China's Unconventional Nationwide CO_2 Emissions Trading System: The Wide-Ranging Impacts from an Implicit Output Subsidy

Lawrence Goulder
Stanford University, Resources for the Future, and NBER

Xianling Long
Stanford University

Jieyi Lu
Resources for the Future

Richard D. Morgenstern
Resources for the Future

November 2019

The authors are grateful to Carolyn Fischer, Charles Kolstad, Mun Ho, Valerie Karplus, Ian Parry, Billy Pizer, Jeremy Schreifels, Rob Williams, Da Zhang, Junjie Zhang, and Xiliang Zhang for helpful advice. We gratefully acknowledge financial support from the Energy Foundation-China and the Stanford Institute for Economic Policy Research's China Program.

ABSTRACT

China has embarked on what has the potential to become the largest CO_2 emissions trading system in the world. To reduce emissions, the nation will rely on a tradable performance standard (TPS), an emissions pricing mechanism that differs in important ways from the emissions pricing instruments used in other countries, such as cap and trade (C&T) and a carbon tax. We employ matching analytically and numerically solved models of China's power sector (the first sector to be covered under China's TPS) to assess, under alternative designs, the cost-effectiveness and distributional implications of the TPS and to compare these impacts with those of an equally stringent C&T program.

We find that achieving given aggregate CO_2-reduction targets is more costly under the TPS than under C&T. This reflects several consequences of the TPS's implicit subsidy to electricity production. The subsidy causes producers to make less efficient use of output-reduction as a way of reducing emissions (indeed, it induces some producers to increase production). It also reduces the extent to which allowance trading can lower costs. And when the TPS employs multiple benchmarks (maximal emission-output ratios consistent with compliance), it distorts the relative contributions of different power plants to emissions reductions. In our central case, the costs of the TPS are about 47 percent higher than under C&T.

Although the TPS has some disadvantages in terms of overall cost, it also has some attractions relative to C&T. Its rate-based structure allows overall policy stringency to adjust automatically to changes the business cycle, and because the TPS causes smaller increases in electricity prices than C&T, it is likely to lead to less emissions leakage. Also, the use of multiple (i.e., varying) benchmarks, while tending to raise aggregate costs, can reduce disparities in the policy's costs across technology types and regions of the country.

Despite its higher overall costs than C&T, the TPS can generate significant net gains once its environmental benefits are counted. If emissions reductions are valued at 290 RMB (or about 44 U.S. dollars) per ton, our central case results indicate that the environmental benefits from the TPS exceed the policy costs by a factor of about 3.

1. Introduction

China has embarked on what promises to be the world's largest carbon dioxide (CO_2) emissions trading system (ETS). When fully implemented, this nationwide system will more than double the amount of CO_2 emissions covered worldwide by some form of emissions pricing.

China will rely on a tradable performance standard (TPS) as its emissions pricing instrument for reducing emissions. This mechanism differs in important ways from the emissions pricing instruments used in other countries, such as cap and trade and a carbon tax. A TPS is a rate-based instrument: the number of emissions allowances granted to a facility depends on the ratio of its emissions to output over the compliance period. Since compliance depends on a ratio, covered facilities can influence the number of emissions allowances they are allocated by changing their output levels during the compliance period. In contrast, under cap and trade (C&T), a covered facility's allocation of allowances is not influenced by within-period production changes. The dependence under the TPS of the allowance allocation on within-period output decisions has important implications for incentives and associated system performance. It significantly affects production levels, overall emissions abatement, and the levels and distribution of costs.

This paper employs matching analytically and numerically solved models to evaluate China's new TPS, focusing on the impact on the nation's power (electricity) sector, the first sector to be covered by the TPS.[1] The power sector includes more than 2,000 coal-fired power plants and is critical to China's climate policy effort, as it currently accounts for over 40 percent of the country's total CO_2 emissions (Yang and Lin, 2016). The sector has been undergoing virtually continuous reform since 1985, when the state monopoly ended (Ho *et al.,* 2017). While electricity prices were set by the government a decade ago, recent reforms allow for market-determined prices of production. The fraction of electricity output sold at market prices has grown steadily over the last decade and now approximates 31 percent.

We apply the analytical and numerical models to assess the TPS's impact on the production costs and CO_2 emissions of power plants of differing technologies, as well as its implications for aggregate costs (lost producer and consumer surplus) and aggregate emissions reductions. We also

[1] Ultimately, the TPS will cover nine major sectors. The cement and aluminum sectors are next in line to be covered, to be followed by iron & steel, nonferrous metals, petroleum refining, chemicals, pulp and paper, and aviation. China's TPS design calls for emissions trading across all facilities and all covered sectors.

examine how costs are distributed across different types of power plants and regions of the country. Throughout, we compare the TPS's impacts with those of a C&T program with similar coverage and stringency.

The TPS's rate-based approach, according to which compliance requires avoiding exceeding a given ratio of emissions to output, contrasts with the mass-based approach of C&T, under which compliance requires avoiding exceeding a given level (mass) of emissions. Under the TPS, the number of emissions allowances the regulator offers to a facility in each compliance period is the product of the maximum emissions-output ratio (or benchmark) assigned to the facility and the facility's level of output in that period. [2] Fischer (2001) and Fischer and Newell (2008) have shown that a rate-based system like the TPS implicitly subsidizes output, since additional output increases the number of (valuable) allowances a facility will receive from the regulator. These authors point out that because of this implicit output subsidy, a TPS tends to be less cost-effective than an equivalent C&T system.[3]

Our theoretical model builds on this earlier theoretical work. We advance the theory by exploring the implications of multiple (i.e., varying) benchmarks – an important feature of China's planned TPS. Differing benchmarks can help serve distributional goals, since higher (that is, less stringent) benchmarks can be assigned to facilities that otherwise would face especially high compliance costs. Our theoretical model shows that greater variation of benchmarks, while addressing distributional goals, reduces cost-effectiveness (that is, raises the cost of achieving any given aggregate emissions-reduction target), other things equal. Greater benchmark variation increases costs because it alters the relative magnitudes of the implicit output subsidies across

---

[2] More precisely, a rate-based system's benchmarks are the assigned emissions-output ratios that covered facilities must meet, net of any emissions credits purchased on the allowance trading market.

[3] In keeping with its rate-based nature, the TPS is sometimes referred to as an example of an intensity-based standard. It is equivalent to a subsidy to output and tax on emissions. Other examples of intensity standards include clean fuel standards and clean energy standards. While the TPS is an *output*-oriented intensity standard (since it focuses on the emissions intensity of output), clean fuel standards and clean energy standards are *input*-oriented. In these cases the tax component of the tax-subsidy combination applies to the fuel or energy input rather than pollution emissions. Studies by Kerr and Newell (2003), Fischer and Newell (2008), Holland *et al*. (2009), Parry and Krupnick (2011), Goulder, Hafstead, and Williams (2016), and several others address the efficiency properties of fuel and energy intensity standards. A feebate is another intensity-based standard. Under a feebate, the subsidy applies to facilities with performance better than the standard, and the tax applies to facilities with emission intensities in excess of the standard. In contrast with the TPS, in which both the tax and subsidy apply to all covered facilities, a feebate involves no output subsidy to facilities that fail to meet the standard, and no tax on facilities that exceed the standard. Parry and Krupnick (2011) assess the economic properties and potential political attractions of this instrument. Fullerton and Metcalf (2001) and Goulder and Parry (2008) compare the incentive effects of a range of instruments, including intensity standards and cap and trade.

covered facilities and thereby distorts the relative outputs of these facilities. Cap and trade also can employ multiple benchmarks for determining the initial allocations of emissions allowances across covered facilities and thereby affecting the distribution of policy costs. But in contrast with the TPS, the use of multiple benchmarks under C&T does not reduce cost-effectiveness. Because a typical C&T program does not include the output subsidy,[4] the extent of benchmark variation across facilities (holding total number of allocated allowances fixed) does not affect decisions at the margin; it only has distributional consequences.

A second contribution of the theoretical model is to reveal that the implicit subsidy reduces gains from allowance trading. This is because the subsidy creates wedges between a facility's private marginal costs of abatement and the social marginal cost associated with that abatement. The size of these wedges generally will differ across the electric power plants. Gains from trade are minimized when the social costs of abatement are equated at the margin, but because individual facilities will focus on private marginal abatement costs, individual generators' trading decisions will not lead to such equality. As discussed in Section 4 below, this compromising of the gains from allowance trading occurs even in the case where the TPS applies the same benchmark to all covered facilities.

In addition, the analytical model provides a close look at the implications of the TPS for electricity output decisions in the context involving heterogeneous electricity producers – specifically, where covered facilities differ in terms of their initial emission intensities. The model shows that under the TPS, covered facilities with exceptionally low emissions-output ratios will tend to *increase* both electricity output and emissions relative to their business-as-usual levels. This contrasts with C&T, which generally motivates all covered facilities to reduce both output and emissions.[5] Firm heterogeneity helps expand the differences between the TPS and C&T in terms of cost-effectiveness.

Thus, the TPS's higher costs reflect three channels of impact that do not apply under C&T. All three channels stem from the implicit output subsidy under the TPS.

---

[4] In Section <xx> we address the case where C&T offers output-based allocation for certain covered facilities. In this case the magnitude of a benchmark influences cost-effectiveness.

[5] As discussed in Section 4, C&T generally leads to increases in electricity prices, and this exerts a positive influence on facilities' output and emissions. It is conceivable that for some facilities, this effect will be large enough to cause them to increase output and emissions. However, our numerical simulations indicate that this price effect is second-order and that nearly all facilities reduce output and emissions under C&T.

Our numerical model yields results consistent with the analytical model's predictions, supplementing the qualitative predictions of the theoretical model with a unique quantitative assessment closely geared to China's power sector.[6] Key findings of the numerical model are as follows.

First, this model finds that the TPS involves higher economy-wide costs than a C&T program of the same stringency and scope, a reflection of the TPS's implicit output subsidy.[7] Under a 3-benchmark TPS (an option given focus in discussions by Chinese policy planners), the TPS would yield a 3.1 percent reduction in aggregate CO_2 emissions. This reduction could be achieved at 47 percent lower private cost under a C&T program with similar allowance allocations. Consistent with the analytical findings, in the numerical model the TPS causes some generating units to expand output, while C&T induces most or all units to reduce output.

Second, the TPS's economy-wide costs rise substantially with the number and variability of benchmarks. A 3-benchmark TPS has 18 percent higher private cost per ton of reduced emissions, compared to a single-benchmark TPS with the same number of allowances initially allocated. Greater variation of benchmarks implies higher costs both by introducing larger distortions in the relative contributions of different facilities to emissions reducitons and by reducing the potential gains from allowance trading.

Third, the distributional impacts of the TPS differ significantly from those under C&T. Because of the TPS's implicit output subsidy, reductions in electricity output contribute a much smaller share to overall emissions reductions than under C&T. The less extensive reductions in output imply smaller increases in electricity prices[8] than under C&T. As a result, producers bear a larger share of the economic burden under the TPS than under C&T, which imposes a smaller fraction of regulatory costs on electricity consumers.

Fourth, the TPS has very different cost-impacts across the Chinese provinces, reflecting differences in technologies and emissions intensities of the generators and the associated

---

[6] This quantitative analysis complements a number of recent empirical studies of China's efforts to reduce CO_2 emissions through emissions trading. See, for example, Duan and Zhou (2017), Ho, Wang, and Yu (2017), Teng, Jotzo, and Wang (2017), Karplus and Zhang (2017), and Zhang, Wang, and Du (2017). Our numerical model is unique in its sharp focus on the incentive effects of the TPS and its ability to yield a close comparison of the impacts of the TPS and C&T.

[7] Other factors can mitigate the potential disadvantages of rate-based approaches such as the TPS. Goulder, Hafstead, and Williams (2016) show that pre-existing distortionary taxes can reduce and sometimes eliminate the potential cost-disadvantage of a clean energy standard relative to cap and trade or an emissions tax.

[8] As noted earlier and discussed further in Section 4, a considerable share of China's electricity prices is now market-determined. Our models account for both government-controlled and market-determined prices.

differences in compliance costs. Under the 3-benchmark central case specification for the TPS, among the generating units that experience losses of profit the largest losses in percentage terms are to generators in provinces in the northern and northeastern regions of the country. We consider an alternative, 4-benchmark policy specification designed to avoid the large cost-impacts in these provinces. In this case, the technologies on which these regions disproportionately rely, and which involve especially high emissions intensities, are given less stringent benchmarks. We find that achieving the distributional objective lowers profits in other regions of the country and gives rise to a significant increase in aggregate policy costs.

Although the TPS is less cost-effective than C&T, it has important offsetting attractions. One is that the TPS's rate-based structure causes policy stringency to adjust automatically in response to current macroeconomic conditions. When the economy is booming, and demand for electricity is relatively high, the expanded output of electricity entitles generators to a larger number of allowances, since allowance allocations are a function of output. Cap-and-trade programs do not have this attribute.

A second potential attraction is that the TPS implies smaller electricity price increases than would occur under an equally stringent C&T program. Smaller price increases suggest less "emissions leakage." To the extent that regulation of China's pollution raises the prices of China's goods relative to foreign goods, consumers will shift toward the imports. If production of the imported goods involves more pollution, this would offset the pollution-reducing goals of the domestic regulation. Thus, to the extent that the TPS yields smaller price increases than C&T, emissions leakage can be reduced.[9] Smaller price increases might also have some political attractions.

A third attraction is familiarity. The TPS's rate-based structure matches that of several of the previous provincial- and regional-level pilot programs for reducing CO_2 emisions. The structure also is in keeping with other rate-based regulations with which China is familiar.

[9] In China, relatively little electricity is imported. Hence the smaller price increases from the TPS relative to the increases under C&T would not likely make much difference in terms of imports of electricity. However, the TPS would also lead to smaller price increases of downstream goods and services, and thus could reduce leakage in the form of shifts to imported downstream goods. The issue of leakage is likely to be more important in the later phase of China's TPS program, when coverage is extended eight industrial sectors – sectors in which domestic production faces more competition from imports.

Thus, although the TPS less cost-effective than a C&T system with the same coverage and stringency, it has important attractions along other dimensions. All of the TPS's limitations in terms of cost-effectiveness derive from its implicit output subsidy.

Despite its higher overall economic costs, the TPS can generate significant aggregate gains once environmental benefits are accounted for. In our central case, the environmental benefits from the TPS exceed the policy costs by nearly a factor of three when emissions reductions are valued at 290 RMB (or about 44 U.S. dollars) per ton.

These issues have significance in other contexts. In many countries, policy makers are making important choices that include whether to adopt a rate-based or a mass-based approach to pollution control. The results shown here for China are highly relevant to their choices.

The rest of the paper is organized as follows. The next section briefly describes key features of the power sector. Section 3 then presents the basic structure of the TPS program, including the implications of alternative approaches to benchmarking. Subsequent sections examine analytically and numerically the potential impacts of the program, considering the emissions reductions achieved, the effects on the pattern and overall level of electricity generation, and the policy costs. Section 4 develops and applies an analytical model to assess qualitatively, within the current regulatory structure, the cost and distributional impacts of the TPS, and compare these impacts with those under C&T. Section 5 lays out the structure, inputs, and solution method of the numerical model. Section 6 then applies the numerical model to assess the cost-effectiveness and distributional impacts of the TPS and C&T. Section 7 offers conclusions.

