancy, Paulette Middleton, and Hongjun Kan are gratefully acknowledged. I am especially indebted to Lu Shen, candidate for master's degrees from the Sloan School and the John F. Kennedy School of Government at Harvard University, for assistance in the research underlying this paper, and to Hongjun Kan for the translation of this paper into Chinese.

This paper focuses on the two main issues that will have to be considered in designing a tradable permits program for the control of SO_2 emissions in China: the transition from non-tradable facility permits to tradable emission permits and the integration of a tradable permits program with the preexisting pollution levy system. These two issues are discussed in the third and fourth sections of the paper. The second section briefly summarizes what must be the starting point for the implementation of a tradable permits system, namely, the existing structure of Chinese policy for controlling SO_2 emissions. Reference is made throughout the paper to the experience with SO_2 emissions trading under the U.S. Acid Rain Program (Ellerman *et al.*, 2000), which offers an instructive example as the first large-scale application of tradable permits for controlling emissions of any kind.

Certain issues have been glossed over here in order to keep the discussion and the length of the paper manageable; three of these are sufficiently important to be noted. First, no attempt is made to consider comprehensively the merits of tradable permits relative to alternative instruments for controlling pollution, such as taxes or command-and-control regulation (technology mandates, efficiency standards, emission rate limits, etc.). Comparisons between the requirements of a tradable permits system and those of tax or more conventional regulatory systems will be unavoidable here, but these comparisons do not provide a complete discussion of the relative merits of these instruments. Readers interested in such a discussion are referred to virtually any standard textbook on environmental economics (for instance, Tietenberg, 2000).

Second, this paper does not address in full detail the standard-setting, monitoring, and enforcement capabilities that will be required to successfully control SO_2 emissions in China, regardless of instrument choice. Instead, a bold simplifying assumption is made: that China will successfully control SO_2 emissions at an appropriate level regardless of the instrument used. This optimistic assumption allows the discussion to focus on the unique requirements of a tradable permits system and how they differ from those of other instruments for controlling emissions. A major theme throughout the paper is that the requirements for establishing an effective tradable permits system do not greatly differ from those for an equally effective tax or command-and-control regime. Although each instrument has distinctive features, the differences are mainly ones of form. All require that the same fundamental problems be solved: How to allocate the cost burden of reducing emissions, what specific requirements to place on emitting sources, and how to ensure compliance.

Third, the transitional nature of the Chinese economy is recognized but not discussed in detail. China's transition is twofold, from a pre-industrial and socialist economy to one that is industrial and market-oriented. As suggested by their name, market-based instruments presume markets; and, where the set of laws, institutions, and practices that are associated with market economies are not fully developed, these instruments, whether they be taxes or tradable permits, will experience a transitional phase. The conditions characterizing this transition do not require that the goal of efficient and effective environmental control be abandoned, but it should be recognized that successful implementation of market-based instruments will depend on the pace of the more general economic transformation.

2 THE EXISTING FRAMEWORK FOR SO_2 CONTROL IN CHINA[2]

2.1 General Context

Responsibility for the development and implementation of environmental policy in China is split between national and local levels. In general, the central government provides policy direction and the general legal and institutional framework, while local levels of government are responsible for implementation and enforcement, often including the choice of appropriate measures to achieve national goals. At the national level, the two principal bodies are the State Council, which provides broad policy guidance, and the State Environmental Protection Agency (SEPA), which is the administrative agency charged with the development and elaboration of this policy. At the local level, the Environmental Protection Bureau (EPB) is responsible. This division of responsibility is similar in some ways to that in the United States, where primary and secondary standards for criteria pollutants are established at the national level and states are expected to come into compliance with these standards through State Implementation Plans.

Despite this similarity in structure, three salient differences should be noted in any comparison with the United States. First, the devolution of authority to the local level is greater in China than in the United States, at least at this stage of policy development.

[2] This section draws heavily from Benkovic (1999) and Luo *et al.* (2000).

Second, China permits more experimentation at the local level, typically with pilot programs, than the United States ever did. Moreover, the central authorities in China encourage this experimentation to determine effective measures that can be adopted and propagated on the national level. Third, the significant devolution of responsibility to the local level and the emphasis on experimentation result in a more incremental approach to policy development in China than in the United States. One would search in vain for a Chinese analogue to the 1970 Clean Air Act, which has provided the basic concepts and structure for the elaboration of air emissions policy in the U.S. for the past three decades. Instead, in a process that can be described as both pragmatic and *ad hoc*, Chinese policy emerges from successive steps, each of which provides guidance and a progressively stronger legal basis for specific actions at the local level.

Initially, in the late 1970s and 1980s, environmental policy in China could be characterized as taking a selective and direct project- or process-specific approach. Certain projects, known as National Environmental Pollution Treatment Projects, new projects, and some highly polluting processes were designated for controls and for special allocations of resources in the national five-year plans; but, unless singled out by process or project, most existing facilities were unaffected. During the 1990s, environmental policy in China assumed an increasingly comprehensive and indirect character by which broad policy goals and authority were developed at the national level and implementation was delegated to local authorities.

This evolution can be seen in the emergence of the policy framework for controlling SO_2 emissions. The first general measure to address SO_2 emissions dates back to 1982 when the pollution levy, which was introduced in the late 1970s, was applied to industrial SO_2 emissions. More serious attention began to be given to SO_2 emissions in 1990 with the State Council's "Suggestions on the Development of Acid Rain Control." This procedural document provided the guiding concept of Two Control Areas, designating the areas to receive priority in controlling emissions, and it generally cleared the way for more concrete actions in subsequent years. The first of a series of extensions in coverage and increases in the rate of the SO_2 pollution levy took place in 1992. The Ninth Five-year Plan in 1996 introduced the concept of Total Emissions Control, which shifted the focus of regulatory attention from emission rates to total emissions. Finally, in April 2000, the People's Congress adopted sweeping changes to the 1987 Air Pollution Prevention and Control Law (APPCL) that incorporate the policies and measures developed during the 1990s and provide a stronger legal basis for their implementation. These changes focus regulatory efforts on the most polluted areas, changing the emphasis of control from emission rates to total emission discharges, shifting the base of the pollution levy from excess emissions to total emissions, and establishing emission permits as the vehicle by which national policy would be implemented at the local level.

This last provision was an important step away from a centrally directed, project-specific approach to a more decentralized and comprehensive structure for controlling SO_2 emissions. The conditions governing emissions are to be specified in permits issued to individual facilities by the local authorities. Although frequently called emission permits in China, these non-tradable permits governing emissions are called *facility permits* in this paper to distinguish them from tradable (emission) permits. Facility permits were tried on a pilot project basis in sixteen cities starting in 1991, and the basis for generalizing their use was provided by the year 2000 revisions of the APPCL. Although non-tradable, facility permits are correctly seen as a precondition for emissions trading, and in fact some limited trading has occurred in the sixteen trial cities. The transition from facility permits to tradable permits will be the subject of the next chapter, but first we turn to a more extended discussion of the three principal components of existing SO_2 emissions control policy: the Pollution Levy System (PLS), Two Control Areas (TCA), and Total Emissions Control (TEC).

2.2 The Pollution Levy System

The Pollution Levy System is the most long-standing component of China's regulatory structure for controlling emissions and discharges of all types. In principle, it imposes a penalty on emissions or discharges *in excess* of some standard applying to a particular process or plant. The use of the term "levy" conveys an important legal distinction, denoting that the payment is not a tax falling within the jurisdiction of the national tax authorities, but rather is a fee imposed and collected at the local level to fund both administrative expenses of the local Environmental Protection Bureau (EPB) and investment in abatement projects. As such, the PLS well illustrates the nature of relations between the national and local government levels in the development and implementation of environmental policy: The basic guidance and legal authority to impose the pollution levy derive from the national level, while assessment, collection, and use of funds reside at the local level.

Perhaps the most important effect of the PLS has been to fund the local Environmental Protection Bureaus, of which more than 1,600 have been established throughout China, employing more than 20,000 persons to implement environmental policy in accordance with the guidelines set by national authorities. The Chinese pollution levy has empowered local regulatory authorities and created a unique and considerably decentralized administrative structure that is developing the capacity to implement and enforce national policy at the local level. As noted by other authors (Wang, 2000; Wang and Wheeler, undated), this decentralization lends itself well to endogenous enforcement, in which community pressures reflecting differences in economic development and environmental quality explain differences in effectiveness.

