China’s Rate-based Approach to Reducing CO_2 Emissions: Strengths, Limitations, and Alternatives

Lawrence H. Goulder*

Richard D. Morgenstern**

We are very grateful to Dallas Burtraw, Valerie Karplus, Jeremy Schreifels and Xiliang Zhang for providing very useful information and advice as we prepared this paper.

* Stanford University, Resources for the Future, and National Bureau of Economic Research
** Resources for the Future

ABSTRACT

China is launching what is expected to become the world's largest carbon dioxide emissions trading system (ETS). A key feature of this new system – and one that distinguishes it from other ETSs around the world – is its rate-based structure. This structure governs the way emissions allowances are allocated and the conditions for compliance. Relative to the more common, mass-based ETSs, it generates different incentives to producers regarding their levels of output and extent of emissions abatement. These differences have important implications for the performance of the program.

This paper looks closely at the rate-based aspect of China's new program and suggests what it implies for the efficiency, cost-effectiveness, and distributional equity of the program. After acknowledging some attractions of the rate-based system, we identify three key limitations in such a system in terms of cost-effectiveness and efficiency. First, we show the rate-based system implies a gap between the marginal costs of abatement as perceived by firms and the marginal cost to society of such abatement. This discrepancy limits the ability of allowance trading to reduce system-wide costs of abatement. Second, we indicate that there is substantial heterogeneity across facilities in the assigned benchmark emissions-output ratios. We show that in the rate-based system, perfectly fluid allowance trading (i.e., trading without restrictions or transaction costs) does not eliminate the adverse impact of such heterogeneity on cost-effectiveness. Third, relative to a mass-based system with the same assignments of benchmark emissions-output ratios to covered facilities, the rate-based system leads to higher output and emissions, all else equal. This reflects the implicit subsidy to output under a rate-based system. The implicit subsidy implies inefficiently low output prices even when the stringency of the benchmarks implies an optimal level of emissions reductions in the aggregate.

These findings suggest that the conversion of China's system to a mass-based system would yield significant gains in terms of efficiency and cost-effectiveness. We indicate that these gains could be realized while addressing concerns about distributional equity.

I. Introduction

China is in the process of launching what is expected to become the world's largest carbon dioxide (CO_2) emissions trading system (ETS). Focusing initially on the electric power sector – which currently is responsible for a third of China's CO_2 emissions -- the program aims to add major industrial sectors to the trading regime, eventually covering more than half of the country's total emissions. The new system will help China meet its Paris pledge of a 60-65% reduction in CO_2 emissions by 2030, in line with its goal of achieving peak CO_2 emissions in the same year.

Internationally, much is riding on this program. If successful, it could serve as a positive model and encourage other nations' climate policy efforts. Failure could impede the adoption of emissions trading programs in many parts of the world.

Virtually all of the existing national and sub-national cap-and-trade systems are *mass-based* programs, meaning that they impose absolute caps on emissions in each compliance period. Examples of such programs include the European Union's emissions trading system (EUETS) and the linked California and Canadian provincial programs. A key distinguishing feature of the Chinese ETS is that it is *rate-based*: to comply in a given period, a facility must achieve a specified ratio of emissions to output rather than reduce emissions to some absolute level. An ETS of this type is often called a tradable performance standard, where "performance standard" refers to the maximum allowed emissions-output ratio. Heretofore, the most economically significant example of a tradable performance standard was the lead phasedown program established in the U.S. in 1985.

The rate-based feature of China's new national program has a major influence on the program's performance. It governs the way emissions allowances are allocated and the conditions for compliance. Relative to the more common, mass-based ETSs, it generates different incentives to producers regarding their levels of output and extent of emissions abatement. This paper looks closely at the rate-based aspect of China's new program and suggests what it implies for the efficiency, cost-effectiveness, and distributional equity of the program. As part of our evaluation, we consider the merits or limitations of China's design relative to those of the more typical, mass-based programs operating in other nations.[1]

On December 15, 2017, the Chinese government announced some important details about its new system.[2] In the initial phase, which concentrates on the power sector, almost all of the nation's 1700 mostly coal-fired electricity generators will be covered. Once the program has expanded, it is expected to include an additional six sectors (including aluminum, cement, iron and steel, chemicals, paper, and nonferrous metals) and to cover about four times the number of facilities in the seven provincial- and municipal-level pilot emissions trading programs currently operating in China. We consider the program not in its initial one-sector form but after it has achieved broader coverage.