2. Key Features of the Electricity Sector

Almost 72 percent of electricity produced in China's power sector comes from its fossil-based plants.[10] The sector contained 2,392 coal-fired, circulating fluidized bed , and natural-gas-fired generating units in 2016. Table 1 groups the the units into three main technology categories – coal-fired units other than circulating fluidized be units, circulating fluidized bed units, and gas-

[10] National Bureau of Statistics, "Steady Increase in Power Production, Adjustment and Optimization of Generation Structure", *National Bureau of Statistics*, March 19, 2018, http://www.stats.gov.cn/tjsj/zxfb/201803/t20180319_1588744.html (accessed August 11, 2019)

fired units – and into 11 more specific technology classifications. The table also provides information on outputs, costs and CO_2 emissions intensities for the different technologies.

Among these units, the 300 MW subcritical coal units account for the largest share of electricity production and CO_2 emissions. The 600 MW supercritical coal units, which operate at a slightly lower emissions intensity, are the second largest producers of electricity and CO_2 emissions. As one might expect, the quite limited gas-fired capacity has much lower emissions per mWh.

Regulations imposed by the central government affect electricity output decisions and pricing. For almost every generating unit, the pattern in recent years is that some of the unit's electricity output is sold at prices fixed by the government while some is sold at market prices. Generating units can choose levels of production, but a three-tiered system determines the prices at which the production can be sold. The first tier applies to electricity output up to the amount associated with a government-assigned number of "guaranteed annual utilization hours" of operation. This output faces fixed prices set by the government. The second tier applies to production in excess of the guaranteed-hours (GH) level and up to another level set by the government. This output also faces fixed prices; these prices generally differ from the first-tier prices. Electricity output in excess of the first- and second-tier production is sold at market prices. The principal markets are a "residual local market," to which the generators in the unit's province are the main suppliers, and a "zonal" market, to which units in the several provinces in a given zone contribute. The purchasers in the zonal market are grid companies. As discussed further in Section 6, the market prices generally are below the fixed prices. Forward markets exist for both the residual local and the zonal markets.

A decade ago, nearly all production was in the first or second tier and therefore faced fixed prices. However, the situation has changed in recent years. In 2018, approximately 30 percent of the electricity produced in China was sold at market-clearing prices. The increased importance of market prices reflects the gradual narrowing of the first and second tiers as well as the significant growth in total electricity demand.[11] These developments are consistent with the central government's efforts to expand the role of market-driven prices in the power sector.

[11] Department of Industrial Development and Natural Resources, "A Brief Analysis of National Electricity Trading in 2017", *China Electricity Council,* February 7, 2018, http://www.cec.org.cn/guihuayutongji/dianligaige/2018-02-07/177779.html (accessed August 11, 2019).

Thus, the nature of China's regulation of the power sector implies that individual generators may choose endogenously their production levels, while their ability to sell output at market prices depends on their production levels. These aspects are captured in our models.

3. Structure of the TPS

Emissions trading systems (ETSs) include both tradable performance standards and cap and trade. Allowance trading, a central feature of these programs, promotes a reallocation of abatement activity, leading to greater effort by facilities that can reduce emissions at lower cost. This helps reduce the economy-wide cost of achieving aggregate emissions reductions. China's system allows for trading across regions in the power sector. It is expected that the system will allow for intersectoral trading once the system extends beyond the sector.

In the first two trading periods of the EUETS, which spanned 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 Quebec, as well as the revised third-period program in the EUETS, have relied on benchmarking, according to which the number of allowances received by a facility is based on a technology- or industry-specific emissions-output ratio rather than on historical levels of emissions.

A key difference between C&T and China's TPS relates to the allocation of emissions allowances. Under C&T, in most cases each covered facility's allowance allocation at a given point in time is exogenous to the firm. The number of allowances a firm receives is the product of the pre-established benchmark emissions-output ratio and some fixed reference quantity (usually an historical level of production). To achieve compliance, a facility's emissions, minus any allowances it purchases from other facilities, must not exceed this product.[12]

There are some exceptional cases where the allocation under C&T is endogenous. This occurs where C&T offers "output-based allocation" to certain facilities. Under output-based

[12] 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.

allocation, a facility's allocation in a given is the product of the benchmark and the facility's output in the previous period. In this case, a firm's output choice in a given period affects its allocation in the next period, and thus the allocation endogenous to the firm, although the impact on the allowance allocation comes with a one-period lag. In the EU-ETS, California's C&T, and some other C&T systems, output-based allocation has been applied to certain firms in the manufacturing sector that are designated as the most "emissions-intensive trade-exposed" and thus the most vulnerable to import-competition. Output-based allocation is a way of helping these firms compete internationally: it effectively subsidizes output, since additional output leads to larger allocations of allowances.[13] In reality, output-based allocation tends to be applied only to a small subset of covered firms and not to the power sector.[14]

In contrast with C&T, under China's nationwide TPS the allocation of allowances to each covered facility is endogenous within each compliance period; it depends on the product of the benchmark β_i assigned to each generator i and the level of electricity output q_i chosen by the generator in that period. Because the number of allowances allocated to each generator is endogenous, the aggregate emissions associated with the government-chosen benchmarks is endogenous as well. Thus, unlike C&T, under the TPS the regulator will not know the total number of allowances to be circulated and the aggregate level of emissions until the end of the compliance period, after firms' production decisions over the period have been made.[15] Reflecting the differences in structure, C&T systems are categorized as *mass-based*, since in each period the regulator sets the aggregate level (or total mass) of emissions, while the TPS is categorized as *rate-based*, since the regulator sets emissions intensities but not total emissions.

Under the TPS, China plans to allocate allowances through a two-step process. At the start of the compliance period, a covered facility receives a number of allowances equal to the product of its designated benchmark emissions-output ratio, β, an "initial allocation factor," α, and some

---

[13] Haites (2003), Fowlie (2012), and Fischer and Fox (2012) offer excellent discussions of output-based allocation.
[14] Californis's ETS does not apply output-based allocation to the power sector. The EU ETS applies such allocation to the power sector only in a few exceptional cases.
[15] In C&T systems that include some output-based allocation, the total number of allowances to be issued – the aggregate cap – is set in advance and remains exogenous. Although firms enjoying output-based allocations can affect their allocations through changes in output, these changes do not alter each period's total allocations. Increased allocations to firms enjoying output-based allocation correspond to reductions in allocations to other firms. Thus, the aggregate cap does not change.

measure of output, q_0 (e.g., some recent level of production).[16] 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 its total allocation into conformity with the sector-specific benchmark emissions-output ratio.[17]

The extent to which China's program will reduce CO_2 emissions depends crucially on the choice of benchmarks. It appears that three benchmarks will be employed in the power sector. The three benchmarks apply to three technology *categories*: coal-fired, circulating fluidized bed (CFB), and gas-fired units.[18] We use the term "technology *class*" to refer to more specific technology types. The Ministry of Ecology and Environment distinguishes 11 technology classes. Table 2 displays the 11 technology classes and their groupings into the three technology categories. We use the same groupings in applying benchmarks in the numerical simulations below.

This section emphasizes three key aspects of the structure of China's forthcoming nationwide ETS. First, the program authorizes trading of emissions allowances across regions and (once it expands beyond the power sector) across sectors. Second, in contrast with a C&T system, under the TPS the number of allowances allocated to a covered facility depends on the facility's chosen production level over the compliance period. Thus, the number of allowances allocated is endogenous to firms' production decisions and the aggregate number of allowances introduced in any given compliance period – the aggregate cap – is endogenous as well. Third, the planners

---

[16] 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 q_0.

It is theoretically possible for a facility to receive more allowances at the beginning of the period than the amount it is will be entitled to have received by the period's end. This happens when end-of-period output is lower than αq_0. This could put the government in an awkward position at the end of the compliance period of needing to take away from the facility some of the allowances it had given out at the beginning of the period. The likelihood of this outcome depends on the value of the initial allocation factor α. It appears that the program will utilize a value for α sufficiently below 1 to make it unlikely that the government would encounter this problem with any facility that remains in operation. As discussed below, any facility that shuts down during the compliance period must relinquish its allowances.

[17] 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 Ministry of Ecology and Environment sets national benchmark emissions-output ratios, but the provincial government can reduce them. 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.

[18] Historically, benchmarks have reflected technological, economic and institutional factors. In California's cap-and-trade system, uniform benchmarks are set for all facilities in an industry at the emissions rate corresponding to the best (i.e., lowest) decile emissions-output ratio experienced historically among facilities in the industry. In some cases, broad industrial categories are subcategorized depending on the predominant technologies in use.

seem to be centering on employing three benchmarks in the first (power-sector) phase of the program, one for each of three main technology categories. Differential benchmarking offers a channel for achieving distributional goals. At the same time, as indicated below, it can compromise cost-effectiveness.

The next section develops an analytical model to examine the impacts of the TPS in the power sector and to contrast these impacts with those of C&T. The subsequent two sections present the structure of and results from the corresponding numerical model.

4. Impacts of the TPS: An Analytical Treatment

In the presence of the TPS, managers of a generating unit need to make several interconnected decisions. One is whether to remain in operation or shut down. Generators that remain in operation need also to decide how much electricity to produce and how much to reduce the emissions intensity of production. These decisions depend on the stringency of the benchmark applied to the generating unit, the price of emissions allowances, and the administered and market prices of electricity. The analytical model considers these elements. For transparency this model assumes just one tier to which administered prices apply and does not separate the residual and zonal electricity markets. The key insights from this model are preserved in the results from the more disaggregated numerical model.

a. Net Revenue, Conditional on Remaining in Operation

Let:

q_{ij}	$\equiv$	total end-of-period electricity output of generator *i* in technology class *j*
$\overline{q}_{ij}$	$\equiv$	guaranteed-hour electricity output of generator *i* in technology class *j*
e_{ij}	$\equiv$	CO_2 emissions by generator *i* in technology class *j*
C_{ij}	$\equiv$	total cost of production by generator *i* in technology class *j*
$\overline{p}_{ij}$	$\equiv$	admininistered wholesale price applying to first-tier production of electricity by generator *i* in technology class *j*
p_{ij}	$\equiv$	market equilibrium wholesale price applying to electricity output by generator *i* in technology class *j* in excess of first-tier production

β_j $\equiv$ benchmark emissions-output ratio assigned to generators in technology class *j*

t $\equiv$ market price of emissions allowances

Consider first the choices of a generating unit conditional on its remaining in operation. The generator's[19] choice variables are *q* and *e*. Net revenue π for operating generator *ij* is given by:

$$\pi_{ij} = \bar{p}_{ij}\bar{q}_{ij} + p_{ij}\left(q_{ij} - \bar{q}_{ij}\right) - C\left(q_{ij}, e_{ij}\right) - t\left(e_{ij} - \beta_j q_{ij}\right) \tag{1}$$

The first right-hand term in (1) is the revenue from production of electricity up to $\bar{q}_{ij}$, the highest level of output subject to the administered tier 1 price $\bar{p}_{ij}$. The second right-hand term is the revenue from electricity output in excess of $\bar{q}_{ij}$. The third and fourth terms refer to total production cost and the expense or revenue associated with allowance purchases or sales. We assume $\partial C_{ij} / \partial q_{ij} > 0$ and $\partial C_{ij} / \partial e_{ij} < 0$. We also assume that each generator's objective is to maximize net revenue.[20] For simplicity of exposition, equation (1) and subsequent equations in this section reflect the assumption that $q_{ij} > \bar{q}_{ij}$. This is the most frequent case in our data. In the infrequent cases where $q_{ij} < \bar{q}_{ij}$, $\bar{p}_{ij}$ replaces p_{ij} throughout.[21]

The endogeneity of q_{ij} in the far-right term in (1) is critical to the impact of the TPS. To be in compliance, the generating unit's ultimate (end-of period) allocation of allowances βq_{ij}, plus (minus) any allowances it purchases (sells) on the trading market, must be at least enough to justify it emissions during the period. The far-right term in (1) represents the additional needed purchases (or potential sales) of allowances consistent with compliance.

---

[19] For brevity, we will let "generator" refer to both the physical unit and the unit's decision-maker (manager). The intended reference will be clear from the context.

[20] See Ho *et al.* (2017). This assumption seems reasonable for the approximately 50 percent of the generators that are privately owned.

[21] Thus, when $q_{ij} < \bar{q}_{ij}$, the equation for net revenue reduces to $\pi_{ij} = \bar{p}_{ij} q_{ij} - C\left(q_{ij}, e_{ij}\right) - t\left(e_{ij} - \beta_j q_{ij}\right)$. This squares with the fact that in this case $\bar{p}$, not the endogenous price, is the price that applies to each unit of electricity sold.

Let $u_{ij}(\equiv e_{ij} / q_{ij})$ represent the generator's end-of-period emissions-output ratio.[22] Then we can rewrite the far-right term as $t(u_{ij} - \beta_j)q_{ij}$. In the absence of purchases of additional allowances, a unit that produces output q will be in or out of compliance depending on whether its emissions-output ratio is less than or greater than β_j.

Let u_{ij0} represent the generator's beginning-of-period emissions-output ratio. A generator with $u_{ij0} > \beta_j$ can come into compliance by purchasing additional allowances, reducing its emissions rate, or both. A generator with $u_{ij0} < \beta_j$ will not need to purchase allowances[23] and will benefit from the sale of its excess allowances. Indeed, once a generator with an initial emissions ratio less than β_j has achieved its optimal emissions ratio, its best option is to sell its excess allowances, since such allowances have no other beneficial use for the facility; selling them involves no opportunity cost.[24]

This suggests some of the potential distributional implications of the TPS. Generators in the $u < \beta$ category can benefit from the TPS by selling their excess allowances, while generators in the $u > \beta$ category face compliance costs, as they will need to reduce emissions intensity and/or purchase additional allowances to come into compliance.[25] Below we explore further the distributional impacts and consider the cost-effectiveness dimension.

b. The Shutdown Decision

[22] By "end-of-period" emissions-output ratio we mean the ratio of cumulative emissions to cumulative output over the compliance period. This is the ratio relevant to ascertaining compliance.

[23] This assumes the generator does not increase its emissions-output ratio during the compliance period enough to cause its ratio to exceed β. There is no reason to expect this to occur, since the TPS gives all generators incentives to reduce their emissions-output ratios, as discussed below.

[24] The National Development and Reform Commission 2017 document, *Guidelines of National Carbon Emissions Trading System (Power Generation Sector),* did not include provisions for intertemporal banking or borrowing of emissions allowances. Correspondingly, the model assumes no such provisions. As a result, the allowances available to generators needing additional allowances are restricted to the excess allowances offered by the generators with $u_{ij}<\beta_j$. In China's pilot trading programs, intertemporal borrowing was not permitted, although intertemporal banking was an option.

[25] Although China's TPS does not cover renewable sources of electricity such as wind and solar, it will encourage production from these sources by increasing the cost of supplying fossil-based generated electricity. A further boost to renewables production would occur if the TPS were to cover these sources, since presumably these sources would have emissions-output ratios well below the benchmarks and thus could benefit significantly by selling excess allowances.

In considering whether to shut down, the generator will compare the revenue from continued operation with the revenue associated with shutting down. In the case of shutting down, the revenue consists solely of the liquidation value[26] of the abandoned capital. Note that the generator's owners cannot earn additional revenue by selling any of the allowances it was allocated at the beginning of the compliance period; the program requires that such allowances be returned to the government.