The incremental nature of policy development in China is illustrated in the application of the PLS to SO_2 emissions. The PLS was first applied to SO_2 emissions in 1982 as a fee of 0.04 RMB/kg (≈ $4.50/short ton @ 8 RMB/US$) on excess emissions from industrial processes only (excluding electric utilities). With the increased emphasis on SO_2 emission control in the 1990s, a trial program was begun in nine cities in which the PLS rate was increased fivefold to 0.20 RMB/kg (≈ $23/short ton) and the PLS became a tax applied to *total* SO_2 emissions from utility as well as industrial sources. Starting in 1996, this trial program was expanded to include all jurisdictions within the Two Control Areas. Higher levy rates have been tried in two instances: new sources face a double levy rate of 0.40 RMB/kg (Meng *et al.*, 1999), and a pilot program with a higher tax rate of 0.63 RMB/kg (≈ $70/short ton) was initiated in 1998 in three cities. Finally, and most significantly, the 2000 revisions to the APPCL formally changed the base for the pollution levy from excess emissions to total emissions.

Despite its success in funding the requisite administrative structure, the PLS has exhibited a number of problems. First, it applies only to medium-sized and large sources; with rare exception, smaller enterprises, particularly town and village enterprises, are not included. Even so, collections are far below what emissions data indicate they should be.[3] Second, the levy is set too low to effectively encourage significant SO_2 abatement. A level frequently cited as being more likely to provide an effective tax, 1.26 RMB/kg (≈ \$140/short ton), is roughly six times the current level in most jurisdictions. Third, the utilization of the reinvestment portion of the PLS leaves much to be desired. The target level for recycling revenues to local enterprises for abatement projects is 80%, the remainder being reserved for EPB administration; the actual percentage recycled, however, is estimated to be more nearly 50–60%. The recycled portion is then often given back to the firm paying the levy to defray its own abatement expenditures, and this practice has led naturally enough to firms' withholding the recycled amount. At best, this recycling of PLS revenue to firms as a proportion of their assessment does not direct funds to the most economically attractive abatement projects; at worst, the effective pollution levy rate is reduced from an already low level. Finally, deficiencies in emission measurement often lead to negotiated payments only roughly—if at all—related to actual emissions. Negotiated payments are an effective and perhaps initially unavoidable way to raise revenue; however, for the pollution levy to significantly affect abatement behavior, the incidence of the levy must be closely correlated with actual emissions.

2.3 Two Control Areas (TCA)

The State Council's 1990 "Suggestions" introduced the concept of two control areas, one for acid rain and the other for ambient concentrations of SO_2. This concept was embedded in the actual planning that culminated in the Ninth Five-year Plan, adopted in 1996. The TCA component of SO_2 control policy is not an instrument like the pollution levy for affecting abatement behavior, but rather a means for prioritizing SO_2 control efforts. It designates the standard, and thereby the cities and regions that should receive extra attention and resources. The SO_2 Control Zone comprises cities in North China where the ambient SO_2 concentration exceeds 60 $\mu g/m^3$. The Acid Rain Control Zone includes areas in South China where the pH value of precipitation is lower than 4.5 and sulfur deposition exceeds the critical load. These two areas contain about 76% of the country's population. Within the two control areas, certain municipalities are designated as "key," and are slated to receive more aggressive emissions control targets.

[3] For instance, collections from electric utility sources, which are relatively large and more easily monitored than smaller sources, have been estimated to be about 25% of what could be expected, based on actual utility emissions (Benkovic, 1999).

As a device for prioritizing abatement efforts, this component of Chinese environmental policy offers no guidance regarding instrument choice. It is intended to work in conjunction with the pollution levy, Total Emissions Control, emissions trading, and any other measure that might prove useful and effective. As can be seen with the evolution of the PLS for SO_2 emissions, changes in either coverage or levy rate are tried first in key municipalities; then, if proven successful, the changes are applied more broadly within the Two Control Areas and eventually in the non-TCA areas of China, as environmental priorities and political limitations dictate. The changes in the APPCL of April 2000 and their implementation in the Tenth Five-year Plan have maintained and intensified the focus on the Two Control Areas by setting specific targets for total SO_2 loadings and expanding the number of key cities slated for priority efforts from forty-seven during the Ninth Five-year Plan to one hundred in the Tenth Five-year Plan.[4]

2.4 Total Emissions Control

Total Emissions Control is the newest and, since 2000, the most important element of Chinese environmental policy. Specifically, it places a ceiling on total emissions for twelve major pollutants, including SO_2. The concept arose in 1996 from a series of State Council documents and SEPA action plans, and became incorporated by amendment into the Ninth Five-year Plan to work in concert with the existing Two Control Area component. The 2000 revision of the APPCL embedded TEC in the fundamental law and thereby shifted the focus from controlling concentrations to controlling total loadings. In the absence of clearly specified and credible sanctions against exceeding these ceilings, the TEC provides a target more than a cap, but it still provides a measure for judging the effectiveness of control measures. The 2000 amendments also strengthen enforcement powers.

The TEC limit for national SO_2 emissions in the Ninth Five-year Plan was 24.5 million metric tons (approximately the 1995 level of emissions),[5] which was then allocated among the thirty-one regions. These regional targets were intended to limit emissions to the 1995 level for all areas within the TCA zones, except for a group of

[4] Ambient air concentrations of SO_2 emissions in Chinese cities have been falling slowly but steadily over the years, but much remains to be done (Wu, Wang, and Meng, 2000). Monitoring results in 1999 indicate that approximately one-third of 338 cites in China did not meet National Class 2 standards for normal residential areas (0.06 mg/m^3) and 15% did not meet the less stringent Class 3 standards applicable to special industrial zones (0.10 mg/m^3).

municipalities in Eastern China that received target allocations at a level 10% below the actual 1995 baseline because of their greater population density, higher level of economic development, and more abundant resources for implementing policy.

As of 1999, SO_2 emissions in China totalled 18.6 million tons, 25% below the national TEC target, and all but three regions met the regional targets in the Ninth Plan. Although local efforts to control emissions from major sources contributed to this remarkable decline in SO_2 emissions, two other aspects of the restructured Chinese economy played a large role. First, economic changes and explicit policy succeeded in closing down small enterprises, particularly small coal mines producing relatively high-sulfur coal, and small, inefficient, thermal electricity-generating units. Second, China's ongoing economic transformation has diminished production from large and medium-sized state-operated enterprises (SOEs) whose output is typically emissions-intensive.

The Tenth Five-year Plan has set a national TEC ceiling for 2005 of 18 million tons for all of China (approximately the 2000 level), and a more restrictive total of 10 million tons for the Two Control Areas. Allocation of this target to subordinate jurisdictions is currently taking place in conformance with the Two Control Area criteria, with special focus on the one hundred key municipalities characterized by high existing SO_2 pollution, high SO_2 emissions, or failure to achieve prior emissions targets.

3 FROM FACILITY PERMITS TO TRADABLE PERMITS

The development of a tradable permits system for SO_2 emissions in China can be usefully seen as a movement from non-tradable facility permits to tradable emission permits. This distinction, while subtle, is important to understand the likely pattern of development in China. A facility permit imposes conditions that limit emissions from the facility in some manner. A tradable permit, often called an allowance, entitles the holder to emit a specified amount without further conditions; and allowances are distributed to facilities in a total amount that will limit emissions.

Both types of permit effectively reduce emissions, although one does so by imposing emission-reducing conditions on the operation of individual facilities while the other requires that emissions from all affected facilities be "covered" by allowances equal

[5] Actual SO_2 emissions for the 1995 base year were 23.7 million tons. The 2000 target reflects additions to the base year that are said to reflect unexpectedly low SO_2 emissions during 1995.

in number to emissions. Emissions trading can occur with both forms of permit: in the one case, a tradable credit is granted after appropriate administrative review for emission reductions beyond what is required by the facility permit, while in the other, trading is a matter of right. Both systems create rights to emit: in one case, the right attaches to the facility meeting the specified conditions and can be traded only after administrative review and approval, while in the other, the right to trade is explicit and inherent in the permit, and therefore readily separable from any particular facility.