The national program is being developed by the National Development and Reform Commission (NDRC), with implementation responsibilities generally delegated to the provinces. Important provincial responsibilities include the distribution of allowances based on national government guidelines, and the validation of the emissions and production data necessary to assess and ensure compliance. The program will cover CO_2 emissions associated with the burning of coal and other fossil fuels.[3] Emissions trading may occur across all provinces and all covered sectors.

---

[1] Use of rate-based versus mass-based trading systems is just one of the issues involved in China's ETS. Other important issues, including the measurement, reporting and verifications of emissions (and output); compliance; administered pricing; and the overall stringency of the system, are discussed in a recent issue of *Economics of Energy & Environmental Policy* (Goulder and Morgenstern, *eds*.)

[2] See Chemmick (2017) for details.

[3] Since electricity prices are subject to regulation in China, the marginal abatement costs of CO_2 abatement are unlikely to be passed on to consumers. Accordingly, to create incentives for greater efficiency by large consumers of electricity or heat, China's ETS will require these consumers to submit emissions allowances based on the emissions associated with electricity production. These emissions are referred to as indirect CO_2 emissions.

II. Structure

Allowance Allocation

Every ETS – whether mass- or rate-based – requires a mechanism for allocating allowances to covered facilities. To the extent that allowances are given out free[4], rules are needed to allocate them across covered entities. In the first two trading periods of the European EUETS, spanning the period 2005-2012, free allowances were given to individual facilities on the basis of their historical emissions. More recently, the trading programs in California and various Canadian provinces, as well as the revised third-period program in the EUETS, have all relied on benchmarking, which bases the number of allowances received by a facility on a technology- or industry-specific emissions-output ratio rather than on historical levels of emissions.

China's nationwide ETS also relies on benchmarking to allocate emissions allowances. However, its system applies benchmarking in a way that differs from its use in other systems. A key difference pertains to whether the initial (beginning-of-period) benchmark-based allocation of allowances is updated at the end of the compliance period. Under mass-based ETSs that employ benchmarking, the initial allocation is not adjusted over the compliance period. This implies, in particular, that a covered firm cannot increase its entitlement of government-provided allowances through changes in its level of production during the period. Hence, for example, if the firm's assigned benchmark emissions-output ratio is β and its original allocation was based on the product of β and its original level of output, the firm cannot entitle itself to an increase in allowances by increasing its level of output over the period. And since the number of allowances allocated to a given firm does not change over the compliance period and trading of allowances within the compliance period does not change the total in circulation, the system-wide cap is determined strictly by the number of allowances distributed at the beginning of the period.[5]

---

[4] Introducing the allowances via an auction is an alternative to free allocation. Historically, regulators have relied principally on free allocation, though auctioning is now gaining importance among ETS systems. Several systems, including California, use a combination of free allocation and auctioning.

[5] Some ETSs include provisions that allow entities to borrow the allowances that it has been promised for future compliance periods, or bank some of its current allowances for use in future periods. In this case, aggregate emissions can exceed (if there is net borrowing) or must fall short of (if there is net banking) the sum of currently issued allowances. When there are provisions for intertemporal borrowing or banking of allowances, the effective cap is on cumulative emissions, and this cap is equal to the sum of the allowances introduced over time.

In contrast, China's nationwide ETS employs a two-step process that allows for end-of-period updating of the allowance allocation. This has important implications for firms' incentives and for program performance. Specifically, at the start of the compliance period, a covered facility in China's ETS receives a number of allowances equal to its output from the previous compliance period, y_0, multiplied by the facility's designated benchmark emissions-output ratio, β, and an "initial allocation factor," α, which is set by the regulatory authorities. At the time of this writing, China has not yet specified the value it will employ for α, although a 0.6 value has been widely discussed. With a value of 0.6 for α, the facility would initially receive 60 percent of the allowances it would need to justify the emissions-output ratio β if its level of output did not change from that of the previous period.

The second step in the process comes at the end of the compliance period, at which time a covered entity receives the quantity of additional allowances needed to bring the ratio of total allowances to end-of-period output in conformity with the sector-specific benchmark emissions-output ratio. The needed quantity of additional allowances is given by the formula:

$$(1) \quad a_1 = \beta y_1 - \alpha\beta y_0$$

where y_1 is end-of-period output and a_1 represents the additional allowances given to the firm at the end of the period. The first term on the right-hand side is the total number of end-of-period allowances that would be consistent with the benchmark ratio. The second term is the number distributed to the covered facility at the beginning of the period. Thus, the difference between the two is the number of additional allowances needed.[6]