It is useful to rewrite (1) as:

$$\pi_{ij} = p_{ij}q_{ij} + \left(\overline{p}_{ij} - p_{ij}\right)\overline{q}_{ij} - C\left(q_{ij}, e_{ij}\right) - t\left(e_{ij} - \beta_j q_{ij}\right) \tag{2}$$

This expression divides the gross revenue from electricity production into $p_{ij}q_{ij}$, a component that depends on q_{ij}, the level of production, and $\left(\overline{p}_{ij} - p_{ij}\right)\overline{q}_{ij}$, a fixed component.[27] The fixed component is the revenue associated with output up to the maximal level to which the administered first tier price applies. This revenue is inframarginal. It affects the level of profit and the decision whether to shut down, but because it is inframarginal it does not affect the optimal level of production for firms that do not shut down. Recall that the equations in this section assume $q_{ij} > \overline{q}_{ij}$. When $q_{ij} < \overline{q}_{ij}$, the corresponding profit equation is $\pi_{ij} = \overline{p}_{ij}q_{ij} - C(q_{ij}, e_{ij}) - t(e_{ij} - \beta_j q_{ij})$ and p_{ij} is the price at the margin.

From (2), a generator will remain in operation if and only if

$$pq + \left(\overline{p} - p\right)\overline{q} - C\left(q, e\right) - t\left(e - \beta q\right) > L \tag{3}$$

where L represents the liquidation value (subscripts have been suppressed for convenience).

We can rewrite (3) as

$$pq + \left(\overline{p} - p\right)\overline{q} - C\left(q, e\right) - L > te - t\beta q \tag{4}$$

Define $\hat{t}$ as the allowance price t that equates the left-hand and right-hand sides of (4):

[26] In discussions with the ETS planers, we have learned that the market for abandoned electricity generation capital is quite limited, so that the liquidation value is very low. Also, it should be noted that in this one-period model, the relevant "liquidation value" is the avoided one-period rental on the capital that is no longer employed.

[27] Note that p_{ij} as well as $\overline{p}_{ij}$ and $\overline{q}_{ij}$ are exogenous to the individual generator.

$$\hat{t} = \frac{pq + (\bar{p} - p)\bar{q} - C(q,e) - L}{e - \beta q} \tag{5}$$

$\hat{t}$ is a critical value of t: the generator will shut down or remain in operation depending on whether the allowance price is above or below this value. Other things equal, $\hat{t}$ will be lower for generators facing a lower (more stringent) β: they will shut down first.[28]

c. Equilibrium Conditions

1. The Allowance Price

Let RP_j refer to the set of generators in technology class j that *remain* in operation and *purchase* allowances – the generators in technology class j with $u_{ij} > \beta_j$ (or equivalently, $e_{ij} > \beta_j q_{ij}$) for which condition (3) above is satisfied. Then the total market demand for allowances, $D(t)$, is expressed by:

$$D(t) = \sum_j \sum_{i \in RP_j} (u_{ij} - \beta_j) q_{ij} \tag{6}$$

Demand is a function of the allowance price t because this price influences the number of generators that remain in operation (the number for which t is below $\hat{t}$). The allowance price also affects demand through its influence on the output levels and emissions intensities of the generators that remain in operation.

The supply of allowances on the trading market comes from generators that remain in operation and have excess allowances to sell. Let RS_j represent the set of generators in technology group j that *remain* in operation and *sell* allowances – the generators in technology group j for which $u_{ij} < \beta_j$.[29] The total supply of allowances into the emissions trading market is:

---

[28] Under cap and trade, the expression for profit is $\pi = pq - C - t(e - a_0)$, where a_0 represents the facility's allocation of (free) allowances. From this it follows that under cap and trade, $\hat{t}$ is equal to $(pq - C - L)/(e - a_0)$. A larger initial allocation of free allowances raises $\hat{t}$.

[29] Recall that u_{ij} is endogenous. We assume that generating units in the group RS_j undertake expenditure on process change to the extent that this will increase net revenue (by increasing the number of excess allowances).

$$S(t) = \sum_{j} \sum_{i \in RS_j} (\beta_j - u_{ij}) q_{ij} \tag{7}$$

The allowance price affects allowance supply by influencing the electricity production levels of the generators with $u < \beta$: this affects the number of excess allowances they have to sell. This price also affects supply by influencing the emissions intensities of these generators.

The market equilibrium price of allowances is the price t that satisfies $D(t) = S(t)$.

2. Electricity Prices

Generators whose production does not exceed $\overline{q}$ face only the administered electricity price $\overline{p}$, while generators that produce more than $\overline{q}$ face both the administered price and the market price p for production beyond $\overline{q}$. In each province, the total demand for electricity is assumed to be a negative function of total supply. The equilibrium market price equates total supply with the total demand.

d. Cost-Effectiveness Considerations

1. TPS and C&T electricity outputs relative to the cost-minimizing output level

Consider the profit-maximizing choices made by an individual generating unit under the TPS. As indicated in expression (2) above, the profit function for a generating unit is $\pi = pq + (\overline{p} - p)\overline{q} - C(q,e) - t(e - \beta q)$, where subscripts are suppressed for simplicity. This function yields the following first-order conditions for the profit-maximizing levels of q and e, given the allowance price t and applicable benchmark β :

$$\partial \pi / \partial q : p - C_q = -\beta t \tag{8}$$

$$\partial \pi / \partial e : -C_e = t \tag{9}$$

where $C_q \equiv \partial C / \partial q$ and $C_e \equiv \partial C / \partial e$. The left-hand side of (8) is the marginal net revenue from output, excluding any change in costs of needed allowances. The right-hand side is the marginal cost of output in terms of the additional allowance costs associated with that increment to output since each unit of output raises allowance payments by βt (holding fixed the emissions-output ratio). Expression (8) states that a generator maximizes profit by equating the marginal net revenue with the marginal allowance cost.

To assess the cost-effectiveness of the TPS, we compare these first-order conditions with those from the following optimization problem:

$$max \prod = \sum_i pq_i + \left(\overline{p} - p\right)\overline{q}_i - C\left(q_i, e_i\right) \tag{10}$$

$$s.t. \ \sum_i e_i \leq \overline{E}$$

where Π represents the net surplus produced by the generators in the aggregate[30] and $\overline{E}$ is a given aggregate emissions target. The solution to (10) is the maximal surplus that can be obtained when emissions are kept within the given target or, equivalently, the minimum cost of reducing emission to the amount indicated by the target. The Lagrangean expression associated with (10) is

$$\mathcal{L}\text{:} \sum_i pq_i + \left(\overline{p} - p\right)\overline{q}_i - C\left(q_i, e_i\right) - \lambda\left(\sum_i e_i - \overline{E}\right) \tag{11}$$

The first-order conditions associated with this expression are

$$\partial\mathcal{L}/\partial q_i : \ p - C_{q_i} = 0 \tag{12}$$

$$\partial\mathcal{L}/\partial e_i : -C_{e_i} = \lambda \tag{13}$$

$$\partial\mathcal{L}/\partial\lambda : \ \sum_i e_i = \overline{E} \tag{14}$$

Equation (12) indicates that social costs are minimized when generators' production levels equate the marginal revenue (p) and the marginal private cost C_{q_i} of production. This condition differs from expression (8), the condition determining generators' choices of q under the TPS. The difference reflects the implicit subsidy to output under the TPS. From equation (2), other things equal[31] each unit of q under the TPS reduces by $t\beta$ the cost of additional allowances

[30] This implicitly assumes no externalities or taxes, and pure competition. Under these conditions, social surplus (the sum of producer and consumer surplus) is maximized when the sum of net revenues to firms is maximized.
[31] In keeping with the fact that (8) is a partial derivative, this condition is calculated holding *e* constant. In fact, the TPS affects both *q* and *e*. The connections between *q* and *e* are important for explaining the impacts of the TPS on

needed for compliance. Thus, condition (8) means that the TPS leads generators to produce more output, for given output prices *p,* than would be the case if equation (12) applied.[32]

Equation (13) is the first-order condition associated with the choice of emissions levels consistent with minimizing the cost of achieving a given emissions-reduction target. The Lagrangean multiplier λ is the shadow value of the constraint on emissions; in an emissions trading market, this is the market price of allowances. Thus, we can interpret λ as equal to *t*. This means that the first-order condition (13) for cost-minimization matches equation (9), the first-order condition regarding emissions under the TPS. Both equations express the condition that the marginal benefit from emissions (or the negative of the marginal cost) should be equated to *t*. Note that the similarity of conditions (9) and (13) does not mean that the level of emissions under the TPS will match the first-best level. This is because C_{e_i} depends on the level of output, and output under the TPS differs from first-best output. For a given value of t, the level of emissions under the TPS will exceed (fall short of) the first-best level if $\partial C_{e_i} / \partial q_i$ is negative (positive).

Consider now the impacts under cap and trade. The expression for profit under C&T is:

$$\pi_{ij}^{CT} = pq + \left(\overline{p} - p\right)\overline{q} - C\left(q, e\right) - t\left(e - a_0\right) \tag{15}$$

a_0 is the initial allocation of (free) allowances and the superscript "C&T" designates the case of C&T. It is straightforward to show that the associated first-order conditions for a generator's optimal choice of *q* and *e* match expressions (12) and (13) for the planner's cost-minimization problem above. This implies that the output and emissions levels under C&T are such as to minimize the cost of achieving the specified aggregate emissions limit.[33] The cost-effectiveness advantage of C&T over the TPS reflects the absence of the output subsidy: the level of output does not appear in the far-right term in the C&T profit expression.

The difference in the impacts of the TPS and C&T become smaller, the lower is the price elasticity of output supply. One way to see this is to compare the TPS first-order condition for optimal output, given by equation (8) with the corresponding condition for C&T (which, as noted

levels of electricity supply and emissions relative to the baseline (no policy) case. We address these connections below.

[32] Generators with $u_0 > \beta$ will reduce output relative to the business-as-usual level, but the reduction will fall short of the optimal amount.

[33] Of course, this assumes the absence of transactions costs and other possible impediments to trading. Such limitations might well exist, but they could apply under the TPS as well.

above, is the same as (12)). The former can be rewritten as $C_q = p - \beta t$, while the latter can be rewritten as $C_q = p$. The difference between these two conditions is βt, which does not depend on q. Note that C_q is inversely related to the supply elasticity, implying that as C_q approaches infinity the supply elasticity approaches 0. Suppose that q satisfies the TPS first-order condition. Since βt is a constant, as C_q approaches infinity (or as the supply elasticity approaches zero) the change in q needed to satisfy the C&T first-order condition becomes infinitely small. In the limiting case of a zero supply elasticity, optimal q is in the same for the TPS and C&T, and since the first-order conditions for optimal emissions are also the same, both policies are the same in terms of cost-effectiveness. A comparison of equations (2) (for the TPS) and (15) (for C&T) indicates that with a zero supply elasticity the two policies will also have identical distributional conseqences so long as the initial allowance allocations $\beta_j q_{ij}$ (for the TPS) and a_0 (for C&T) are the same.

2. TPS and C&T electricity outputs relative to business-as-usual levels

Here we consider how outputs of the TPS and C&T differ from their baseline (or business-as-usual) values. We will see that while C&T induces all generators to reduce production relative to the baseline level, keeping price constant, the TPS can cause some generators to increase output relative to the baseline. We start with a focus on the TPS. To determine the relationship with baseline output, we examine the *total* derivative[34] of the TPS profit expression (2):

$$d\pi = pdq - \frac{\partial C}{\partial e}de - \frac{\partial C}{\partial q}dq - tde + t\beta dq \qquad (16)$$

Dividing the above expression by dq yields:

$$\frac{d\pi}{dq} = p - \frac{\partial C}{\partial e}\frac{de}{dq} - \frac{\partial C}{\partial q} - t\frac{de}{dq} + t\beta \qquad (17)$$

Setting $d\pi / dq$ equal to 0 and rearranging give:

[34] In contrast with the partial derivative condition shown in expression (8), the total derivative considers at one time the impact of changes n both q and e on profit.

$$p = \frac{\partial C}{\partial q} + \frac{\partial C}{\partial e}\frac{de}{dq} + t\frac{de}{dq} - t\beta \tag{18}$$

The left-hand side is marginal revenue from output, while the right-hand side is the marginal cost. The first two right-hand side terms are the direct cost of an increase in output and the indirect cost via the output's impact on emissions. The third and fourth right-hand-side terms represent the change in compliance costs associated with change in emissions, net of the implicit subsidy $t\beta$. Under business as usual, the allowance price is zero, and the above expression reduces to:

$$p_{BAU} = \frac{\partial C}{\partial q} + \frac{\partial C}{\partial e}\frac{de}{dq} \tag{19}$$

where p_{BAU} is the business-as-usual electricity price. Assume for the moment (and counter to fact) that the TPS does not affect electricity prices, so that $p = p_{BAU}$. Define $A(q)$ as $\frac{\partial C}{\partial q} + \frac{\partial C}{\partial e}\frac{de}{dq}$.

Under business-as-usual, q is chosen so that $A(q)$ satisfies equation (19) above. Under the TPS, $A(q)$ must change to offset the presence of the extra elements $t\frac{de}{dq} - t\beta$ in (18). The extra elements capture the impact of q on profits by way of q's effect on compliance costs. The extra elements can be rewritten as $t(de / dq - \beta)$. Under the TPS, this term is either positive or negative depending on whether de / dq is greater or lower than β. Since $A(q)$ increases with q, satisfying (18) requires electricity supply to decline or increase (assuming no change in electricity prices) depending on the sign of $\frac{de}{dq} - \beta$. But de / dq is the generator's emissions rate at the margin. So a generator's electricity supply under the TPS is either below or above the baseline level of output, depending on whether the emissions rate at the margin is greater or less than β. Up to now we have referred mainly to average emissions rates (u, or e/q) rather than the marginal rate. The TPS encourages firms to undertake various process changes to reduce the average rate in order to come into compliance. If reducing these efforts to reduce the average emissions rate have little impact on the marginal emissions rate, then a firm with initial average emissions rate above (below) its benchmark will also have a marginal rate above (below) the benchmark. This would imply that a facility will have incentives to expand or reduces it output depending on whether its initial emissions intensity is below or above its benchmark.

Thus far we have assumed no change in electricity price as a result of the TPS. Define Δ_{TPS} as $p - p_{BAU}$. Applying equation (18) and our definitions of $A(q)$ and Δ_{TPS}, we can write:

$$p_{BAU} = A(q) + t\left(de/dq - \beta\right) - \Delta_{TPS} \tag{20}$$

This expression implies that the impact of the TPS on output is modified by the change in electricity prices. If the TPS causes an increase in electricity prices (as it usually does), Δ_{TPS} is positive and satisfying the above expression requires a higher value of q than would be the case if prices did not increase. In our numerical simulations, we find that this price-effect is second-order, and that the principal factor influencing whether output rises or falls relative to baseline values is the relationship between a generator's emissions-intensity and its benchmark.

In contrast, under cap and trade the corresponding equation to (20) is:

$$p_{BAU} = A(q) + t\left(de/dq\right) - \Delta_{C\&T} \tag{21}$$

where $\Delta_{C\&T} \equiv p_{C\&T} - p_{BAU}$ and $p_{C\&T}$ refers to the electricity price under cap and trade. In contrast with equation (20) for the TPS, β does not appear in (21). As a result, the middle right-hand-side term is always positive. This implies that, in contrast with the TPS, C&T causes all generators to reduce output relative to the business-as-usual levels when $\Delta_{C\&T} \leq 0$, that is, when the policy does not cause the electicity price to rise. However, to the extent the C&T leads to a higher electricity price (and it usually does) the impact on the electricity price counters the effect exerted by the allowance price. As was the case under the TPS, in our numerical simulations of C&T this electricity price effect is usually second-order. As indicated below, in many simulations C&T causes all generators to reduce output relative to the baseline levels.