The process of moving from non-tradable facility permits to tradable emission permits in China will differ from the process that has occurred in the U.S. and other OECD countries where tradable emission permit systems have been implemented or are being proposed. In these latter countries, facility permits with all associated monitoring and enforcement capabilities are usually already in place, and the cap-and-trade system is simply added. In contrast, China has had little experience with facility permits, and the two types of permit will be developed more or less concurrently.

Avoiding facility permits altogether and moving directly to allowances is tempting to contemplate, but the practical reality is that facility permits are indispensable in the early stages of implementing emission control measures, particularly in countries where the markets that are assumed by market-based instruments exist in rudimentary form. Moreover, wherever environmental goals are multiple, such as to improve ambient air quality in cities and reduce acidic deposition, more than one instrument is desirable. For instance, in highly market-oriented economies, such as the United States, facility permits coexist with tradable permits, effectively allowing multiple goals to be achieved.[6]

Even so, there is an advantage in moving quickly to primary reliance on a cap-and-trade system. Gaining that advantage requires that facility permits be designed to be consistent with tradable emission permits in order to lead naturally and quickly to the latter's use. To that end, this chapter discusses the distinctive requirements of a tradable emission permit system and how these requirements can be incorporated in the design of facility permits to encourage the development of allowance trading.

---

[6] In the United States, ambient air quality standards are met through the use of conventional command-and-control regulations, and these regulations can and in some cases do restrict the extent to which facilities can trade SO_2 emissions under the Acid Rain Program.

3.1 *Setting the Cap*

The distinguishing feature of a tradable permits system is the “cap,” or limit on the total amount of emissions. Assuming enforcement, this constraint makes emissions scarce, imparting value to allowed emissions. As discussed earlier in connection with the Total Emissions Control component of Chinese environmental policy, this feature of a tradable permits system is already in place in China.

Not only has a national TEC cap been decided, but that cap has been distributed to subordinate levels of government in what can be seen as regional caps, the sum of which equals the national cap. Such was done in the Ninth Five-year Plan, when the 24-million-ton national cap was distributed to the thirty-one regions. The process of distributing the 18-million-ton cap in the Tenth Five-year Plan to regions is underway currently;[7] however, in contrast to the Ninth Five-year Plan, the regional caps are to be broken down further and distributed to even lower levels of government, and ultimately to emitting facilities.

At the level of the local Environmental Protection Bureau (EPB), the most important issue will be how to translate the cap into instructions to firms in a way that will ensure that the cap is not breached. The form these instructions take will determine not only how costly it will be for each EPB to achieve its cap, but also whether a tradable permits system will emerge. For this system to develop, the further allocation of the cap to emitting sources within each jurisdiction must take the form of emission permits that are potentially tradable.

3.2 *Allocating the Cap and Establishing Tradability*

The EPBs enjoy considerable discretion in the specific requirement they impose on emitting facilities within their jurisdictions. For example, firms might be instructed to pay a tax per unit of emissions that is sufficiently high to motivate the firms to reduce emissions enough to stay below the local cap. Alternatively, sources might be required to install equipment or undertake practices that will reduce aggregate emissions to the required level. Finally, as a logical extension of the process by which the national TEC cap is assigned to subordinate levels of government, each facility might be allocated its share of the local cap and told simply to make sure that emissions do not exceed this amount, as was done in some of the emission permit pilot projects in China during the 1990s.

A facility permit with an embedded quantity limit is similar to the allocation of tradable emission permits in like quantity to the facility, with two important differences. First, in a tradable permits system, the aggregate quantity limit is broken down and issued in many small units, as “allowances.” Second, each firm is instructed to give back to the regulating authority an allowance for every ton of emissions. In the case of *facility permits*, a facility would receive a cap of, say, one hundred tons and be told to stay within that limit. In the case of *tradable emission permits*, the facility would receive one hundred allowances to emit one ton, and be told to return an allowance for every ton emitted. When trading occurs, the facility might actually emit more than the number of allowances received, if it can acquire allowances from other sources at less cost than would be required for it to reduce its emissions, but it might also emit less in order to sell its unused allowances, as it would do if its marginal cost of abatement were less than that of other facilities receiving tradable permits.

If no market for allowances develops, the two sets of instructions are equivalent. All that would distinguish the tradable permits case then would be that the allowances issued would be returned in an amount equal to emissions, and any not returned would have no further value. Nevertheless, the potential value of a tradable permit provides an important reason for issuing the quantity limit in this form. It is virtually certain that the distribution of the EPB cap to firms, whether in the form of an aggregate limit or of an equal number of allowances denominated in tons, will not be such as to equalize marginal costs among all emitting units. Initially (and perhaps especially in transitional economies), firms may treat their allowance allocations as quantity ceilings and incur differing marginal costs of compliance; however, if the ability to sell and to buy is clearly understood to be acceptable, firms will seek trading partners in order to reduce their costs by eliminating these differences in marginal cost. Thus, issuing the quantity restriction in a form that makes trading emission reductions at the margins practicable and easy provides incentive for firms and other intermediaries to create a market in allowances, with the consequent well-known gains in economic efficiency.

[7] Wu *et al.* (2000) provide a proposal for allocating the 18-million-ton national and 10-million-ton TCA targets to the 31 regions.

One of the EPBs' most difficult challenges in issuing allowances will be deciding how many allowances to issue each affected source within their jurisdictions. This allocation is often seen as a highly political and thankless task because it is so explicit and transparent. But the fundamental task—distributing the cost burden of the required emission reduction among firms and sectors within each EPB's jurisdiction—is no different than what an EPB would face if it chose to meet the cap through use of the more familiar command-and-control measures. Imposing a uniform emission rate limit or a technology mandate on all sources sufficient to achieve the cap may appear to provide a more objective basis for the reductions required of all firms, but the burden of emission reduction will have been distributed just the same, and it is unlikely to be both equitable and efficient, since firms face different abatement cost opportunities. Although the command-and-control allocation may have the political advantage of making the distribution of the cost burden less obvious, total costs will be greater.

Even so, the allocation of permits to emitting sources will be easier if some basic principle is used, and that principle could be the standard that might have been applied in the absence of a tradable permits system. This was the case in the U.S. Acid Rain Program. An emission rate of 1.2 pounds of SO_2 per million Btu of heat input (≈ 137 grams per kilocalorie) had become embedded in the air emission regulatory structure in the United States as an appropriately stringent level of restriction, based on estimates of the emission rate that could be attained by the best available control technology in the early 1970s. Accordingly, the initial principle for allocating allowances to firms was this emission rate multiplied by average annual fuel use during a three-year, historic baseline (1985–87).[8] The aggregate amount of emissions allowed by this standard was approximately nine million short tons, or approximately half of 1980 emissions, a level deemed sufficient to meet the environmental objective of eliminating damage from acid rain.

Adjustments from the basic allocative principle can be made, and were in the U.S. SO_2 emissions trading program. The enabling legislation contains more than thirty deviations from the basic allocation principle. Certain types of facilities were given more allowances than they otherwise would have received, and the amount available to all

[8] Sources having an emission rate in 1985 less than 1.2 lb SO_2/mmBtu were issued permits equal to the lower rate multiplied by baseline energy use.

others was “ratcheted down” by a small amount to preserve the cap.[9] For example, the state of Florida successfully argued that the movement of retirees to Florida would raise the cost of electricity under a cap for its citizens more than would be the case for citizens of the states from which the retirees were moving. As a result, utilities in Florida were awarded more allowances than they would have received under the basic principle, at the expense of everyone else.

These adjustments in the allocation of allowances redistributed the cost burden among facilities affected by the SO_2 cap in the same manner as would a regulator who recognized the implications of cost heterogeneity for implementing a uniform standard. Such a regulator might attempt to achieve a more equitable and efficient distribution of the burden by imposing more stringent requirements on firms having lower abatement costs, and less stringent requirements on those facing higher abatement costs. Unfortunately, regulators are unlikely to possess good information about the abatement opportunities and costs facing firms, so later administrative exceptions for facilities able to show undue hardship are inevitable. Besides this, even though a well-informed, well-intentioned regulator may grant truly equitable exceptions, deviations from the standard will all be in one direction, since no firm facing low abatement costs can be expected to volunteer to meet a more stringent standard.