It is theoretically possible for a_1 to be negative – for a facility to receive excessive allowances at the beginning of the period. This happens when y_1 is lower than αy_0. A negative value for a_1 could put the government in an awkward position at the end of the compliance period, since in this case applying the above formula would oblige it to take away from the

---

[6] In fact, each province has the option of reducing the allocation of allowances to facilities within the province if it wishes to make the program more stringent locally. The NDRC sets national benchmark emissions-output ratios, but the provincial government can reduce them. A reduction corresponds to a lower value for $\bar{e}$ in equation 1 and thus a smaller number of allowances allocated to the relevant entities in the province in question. It is also our understanding that the central government will also offer "reserve allowances" to governments in some low-income provinces, additional allowances that these governments can allocate according to their own chosen criteria.

facility some of the allowances it had given out it at the beginning of the period. The likelihood of this occurring depends on the value of the initial allocation factor α, which, as noted, has not yet been officially specified. In any event, it seems clear that the program will utilize a value sufficiently below 1 to make it unlikely that a_1 ends up negative.[7]

Thus, in contrast with more typical mass-based ETSs, under China's new program the number of allowances an entity receives depends on its end-of-period emissions-output ratio. This makes the Chinese system rate-based in a strong sense and means that it is properly regarded as a tradable performance standard.[8] The uncertain connection between allowances a facility initially receives and its end-of-period emissions-output ratio implies that, in contrast to more typical systems, the nationwide cap is not specified in advance by the regulatory authority but instead depends on firms' output decisions during the compliance period. This dependence of the cap on production levels has both attractions and limitations; we discuss these in Section III below.

Determinants of the Benchmarks

The extent to which China's program will reduce CO_2 emissions depends fundamentally on the choice of benchmarks, about which we have only limited information at this writing. Historically, benchmarks have reflected a range of considerations, including technological, economic and institutional factors. In the case of the California ETS, uniform benchmarks are set for all facilities in an industry, based on the top decile of industry performers in terms of their historical emissions per unit of output. In some cases, broad industrial categories are subcategorized depending on the predominant technologies in use. One implication of this is that when technologies differ across sectors or subsectors, benchmarks will differ across these groups as well. It is also possible for benchmarks for a given technology to differ based on the location (e.g., state or province) in which the facility is located, although this is not the case in the California, Canadian or EU systems.

---

[7] Note, for example, that if α is .5, y_1 would need to be 50 percent below y_0 for a_1 to be negative.

[8] A cap-and-trade system is rate-based (and can be categorized as a tradable performance standard) when compliance directly depends on a stipulated ratio of emissions to output. Thus, for example, the Clean Power Plan's default option was rate-based tradable performance standard since compliance by a given state required it to achieve a ratio of emissions to electricity output that did not exceed some given standard.

China's current plan is to distinguish 11 benchmarks for the power sector, differing according to the technology/fuel combination being used. Thus, for example, gas facilities are expected to have different benchmarks from coal units, and gas-boilers will likely have different benchmarks from those of combined-cycle generators.

Beyond the strictly technological considerations, differences in benchmarks may also reflect considerations of regional economic development. Chinese officials have long debated how to deal with cross-provincial differences in economic development across the nation's 31 provinces. In China's ETS, benchmarks for given technologies are likely to differ across provinces as a way of addressing these differences.[9] As shown in Table 1, the cross-provincial differences in development levels are substantial: average per capita income varies by more than three hundred percent across the provinces, while average energy intensity differs by more than a factor of six.

Table 1. Differences among Selected Provincial-level Regions (2015)[10]

Region	GDP	Population(million)	GDP (1000	Energy Consumption (Million tce)	GDP Energy Consumption Intensity (tce/million Yuan)	Industrial Structure (Primary: Secondary: Tertiary)	Reduction target for CO_2 intensity of economy (2016-2020)	Geograph
Beijing	2,301	22	105	69	30	0.6:19.6:79.8	20.5%	East
Jiangsu	7,012	80	88	323	46	5.7:45.7:48.6	20.5%	East
Guangdong	7,281	108	67	301	41	4.6:44.6:50.8	20.5%	East
Jilin	1,427	27	53	81	57	11.2:51.4:37.4	18%	Central
Henan	3,701	95	39	232	63	11.4:49.1:39.5	19.5%	Central
Shanxi	1,277	37	35	173	135	6.2:40.8:53.0	18%	Central
Xinjiang	932	24	39	157	168	16.7:38.2:45.1	12%	West
Ningxia	291	7	42	54	186	4.2:57.9:37.9	17%	West
Yunnan	1,372	47	29	104	76	15.0:40.0:45.0	18%	West

---

[9] It is also our understanding that provincial authorities will have the authority to tighten benchmarks beyond those set by the national government if, for example, they want to demonstrate a stronger commitment to emissions reduction.