3. *Gains from allowance trading*

With a perfectly fluid market for allowance trading, managers of generating units will reduce emissions to the point where the private marginal costs of abatement equal the private marginal benefits. The two elements can be obtained from the total derivative of profit shown in equation (16) above. Dividing both sides by *de* yields:

$$\frac{d\pi}{de} = p\frac{dq}{de} - \frac{\partial C}{\partial e} - \frac{\partial C}{\partial q}\frac{dq}{de} - t + t\beta\frac{dq}{de} \tag{22}$$

Setting $d\pi / de$ to 0 and rearranging yields:

$$\underbrace{p\frac{dq}{de} - \frac{\partial C}{\partial e} - \frac{\partial C}{\partial q}\frac{dq}{de}}_{MB_e^{pvt}} = \underbrace{t\,(1 - \beta\frac{dq}{de})}_{MC_e^{pvt}} \tag{23}$$

The left-hand side is the marginal private benefit from emissions (or marginal private cost of abatement), while the right-hand side is the marginal private cost of emissions (or marginal private benefit from abatement). Importantly, the right side of (23) will generally vary across generators, since $\beta\,(dq / de)$ is specific to individual generators. Hence, even if trading is perfectly fluid, it will not result in the equalization of marginal benefits and marginal abatement costs. Other things equal, the right-hand side of (23) will be higher for generators facing a lower (more stringent) β; hence after trading they will have higher marginal abatement costs than generators facing a higher β. This limits the cost-effectiveness of trading: the total private cost of achieving the same aggregate emissions reduction would be lower if the lower-β generators undertook less abatement and the higher-β generators undertook more. Thus, even though generators will face a common allowance price, their marginal abatement costs after trades will generally differ, even if trading is perfectly fluid.[35] This limits the achievable cost-reductions from emissions allowance trading. The limits to the cost-reductions are a symptom of the presence of the β in equation (23), which leads to a discrepancy between marginal private and marginal social costs of abatement. Note that even when all facilities face the same β, the benefits from allowance trading will often be be compromised, since *dq/de* will often differ across generators.

The key message from equation (23) is that the TPS's implicit subsidy to output reduces the gains from allowance trading. The compromising of the gains is greater, the larger the variation in the benchmarks, other things equal. Across alternative TPS systems of given overall

[35] This result parallels a result in the simpler analytical model in Goulder and Morgenstern (2018).

stringency, the cost-effectiveness gains from trade are likely to be greatest when a single benchmark is employed instead of multiple benchmarks. In the presence of the output subsidy, variation across generators in the value of dq/de also works to limit the gains from allowance trading. C&T has an advantage over the TPS in terms of the cost-effectiveness gains from allowance trading. This is because under C&T with fluid trading the right-hand side element in the $MB=MC$ expression (23) is simply t, which implies that all units equate their marginal private benefits from emissions (marginal private costs of abatement) to the same value. This leads to the maximal reduction in aggregate costs of meeting a given emissions-reduction target.

e. Distributional and Other Considerations

Although the use of multiple (i.e., differing) benchmarks in a TPS compromises cost-effectiveness, it can serve distributional goals. Higher (less stringent) benchmarks can be applied to generators that otherwise would suffer especially high costs of compliance, or would be forced to shut down. This suggests trade-offs between cost-effectiveness and the achievement of certain distributional goals.[36]

Some attractions of the TPS relative to C&T deserve mention. As our numerical simulations will show, because of the TPS's implicit subsidy to output, the TPS leads to smaller increases in costs to electricity producers than does a comparably stringent C&T system. Correspondingly, power plants bear a larger share of the regulation's cost-burden than is the case under C&T. This can help reduce emissions leakage.[37]

The TPS also has an advantage in terms of adaptability to changes in macroeconomic conditions. In boom times, when electricity demand and production are high, the allowance allocations increase automatically. This prevents what otherwise could be very high abatement costs in a cap-and-trade program with a fixed cap on allowances. Likewise, the TPS's allowance

[36] Note that the use of multiple benchmarks in a C&T system does not compromise cost-effectiveness because C&T does not involve an output subsidy. As indicated by the above analysis, under the TPS the use of multiple benchmarks expands the distortions associated with the TPS's inherent output subsidy.
[37] The TPS's tendency to produce smaller increases in output prices can help mitigate domestic producers' potential losses of international competitiveness. This issue is particularly significant to industries that are especially import-competing and/or carbon-intensive. It is not a major issue for producers in China's power sector, since relatively little domestically produced electricity is sold internationally. The issue will be more important once the TPS expands to major industries in China's manufacturing sector.

allocation is lower in slack times, when electricity demand is likely to be lower and a fixed cap could have yielded excessive allowances.

f. Summary and Challenges

Key findings from this analysis are:

- The TPS induces some covered facilities – in particular, those with emissions-output ratios below their required benchmarks – to increase supply beyond their business-as-usual levels. This contrasts with C&T, which tends to cause all facilities to reduce production relative to their BAU levels.

- A TPS generally is less cost-effective than an equivalently scaled C&T program. The difference in cost-effectiveness reflects the implicit subsidy to output under the TPS, which distorts supply decisions and causes generators' electricity output levels to exceed the levels consistent with minimizing the costs of achieving a given aggregate emissions limit. This distortion gains importance the higher the price elasticity of electricity supply.

- When the benchmarks differ, allowance trading under the TPS does not lead to equality in marginal abatement costs across facilities that continue to operate, even when trading is perfectly fluid. This limits the aggregate cost-reductions from allowance trades. The most cost-effective TPS is one involving a single benchmark. Costs increase with variation in the benchmarks, other things equal.

- Multiple benchmarks under the TPS can help serve distributional goals. This implies a trade-off between cost-effectiveness and distributional equity.

The results from our numerical model reinforce these analytically derived findings. They also provide estimates of the magnitudes of the analytical model's predicted qualitative impacts.

5. A Numerical Model

a. Overview

The model considers the 2,392 generating units of Table 1, dividing them into the 11 technology classes shown. Within each technology class, the model allows for heterogeneity in the cost functions and thus considers a large number of generation units in each class.

The numerical model's basic structure matches that of the previously described analytical model. We calibrate the numerical model so that its solution under baseline (status quo) conditions matches the data in terms of costs, production levels and electricity prices.

We compare the baseline outcomes with the results under TPS and C&T policies. TPS policies are defined by the stringency and distribution of the assumed benchmark emissions-output ratios applying to different generators, while C&T policies are defined by assumed initial allocations of allowances to the different generators. All generators within a given technology class receive the same benchmarks under the TPS and the same initial allowance allocations under C&T.

Under each policy, profit-maximizing managers of generating units determine whether to shut down or remain in operation and, conditional on continuing to operate, the optimal level of production, the extent of effort to reduce emissions intensity of production, and the number of allowances to purchase or sell. Under each policy, the model solves for the equilibrium allowance price and for the equilibrium prices of electricity in each provincial and zonal (regional) market. The equilibrium allowance price equates the aggregate supply of allowances with the aggregate demand. The equilibrium electricity prices pertain to the electricity produced in excess of the quantities facing administered prices. Such excess electricity is sold either to residual local electricity markets or to regional grid companies.[38]

In the data, a given generating unit will often sell its electricity in the local market and zonal market at different prices. Transactions costs help explain the difference in equilibrium prices of electricity, a homogenous product. As indicated in subsection 5b below, we model transaction costs as increasing in the quantity of electricity that a given generator sells to the zonal market. We calibrate the parameters of the transactions cost function so that sales to the zonal market in the baseline simulation match the observed data. In both baseline and policy simulations, the equilibrium market price of electricity in the local market equals the price in the relevant zonal market net of the marginal transactions cost.

b. Costs and Supply

[38] In a few unusual cases, the overall demand for electricity at the administered price is less than the GH level of output. In this case, the equilibrium quantity produced is less than the GH output level and all electricity is sold at the administered price.

For each generator, we employ the following specification for total production cost:

$$C(q,h) = (\phi_0 + \phi_1 q^{\phi_2}) + h(z, p_z, t) \tag{24}$$

where q is the supply of output, , ϕ_0, ϕ_1, and ϕ_2 are parameters, and the function $h(z, p_z, t)$ captures the direct[39] cost of compliance. (Subscripts have been suppressed for simplicity.) In the compliance cost function, z is an index of real resources devoted to reducing emissions intensity, p_z is the price of a unit of z according to that index, and t is the allowance price. Specifically, under the TPS,

$$h(z,t) = p_z z + t[u(z/q) - \beta]q \tag{25}$$

and under C & T,

$$h(z,t) = p_z z + t[u(z/q)q - a_0] \tag{26}$$

As in the previous section's analysis, equations (25) and (26), respectively, indicate that the expenditure on the additional allowances needed for compliance depends on the gap between uq and βq (under the TPS) and the gap between uq and a_0 (under C&T). In (25) and (26), the emissions-output ratio u depends on z. In determining how much to spend on emissons abatement, cost-minimizing producers consider the benefits (reduced u) and costs (p_z) of z.

The following function connects z with the emissions-output ratio:

$$u(z) = u_0 / [1 + (z/q)^{\varepsilon}] \tag{27}$$

We employ values above 1 for ε, which implies a diminishing marginal effect of z on the emissions-output ratio and associated increasing marginal abatement costs. Equation (27) has the property that $u(z)$ is u_0 when z is 0. Under business-as-usual, t is 0 and producers will choose z = 0. In this case the $h(z, p_z, t)$ function is zero and the production cost function reduces to $\phi_0 + \phi_1 q^{\phi_2}$.

As was noted, transactions costs explain the observed differences in the market prices of electricity between the local (provincial) market and the zonal market to which the province contributes. The transaction cost function has the form, $\theta_1 q_{i,zone}{}^{\theta_2}$, where $q_{i,zone}$ is the quantity of

[39] The qualifier "direct" is included since the costs of compliance also include the impact that policy-induced changes in q have on production apart from abatement effort. This other cost is captured by the first term in (24).

electricity sold by generator i to the zonal market and θ_1 and θ_2 are parameters calibrated from market data.

c. Producer Heterogeneity

Our data on production costs consist of *average* total costs for each of the 11 technology classes shown in Table 1. We allow for cost heterogeneity within technology classes by assuming that the parameter ϕ_0 in the cost function differs across the units within a class according to a beta distribution. Since ϕ_0 is a constant in that function, it does not affect generators' output supplies or abatement expenditures at the margin. However, it does affect the level of profits and thus, under any given policy scenario, it influences whether profits for a given unit are positive and whether the unit shuts down. Because the values of ϕ_0 are distributed according to the (continuous) beta distribution, the number of units that shut down is a continuous function of policy parameters and the allowance price.[40]

d. Optimal Output and Emissions Choices

After substituting equation (27) and into (25) and (26) for the TPS and C&T cases, respectively, substituting the results into (24), and recognizing from equation (2) that profit is $pq + (\overline{p} - p)\overline{q}$ minus overall cost, we have

$$\pi = pq + (\overline{p} - p)\overline{q} - \phi_0 - \phi_1 q^{\phi_2} - p_z z - t\Big[[u_0 / (1 + (z/q)^{\varepsilon}) - \beta]\Big] q \tag{28}$$

and

$$\pi^{CT} = pq + (\overline{p} - p)\overline{q} - \phi_0 - \phi_1 q^{\phi_2} - p_z z - t\Big[[u_0 / (1 + (z/q)^{\varepsilon})] q - a_0\Big] \tag{29}$$

As was noted, p is the market price – the price that applies at the margin. Throughout, p should be interpreted as the price net of any applicable transactions costs. Thus, p is the same for electricity sold to the local market and the zonal market.

[40] Using a continuous probability distribution function to incorporate heterogeneity within broad technology classes causes the model's aggregate demand functions for allowances to be continuous. This facilitates solving the model.

Differentiating these functions with respect to the choice variables q and z yields:

$$\partial\pi / \partial q : p = \phi_1\phi_2 q^{\phi_2 - 1} + t[u_0 / (1+(z/q)^{\varepsilon}) - \beta] + t\varepsilon u_0 (z/q)^{\varepsilon}(1+(z/q)^{\varepsilon})^{-2} \tag{30}$$

$$\partial\pi / \partial z : \ p_z = \ t\,\varepsilon u_0 (z/q)^{\varepsilon-1}(1+(z/q)^{\varepsilon})^{-2} \tag{31}$$

$$\partial\pi^{CT} / \partial q : p = \phi_1\phi_2 q^{\phi_2 - 1} + t[u_0 / (1+(z/q)^{\varepsilon})] + t\varepsilon u_0 (z/q)^{\varepsilon}(1+(z/q)^{\varepsilon})^{-2} \tag{32}$$

$$\partial\pi^{CT} / \partial z : p_z = t\varepsilon u_0 (z/q)^{\varepsilon-1}(1+(z/q)^{\varepsilon})^{-2} \tag{33}$$

Expressions (30) and (32) equate the marginal benefit from q (left side) with its marginal compliance cost. Expressions (31) and (33) equate the cost of z (left side) with the implied marginal benefit in terms of compliance cost reductions. Under each of the two policies, we solve simultaneously the two relevant first-order conditions to obtain the optimal values for q and z.[41] The optimal responses to changes in prices or other relevant parameters of both output supply and abatement effort are consistent with the analytical model. Under the TPS, optimal q may increase or decrease with β and t. Under C&T, the optimal q declines with t and is not affected by a_0 for interior (non-shutdown) solutions. Optimal z declines with p_z under both policies.

e. Equilibrium Conditions

A given TPS policy is defined by a set of benchmarks applying to each generating unit, while a given C&T policy is defined by a set of initial (exogenous) allowance allocations. The solution approach under a given policy is as follows: Let V represent a vector consisting of an allowance price and a set of province-level and zonal electricity prices. For any given V, the model calculates each generator's revenue-maximizing quantity of output and the optimal emissions intensity of each generating unit, conditional on remaining in operation. For some units – particularly those with emissions-output ratios above the applicable benchmark – production costs can be sufficiently high to imply negative profits. These are the units with exceptionally high

[41] The numerical model obtains the solution by the following iterative procedure. It first posits a value of q and uses equation (31) (or (33)) to solve for the optimal z conditional on the posited value. It then uses equation (30) (or (32)) to obtain a value for q that is optimal conditional on the derived value of z . If the derived q and original posited q do not match, the model posits another value for q and repeats the procedure. This iterative procedure continues until the posited and derived values of q match, at which point both of the applicable first-order conditions are satisfied.

values of ϕ_0 within the distribution of this parameter for the technology class in question. These units will shut down.

The production decisions of individual generators determine the aggregate demand and supply of allowances and they affect the supply and demand for electricity in both the residual local market and the six zonal markets. The model continually alters both the allowance price and the electricity prices in V until three sets of equilibrium conditions are satisfied: (1) the aggregate allowance supply equals the aggregate allowance demand; (2) for each province, the supply of electricity to the residual local market equals the demand in that market; and (3) for each zonal market, the sum of provinces' supplies to that market equals the electricity demand in that market. The equilibrium allowance and electricity prices are closely connected, since electricity prices affect allowance supply and demand through their impact on electricity production, and the allowance price affects electricity supplies through its impact on compliance costs.[42]

f. Data and Calibration

The data for the model are for the year 2016. Table 1 above organized key data by technology class. Table 3 displays baseline administered and market prices and outputs by province. Data were collected from National Development and Reform Commision, China Electric Council, and the Electric Power Development Research Institute. Details on the sources and organization of the information on electricity prices are provided in Appendix A.