The great advantage of tradable permits lies in their allowing such readjustments to take place without intervention of the regulator, and in such a way that each deviation automatically offsets a deviation of equal magnitude in the opposite direction. In effect, the cost burden of emission reduction is distributed in a process consisting of two steps. The first occurs when the regulator, using all knowledge at his disposal at the time of allocation, distributes allowances in what seems to be the most equitable manner. The second step occurs as firms, acting in their own self-interest, eliminate the differences in marginal cost that will inevitably emerge. Unlike the zero-sum game encountered in the allocation of allowances within a fixed total, this second, market-driven reallocation of permits and of abatement effort enables both sellers and buyers to benefit, and total societal costs to be reduced, as well. In contrast, adjusting command-and-control regulations or taxes to meet equity concerns typically involves some departure from either effectiveness or efficiency, if not both.

[9] See Chapter 3 of Ellerman *et al.* (2000) for a discussion of these deviations.

3.3 The Scope of Trading

Tradable emission permits presume some domain in space and time over which the permits can be traded. The issue of *what* space and time period does not arise initially when trading occurs through the creation and transfer of credits within the confines of facility permits, because the scope of trading is determined on a case-by-case basis. Unfortunately, the review process required to make this determination imposes high transaction costs, which generally lead to the poor results associated with this form of emissions trading. In contrast, allowance trading avoids transaction costs by making trading a matter of right; however, the regulatory authority must address the scope of trading at the outset.

3.3.1 THE SPATIAL DIMENSION

The geographic scope of trading can be defined as narrowly as several sources within an industrial facility or as broadly as all sources within an EPB's jurisdiction, or even more broadly to include sources in adjacent EPBs. The basic consideration is the environmental effect of the marginal unit of emissions, but this effect must be interpreted realistically. Obviously, the environmental effect of emissions from several exhaust stacks at a single facility is the same, and this logic can be extended easily to emissions from adjacent facilities. The question is, "how far can the trading zone be extended?", since it is equally obvious that trading with a distant source whose emissions have no effect on local conditions will frustrate the achievement of the local environmental objective embodied in the cap.

An EPB with several geographically distinct urban areas within its jurisdiction has several alternatives. One would be to allow trading within each urban area but not among them. Rarely, however, do emissions affect only the area where a source is located. Typically, some pollution within any given area originates from emissions transported by wind from other areas. One way to address this concern is to assign different redemption values to permits obtained from sources in different areas. Thus, permits issued in area A could be deemed to cover a ton of emissions for all sources in area A, but permits issued in the adjacent area B would be deemed to cover only half a ton if presented for compliance by a source in area A. Given sufficient information about the atmospheric transport and transformation of emissions, schemes of differential pricing can be devised; however, such information is not always available, and even when it is, the added complication and cost may overwhelm the environmental benefit. When wind direction is variable and atmospheric transformation depends upon meteorological conditions that also vary, it is often found that different areas are polluting each other. This reciprocal nature of pollution suggests yet another alternative, which is to broaden the scope of trading to include both areas so that permits issued in areas A and B are equally valid when presented for compliance in either area. This last alternative is especially attractive when there is considerable uncertainty about what happens to emissions from particular sources, or when the policy goal is to effect a general reduction in emissions in order to meet multiple environmental goals.

Still, even with the uncertainties and reciprocal nature of pollution, which argue for geographically wide emissions-trading areas, a potential problem remains: "hot spots." This term refers to the possibility that the spatial pattern of abatement arising from a wide, unrestricted market in allowances may lead to concentrations of emissions in certain areas (hot spots) that will violate local air quality goals. This situation will occur when the emissions creating the worst pollution are more costly to abate than other emissions.

The solution to this problem is not to restrict trading, but to use another instrument to ensure achievement of the purely local goal. This has been the case with the U.S. Acid Rain Program, which was enacted on top of an elaborate command-and-control program, designed to avoid adverse local health effects, which had been in existence for over two decades before the SO_2 cap-and-trade program was enacted. The occasional assertion that the trading of SO_2 emissions is unrestricted within the United States is not true, strictly speaking. A more accurate statement would be that trading in SO_2 allowances is unrestricted within the limits and mandates imposed by the pre-existing regulatory structure. Since nearly all sources were in compliance with the pre-existing ambient air quality standards for SO_2 when the Acid Rain Program was enacted, and the Acid Rain Program required a further fifty percent reduction in aggregate SO_2 emissions, most sources are effectively unrestricted in trading. Nevertheless, they must operate within the limits imposed by local environmental regulations. Thus, hot spots, defined as violations of local ambient air quality standards, did not appear as a result of emissions trading in the United States mostly because of the sequence in which the regulatory structure for controlling SO_2 emissions was developed.

There is no particular reason why the sequence followed in the United States—getting the local conditions right first, then trading to deal with regional problems—should be observed in China. The practical reality in China is that, for the most part, neither local nor regional conditions for avoiding environmental damages are being met. The all-important imperative is to reduce emissions; where and how are, for the moment, secondary considerations. If allowance trading leads more surely to emission reductions, then it should be adopted because it will reduce emissions sooner. The argument is not to ignore local details, but to recognize that there is as much logic, and perhaps more, in implementing broad emission reductions first and then tending to the local details, as there is in making sure that the local details are right before engaging in emission reductions. This sequence of focusing initially on broad emission reduction goals and then tending to local details was recognized in a recent study by the Chinese Academy for Environmental Sciences, which noted that Class 2 ambient air quality standards for SO_2 are not likely to be met unless total SO_2 loadings within the Two Control Areas are reduced to below 12 million metric tons (Wu *et al*, 2000). This sequence is as applicable at the level of the EPB as it is at the national level.

The concern about hot spots also arises with conventional command-and-control instruments, but the presumption usually is that the regulation imposed upon particular firms can be changed later, if necessary, to ensure that local air quality goals are being met. The same presumption can and should be made with tradable permits systems. If the cap is successfully enforced and is adequate in the aggregate, then it is not possible for all or even most areas to be hot spots. Those that do appear can be treated with supplementary instruments—either a higher pollution levy, as discussed in the next section, or other command-and-control restrictions that can be incorporated in a facility permit. These supplementary instruments will create differences in the marginal cost of abatement, but such differences are the logical consequence of ambient air quality standards that may be more costly to attain in some areas than in others.[10]

3.3.2 The Temporal Dimension

The spatial scope of trading is not the only dimension that needs to be considered. Emissions trading can also take place across time periods, resulting in many

[10] Taking these supplementary measures in hot spots will increase emissions in surrounding areas; however, these areas have reduced more than the hot spots and are therefore more likely

advantages. Temporal trading can occur through banking (i.e., saving permits issued but not used in one period for compliance in a later period) and through borrowing (i.e., using in the current period allowances that had been allotted for future periods). In most allowance-based systems, banking is allowed but borrowing is not.

Banking serves several important purposes, as demonstrated by the U.S. Acid Rain Program. The most important one is the incentive to move emission reductions forward in time when greater reductions are required in the future. The phased-in structure of the U.S. Acid Rain Program resulted in the banking of approximately eleven million allowances that had been issued for use in 1995–99, for use after 1999, when the national cap was reduced and the marginal cost of abatement was expected to be higher. In effect, emissions were eleven million tons lower than allowed in the first years of the program, and emissions will be eleven million tons higher than the number of permits issued later in the program. In China also, the intent under TEC policy is to reduce the limit progressively over time, from 19.95 million tons in 2000 to 17.95 million tons in 2005 for all of China, to ten million tons by 2010 for the Two Control Areas. Although the 2000 target for all of China was easily met and the 2005 target will not allow any growth in SO_2 emissions for all of China, the TCA target of 10 million tons will call for more abatement effort than would result from the continuing transformation of the Chinese economy and a policy requiring new sources to offset emissions. Existing firms within the Two Control Areas will be required to make progressively greater emission reductions over time.