[10] Data from various documents published by national and provincial-level authorities.

Source: Maosheng DUAN and Li ZHOU, 2017. "Key Issues in Designing China's National Carbon Emissions Trading System," in Goulder *et. al.* 2017.

For a given technology in a sector, it is expected that China's planners will set the benchmark according to the emissions-output ratio of a facility at a certain point in the distribution of the emissions-output ratios for facilities in the particular fuel/technology/location groups.[11]. Thus, if the target percentile is the 75th percentile in the distribution (ranking the facilities from the highest emissions-ratio to the lowest), the emissions-output ratio for the facility in that percentile becomes the benchmark.[12] Planners have not yet specified the boundaries of the fuel/technology/location subcategories or the targeted percentiles: they are expected to be announced by March 2018.

Impediments to allowance trading can hamper cost-effectiveness in any ETS, whether mass-based or rate-based. China's application of 11 different benchmarks for the power sector might be a cause for concern among analysts concerned with cost-effectiveness. Differences in benchmarks can imply differences in facilities' marginal abatement costs prior to allowance trading. If there are significant impediments to allowance trading, such differences will be sustained, implying a sacrifice of cost-effectiveness. However, in China's rate-based system, heterogeneity of benchmarks would hamper cost-effectiveness even if there were no impediments to allowance trading. We address this important issue in Section III below.

We have emphasized four key aspects of the structure of China's nationwide ETS. First, the allocation of allowances across facilities is based on benchmark emission-output ratios rather than historical levels of emissions. Second, the system will employ 11 differentiated benchmarks for the power sector alone, and potentially dozens or hundreds more once the full system is operational. Third, the program involves end-of-period updating, according to which allowable emissions depend on the covered entities' levels of production during the compliance period. Because of the updating, the aggregate number of allowances introduced in any given compliance period – the aggregate cap -- is not established at the beginning of the period but

---

[11] This is similar to the method used in California, although China seems to be defining a larger number of subcategories for a given industry.

[12] This is roughly similar to the allocation methodology that was developed for the US Clean Power Plan promoted by the Obama administration.

instead depends on covered entities' levels of production during the period. And fourth, the program allows for trading of emissions allowances across sectors and regions.

In the next section, we consider the attractions and drawbacks of these features along the three dimensions noted at the outset: efficiency, cost-effectiveness, and distributional equity.

III. Attractions, Limitations, and Alternatives

Attractions:

China's commitment to a nationwide cap-and-trade system is a major step for climate change policy. As noted, it provides an important signal to other nations and thereby could help encourage carbon pricing policies elsewhere.

Beyond the international implications, the new ETS offers potential domestic rewards, as it improves on China's earlier pilot ETS programs. The broader sectoral and geographical coverage in the fully operational system increases the opportunities to identify and exploit low-cost opportunities for emissions abatement across the country. Broader scope alone could contribute to lower costs even if the program had no provisions for trading allowances, but the new ETSs provisions for trading of allowances across sectors and regions could lower costs further.

In addition, the rate-based structure has certain attractions. A principal advantage of issuing allowances based on a ratio of emissions to actual output during the compliance period is that it can more easily adapt to economic conditions. Because the standard is expressed as a ratio, with output in the denominator, the amount of emission reduction needed for compliance adjusts automatically to current economic conditions. If the economy is booming (and output is high), the allowable level of emissions increases. When the economy is in a slump (and output is low) the allowable emissions are more limited. This flexibility can help avoid exceptionally high allowance prices and abatement costs during boom times, while helping assure that allowance prices do not plummet during economic slumps.[13]

---

[13] Price spikes have occurred in existing national and regional ETSs. Regional Clean Air Incentive Market (RECLAIM) emissions trading program in the Los Angeles area initially had no mechanisms to adjust allowance supplies to changing macroeconomic conditions. In the summer of 2000, the combination of a booming economy

A second attraction is the potential to use benchmarking within a rate-based system to serve distributional objectives. Higher benchmark emissions-output ratios – more easily achieved by the entity to which the benchmark applies -- can be assigned to the lower-income provinces to reduce their regulatory burden and thus soften the differential distributional impacts of a carbon pricing system.[14] It should be noted, however, that mass-based systems also can employ differing benchmark stringencies to achieve distributional goals.