Overall, 42.8% of coal-fired electricity and 6.3% of gas-fired electricity is sold in local market. The average administered price of guaranteed-hour electricity is about 0.3668 RMB or 7.9 percent higher than market price of electricity sold locally. Because of limited data availability, we assume that, for a given technology class, the average total production costs and emission intensity are the same across provinces.

For each technology class in each province, we identify the parameters ϕ_1, and ϕ_2 of the cost function, along with ϕ_{0mean}, the mean value of the cost function's constant term ϕ_0, through a calibration procedure that imposes three requirements on the average generator in each technology

[42] The solution method obtains equilibrium electricity prices for 29 province-level residual electricity markets, equilibrium electricity prices for 6 zonal markets, and one equilibrium price for the national allowance market. We solve for the 36 equilibrium prices by minimizing the differences between supply and demand in each market through gradient descent.

class. The requirements are that, in the business-as-usual simulation: (1) net revenue equals the net revenue from the data, (2) the net-revenue-maximizing level of output (that is, the level at which marginal production cost equals the electricity price) matches q_0, and (3) the implied price elasticity of supply η equals 0.22, our central case value for this elasticity. Details on the calibration method are provided in Appendix B.

A further step is to specify the distribution of the constant term ϕ_0 in the cost function of each technology class. As mentioned, we employ a beta distribution, which involves finite bounds for the parameter, and we assume the distribution is symmetric. We impose two conditions to identify the parameters of this distribution. First, the distribution of ϕ_0 must imply that under baseline conditions the mean total cost C for the technology class equals the average total cost from the data. Second, the largest value for ϕ_0 in the distribution, $\phi_{0\max}$, must imply a value for C in the baseline that makes profit just equal to zero for the generator with that value. This means that the highest cost generator in each technology class is marginal in the sense that it makes exactly zero economic profit. It would be the first to shut down when costs rise for its generator class. Details of the procedure for establishing the distribution of ϕ_0 are in Appendix B.

For the transaction cost function, we set the θ_2 (curvature) parameter equal to 2. Then, for each generator, we identify θ_1 based on the requirement that in the baseline simulation, the amount of electricity sold to the zonal market equals the amount from the data, given the baseline electricity prices in the local market and the applicable zonal market.

6. Numerical Results

We consider a range of TPS and C&T policies. Our central case TPS policy involves three benchmarks: β_{GF} = .374, β_{CF} = .864, and β_{CFB} =1.006 where the subscripts refer to the three technology categories indicated in Table 2 – gas-fired generators, coal-fired (other than circulating fluidized bed) generators, and circulating fluidized bed generators. We selected the benchmarks by first calculating the emissions-weighted average emissions-output ratio under the baseline for the entire population of generators. We then set the three benchmarks. In each category the benchmark is a given percentge below the emissions-weighted emissions-output ratio for all generators in that category. We choose the percentage reduction so that the resulting baseline-emissions-weighted average benchmark for the entire population of generators corresponds to the

60th percentile emissions rate among all of the generators.[43] In discussions with individuals involved in the planning of the TPS, the 60th percentile ratio was often mentioned as a possible basis for determining the overall benchmark stringency.

We also consider alternative benchmark specifications that differ in terms of the number of benchmarks, their variation, and their stringency. We offer the specifics below.

For comparability with the TPS policies, we distribute the initial C&T allowances in a way that matches the initial distribution under the TPS and leads to the same aggregate emissions (total number of allowances in circulation) as under the TPS.[44]

a. Central Case Results

-- Prices, Costs, Emissions, and Outputs

Table 4 displays the results in our central case. With the central-case benchmarks, the TPS prompts a reduction in emissions of 84.65 million tons, or 3.1 percent. An allowance price of 226 RMB (or about 32 U.S. dollars) brings the supply of excess allowances by the $u < \beta$ generators into balance with the demand for allowances by the $u > \beta$ generators. In the allowance market, the $u > \beta$ generators purchase 56.91 million tons of allowances from the $u < \beta$ generators.

The shutdown of some units accounts for an emisions-reduction of about 21 million tons, or about 25 percent of the overall reduction. The generators that remain in operation contribute to emissions-reductions through lowered emissions intensities and (for the $u > \beta$ units) through reduced electricity production. Even the units that increase electricity output contribute to the overall emission reductrions by virtue of their reduced emissions intensities.

The TPS causes aggregate electricity supply to decline by about 0.6 percent although, as expected, the $u < \beta$ units increase their output. The aggregate reduction in supply reflects the fact that the TPS raises production costs at the margin: each additional unit of electricity produced

[43] Specifically, it is the emissions rate that corresponds to the generator at the 60th percentile in the distribution of emissions-output ratios across generators ordered by emissions rates, starting with the highest-rate generator.

[44] The TPS and C&T policies lead to different adjustments in output, including different choices as to whether to shut down, in response to the policy implementation. As a result, the end-of-period distribution of allowances across units differs, although by construction the total number of allowances held at the end of the complicance period (which determines total emissions) is the same.

either requires the purchase of additional allowances or reductions in the number of allowances the unit can sell. Thus, the reduction in supply by u > b units – those that shut down and those that remain in operation – exceeds the increase by u < b units. The reduction in aggregate supply gives rise to an increase of 0.5 percent in the quantity-weighted-average price of electricity. This increase reflects the higher market-clearing prices of electricity sold in the local residual and zonal markets. Administered electricity prices are constant.

The private cost of this central case TPS policy, measured as the negative of the change in producer and consumer surplus, is about 8.4 trillion RMB, or 99 RMB per ton. Seventy-two percent of this cost is borne by consumers, a reflection of the policy-induced increase in electricity prices.

Although assessing the climate-related environmental benefits from emissions reductions involves great uncertainties, it is worth considering how the climate-related benefits from the TPS might compare with these estimated costs. The Interagency Working Group arrived at a central value of about $44 (2016$) (or 290 RMB) per ton for the social cost of carbon. Applying this value to the estimated 84.65 million ton reduction in CO_2 emissions yields a climate-related benefit of 25 trillion RMB, approximately 2.9 times the estimated costs.

Comparison with Cap and Trade

Although both the TPS and C&T are examples of emissions trading policies, their impacts differ in important ways. As was noted, for comparability under the C&T policy (free) allowances are allocated to all of the generators in proportion to their initial allocations under the TPS. The allocations are scaled so that the total end-of-period number of allowances allocated matches the total end-of-period allocations under the TPS. This assures that the aggregate emissions reduction is the same (84.65 million tons) under both policies. At the same time, end-of-period allocations differ from initial allocations because of changes in output and shutdowns.

Table 4 includes results under C&T. Although the C&T allocations parallel those under the TPS, the responses by generators under C&T are quite different. In contrast with the results under the TPS, no units increase electricity supply; none have an incentive to increase supply because the allowance allocation is exogenous and there is no implicit subsidy to increased electricity output. As a result, under C&T emissions changes from generators that remain in operation and reduce output account for 96 percent of the emissions reductions, while they account for only 29 percent under the TPS.

The two pie charts in Figure 1 further illustrate the significant differences between the TPS and C&T in terms of their reliance on the different channels for emissions reductions.[45] The charts decompose the overall reductions into those due to changes in electricity production, changes in the relative outputs among generating units, and changes in emissions intensity. Holding industry composition and emissions intensity fixed, changed electricity output contributes about 22 percent of the emissions reductions, as compared with about 60 percent under C&T. Because the TPS does not exploit the output channel as efficiently as C&T does, to achieve comparable emissions reductions this policy must rely more on reduced emissions intensities. Such reductions account for about 75 percent of the reductions under the TPS, as compared with about 37 percent under C&T. Under both policies, the changes in industry composition contribute to a relatively small fraction of the overall emissions reductions.

The greater reduction in electricity output under C&T yields larger increases in electricity prices than under the TPS: the province-weighted-average price rises to .379 RMB/kWh, as opposed to .374 under the TPS. The higher electricity prices under C&T moderate the profit losses. They also account for the lower rate of shutdowns under C&T. Shutdowns account for about 3.747 billion kWh reduction in electricity supply under C&T, as compared with 20.897 billion kWh under the TPS. Table 5 shows that under both policies, the units that shut down are in technology classes C5, C6, and C7 (within the coal-fired category) and classes C9 and C11 (within the circulating fluidized bed and natural-gas categories, respectively). These are the units with original emissions intensities above the benchmarks for their categories. As we described in subsection 5f above, there is heterogeneity within each technology class in the costs of production, and it is only the highest-cost units within each class that shut down.

The equilibrium allowance price under C&T is 41 percent lower than under the TPS, a reflection of the signfiicantly lower electricity output and allowance demand associated with any given allowance price.

---

[45] The decomposition in the pie charts was accomplished as follows. The contribution from reduced electricity output is the emissions reduction that would occur from the differences between output in the policy case and the baseline, if emissions intensities and sector composition remained the same as in the baseline. The contribution from lowered emissions intensities is the reduction that would occur if the emissions intensities changed but industry production levels remained at baseline levels. The contribution from changed sector composition is the reduction that would occur if the only change from the policy were in the shares of production from the different technology classes.

These differences between the TPS and C&T prices and outputs imply differences in overall costs as well as in the distribution of those costs between producers and consumers. The overall private cost (measured as the sum of the losses in producer and consumer surplus) is 47 percent lower under C&T than under the TPS. As indicated in the analytical model, this reflects the absence of the implicit output subsidy under C&T and the associated more efficient exploitation under C&T of reductions in electricity output as a mechanism for reducing emissions.

The distribution of the costs between producers and consumers is quite different as well. Because electricity prices rise more under C&T, consumers experience a much larger share of the burden under this policy. Indeed, they bear over 100 percent of the burden, as the change in producer surplus is positive. This result is in keeping with earlier studies that show how 100 percent free allocation of emissions allowances can create large rents or windfalls for producers.[46] The basic mechanism is that the limited supply of allowances compels producers to reduce output as one channel for achieving compliance; this boosts electricity prices and creates economic rents for competitive producers in the same way that a cartel's restriction in output would do so.

-- *Regional Impacts*

The numerical model incorporates data on the geographical locations and electricity production levels of each technology class under business as usual. Using this information, the model calculates how the policy costs experienced by each technology class are distibuted across provinces and regions. While the available data include differences in costs across technology classes, we do not have information on how, within a given technology class, the costs might differ across regions. As a result, within a given technology class the model-generated differences in profit impacts are entirely due to regional differences in impacts on electricity prices rather than regional differences in production costs.

---

[46] See Bovenberg and Goulder (2001), Parry (2003), Burtaw et al., (2007), and Goulder, Hafstead, and Dworsky (2010). In the present study, 100 percent of the allowances under C&T are given out free. Previous studies indicate that freely allocating a significantly smaller share of the allowances would be sufficient to prevent a loss of profit. See, for example, Goulder, Hafstead and Dworsky (2010). Note also that if the government were to auction off rather than freely allocate the allowances, what otherwise would be rents to producers take the form of revenues to the government. The recognition that 100 percent free allocation is not needed to preserve profits partly explains the increased reliance on auctioning under the European Union's Emissions Trading System and California's cap-and-trade program over the past decade.

Table 6 indicates how the costs to producers under the TPS are distributed across provinces and regions of the country. One key result is that four of the seven regional provincial categories in the table experience overall *increases* in producer surplus from the TPS. This reflects the rents stemming from the free allocation of allowances under the TPS. It is the North, Southwest and Northeast provinces that experience overall losses of producer surplus, with the largest losses in percentage terms applying to Shandong Province in the North and Heilongjian Province in the Northeast. These provinces are especially reliant on coal-fired generation, and our results indicate that (under the benchmarks we have chosen) coal-fired generators experience the largest cost increases under the TPS.

To some, these results might come as a surprise – some might expect the TPS to impose more widespread losses of profit. We do find that the TPS reduces profits of some generating units – indeed, it causes some units to shut down – but the scope of the profit-losses is much smaller than might have been expected. These results attest to the importance of free allowances to the distribution of impacts between producers and consumers.

b. Impacts under Alternative Benchmark Scenarios

Here we explore the sensitivity of policy impacts to alternative benchmark specifications. We first consider how the the "spread" of benchmarks – the range between the high and low benchmarks – affects the results. Figure 2 displays the overall costs under different specifications for the spreads. The one-benchmark case, where the same benchmark applies to all 11 technology classes, is the limiting case of zero spread. The single benchmark is scaled so that the number of allowances initially allocated matches the initial allowance total from the 3-benchmark case.[47] In the other benchmark cases, the same three benchmark categories apply as in the central case, but the benchmark values are different. To obtain these values, we expand or shrink the spread across the three benchmarks while preserving the total number of initially allocated allowances.[48] This

[47] This benchmark is also the emissions-weighted average of the three central-case benchmarks.

[48] More specifically, for each of the three central-case benchmarks, we calculate the difference between the central-case benchmark and the benchmark in the 1-benchmark case. Let d_j denote the difference for technology category *j*. The new category-*j* benchmark is the value in the uniform-benchmark case plus the product of d_j and a scaling factor. We employ scaling factors of .5 and 1.5 to produce benchmarks with less and more spread. Note that applying a scaling factor of 0 recreates the 1-benchmark case, and applying a scaling factor of 1 reproduces the central-case benchmarks.

preserves overall stringency because it does not alter the total number of initially allocated allowances; however, because of different number of shutdowns, the end-of-period emissions are not equal. Figure 2 shows that greater spread implies higher costs per ton of reduced emissions.

The analytical model indicated that the use of multiple benchmarks under the TPS limits the ability of allowance trading to lower costs. To test this prediction, we performed counterfactual simulations where the TPS did not include provisions for trading. Consistent with the analytical model's findings, the cost-reduction from trades is considerably smaller in the 3-benchmark (central) case than in the equivalently stringent one-benchmark case. Specifically, in the 3-benchmark case, trading reduced overall costs by 75.1 percent, from 33.8 billion RMB to 8.4 billion RMB. In the one-benchmark case, trading reduced costs by 83.7 percent, from 48.1 billion RMB to 7.8 billion RMB.

We also consider how the overall stringency of the benchmarks alters policy costs. Here we scale up or down each of the three central-case benchmarks by a common factor. This alters stringency while maintaining the relative sizes of the benchmarks. In the central case, the sum of the initial allocations is 2.5 percent below the aggregate level of emissions under BAU. In the two alternative stringency scenarios displayed in Figure 3, the overall stringency, as measured by policy-induced emissions reductions, is 60%, and 140% percent of the overall stringency in the central case. Costs increase with the stringency of the TPS, at an increasing rate.

Another important policy consideration is the number of benchmarks. Trade-offs apply here. A larger number of benchmarks can help meet distributional objectives, although cost-effectiveness is sacrificed. As noted, the Heilongjian and Shandong provinces experience the largest percentage losses of producer surplus in the central case. These losses reflect the heavy reliance on coal-fired generation, along with the fact that the emissions-output ratios of the coal-fired generators in these provinces were significantly above the benchmark for that generation category.