Under such conditions, banking provides an incentive to firms to reduce emissions more than required in the early years, because the banked allowances can be used to defer the higher marginal costs associated with the more stringent, later requirements. For example, imagine a firm that faces three abatement options: 1) reducing emissions by 10% at low cost, 2) reducing them by 25% at somewhat higher cost, and 3) reducing emissions by 50% at still higher cost. Suppose further that the firm faces an initial requirement to reduce emissions by 10% and a later requirement to reduce them by 50%. Without banking, the firm would adopt the 10% reduction technology and the costly 50% reduction technology only when required; there would be no incentive to reduce emissions by 25% in the early years. With banking, an incentive is

to have overcomplied with local requirements. Additional actions can be taken, as new hot spots appear, until the complete set of desired environmental goals is achieved.

provided to adopt the 25% abatement technology, which will reduce emissions more in the early years at the expense of delaying the adoption of the 50% technology by some period of time. Over the entire period, cumulative emissions are the same, but they are reduced more rapidly in the early years. For a country like China that seeks to reduce highly polluted areas quickly, such an incentive is highly desirable, even if the incentive implies that a few more years may be required to attain the control levels implied by the eventual aggregate annual TEC level.

This incentive is particularly important when doubt exists about how quickly allowance markets might develop. In the preceding example, a single firm might adopt the 25% reduction technology in the first period, even without banking, if it could find buyers for its unused allowances. In the absence of such a market, however, banking would provide firms with their only incentive to adopt the 25% control technology in the first period since it would allow them to defer (but not avoid) the costs associated with the 50% control technology. Moreover, with phased-in caps, firms in the aggregate have an incentive to bank, even when a market exists in both periods. Firms would still, by their own actions, accelerate the cumulatively required emission reductions so that more would occur earlier, while the cumulative amount of emissions would be the same.

The second main purpose served by banking is to avoid undue volatility in allowance prices. Banking allows a firm to maintain an allowance inventory, just as it would with fuel or any other requirement of its production process. Without such carry-over, prices in each period would be subject to greater fluctuation, reflecting random variations in demand, such as those caused by weather. The supply of allowances for each period would be fixed, and any unexpected change in demand would cause prices to fluctuate more than they would with some carry-over. With banking, an unused allowance retains value in the next period, placing a floor under prices. Similarly, inventory carried over from earlier periods will cushion the effect on the current price of unanticipated increases in demand. Although not usually allowed, borrowing could provide a similar price-dampening effect.[11]

[11] Recent experience in California with the Los Angeles NO_x allowance trading program (RECLAIM) illustrates how useful borrowing could be. Because the spatial dimensions of this market are relatively narrow and virtually no banking or borrowing is allowed, an unusual confluence of events increasing the demand for electricity generation in Los Angeles caused NO_x prices to increase about fortyfold, from $2,000 a ton to $80,000 a ton.

Temporal flexibility serves important purposes, providing incentives for earlier cost-effective emission reduction and moderating otherwise volatile permit prices. It is particularly significant, however, when the spatial dimension of the trading market is limited, either for environmental reasons or as a result of slow market development. For an economy in transition, such as China, this latter consideration is especially important.

3.4 Measurement, Registries, and Compliance

Cap-and-trade systems impose unique measurement and accounting requirements on the regulator. First, emissions must be measured to determine the number of allowances to withdraw from the system. Most command-and-control regulations, such as technology mandates or emission rate limits, do not in fact require that emissions be measured to ensure compliance, although measurement is implied by any assessment of effectiveness. Second, the regulator must have some means of knowing whether allowances submitted for compliance are valid. This second requirement is accomplished through the use of a registry or, as it is called in the U.S. Acid Rain Program, an Allowance Tracking System, that accounts for all allowances issued, transferred, and submitted for compliance.

The cost and difficulty of measuring emissions are two reasons for using command-and-control measures rather than tradable permits or taxes. Compliance with a requirement to burn low-sulfur fuel or to install specified abatement equipment can be presumed to reduce emissions, based on the characteristics of the required fuel or equipment. All that is necessary to determine compliance is to ensure, usually by periodic inspection, that the fuel burned is of the specified quality or that the equipment is installed and operating. The obvious problem is to ensure that the equipment is operating between visits, but where measurement is difficult and expensive, little alternative to command-and-control measures exists.

In China, the adoption of the Total Emissions Control policy, not to mention the long-standing Pollution Levy System, implies an estimate of total emissions that can provide the measurement required by a tradable permits system. For instance, a mandate to burn coal of a sulfur quality below some specified level implies that data on the sulfur quality of coal is available. Demonstrating the use of coal of the requisite quality may suffice when the regulatory requirement is to burn coal of a certain quality or to emit below a certain emission rate. However, for the EPB to demonstrate that emissions within its jurisdiction are less than its assigned cap, the EPB will have to estimate the amount of coal burned or the output of the facility. Similar considerations apply for abatement equipment. In both cases, the EPB will require information that allows it to determine total emissions and that can be used to determine compliance in a tradable permits system.

The present requirement that all newly built, expanded, or transformed thermal power plants within the Two Control Areas install continuous SO_2 monitoring systems will lead to an increasing number of power plants with the requisite monitoring capability; however, the monitoring required for a tradable permits system need not be real-time and continuous. Approximate methods suffice, and material balance calculations based on fuel sampling and engineering specifications can provide measurements of sufficient accuracy. In fact, such alternative methods are used in the U.S. Acid Rain Program for small sources for which the installation of continuous monitoring equipment would be unduly expensive. The important issue is the quality and integrity of data obtained, not the manner by which it is obtained.

The difficulty of measuring emissions is also an important factor determining which sources should be included in a tradable permits system. Some sources will likely never lend themselves to developing data of sufficient quality for use in a tradable permits system, even though the EPB will have to estimate emissions from these sources to determine compliance with its allocation of the TEC policy. Household emissions are an example: it would probably be easier to restrict the use of coal in households, for instance, than to attempt to obtain an accurate reporting of coal use and quality for home heating and cooking. Thus, some sources within an EPB’s jurisdiction may not be included within the allowance trading system, and some portion of the cap received by the EPB will have to be reserved to “cover” these sources.

The requirement to maintain a registry of permits arises from the nature of compliance in a tradable permits system. A source is deemed to be in compliance when it gives up a number of permits equal to emissions. Since the permits deducted need not be the same as those issued initially to a particular source, some means must exist to track permits from the time they are issued until the time they are withdrawn from the system for compliance, and to ensure that permits are not used twice. Thus, if firms A and B are each issued permits totaling 100 tons of emissions, and A buys 25 permits from B to cover 125 tons of emissions, the regulator would know that the permits came from B, which could then emit only 75 tons (unless B had purchased permits from some other source).

Registries are bookkeeping systems that maintain for each emitting source an account into which permits are initially deposited by the regulatory authority and subsequently deducted in an amount equal to emissions at the end of each compliance period.[12] Since no actual certificates are issued to sources, the permits are bookkeeping entries that can be readily transferred from one account to another upon appropriate instruction from account holders or sources, in the same manner as funds are transferred upon appropriate instruction among checking accounts at a bank. Typically, permits are assigned serial numbers to facilitate tracking.

Although a registry is a unique requirement of a tradable permits system, the task of determining compliance is simplified for all. First, no corps of inspectors is needed, as it is with command-and-control regulation, to determine that the required technology is installed and operating, and that the prescribed emission-reducing practices are being followed. The only required on-site inspection concerns the emissions monitor or alternative means of determining the quantity of emissions. Second, there is little discretion in determining whether a source has complied or not. Unlike the situation with technology mandates and best practices, where equipment can break down or specified practices cannot always be followed, the only criterion in a tradable permits system is whether the source has permits equal to the quantity of measured emissions in its unit account. This aspect is identical to that for an emission tax: the only question is whether the tax corresponding to the quantity of emissions has been paid. A third simplifying feature of the registry is that the means of collecting the penalty are at all times in the regulator's hands, since allowances for the next period's compliance are maintained in the registry. There is never any question of collection, as can be the case with taxes. Rather, it is as though polluters are required to place in escrow the estimated amounts of tax due. An enforcement issue exists if insufficient permits are on deposit, but most enforcement can occur automatically.