Some Limitations:

In order to assess the limitations of the rate-based approach, we need to examine closely the incentives faced by firms covered under this system. Here we focus on the incentives of a firm that produces output y as a function of a single input x and emissions e: $y = f(x,e)$, with $\partial y / \partial x > 0$ and $\partial y / \partial e > 0$.[15] The firm is included in a rate-based system and faces the benchmark emissions-output ratio β. It regards its output price and the price of its input as exogenous. The market price of emissions allowances is also exogenous to the firm.

The firm's profit (π) is given by:

$$\pi = p \cdot y - p_x x - p_a (e - a_0 - a_1) \tag{2}$$

where p is the output price, p_x is the price of input x, and p_a is the emissions allowance price. Suppose the firm is awarded a_0 in allowances (free) at the beginning of the period. This is determined by the benchmark, β, and some initial measure of output, as characterized in equation (1) from Section II. Let a_1 represent the additional allowances the firm receives at the end of the period based on its assigned benchmark, its end-of-period output, and the allowances already received, a_0. The firm's profit-maximization problem is to choose the optimal levels of output and emissions, taking into account the fact that it would need to hold $\beta * y$ in allowances

---

and other factors caused allowance prices to jump dramatically. Average price per ton in 2000 exceeded $20000, up from $257 in 1999 (SCQMD, 2002).

[14] As mentioned earlier, it is also the case that China's central government has given the provincial governments authority to lower the benchmarks in their area if they want to demonstrate a stronger commitment to emissions reduction.

[15] The main results would be the same if a vector of inputs were considered.

at the end of the period to be in compliance. Thus, as indicated in equation (2), if the firm was initially given a_0 in allowances and its level of output entitles it to another a_1 in free allowances, then it will need to pay $p_a*(e - a_0 - a_1)$ for additional allowances by the end of the period to be in compliance.[16]

The first-order condition for profit-maximization is:

$$(3) \qquad d\pi / de = p\left(\frac{\partial f}{\partial e} + \frac{\partial f}{\partial x}\frac{dx}{de}\right) - p_x \frac{dx}{de} - p_a\left(1 - \frac{da_1}{de}\right)$$

As shown in the appendix, and based on the definition of the benchmark, the above expression implies:

$$(4) \qquad \underbrace{\underbrace{p\left(\frac{\partial f}{\partial e} + \frac{\partial f}{\partial x}\frac{dx}{de}\right) - p_x \frac{dx}{de}}_{MBsoc} + \beta p_a \frac{dy_1}{de}}_{MBfirm} \ = \ p_a$$

The left-hand side of (4) is the marginal benefit to the firm from an increment of emissions. It consists of the value of the change in output y (first term) minus the change in the value spent on input x (second term) plus the value to the firm of the additional allowances offered as a result of the increase in output (third term). The right-hand side is the marginal cost of the additional unit of emissions. The firm maximizes profits by choosing the level of emissions that equates the marginal benefit and marginal cost.

As shown in the appendix, the difference between equation (4) and the corresponding equation under a mass-based system (in which the allocation of allowances does not depend on current-period output) is the presence in the above equation of the far-right term on the left-hand side ($\beta p_a (dy_1/de)$). This term represents the marginal benefit to the firm from the induced increase in emissions allowances associated with the increase in output. This element of the marginal benefit expression, representing the value of the increased allowances generated by the

---

[16] If $e - a_0 - a_1$ is negative, it will be able to sell $e - a_0 - a_1$ of its allowances and remain in compliance.

rate-based system, is a subsidy that accrues only to the firm, not to society at large. As a result, the marginal benefit function is higher in the case of a rate-based as opposed to a mass-based allocation regime: when evaluated at any given value for *e*, the marginal benefit of an increment to emissions is higher than in the case where the firm's allowance allocation is independent of its output. Equivalently, when evaluated at any value for *e*, the marginal cost of an incremental emission reduction is higher in this case, since the opportunity cost (foregone marginal benefit) is greater. The subsidy component, represented by the term $\beta p_a (dy_1/de)$, creates a wedge between the marginal benefit to society from emissions (shown as *MBsoc* above) and the marginal benefit to the firm from emissions (shown as *MBfirm* above). The subsidy term is not an element of marginal social benefit because it is a transfer rather than a resource cost.

Equivalently, one can write equation (4) as:

$$(4') \qquad p\left(\frac{\partial f}{\partial e}+\frac{\partial f}{\partial x}\frac{dx}{de}\right)-p_x\frac{dx}{de} \;=\; p_a-\beta p_a\frac{dy_1}{de}$$

which indicates that, from the firm's point of view, the effective price of an emissions allowance (right-hand side) is lower than p_a. Thus, firms will prefer to purchase more allowances (for a given market-equilibrium price p_a) than a mass-based system involving the same benchmarks.