In an alternative sensitivity analysis, we introduce a TPS policy involving four benchmarks, with the extra benchmark designed to reduce the cost-burden on these provinces. Here we split the coal fired generation category into two sub-categories, with technology classes 1-5 in one and classes 6 and 7 in the other. In the Heilongjian and Shandong provinces, an especially large share of production is by class 6 and 7 generators. In this alternative benchmark scenario, we increase (i.e., loosen) by a common factor the benchmark that applies to technology classes 6 and 7, and reduce (i.e., tighten) by a common factor the benchmark applicable to technology classes 1-

5. These changes are defined by the following two conditions: (1) the baseline-emissions-weighted average benchmark for the coal-fired generation category is unchanged, and (2) the increase in the benchmark for the class 6 and 7 sub-category is just large enough to limit profit losses to all provinces to no more than 5 percent. As noted, the central case benchmark for the coal-fired generators is .864 tCO_2/mWh. Meeting the two conditions requres changing the benchmarks for technology classes 1-5 and 6-7 to 0.820 and 0.925 respectively.

The right-hand pair of columns in Table 6 displays the results in this alternative "subcategorization" case. In this case, the percentage reduction in profit in Heilongjiang is five percent (the maximum allowed under this scenario), as compared with 9.3 percent in the central case. The percentage reduction in profits in Shandong Province is also reduced considerably. Several provinces that would experience profit increases in the central case have lower profits under this alternative TPS policy, a consequence of the tighter benchmarks applied to technology classes 1-5. As indicated in the final row of the table, the overall loss of profit is larger by about about 3,013 million RMB, or 0.6 percentage points, under the alternative benchmarking. Also, the overall economic cost (not shown in the table) is 8,804 million RMB, as compared with 8,387 million RMB in the central case (Table 4) .

c. Further Sensitivity Analysis

Table 7 indicates how alternative parameter values for generators' supply or demand elasticities affect the results. Consider first the impact of alternative values for the supply elasticity. The analytical results from Section 4 indicate that the cost-effectiveness disadvantage of the TPS relative to C&T depends on the extent to which producers respond to the TPS's implicit subsidy to output. This disadvantage is muted, the lower the value of the supply elasticity. Table 7 shows that in the limiting case of zero for this elasticity, the results under the TPS match those of C&T, while the differences across policies in the high-elasticity case are greater than in the central or zero-elasticity cases. With a lower supply elasticity, the TPS's implicit subsidy to output has less force and does less to counteract the tendency of the regulation-induced higher production costs to cause a reduction in output. As a result, a lower supply elasticity causes the TPS to function more and more like cap and trade, occasioning larger reductions in electricity output, prompting larger increases in electricity prices, and shifting more of the policy burden to consumers.

A higher absolute value for the demand elasticity moderates the differences in the price impacts of both the TPS and C&T. Specifically, the difference between the policies in terms of the percentage increase in the average electricity price is 2.47 percentage points when the demand elasticity is -0.15, versus 0.97 percentage points when this elasticity is -0.45. Correspondingly, the larger elasticity narrows the differences in the impacts across the two policies. In particular, the ratio of the overall economic cost of the TPS and to that of C&T is about 2 in the low elasticity case, as compared with 1.8 in the high elasticity case.

Although the alternative parameter values affect the magnitudes of impacts, the fundamental differences between the policies in terms of their relative price impacts, their relative costs, and their distributional impacts are robust across these scenarios.

7. Conclusions

China, the world's largest emitter of CO_2, is poised to introduce what will be the most extensive CO_2 emissions trading system, with coverage more than twice as great as all other CO_2 trading systems in the world combined. The new system has the potential to make a very substantial contribution to the world's efforts to confront global climate change.

Important specifics of China's planned emissions trading system – which will take the form of a tradable performance standard -- are still being worked out. This paper provides unique insights of the environmental, cost-effectiveness, and distributional consequences of alternative designs of the TPS, using matching analytically and numerically solved models. It also compares the TPS's impacts with those of an equivalently scaled cap-and-trade program.

We find that achieving given aggregate CO_2-reduction targets is more costly under the TPS than under C&T, a reflection of the TPS's implicit subsidy to electricity production. The implicit subsidy compromises cost-effectiveness through several channels. First, it implies that the TPS makes less efficient use of electricity output reduction as a way of reducing emissions. While C&T induces all covered power-generation facilities to reduce electricity output, the TPS causes covered facilities with relatively low emissions intensities to increase both electricity output and emissions relative to their levels under business as usual. Second, the implicit subsidy reduces the extent to which emissions allowance trading can reduce costs. This reflects the wedge that the subsidy drives between producers' private marginal costs of emissions abatement (which determine their decisions for purchases and sales of allowances) and society's marginal costs of abatement. No such wedge is introduced under C&T, and thus there is no equivalent limitation to the gains from trading under C&T. Third, the implicit subsidy further compromises cost-effectiveness when multiple benchmarks are employed. Multiple benchmarks add to costs by affecting the relative strength of the subsidy across different covered facilities, distorting the relative contributions of different facilities to emissions abatement. The TPS's costs are about 18 percent higher in our central case's 3-benchmark system than in an equally stringent single-benchmark system.

These impacts combine to produce the higher overall costs of the TPS. In our central-case numerical simulation, the costs of the TPS are 47 percent higher than under C&T. To our knowledge, this study is the first to make identify these three channels and quantify their impact.

In addition to yielding different overall cost impacts, the TPS and C&T produce quite different distributional consequences. Because producers make less use of the output-reduction channel under the TPS, aggregate output is reduced less under the TPS than under C&T and electricity prices rise by a smaller amount. Hence electricity producers shift less of their compliance costs to consumers, and the share of the overall economic burden borne by producers is considerably larger under the TPS than under C&T.

To address distributional concerns, China's TPS will apply different benchmarks to different power plants. The especially emissions-intensive coal-fired power plants will receive higher (less stringent) benchmarks in order to avoid what would be exceptionally high compliance costs if they faced the same benchmarks as other generators. The planners seem to be focusing on a 3-benchmark system. We find that although this system would reduce the TPS's cost-disparities significantly relative to a 1-benchmark system, it would still produce quite different cost-impacts across the Chinese provinces, reflecting regional differences in the composition of generation technologies. Provinces in the northern and northeastern regions of the country would face the largest percentage reductions in profits. An alternative, 4-benchmark system that "customizes" the benchmarks successfully avoids exceptional cost-impacts in some areas. However, achieving this distributional objective lowers profits in other regions of the country and involves higher aggregate policy costs.

Economists have reason to applaud China's decision to reduce CO_2 through an emissions pricing instrument as well as its plan to move from a group of provincial or municipal pilot programs to an integrated nationwide program. The TPS may not be as cost-effective as C&T, but its reliance on emissions pricing and its broad geographical scope can help achieve emissions reductions at low cost. Also, the TPS has certain attractions relative to C&T. Its rate-based structure implies that policy stringency adjusts automatically in response to changes in macroeconomic conditions. And the fact that it causes smaller increases in electricity prices implies that it could lead to less emissions leakage. The smaller price increase could also be an attraction in terms of fairness and political feasibility. Another potential attraction – though analysts might disagree on this -- is the fact that Chinese planners are more familiar with intensity-based regulation.

It is important to note that despite the fact that its costs are higher than those of C&T, we find that the TPS can generate significant net gains once environmental benefits are counted. If CO_2 emissions reductions are valued at 290 RMB (or about 44 U.S. dollars) per ton, our central case results indicate that the environmental benefits from the TPS would exceed the policy costs by a factor of about 2.9.

Some caveats area in order. First, although we have been fortunate to gain access to important data through our contacts in China, we have faced some limitations in available data, and have needed to calibrate or borrow others' estimates of important parameters rather than estimate them econometrically. Still, the robustness of our results leads us to believe that the key insights from this study would not change significantly with better data. Second, ours is a one-period model. Hence it does not capture investment decisions and associated changes to capital stocks, though it accounts for shutdowns. We would expect that in a multi-period model, the results would have a similar pattern but be amplified. Specifically, we would expect that the use of multiple benchmarks would imply larger sacrifices of cost-effectiveness, as multiple benchmarks would distort investment decisions in addition to output decisions captured in the current model. We would also expect that the differences between the costs of the TPS and C&T would be widened in a model with investment decisons, as the implicit output subsidy of the TPS would cause investment decisions to be less efficient than those under C&T.

We believe this study's findings can significantly help Chinese planners arrive at designs for the TPS that achieve distributional goals with the least additional aggregate cost. The findings should be useful to the broader policy community as well. They bring out hitherto unrecognized channels of impact of the TPS, and they offer unique quantitative estimates of the wide-ranging impacts of China's planned TPS system and the magnitude of these impacts relative to those of

C&T. These results should prove useful to the many regional and national jurisdictions that are considering rate-based, mass-based, and other ways to achieve reductions in emissions of CO_2 and other pollutants.

References

Bovenberg, A. Lans, and Lawrence H. Goulder, 2001. "Neutralizing the Adverse Industry Impacts of CO2 Abatement Policies: What Does It Cost?" in C. Carraro and G. Metcalf, eds., *Behavioral and Distributional Effects of Environmental Policies*, University of Chicago Press.

Burtraw, Dallas, Karen Palmer, Ranjit Bharvirkar, and Anthony Paul, 2002. "The Effect on Asset Values of the Allocation of Carbon Dioxide Emission Allowances." Discussion Paper 02-15, Resources for the Future, Washington, DC.

Burtraw, Dallas, and Karen Palmer, 2007. Compensation Rules for Climate Policy in the Electricity Sector. Discussion Paper 07-41, Resources for the Future, Washington, DC.

Duan, Maosheng and Li Zhou, 2017. "Key Issues in Designing China's National Carbon Emissions Trading System." *Economics of Energy & Environmental Policy* 6(2).

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

Fischer, Carolyn, Biao Mao, and Daniel Shawhan, 2018. "Trade between Mass- and Rate-based Regulatory Regimes: Bad for Emissions?" *Energy Economics,* vol. 73, pp. 326-336.

Fischer, Carolyn, and Richard G. Newell, 2008. "Environmental and Technology Policies for Climate Mitigation." *Journal of Environmental Economics and Management 55*:142-162.

Fischer, Carolyn, and A. K. Fox, 2012. "Comparing Policies to Combat Emissions Leakage: Border Tax Adjustments versus Rebates." *Journal of Environmental Economics and Management* 64(2): 199-216.

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 U.S. Climate Policy*. University of Chicago Press.

Fullerton, Don, and Gilbert Metcalf, 2001. "Environmental Controls, Scarcity Rents, and Pre-Existing Distortions." *Journal of Public Economics* 80(2):249-267.

Goulder, Lawrence H., Marc A.C. Hafstead, and Michael Dworsky, 2010. "Impacts of Alternative Emissions Allowance Allocation Methods under a Federal Cap-and-Trade Program." *Journal of Environmental Economics and Management* 60, pp. 161-181.

Goulder, Lawrence H., and Ian W.H. Parry, 2008. "Instrument Choice in Environmental Policy," *Review of Environmental Economics and Policy* 2(2), Summer 2008, pp. 152–174.

Goulder, Lawrence H., Marc A. C. Hafstead, and Roberton C. Williams III, 2016. "General Equilibrium Impacts of a Federal Clean Energy Standard." *American Economic Journal: Economic Policy 8(2):*186-218.

Goulder, Lawrence H. and Richard D. Morgenstern, 2018. "China's Rate-Based Approach to Reducing CO_2 Emissions: Attractions, Limitations, and Alternatives, *AEA Papers and Proceedings,* May 2018 vol. 108, pp. 458-62.

Haites, Erik, 2003. "Output-Based Allocation as a Form of Protection for Internationally Competitive Industries." *Climate Policy 3*:S29-S41.

Ho, Mun S., Zhongmin Wang, and Zichao Yu, 2017. "China's Power Generation Dispatch." RFF Report, April.

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).

Karplus, Valerie and Xiliang Zhang, 2017. "Incentivizing Firm Compliance with China's National Emissions Trading System." *Economics of Energy and Environmental Policy* 6(2).

Kerr, Suzi, and Richard G. Newell, 2003. "Policy-Induced Technology Adoption: Evidence from the US Lead Phasedown." RFF Discussion Paper 01-14, May.

Parry, Ian W. H., 2003. "Fiscal Interactions and the Case for Carbon Taxes over Grandfathered Carbon Permits," *Oxford Review of Economic Policy* 19, pp. 385-399.

Parry, Ian W.H., and Alan J. Krupnick, 2011. "Is a Clean Energy Standard a Good Way to Move U.S. Climate Policy Forward?" Resources for the Future *Issues Brief*, 11-04, April.

Pizer, William A. and Xiliang Zhang, 2018. "China's New National Carbon Market", *AEA Papers and Proceedings* 108:463-67.

Teng, Fei, Frank Jotzo, and Xin Wang, 2017. "Interactions between Market Reform and a Carbon Price in China's Power Sector." *Economics of Energy & Environmental Policy* (2).

Yang, L.; Lin, B., 2016. "Carbon Dioxide Emission in China's Power Industry: Evidence and Policy Implications." *Energy Rev.* 60, pp. 258–267.

Zhang, Junjie, Zhenxuzn Wang, and Xinming Du, 2017. "Lessons Learned from China's Regional Carbon Market Pilots." *Economics of Energy & Environmental Policy* (2).