[12] Such an arrangement assumes grandfathering of permits, which is the prevailing practice in nearly all existing and proposed cap-and-trade systems. If permits were auctioned, firms would purchase the required permits by bidding into the open auction and placing the permits in appropriate unit accounts.

4 INTEGRATION OF TRADABLE PERMITS WITH A POLLUTION LEVY SYSTEM

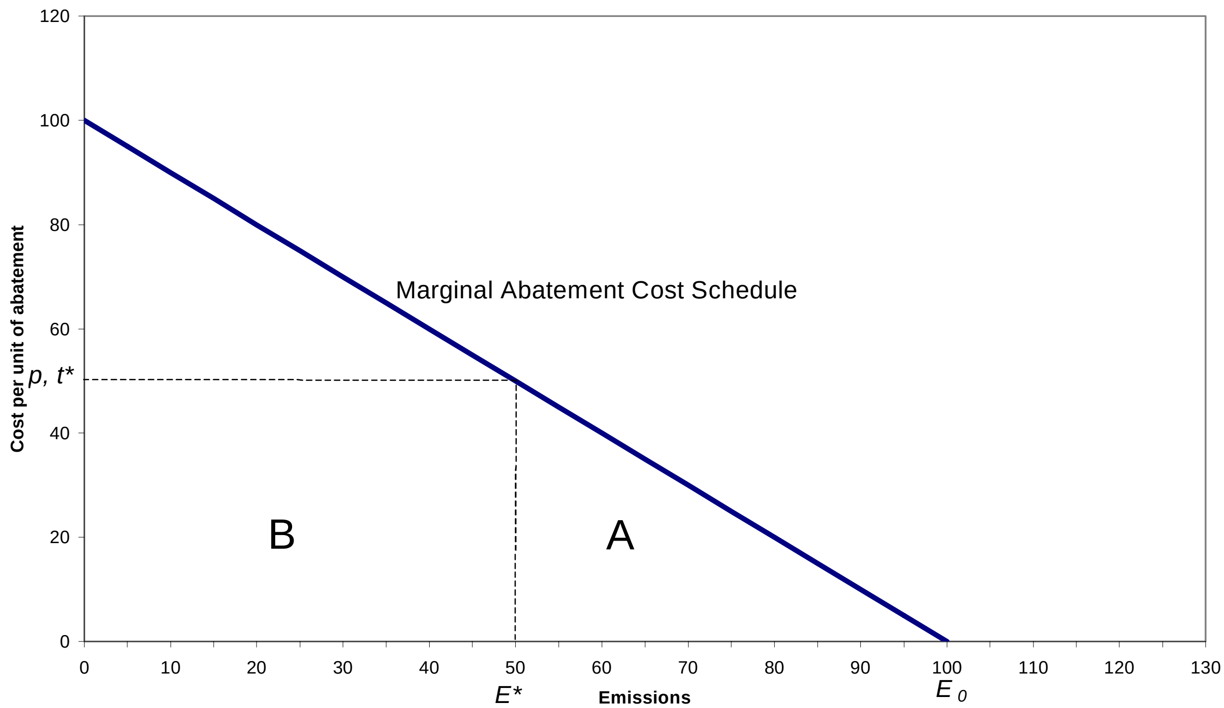
The theoretical and applied literatures in environmental economics often present tradable permits and taxes as alternative, mutually exclusive instruments for achieving environmental goals. This treatment could be interpreted as implying that a tradable permits system should replace China's pollution levy on SO_2 emissions, but doing so would pose many practical problems. Not only is the PLS the most well established instrument for meeting environmental goals in China, but, more importantly, it provides funding for the EPBs, which are the critical level of government for implementing effective control of SO_2 emissions regardless of instrument choice. Fortunately, the choice is not so stark.

Replacing the PLS on SO_2 emissions is neither necessary nor desirable in implementing a tradable permits system to control SO_2 emissions in China, but a choice must be made concerning which instrument is to be the primary one for achieving environmental goals. If tradable permits are to be primary, the PLS can continue to exist as a subordinate instrument to achieve other goals. The cap embodied in the tradable permits system would determine the aggregate emissions against which the PLS would be levied, and the price of permits would be reduced by the amount of the tax; otherwise the two instruments could coexist without adverse effect. The converse is, however, not true. If the PLS is to be the primary instrument, tradable permits would be redundant and there would be no point in maintaining a registry, allocating permits, or determining the scope of trading.

4.1 Taxes and Tradable Permits?

Some basic features of taxes and tradable permits will help to explain how the pollution levy and a tradable permits system could work together. Figure 1 presents the relationship between emission levels and the cost of emission control.

Figure 1



The horizontal axis represents total emissions and the vertical axis indicates the cost per unit of abatement. The downward sloping line plots the marginal cost of reducing emissions against the corresponding level of total emissions. In Figure 1, the numbers are purely illustrative; the marginal abatement cost schedule could be that for a single firm, for all firms within the jurisdiction of an EPB, or for China as a whole. The emission level E_0 indicates uncontrolled emissions, for which the marginal cost of abatement is zero, by definition. Any lower level of emissions requires some cost to be incurred for abatement, and the cost of the last unit of abatement is reasonably presumed to steadily increase as more emission reduction is undertaken.[13] At some relatively high marginal cost, such as that required to switch all sources to natural gas, SO_2 emissions would be zero.

Figure 1 can also be used to illustrate the fundamental equivalence between tax and quantity instruments (i.e., between the PLS and tradable permits). Suppose the

[13] In reality, the marginal abatement cost schedule may not rise smoothly and monotonically as depicted here, but may instead be characterized by horizontal segments, or steps, of varying length that rise from the least costly segment to the most costly one.

regulatory authority wishes to limit total emissions to *E**. One means would be to issue tradable permits in the amount *E** and impose a very large penalty for emitting without a permit. With a functioning permit market, the clearing price would be *p* and the marginal cost of achieving the last unit of abatement required to meet the cap, *E**. Alternatively, the regulatory authority could impose a tax, *t**, equal to *p*, which would cause firms to undertake abatement costing less on the margin than *t**, or in the same amount as would occur with the issuance of permits equal to *E**. As presented so far, there would appear to be no difference between imposing a tax in the amount *t** and distributing permits in the amount *E**. The aggregate level of emissions and the total and marginal costs incurred by firms would be the same under taxes or tradable permits. The two instruments differ significantly, however, in their informational and distributional characteristics, to which we now turn.

The informational aspect concerns the regulator's knowledge of the backward-sloping line in Figure 1. As presented there, the regulator is assumed to know the price and quantity relationship for marginal abatement cost with certainty, in which case the choice between *t** and *E** makes little difference. In practice, the regulator will not have this information, or at best only a vague idea of the relationship. As a result, the regulator faces a choice: either to fix the quantity of emissions by issuing permits in the amount *E** while remaining uncertain about the marginal cost of abatement, or to fix marginal cost (price) by imposing a tax *t** on emissions, remaining uncertain whether the desired level of emissions E* will be achieved.[14]

The distributional aspect of the choice between permits and taxes can be described by reference to the areas **A** and **B** in Figure 1. Area **A** is the total cost of abatement incurred by firms in reducing emissions from E_0 to *E**, that is, the integration beneath the marginal abatement cost curve over this distance. Firms incur this cost regardless of whether the tax *t** is imposed or tradable permits are issued in the amount *E**. The area **B** represents the scarcity rent associated with constraining total emissions to *E**; the distribution of this rent constitutes a distinguishing difference between tax and permit systems.

---

[14] The now classic answer to this problem depends on the relative slopes of the marginal benefit and marginal cost curves (Weitzman, 1974); however, reasons of political economy seem often to result in the choice of quantity instruments even when price instruments would be preferable.

When a tax *t** is imposed, area **B** is the amount of tax revenue paid to the government for the right to emit E*. In effect, the government receives the scarcity rent by charging an appropriate amount for the right to emit. The emitting firm will naturally avoid all emissions for which the marginal cost of abatement is less than *t** and thereby incur abatement costs equal to area **A**. The full cost to the firm is then **A** + **B**, and the additional total and marginal costs imposed by the tax will be passed on to customers in the price of output.