Three key results emerge from the analysis above:

1. *For a given array of assigned benchmarks, the rate-based system leads to higher output and emissions, all else equal.* The higher marginal abatement cost function in rate-based case implies that, for a given p_a, firms will reduce emissions by less than under a mass-based system. Taking advantage of the implied subsidy to output, firms will increase output and receive the associated additional allowances. Because emissions are higher, the demand for allowances is greater than in Case 1 (for the same array of benchmarks); the supply of allowances is higher as well. The market equilibrium is characterized by greater output by firms and greater emissions than under the mass-based system.

2. *Complete trading of allowances generally does not yield maximal cost-effectiveness.* The discrepancy between the marginal costs of abatement to the firm and to society implies that

trading of allowances will not bring about maximal cost-effectiveness, even if there are no transactions costs or other impediments to trading. Although perfectly free allowance trading can bring about equality across marginal abatement costs experienced by all trading firms, this cannot be relied upon to equate the social marginal costs of abatement. As long as there are any differences across firms in the subsidy component $\beta\,(dy_1/de)$, firms' marginal abatement costs will differ from society's marginal abatement costs. Since equality of social and private marginal abatement costs is required for maximal cost-effectiveness, the maximal cost-effectiveness level will not be achieved. A corollary: unless β and dy_1/de are perfectly negatively correlated, greater variation in benchmarks β across firms implies greater discrepancies in social marginal abatement costs after trades, and lower cost-effectiveness.

3. *The subsidy to abatement implies inefficiently low output prices, all else equal.* Suppose that the benchmarks were just right in the sense that they impose the efficiency-maximizing ratios of emissions to output. The subsidy still implies that the output prices from the covered sectors will be too low from an efficiency point of view. This implies a distortion in industrial, commercial and residential consumers' choices between the output from the covered (carbon-intensive) industries and the outputs from other industries. This result has been obtained in Fischer (2001), Holland *et al.* (2009), and Fowlie (2012).[17]

It should be noted that, compared with the mass-based system, the rate-based system does not necessarily sacrifice efficiency in terms of the *aggregate* level of emissions. Suppose the optimal level of aggregate emissions in the absence of the subsidy (that is, under a mass-based system) is E. And suppose the array of benchmarks $\boldsymbol{\beta}$ gives rise to an aggregate level of emissions in the mass-based system below E. Then the rate-based system could be more efficient (yield higher net benefits relative to the unregulated status quo) by leading to a higher E. At the same time, if the array of benchmarks $\boldsymbol{\beta}$ in the mass-based system implied the optimal level of E, then under the rate-based system the aggregate emissions level would be inefficiently high.

A further implication from the above is that that the more narrowly benchmarks are defined – that is, the more narrow the technology to which a benchmark applies – the bigger is

---

[17] The implicit subsidy to emissions also distorts the firm's choice between input x and emissions e, which has efficiency consequences.

the challenge posed for trading. Consider two scenarios involving the same total allowances allocated initially, but differing in that the benchmarks are defined more narrowly in one than in the other. In the case with the more narrowly defined benchmarks, a larger number of "separate" schedules for marginal abatement costs would need to be reconciled through trading in order to promote cost-effectiveness. As has been discussed, both transactions costs as well as the wedge, for each facility, between *MBfirm* and *MBsociety* limit the extent to which such a reconciliation will occur. Hence cost-effectiveness can be compromised further, the more narrowly the benchmarks are defined.

Finally, a rate-based system can involve higher information costs than a mass-based system. The joint production or multi-product nature of many manufacturing processes creates added challenges in setting benchmarks for some industries. Electric generation has only one principal output (kwh) and one secondary product (steam), and the production of leaded gasoline was the sole product of refining operations involving lead. However, most other industries likely to be subject to regulation for their CO_2 emissions produce multiple products. Determining the CO_2 emissions associated with each of those products is a daunting task. Sectors with a particularly large, intertwined set of products may not be suitable for benchmarks at all. In contrast, if allowances were to be allocated based only on historical levels emissions, the information requirements would be less demanding.[18]

Alternatives:

Below we offer suggestions for modifications to the design of China's ETS that could have advantages in terms of efficiency and cost-effectiveness without loss of distributional equity. However, it is worth mentioning that our discussion ignores many important institutional and political constraints. We do not suggest that the reforms mentioned here would be easily introduced.