Table 1: Production Levels, Production Costs, Emissions Intensities and Emissions by Technology Class, 2016

Technology Category	Technology Class	Number of Units	Annual Electricity Production (million mWh)	Annual Electricity Production as Percent of Total	Representative Facility's Production Cost (million RMB)	Average Emissions Intensity (tCO_2/mWh)	Annual CO_2 Emissions (million tCO_2)
Coal Fired Units							
	C1 - 1000MW Ultra-supercritical Units	74	363.8	11.6	923.27	0.802	291.8
	C2 - 600MW Ultra-supercritical Units	55	187.4	6.0	971.36	0.827	155.0
	C3 - 600MW Supercritical Units	210	641.5	20.5	779.41	0.867	556.1
	C4 - 300MW Supercritical Units	63	98.1	3.1	324.67	0.868	85.1
	C5 - 600MW Subcritical Units	130	359.0	11.5	626.62	0.907	325.6
	C6 - 300MW Subcritical Units	499	836.7	26.7	479.57	0.894	748.0
	C7 - High/Ultra-high Pressure and Lower Pressure Units (with installed capacity less than 300MW)	930	353.3	11.3	120.00	1.006	355.4
Circulating Fluidized Bed Units							
	C8 - Circulating Fluidized Bed Units (with installed capacity greater than or equal to 300MW)	57	71.1	2.3	400.00	0.971	69.0
	C9 - Circulating Fluidized Bed Units (with installed capacity less than 300MW)	229	89.2	2.8	106.57	1.081	96.5
Gas-Fired Units							
	C10 - F-class Gas-fired Units	73	99.7	3.2	382.10	0.372	37.1
	C11 - Gas-fired Units with Pressure Lower than F-class	72	31.4	1.0	106.23	0.422	13.3
All Units		**2392**	**3,131.1**	**100.0**			**2,732.9**

Table 2: Baseline Emissions Intensities and Levels by Technology Class

Technology Category	Technology Class	Average Emissions Intensity (tCO_2/mWh)	Average Emissions Intensity Rank	Baseline Emissions (million tCO_2)	Emissions as Percent of Total Baseline Emissions
Coal-Fired Units					
	C1 - 1000MW Ultra-supercritical Units	0.802	9	292	10.7
	C2 - 600MW Ultra-supercritical Units	0.827	8	155	5.7
	C3 - 600MW Supercritical Units	0.867	7	556	20.4
	C4 - 300MW Supercritical Units	0.868	6	85	3.1
	C5 - 600MW Subcritical Units	0.907	4	326	11.9
	C6 - 300MW Subcritical Units	0.894	5	748	27.4
	C7 - High/Ultra-high Pressure and Lower Pressure Units (with installed capacity less than 300MW)	1.006	2	355	13.0
Circulating Fluidized Bed (CFB) Units					
	C8 - CFB Units with installed capacity greater than or equal to 300MW	0.971	3	69	2.5
	C9 - CFB Units with installed capacity less than 300MW	1.081	1	96	3.5
Gas-Fired Units					
	C10 - F-class Gas-fired Units	0.372	11	37	1.4
	C11 - Gas-fired Units with Pressure Lower than F-class	0.422	10	13	0.5
All Units				**2,733**	**100.0**

Table 3: Baseline Production and Prices by Province, 2016

		Administered-Price Production				Market-Priced Production				
Province	Number of Units	Guaranteed-Hour Production	Guaranteed-Hour Price	Production Sold to Zone	Administered Price	Direct Contracting Production	Direct Contracting Price	Production Sold to Grid Companies	Grid Company Price	Total Production
Anhui	62	85544	0.369	21216	0.448	30847	0.347	8290	0.426	145897
Beijing	16	10299	0.352	2554	0.411	1009	0.329	271	0.388	14133
Chongqing	24	26957	0.380	6686	0.456	10007	0.368	2689	0.434	46339
Fujian	59	68425	0.374	16970	0.448	23769	0.369	6388	0.426	115552
Gansu	44	20125	0.298	4991	0.372	30712	0.276	8254	0.350	64082
Guangdong	214	156817	0.451	38892	0.458	85435	0.357	22961	0.435	304105
Guangxi	24	2069	0.414	513	0.458	33420	0.369	8982	0.435	44984
Guizhou	46	51657	0.336	12811	0.458	22888	0.335	6151	0.435	93507
Hainan	14	10720	0.420	2659	0.458	3480	0.398	935	0.435	17793
Hebei	134	91937	0.357	22801	0.411	33056	0.348	8884	0.388	156679
Heilongjiang	77	28220	0.372	6999	0.423	10176	0.345	2735	0.401	48130
Henan	120	91432	0.355	22676	0.456	46428	0.338	12477	0.434	173013
Hubei	28	29087	0.398	7214	0.456	10296	0.381	2767	0.434	49364
Inner Mongolia	191	126746	0.290	31434	0.423	72285	0.268	19427	0.401	249892
Jiangsu	243	140888	0.378	34942	0.448	86395	0.369	23219	0.426	285443
Jiangxi	28	33234	0.399	8242	0.456	11984	0.396	3221	0.434	56681
Jilin	44	33632	0.372	8341	0.423	12128	0.363	3259	0.401	57360
Liaoning	78	49270	0.369	12219	0.423	26034	0.352	6997	0.401	94520
Ningxia	38	39708	0.260	9848	0.372	14319	0.237	3848	0.350	67723
Qinghai	4	2017	0.325	500	0.372	1258	0.303	338	0.350	4113
Shaanxi	75	50122	0.335	12431	0.372	24668	0.312	6630	0.350	93851
Shandong	312	146643	0.373	36369	0.411	52880	0.351	14211	0.388	250103
Shanghai	42	52668	0.405	13062	0.448	15889	0.383	4270	0.426	85890
Shanxi	111	76604	0.321	18998	0.411	26923	0.306	7236	0.388	129762
Sichuan	39	31992	0.401	7934	0.456	11536	0.379	3100	0.434	54564
Tianjin	26	23068	0.351	5721	0.411	7824	0.329	2103	0.388	38716
Xinjiang	72	46017	0.262	11413	0.372	16401	0.240	4408	0.350	78237
Yunnan	33	31216	0.336	7742	0.458	11257	0.314	3025	0.435	53240
Zhejiang	194	156698	0.415	38862	0.448	48802	0.389	13116	0.426	257478
National	2392	1713813	0.367	425040	0.433	782105	0.340	210191	0.411	3131149

Table 4: Impacts of Tradable Performance Standard and Cap & Trade – The Central Case

	Baseline	TPS	C&T
Benchmarks (tCO_2/mWh)			
-- Coal-fired (technology classes 1-7)		0.864	Allowance Allocations Scaled to Match TPS Emissions
-- CFB (technology classes 8 and 9)		1.006	
-- Gas-fired (technology classes 10 and 11)		0.374	
Emissions (million tCO_2)	2732.88	2648.22	2648.22
-- change from baseline		-84.65	-84.65
-- change from units that shut down		-20.98	-3.76
-- change from units that remain and increase supply		-38.91	0
-- change from units that remain and reduce supply		-24.76	-80.89
-- percentage change from baseline		-3.1	-3.1
Allowance Price (RMB)		225.89	133.39
Allowances Traded (million tCO_2)		56.91	56.88
Aggregate Electricity Supply (million mWh)	3131149	3112288	3073229
-- change from baseline		-18860	-57919
-- change from units that shut down		-20897	-3747
-- change from units that remain and increase supply		3578	0
-- change from units that remain and reduce supply		-1540	-54172
-- percentage change from baseline		-0.60	-1.85
Electricity Price			
Average Electricity Price (RMB)	0.372	0.374	0.379
-- marketed electricity in intraprovincial market	0.340	0.346	0.363
-- administered electricity in intraprovincial market	0.367	0.367	0.367
-- marketed electricity in interprovincial market	0.411	0.417	0.432
-- administered electricity in interprovincial market	0.433	0.433	0.433
Private Cost (million RMB)		8387	4445
-- change in Consumer Surplus		-6062	-20465
-- change in Producer Surplus		-2325	16019
Private Cost per Ton of Reduced Emissions (RMB/tCO_2)		99.07	52.52
Environmental Benefit		24548.5	24548.5

Table 5: Impacts on Generators' Market Status

Technology Category	Technology Class		Market Status, TPS					Market Status, C&T			
			Initially in Compliance?	Policy Response *(percentage of generators in each category)*				Initially in Compliance?	Policy Response *(percentage of generators in each category)*		
				Shut Down	*Operate and Purchase Allowances*	*Operate and Sell Allowances*			*Shut Down*	*Operate and Purchase Allowances*	*Operate and Sell Allowances*
Coal-Fired Units											
	C1		Y	0	0	100		Y	0	0	100
	C2		Y	0	0	100		Y	0	0	100
	C3		Y	0	0	100		N	0	0	100
	C4		Y	0	0	100		N	0	0	100
	C5		N	0.07	99.93	0		N	0	100	0
	C6		N	0.02	99.98	0		N	0	92.99	7.01
	C7		N	5.62	94.38	0		N	1.05	98.95	0
Circulating Fluidized Bed (CFB) Units											
	C8		Y	0	0	100		Y	0	0	100
	C9		N	0.60	99.40	0		N	0.01	99.99	0
Gas-Fired Units											
	C10		Y	0	0	100		N	0	0	100
	C11		N	0.18	99.82	0		N	0.05	99.95	0

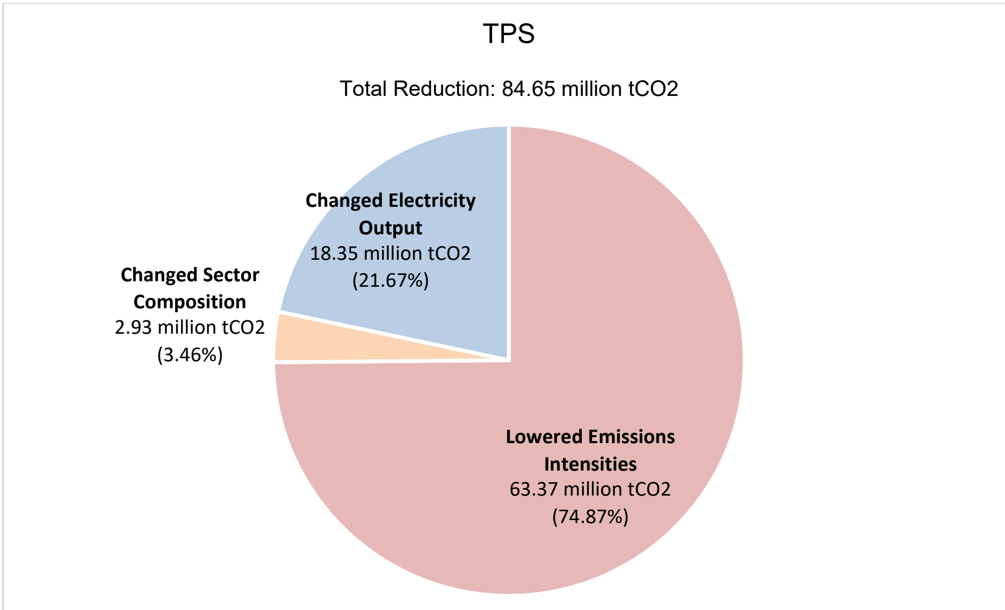
Table 6: TPS Cost Impacts by Region and Province

Region/Province	3-Benchmark (Central) Case		4-Benchmark (Subcategorization) Case	
	Change in Profit (million RMB)	*Change as Pct of Baseline Profit*	*Change in Profit (million RMB)*	*Change as Pct of Baseline Profit*
East	**1273**	**0.78**	**-1257**	**-0.77**
Anhui	590	2.29	-277	-1.07
Shanghai	343	1.87	203	1.11
Jiangsu	26	0.05	-472	-0.91
Jiangxi	258	2.4	-94	-0.88
Zhejiang	54	0.1	-616	-1.1
North	**-3615**	**-3.32**	**-2436**	**-2.24**
Beijing	-46	-2.4	-27	-1.42
Tianjin	33	0.53	-89	-1.39
Shanxi	-430	-2.85	-239	-1.59
Shandong	-1889	-4.98	-856	-2.26
Hebei	-683	-2.99	-118	-0.52
Inner Mongolia	-599	-2.42	-1105	-4.45
Central	**45**	**0.12**	**-36**	**-0.1**
Hubei	34	0.35	-15	-0.16
Henan	11	0.04	-21	-0.08
South	**489**	**0.49**	**-482**	**-0.49**
Guangdong	190	0.28	-240	-0.36
Guangxi	2	0.04	-42	-0.54
Fujian	325	1.61	-170	-0.85
Hainan	-29	-0.73	-29	-0.73
Southwest	**-581**	**-1.39**	**-1143**	**-2.73**
Chongqing	67	0.75	-31	-0.35
Sichuan	-269	-2.53	-284	-2.67
Guizhou	-240	-1.7	-447	-3.16
Yunnan	-138	-1.7	-380	-4.67
Northwest	**863**	**3.22**	**402**	**1.5**
Shaanxi	-12	-0.1	-466	-4.02
Gansu	364	7.55	520	10.8
Ningxia	232	4.73	104	2.13
Qinghai	-8	-2.13	32	8.5
Xinjiang	286	5.62	211	4.14
Northeast	**-801**	**-2.62**	**-383**	**-1.25**
Heilongjiang	-653	-9.32	-351	-5
Jilin	-107	-1.17	7	0.08
Liaoning	-40	-0.28	-39	-0.27
Total	**-2325.0**	**-0.5**	**-5338.0**	**-1.1**

Table 7: Impacts under Alternative Supplly and Demand Elasticities

		Supply Elasticity				Demand Elasticity		
		0	.22 (central case)	0.44		-0.15	-0.3 (central case)	-0.45
TPS								
	Percent Reduction in Emissions	-3.15	-3.1	-3.06		-2.95	-3.1	-3.18
	share from reduced emissions intensities (%)	74.2	74.9	75.2		79.2	74.9	72.8
	share from compositional changes (%)	3.1	3.5	3.8		3.2	3.5	3.6
	share from reduced output (%)	22.7	21.7	21.0		17.7	21.7	23.6
	Percent Increase in Average Electricity Price	0.67	0.01	0.56		0.91	0.62	0.48
	Private Cost	8389	8387	8374		8459	8387	8360
	change in Consumer Surplus	-6635	-6062	-5645		-9674	-6062	-4399
	change in Producer Surplus	-1754	-2325	-2728		1215	-2325	-3961
C&T								
	Percent Reduction in Emissions	-3.15	-3.1	-3.06		-2.95	-3.1	-3.18
	share from reduced emissions intensities (%)	74.2	37.3	27.3		43.4	37.3	35.0
	share from compositional changes (%)	3.1	2.4	2.5		2.0	2.4	2.5
	share from reduced output (%)	22.7	60.3	70.2		54.6	60.3	62.5
	Percent Increase in Average Electricity Price	0.67	1.99	2.29		3.39	1.99	1.45
	Private Cost	8389	4452	2560		4070	4452	4658
	change in Consumer Surplus	-6637	-20465	-23528		-36482	-20465	-14286
	change in Producer Surplus	-1751	16019	20967		32411	16019	9628