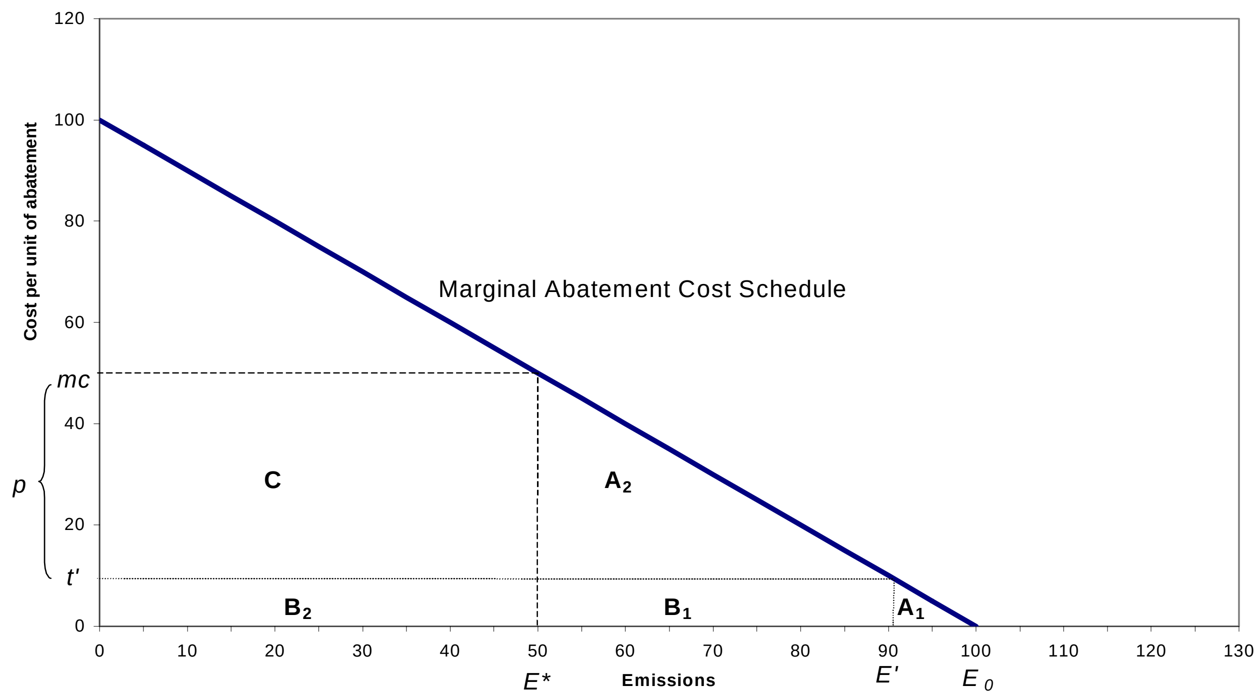
When a tradable permits system is used, the distributional consequences for an SO_2-emitting firm depend on the allocation of the permits. If the government were to auction the permits, the consequences to the firm would be largely the same as with a tax. The regulatory authority would receive the scarcity rent, firms would incur a total cost equal to **A** + **B**, and the marginal cost would be passed on in the price of output. If the permits were issued to firms without payment, or "grandfathered," these firms would receive the scarcity rent. Trading of these allowances among firms would ensure that each firm faces marginal costs equal to *p* so that the total abatement cost equals **A**, as before; however, the total cost to regulated firms would be **A** – **B**. Although the marginal cost will always equal *p*, the net total cost of the environmental constraint to regulated firms could be positive (**B** < **A**) or even negative (**B** > **A**), depending on the shape of the marginal abatement cost schedule and the level of allowed emissions. In a similar manner, the outcome for individual firms will depend upon the permits received by each firm and its own abatement cost schedule. Inevitably, some firms will be better off than others, but all will be better off than with a tax or with auctioned permits.

The ultimate distributional consequences of these alternatives depend on how firms or governments further recycle the rent represented by area **B**. A large literature exists on the possibilities of government beneficially recycling the revenues from environmental taxes or permit auctions to obtain a "double dividend," but the more immediate and practical problem for environmental regulators is likely to be firms' attitudes toward the use of these alternatives for controlling emissions. While taxes promise only higher costs, grandfathered permits offer firms the prospect of additional profit, which can operate as an incentive (or bribe) for accepting an enforceable constraint on emissions. For instance, experience in Chile has shown that grandfathered permits encourage firms to come forward with information about their emissions and their means of measuring them in order to obtain a share of the limited permits (Montero *et al.*, 2000). In contrast, tax systems offer no comparable inducement for reporting emissions or accepting the emission constraint. This consideration may be set aside if the regulatory authority is strong and knowledgeable; however, if successful implementation of the environmental constraint depends on the consent and cooperation of those regulated, this incentive becomes more important.

4.2 The Interaction between Tradable Permits and a Pre-existing PLS

The discussion accompanying Figure 1 treated price and quantity instruments as though only one or the other would be chosen; Figure 2 illustrates how the PLS and tradable permits might work together.

Figure 2



The diagram is the same as Figure 1 except that the PLS is represented by the relatively low tax *t′,* which leads to a small amount of abatement and a level of emissions *E′* that is greater than the desired level of emissions, *E**. Assuming that the PLS is effectively collected from all sources, the total cost of abatement is represented by area $\mathbf{A_1}$, and levies paid to the government equal the sum $\mathbf{B_1} + \mathbf{B_2}$.

The imposition of an effective grandfathered tradable permits system with E^* permits on top of the PLS has several consequences. The amount of abatement and the marginal cost of abatement are necessarily greater than they would be under the levy system alone in order to reduce emissions from E' to E^*. The further reduction of emissions reduces PLS revenue by the amount $\mathbf{B_1}$, which the firm will use for abatement expenditure, as well as an additional amount, represented by $\mathbf{A_2}$. The scarcity rent received by the firm is the area **C**, which is smaller than the amount **B** in Figure 1 by the amount of the PLS revenues $\mathbf{B_2}$. In effect, the scarcity rent is shared between the firm and the government in an amount determined by the relationship between marginal cost, *mc*, and t'. Finally, the clearing price of permits in the market, *p*, is not *mc*, as it would be if there were no PLS, but rather the difference between *mc*, the level of marginal cost required to meet the constraint E^*, and t', the level of the pollution levy.

The reduction of potential PLS revenue by the amount $\mathbf{B_1}$ may be a concern to the EPBs. Several responses can be made. First, the word "potential" should be emphasized. To the extent that the PLS is not collected from all sources or is incompletely collected from those that do pay something, the amount $\mathbf{B_2}$ may exceed the partial amount currently collected on E' emissions. A second response to concerns about reduced PLS revenues could be an increase in the PLS levy rate by some amount that would replace part or all of the revenue loss $\mathbf{B_1}$. Such an increase in the levy rate would change the split of the scarcity rent between the EPB and the polluting firms, but the amount of the increase in the EPB share may not need to be large. Yet a third response would be to change the split between the EPB's administrative expenses and the amounts reinvested in abatement activities. Instead of allocating 80% of PLS revenue to reinvestment in abatement and retaining 20% for EPB administrative expenses, this ratio could be altered. For instance, using the numbers in Figure 2, if 20% of revenues collected on 90 units of emissions is currently required to fund the EPB, then 36% of the revenues collected on 50 units of emissions must be retained for this purpose. The same amount as currently required will be kept to fund the EPB, and an identical amount will be effectively recycled into abatement projects, since the amount of reduced pollution levy collections, $\mathbf{B_1}$, will be allocated to abatement expenditure by firms as a result of the more stringent cap.

4.3 *Using the Pollution Levy to Reinforce a Tradable Permits System*

The role of the pollution levy need not be limited to that of a convenient pre-existing feature of SO_2 emissions regulation that can continue to fund the EPBs. In the spirit of taking full advantage of what is already available and familiar, the pollution levy can be used to reinforce the tradable permits system, constituting an integrated package of instruments working together to achieve environmental goals. Two such reinforcing uses can be envisaged.

4.3.1 A Second-tier penalty rate

Like any other regulatory system, a tradable permits system must penalize noncompliance. The pollution levy is an obvious candidate for performing this role because of its former use as a charge for emissions in excess of some required level of control. Given the recent decision to extend the existing pollution levy to cover all emissions from a facility, such a use implies a second, higher tier so that an emitting firm would surrender permits and pay the low, first-tier pollution levy for all covered emissions and a higher, penalty rate for any emissions not covered by allowances.