---

[18] Of course, accounting only for emissions might raise concerns about fairness. Firms that were especially "good actors" in the past, achieving low emission relative to output, would not be rewarded for the past good behavior.

That said, the key alternative to the rate-based approach is a mass-based system. This alternative would forfeit the ability of China's current rate-based system to adjust emissions allocation levels automatically in response to macroeconomic booms and busts.

At the same time, moving to a mass-based system would offer several potential benefits. First, it would avoid the cost-effectiveness problems identified above. A mass-based system has greater potential to reduce discrepancies between what the firm regards as its marginal abatement costs and the marginal social costs of its abatement. As a result, a mass-based system increases the ability of trades to reduce disparities across firms in the marginal social cost of abatement, thereby lowering society's overall costs of abatement.

A second advantage of a mass-based system is that it is not dependent on the measurement of emissions-output benchmarks. Thus, the regulator can focus on the task of measuring emissions without the added burden of measuring output levels.[19] Shifting to a mass-based system has the additional attraction of making it easier to (eventually) link China's system to mass-based cap-and-trade systems around the world.

A more fundamental change to China's ETS would be to replace the current structure with a carbon tax. A carbon tax would share many of the positive cost-effectiveness attributes of a mass-based cap-and-trade system, e.g., the tendency to equate marginal abatement costs. Distributional concerns could be addressed through infra-marginal exemptions to the carbon tax – which would function much like free allowances function under cap and trade.[20] A potential drawback of a carbon tax is that would make it more difficult to link with other nations' cap-and-trade systems.

Within the context of China's currently envisioned rate-based design, there might be ways to improve the cost-effectiveness of the system. As discussed previously, reducing the heterogeneity of benchmarks and the narrowness of their definitions would improve the system's cost-effectiveness: the smaller the variation among benchmarks the smaller the disparities in marginal abatement costs to society across firms, and the more cost-effective the trading

---

[19] In a retrospective analysis of the US lead phasedown program, Newell and Rodgers (2006) found that the flexibility introduced by the tradable performance standard increased the likelihood of both intentional and unintentional violations, especially on the part of smaller refiners and fuel blenders. This placed additional administrative burdens on the EPA's monitoring and enforcement staffs. The most common violations involved errors in the quantities of lead used and the volumes of gasoline produced and imported.

[20] See Goulder and Schein (2013).

system.[21] China could offset any unwanted distributional implications of such a change by introducing compensatory direct income transfers.

In terms of potential reforms of China's ETS, it might be possible to establish a process to adjust the benchmarks over time to reflect a long-term goal of reducing benchmark heterogeneity. The government could offset any unwanted distributional implications of such a change by introducing compensatory direct income transfers. It might also be possible to apply a mass-based system to some or all of the industrial sectors as coverage of China's ETS is expanded.[22]

Another potential reform would be the promotion of more fluid trading across sectors and provinces. This would benefit China even if the system were to remain fully rate-based.[23] China could take certain steps to encourage fluid trading across sectors and provinces, including the establishment of a strong legal foundation for the program; credible and consistent measurement, reporting, and verification; and comprehensive program review and adjustment.

IV. Conclusions

China's introduction of a national ETS is a major step forward toward the global goal of addressing global climate change. Even at this early stage, it is clear that the system has a solid foundation and addresses important efficiency and distributional issues in a serious way.

At the same time, some major changes to its structure could yield gains in terms of efficiency and cost-effectiveness while addressing concerns about distributional equity. We have found that a rate-based system yields a gap between what the firm regards as its marginal abatement costs and what are the marginal abatement costs society. This discrepancy hampers

---

[21] Under the US Clean Power Plan only two benchmarks were used to calculate emission standards, as opposed to the 11 power sector benchmarks in the Chinese ETS. Simulation modeling could be used to estimate the potential gains in cost-effectiveness from reducing the disparities in marginal abatement costs to society across firms, measured by the heterogeneity of the benchmarks.

[22] Of course, linking the rate- and mass-based system would raise concerns about potential leakage across sectors – a topic beyond the scope of this paper.

[23] In previous pilot programs for SO_2 cross-provincial trades were sometimes restricted out of a concern that such trades could limit a provincial government's ability to meet other emissions-related requirements, or out of a concern that sufficient allowances might not be available within a province to meet future demand (B. Zhang et al. 2016; Tao and Mah 2009).

the ability of allowance trades to lower society's overall abatement costs. Converting to a mass-based system would help promote cost-effectiveness by eliminating this discrepancy.[24]

The move to a mass-based system would also promote greater efficiency. The rate-based system implicitly subsidies output, leading to inefficiently low output prices. This distorts consumer decisions between the output of covered (carbon-intensive) industries and the outputs from other industries. The conversion to a mass-based system would promote efficiency by eliminating this subsidy.