Figure 1: Sources of Emissions Reductions under the TPS and C&T



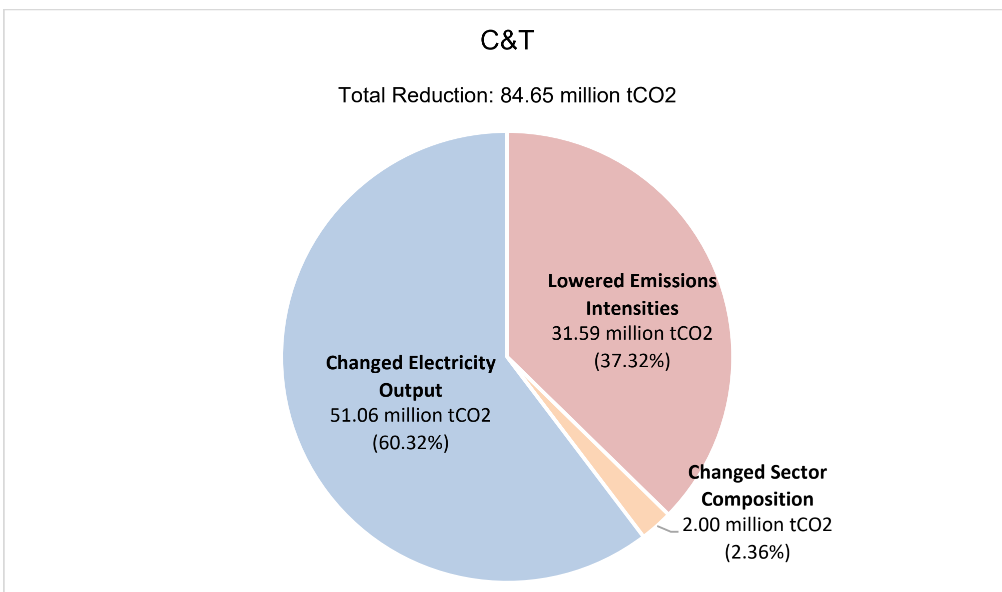


Figure 2: Benchmark Variaton and TPS Cost per Emission Reduction

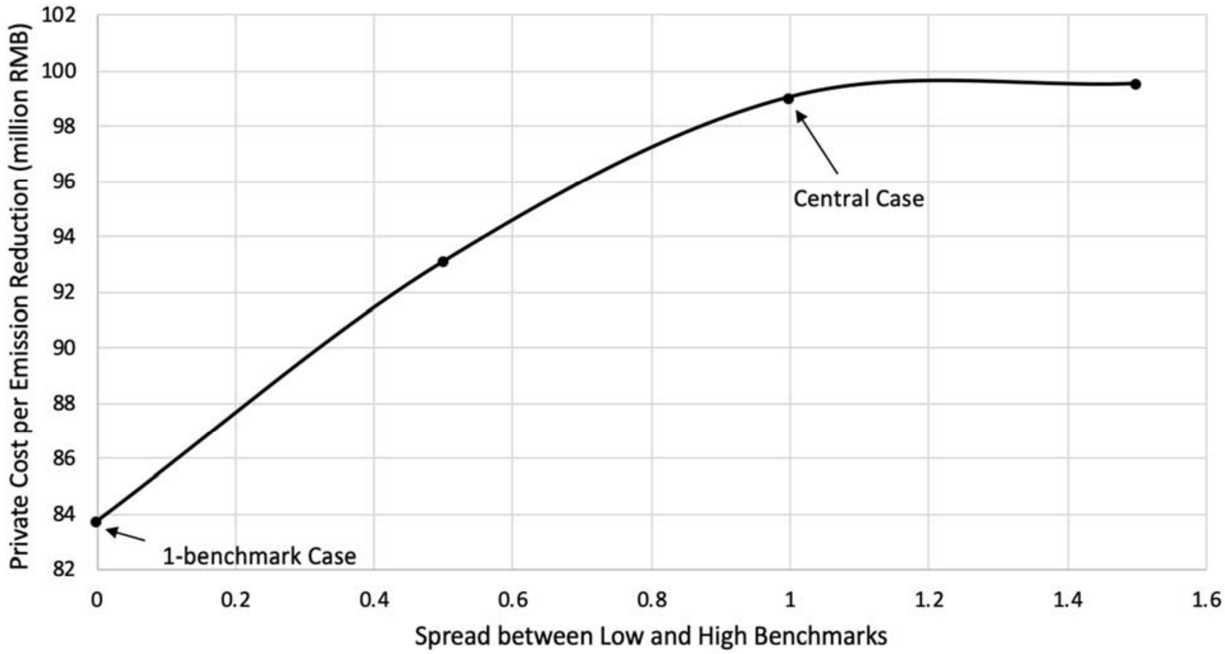
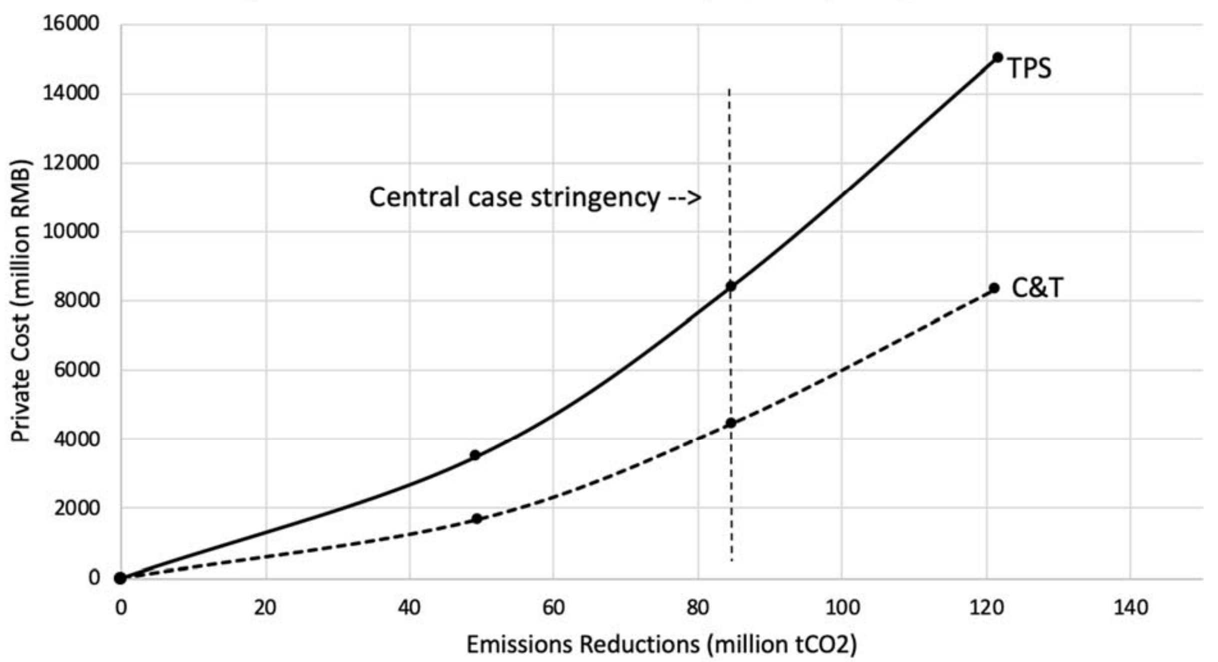


Figure 3: TPS and C&T Costs under Varying Policy Stringencies



Appendix A: Sources and Organization of Electricity Price and Production Data

We distinguish four categories of electricity. For electricity produced in a given province, we distinguish:

- electricity sold in the province at administered prices
- electricity sold in the local (provincial) market
- electricity sold outside the province at administered prices
- electricity sold in the zonal market

This appendix describes how we organize the price and output data for electricity in each of these categories.

1. Electricity Prices

For p_1, the administered prices of electricity sold within a province, we obtained data from China's National Development and Reform Commission. All prices are net of subsidies for desulfurization, denitrification, and soot removal.

Market prices of electricity are collected from China's Electric Power Development Research Institute.

For the market prices of electricity sold within a province, p_2, we calculated the production-weighted average prices in each province using the original data and assigned them to all units in a province. For those provinces that have a higher market prices of electricity sold in the province than its administered prices, we took a further step, summing up their production-weighted average market prices and the average of differences in p_2 and p_1 from provinces with $p_2 < p_1$.

For the market prices of electricity sold in a zonal market, p_3, we calculated the production-weighted average prices in each province, $p_{3\_weighted}$, using the original data. Assuming the existence of transaction costs in zonal market, we calculated p_3 by adding the average of differences between $p_{3\_weighted}$ and p_2 from provinces with $p_2 < p_{3\_weighted}$ to the maximal p_2 in each zone.

Data on administered prices of electricity sold to the zonal market, p_4, were not available. To construct these data, we assume that the gap between the administered price and the market price is the same as in the zonal market. Thus, p_4 is calculated by summing up p_3 and the average of differences in p_1 and p_2.

We do not have data indicating how, for electricity produced in a given province, the administered prices p_1 and p_4 might differ by technology class. Thus we assume that these prices are the same across technology classes of the units within a given province.

2. Electricity Production Levels

Data on the amount of electricity production are collected from China Electricity Council and National Energy Administration. For 2016, we have the total amount of electricity production by province, and national amount of electricity production for the four categories. For 2018, we have the ratio of administered electricity to marketed electricity by province, and the national ratio of administered to marketed electricity. Table A1 summarizes the data we collected.

Table A1: Production Data

	Provincial	National
2016	$X_{1,P}^{2016}$	$X_{1,N}^{2016}$, $X_{2,N}^{2016}$, $X_{3,N}^{2016}$, $X_{4,N}^{2016}$, $X_{5,N}^{2016}$
2018	m_{P}^{2018}	m_{N}^{2018}

Notation:

X_1: total production
X_2: quantity of electricity sold within the province at the province-level administered price
X_3: quantity of electricity sold within the province at the province-level market price
X_4: quantity of electricity sold in the zonal market
X_5: quantity of electricity sold outside the province at the outside-of-province administered price

m: ratio of administered electricity to marketed electricity, i.e., $m=(X_2 + X_5) / (X_3 + X_4)$

Subscript P is for province, N is for national.

The 2016 data do not distinguish the four categories. We use the 2018 ratio data to generate the production levels by category in 2016.

Given the national data for 2016, we compute:

$$m_N^{2016} = (X_{2,N}^{2016} + X_{5,N}^{2016}) / \left(X_{3,N}^{2016} + X_{4,N}^{2016} \right)$$

So, we know the change in m from 2016 to 2018:

$$\%m = m_N^{2018} / m_N^{2016}$$

Assuming change in m from 2016 to 2018 is the same for all provinces:

$$m_P^{2016} = m_P^{2018} / \%m$$

This implies:

$$X_{2,P}^{2016} + X_{5,P}^{2016} = \frac{m_P^{2016}}{m_P^{2016} + 1} X_{1,P}^{2016}$$

$$X_{3,P}^{2016} + X_{4,P}^{2016} = \frac{1}{m_P^{2016} + 1} X_{1,P}^{2016}$$

Assuming ratio of X_2 to X_5, ratio of X_3 to X_4 are the same to the national ratios in 2016.

$$r_{25} = X_{2,N}^{2016} / X_{5,N}^{2016}$$

$$r_{34} = X_{3,N}^{2016} / X_{4,N}^{2016}$$

We then compute amount of production for the four categories:

$$X_{2,P}^{2016} = \frac{r_{25}}{r_{25} + 1} \left(X_{2,P}^{2016} + X_{5,P}^{2016} \right)$$

$$X_{5,P}^{2016} = \frac{1}{r_{25} + 1} \left(X_{2,P}^{2016} + X_{5,P}^{2016} \right)$$

$$X_{3,P}^{2016} = \frac{r_{34}}{r_{34} + 1} \left(X_{3,P}^{2016} + X_{4,P}^{2016} \right)$$

$$X_{4,P}^{2016} = \frac{1}{r_{34} + 1} \left(X_{3,P}^{2016} + X_{4,P}^{2016} \right)$$

In this way, we obtain the amount of electricity production by category by province in 2016.

Appendix B: Determining Values and Distribution of Cost Function Parameters

We obtain parameters for the generators' cost functions in two steps. First, we calibrate all of the parameters for the cost function of the median-cost generator in each technology class. We then derive parameters that determine the distribution of the constant term for the family of cost functions of each technology class.

1. Parameters for Median Generator in Each Technology Class

For each of the 11 generator technologies, we specify the following functional form for the cost function:

$$C = \phi_0 + \phi_1 q^{\phi_2}$$

where ϕ_0, ϕ_1, and ϕ_2 are parameters. We impose the following two restrictions to identify the parameters for each generator technology.

1. At the benchmark level of output, q_0, the cost function yields the observed benchmark cost, C_0: $\phi_0 + \phi_1 q^{\phi_2} = C_0$.

2. At q_0, marginal production cost C' C' is equal to the benchmark market price of electricity p_0: $\phi_2 \phi_1 q^{\phi_2 - 1} = p_0$.

The two conditions together imply:

$$\phi_2 = \frac{p_0 q_0}{C_0 - \phi_0}$$

and

$$\phi_1 = (C_0 - \phi_0) q_0^{-\phi_2}$$

With three parameters and two identifying conditions, there is an infinite number of combinations of the parameters could meet the conditions. In particular, for any given choice of ϕ_0 there is a combination of ϕ_1 and ϕ_2 that meets the two conditions. We choose ϕ_0 so that, applying the two equations immediately above, it yields values of ϕ_1 and ϕ_2 that generate the desired target price elasticity of supply under business as usual. This cost function implies the following formula for the price elasticity of supply, η:

$$\eta = (\phi_2 \phi_1)^{1/(1-\phi_2)} \frac{1}{\phi_2 - 1} p^{1/(\phi_2 - 1)} / q$$

2. Distribution of Costs within Technology Classes

To incorporate cost heterogeneity within each technology class, we vary the parameter ϕ_0 that applies to each class. ϕ_0 is a constant term in the cost function for each class. We assume that this parameter is distributed according to a beta distribution. This is a bounded distribution. The probability density function (pdf) of beta distribution has the general form:

$$\frac{x^{\omega-1}(1-x)^{\delta-1}}{\int_0^1 v^{\omega-1}(1-v)^{\delta-1} dv}$$

We impose symmetry on this pdf by setting ω equal to δ.

a. Determining values for the maximal, minimal, and mean values of ϕ_0

For consistency with the data, we require that ϕ_{0mean} the mean value of ϕ_0 from the distribution, be equal to the value obtained in the calibration procedure above for the representative generator in the given technology class.

Economic considerations imply that the upper bound of the distribution should have a value that makes profit just equal to zero for the generator with that value in the baseline. The generator with $\phi_0 = \phi_{0max}$ is a marginal producer. It makes zero economic profit and thus even a slight increase in cost implies negative profit and induces this unit to shut down. Thus ϕ_{0max} must have a value that makes profit equal to zero under business-as-usual (or baseline) conditions. Hence ϕ_{0max} satisfies:

$$pq_{BAU} + (\overline{p} - p)\overline{q}_{BAU} - \phi_{0\max} - \phi_1 q_{BAU}{}^{\phi 2} = 0$$

Hence,

$$\phi_{0max} = pq_{BAU} + (\overline{p} - p)\overline{q}_{BAU} - \phi_1 q_{BAU}{}^{\phi_2}$$

b. Translating the Beta Distribution into a Distribution for ϕ_0.

In its standard form, the beta distribution is defined over the interval (0,1). We need to shift and scale the standard distribution over the interval (0,1) translates to the interval (ϕ_{0min}, ϕ_{0max}) with mean value ϕ_{0mean}.

Let a and b denote the scale factor and the shift factors that translate the pdf's initial (0,1) distribution into the desired distribution for the model. And let x be the mean of the initial beta distribution. We choose a and b to satisfy:

$$\phi_0 = a(x - 0.5) + b$$

s.t.

when x = 0.5, $\phi_0 = \phi_{0mean}$

when x = 1, $\phi_0 = \phi_{0\max}$

The solution is: $a=2(\phi_{0max} - \phi_{0mean})$ and $b=\phi_{0mean}$. Under the TPS or cap and trade, there will exist values of ϕ_0 that are critical in the sense that any generator with ϕ_0 greater than that value will have zero profit. This translation enables us to determine the fraction of generators in the technology class involved that have ϕ_0 above this value, and thus the number of generators that must shut down. This enables us to calculate the loss of profits to the generators that shut down. In addition, from the distribution of costs for the generators that remain in operation, we can calculate the changes in profit to the remaining generators.

The calculations rely on the pdf and cumulative distribution functions defined on the distribution of ϕ_0 . These distributions can be derived from the pdf for the x translation of ϕ_0. Because the translation is linear, the cdf for ϕ_0 is identical to the cdf for $x(\phi_0)$. The probability density functions $pdf_{\phi_0}\left(\phi_0\right)$ and $pdf_x\left(x\right)$ are not identical, however. The relationship between the two can be derived as follows. We start with the recognition that $cdf_{\phi_0}(\phi_0) = cdf_x(x)$. Then we take the full derivative with respect to x on both sides:

$$\frac{d}{d\phi_0} cdf_{\phi_0}(\phi_0) \frac{d\phi_0}{dx} dx = \frac{d}{dx} cdf_x(x) dx$$

$$pdf_{\phi_0}\left(\phi_0\right) \frac{d\phi_0}{dx} = pdf_x\left(x\right)$$

Since $\phi_0 = a(x-0.5)+b$, we have $\frac{d\phi_0}{dx} = a$. As a result,

$$pdf_{\phi_0}(\phi_0)a = pdf_x(x)$$

or

$$pdf_{\phi_0}\left(\phi_0\right) = \frac{pdf_x\left(x\right)}{a}$$