The critical issue in such a system is the level of the second (penalty) tier. A penalty rate that is too low provides firms little incentive to reduce emissions beyond what is justified by the existing pollution levy, or to seek trading partners. A rate that is too high can lead to exorbitant costs or, more likely (and worse), exemption from the tradable permits system in the absence of trading partners. The critical issue is the existence of a market in which sources facing relatively high marginal abatement costs can find sources facing relatively low marginal abatement costs, with which to trade allowances. When such a market can be assumed, a high penalty rate is desirable, both to ensure that the cap is observed and to encourage trading. If a market for allowances is nonexistent or slow to develop, however, a high penalty rate can impose an unjustifiably high cost; a lower, transitional penalty rate may be preferable to provide an initial incentive for further emission reduction and some emissions trading.

Cap-and-trade systems with a penalty for uncovered emissions many times higher than any conceivable market price are characterized as having “tight” or “hard” caps because no rational agent would consider emitting without an allowance. The U.S. Acid Rain Program, in which the penalty is over ten times the highest market price observed, is such a system, but its trading is nationwide in scope, and unlimited banking is allowed. The situation facing a Chinese EPB attempting to implement measures to meet its cap is very different. At best, trading will be local, and, for an economy transitioning to market institutions, local allowance markets may be slow to develop. Under these circumstances, a penalty rate that is higher than the first-tier levy rate but lower than the intended eventual level would be appropriate.

The current pollution levy is widely recognized as being too low to provide much incentive to reduce emissions, but increasing the levy is resisted. Tradable emission permits offer a way out of this impasse, but a “hard” cap with the accompanying high penalty rate could be problematic in an economy without well-developed, long-standing market institutions. An initially lower penalty rate avoids the undesirable effects of a high penalty during the transition to markets, while preserving the needed incentives both to abate and to trade. An initially lower penalty rate also recognizes the likelihood that an EPB’s allocation of the national TEC limit will be more a “soft” cap at first than a “hard” cap, and that the latter will be possible only as allowance markets develop. Over time, and as markets develop, the penalty rate can be raised to a level that would ensure the TEC target becomes a “hard” cap.

4.3.2 Using the first tier as a second instrument for achieving local environmental goals

The other reinforcing use of the pollution levy is as a second instrument for achieving local air emission goals. As explained in Section 3.3.1, a potential conflict exists between meeting local ambient air quality goals and trading emissions over broad areas. Experience with the U.S. SO_2 allowance trading program suggests that problems with “hot spots” may not be as severe as is often feared. The potential remains, however, and the emergence of a hot spot depends only on the right combination of circumstances. Where those circumstances obtain, a second instrument is needed to ensure that local air quality goals are met. That second instrument could be some form of command-and-control regulation operating independently of the permit trading system, as discussed above, or it could be achieved by raising the level of the pollution levy’s first tier within the local area of concern to make the purchase of allowances from outside the local area less attractive.

The first tier of the pollution levy becomes a potential instrument because of its interaction with a tradable permits market. As explained above, the price of allowances in a fully developed market will equal the marginal cost of abatement that firms incur to meet the aggregate quantity restriction less the first-tier pollution levy. Increasing the levy for any subset of firms creating a hot spot will reduce their willingness to purchase permits and thereby increase their incentive to abate more. As a result, the price of permits in the market will decline somewhat, causing other firms not subject to the higher first-tier levy rate to abate less, so that emissions within the cap will be redistributed away from the hot spot. There will be a difference in marginal abatement costs between the two zones, but that difference will reflect no more than the differing value placed on emission reduction in the two areas. More importantly, trading between the two zones could still occur so that the efficiency advantages of tradable permits could be obtained, within the limits imposed by different air quality goals. Eventually and ideally, a broad geographic market, perhaps China-wide, can be envisaged in which the allowed TEC emissions are distributed by a combination of allowance trading and local tweaking of the first-tier pollution levy rate to ensure that local ambient air quality requirements are achieved.

5 CONCLUSION

The creation of an effective SO_2 cap-and-trade system in China will not be easy, but the difficulty should not be overstated. Most of what may seem difficult about a tradable permits system will be required for any regulatory system that can effectively control SO_2 emissions. Tradable permits systems impose distinctive and unfamiliar requirements for how a regulator goes about setting up the regulatory structure, but the differences from alternative systems are more in form than in substance. Where the alternative is a command-and-control system, as is usually the case, the gains in economic efficiency and environmental effectiveness that are possible through tradable permits would seem to justify the greater effort that may be required to overcome the novelty of tradable permits.

Chinese regulatory authorities do not have the luxury of creating a tradable permits system without regard to either the environmental measures already in place or the market institutions that are presumed by market-based environmental instruments. Fortunately, the policies that have already been developed for controlling SO_2 emissions in China are not inconsistent with developing a tradable permits system. In particular, the national policy of Total Emissions Control implies the use of instruments that fix quantities; however, implementation depends on regulatory entities at the local level, where discretion about instrument choice is great and where decisions are likely to be dominated by very practical considerations. The two most important considerations have been emphasized in this paper: 1) how to issue facility permits that will allow and even encourage the development of a tradable permits system, and 2) how to integrate tradable permits with the pre-existing pollution levy system. Neither issue poses an insuperable problem to the creation of a tradable permits system.

Although much can be learned from experience using tradable permits in the U.S. and elsewhere, conditions in China are unique. China's simultaneously top-down and bottom-up structure of environmental regulation places a premium on local experimentation and incremental progress, and this resolutely pragmatic approach precludes the all-at-once, top-down implementation that characterized the U.S. Acid Rain Program. Instead, the process seems likely to be one in which Total Emissions Control targets will be progressively transformed from "soft" caps embedded in facility permits into "hard" caps using tradable permits as markets develop and appropriate regulatory mechanisms are put in place. The pace of this transformation in regulatory approach will depend inevitably on the more general transformation of the Chinese economy. Recognizing this dependence is, however, no reason to delay implementation of instruments that presume market institutions. Indeed, their adoption will anticipate and facilitate this more important transformation without adversely affecting the achievement of environmental goals.

REFERENCES

Benkovic, Stephanie, "SO_2 Emissions Control Policy in China," staff paper, Clean Air Markets Division, U.S. Environmental Protection Agency, July 1999.

Ellerman, A. Denny, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth M. Bailey, *Markets for Clean Air: The U.S. Acid Rain Program*, Cambridge University Press, 2000.

Luo, Hong, Jinnan Wang, Jintian Yang, and Zi Liou, "Recent Developments of Cleaner Air Legislation and its Implications for SO_2 Emissions Trading in China," paper presented at the Second EPA-SEPA Workshop on SO_2 Emissions Trading, Washington, D.C., October 2000.

Meng, Fan, Fahe Cai, Jianxiang Yang, and Yifen Pu, "Management and Monitoring of SO_2 Emissions Sources in China," paper presented at the First SEPA-EPA Workshop on SO_2 Emissions Trading, Beijing, China, November 1999.

Montero, Juan-Pablo, Jose Miguel Sanchez, and Ricardo Katz, "A Market-based Environmental Policy Experiment in Chile," MIT CEEPR Working Paper 2000–005, August 2000.

Tietenberg, Thomas H., *Environmental and Natural Resource Economics,* 5th edition, Addison Wesley Longman, 2000.

Wang, Hua, "Pollution charge, community pressure and abatement cost: An analysis of Chinese industries," working paper of the Development Research Group, World Bank, January 2000.

Wang, Hua, and David Wheeler, "Endogenous enforcement and effectiveness of China's Pollution Levy System," working paper of the Development Research Group, World Bank, undated.

Weitzman, Martin L., "Prices vs. Quantities," *Review of Economic Studies*, XLI, pp. 477–91, October 1974.

Wu, Zuefang, Jinnan Wang, and Fan Meng, "Proposed Scenarios for Total Amount Control of SO_2 during the Tenth Five-years in China," paper presented at the Second EPA-SEPA Workshop on SO_2 Emissions Trading, Washington, D.C., October 2000.