Within the rate-based structure, reducing the heterogeneity of benchmarks would promote cost-effectiveness. The gap between firms' perceived and society's actual marginal abatement costs depends on the benchmarks. Reducing heterogeneity of benchmarks promotes cost-effectiveness by reducing disparities in these gaps. To the extent that reducing benchmark heterogeneity would run counter to some distributional objectives, the government could help meet the distributional goals through direct income transfers.

Promoting more fluid trading across sectors and provinces would also help expand cost-effectiveness, whether or not the system remained rate-based.

It is important to put the issues in this paper in perspective. China's ETS faces very substantial challenges along dimensions beyond those connected with its rate-based structure. One key additional challenge is the need for pricing reform, since for many important sectors (including the power sector) output prices are administered rather than market-determined. A further challenge is the need to improve the capabilities for emissions measurement, reporting and verification, and to strengthen incentives for compliance. China is confronting these important challenges as well. Overcoming them may be as important to the ETS performance as the reforms proposed in this paper.

---

[24] A uniform mass-based system is clearly preferred. Use of a mass-based system for some or all of the manufacturing industries to be added at a later date is also a possibility, although such a mixed system raises other issues not addressed in this paper.

References

Chemmick, Jean. 2017. "World's Largest Carbon Market Poised to Start Slowly," *Climatewire,* December 15, 2017. https://www.eenews.net/climatewire/2017/12/15/stories/1060069093

Duan, Maosheng and Li Zhou., 2017. "Key Issues in Designing China's National Carbon Emissions Trading System," in Goulder, Lawrence H., Richard D. Morgenstern, Clayton Munnings, and Jeremy Schreifels. 2017. "China's National Carbon Dioxide Emission Trading System: An Introduction," *Economics of Energy & Environmental Policy,* Vol. 06 02. doi://10.5547/2160-5890.6.2.lgou.

Fischer, Carolyn. 2001."Rebating Environmental Policy Revenues: Output-Based Allocations and Tradable Performance Standards," RFF Discussion Paper 01–22.

Fowlie, Meredith. 2012. Updating the Allocation of Greenhouse Gas Emissions permits in a Federal Cap-and-Trade Program. In Don Fullerton and Catherine Wolfram, eds., *The Design and Implementation of a U.S. Climate Policy.* University of Chicago Press.

Goulder, Lawrence H. and Richard D. Morgenstern (*eds.*). 2017. "Special issue on China's National Carbon Dioxide Emission Trading System," *Economics of Energy & Environmental Policy,* Vol. 06 02. doi://10.5547/2160-5890.6.2.lgou.

Goulder, Lawrence H., Richard D. Morgenstern, Clayton Munnings, and Jeremy Schreifels. 2017. "China's National Carbon Dioxide Emission Trading System: An Introduction," *Economics of Energy & Environmental Policy,* Vol. 06 02. doi://10.5547/2160-5890.6.2.lgou.

Goulder, Lawrence H., and Andrew R. Schein. 2013. "Carbon Taxes Versus Cap and Trade: A Critical Review." *Climate Change Economics* 4(3).

Holland, Stephen P., Jonathan E. Hughes, and Christopher R. Knittel. 2009. "Greenhouse Gas Reductions under Low Carbon Fuel Standards?" *American Economic Journal: Economic Policy* 1(1).

Newell, Richard G. and Kristian Rodgers. 2006. "The Market-Based Lead Phasedown" in *Moving to Markets in Environmental Regulation* (Jody Freeman and Charles D. Kolstad, editors), Oxford University Press.

South Coast Air Quality Management District. 2002. "Annual RECLAIM Audit Report for 2000 Compliance Year," Board Agenda Item No. 42 Board Meeting.

Tao, Julia, and Daphne Ngar-yin Mah. 2009. "Between Market and State: Dilemmas of Environmental Governance in China's Sulphur Dioxide Emission Trading System." *Environment and Planning C: Government and Policy* 27 (1): 175–188. doi:10.1068/c0768.

Zhang, Bing, Hanxun Fei, Pan He, Yuan Xu, Zhanfeng Dong, and Oran R. Young. 2016. "The Indecisive Role of the Market in China's SO2 and COD Emissions Trading." *Environmental Politics* 25 (5): 875–98. doi:10.1080/09644016.2016.1165951.