The Welfare Implications of Carbon Price Certainty

Joseph E. Aldy, Sarah Armitage

Abstract: Experiences in real-world pollution markets suggest that firms make persistent errors in forecasting allowance and credit prices that inform their investment decisions. The residual uncertainty characterizing allowance and credit trading means that pollution markets may fail to deliver cost-effective abatement. This contrasts with price-based policies under which firms make investments that equate marginal abatement cost to an emission tax. We incorporate the additional cost of forecast errors under quantity-based programs into a standard Weitzman-style prices versus quantities framework. We distinguish between individual firms' uncertainty over competitors' private information and systemic uncertainty over future cost shocks. We show that a welfare-maximizing regulator would favor price instruments in response to the prospect of firm-specific forecast errors under quantity instruments, ceteris paribus, and the relative benefit of price instruments increases with forecast error variance. We discuss the role of policy design, such as incorporating price collars, in mitigating cost inefficiencies from price forecast errors.

JEL Codes: Q52, Q54, Q58

Keywords: instrument choice, investment under uncertainty, climate change, carbon tax, cap and trade, tradable performance standard

PUBLIC POLICIES—and uncertainty over public policies—play critical roles in firms' investment and strategic decision making (Arrow and Fisher 1974; Rodrik 1991; Hassett and Metcalf 1999; Baker et al. 2016). Even after the regulator selects a policy

Joseph E. Aldy is at Harvard Kennedy School, Resources for the Future, National Bureau of Economic Research, and Center for Strategic and International Studies (joseph_aldy@hks.harvard.edu). Sarah Armitage is at Harvard University (saraharmitage@g.harvard.edu). Ken Norris provided excellent research assistance. We have benefited from constructive feedback provided by Ed Glaeser, Jerry Green, Sam Hanson, Garth Heutel, Derek Lemoine, Robert Pindyck, Jim Stock, two referees, and the editor, as well as participants at the AERE summer conference, Northeast

Dataverse data: https://doi.org/10.7910/DVN/KVNPGY

Received August 13, 2020; Accepted May 6, 2022; Published online June 27, 2022.

Journal of the Association of Environmental and Resource Economists, volume 9, number 5, September 2022.

https://doi.org/10.1086/720768

instrument, thereby resolving uncertainty around policy design, firms may continue to face residual uncertainty over returns to investment due to the characteristics of policy implementation. In the context of carbon pricing, cap-and-trade allowance prices are inherently more uncertain than a carbon tax.

Much of the literature comparing cap-and-trade programs and emission taxes has focused on the regulator's information deficit (Weitzman 1974; Stavins 2003). Firms typically know their own marginal abatement costs with greater precision than the regulator, and firms lack the incentive to reveal their true marginal abatement costs to the regulator (and to their competitors). A cap-and-trade program enables the regulator to circumvent its information problem by providing incentives for firms to pursue cost-minimizing strategies for reducing pollution. This approach does not eliminate the information problem, however; it simply shifts the information deficit onto firms. To illustrate, suppose the regulator faces a choice between a pollution tax or a cap-and-trade policy. If the regulator chooses a tax, the firm learns the tax rate, identifies its abatement options, and invests in abatement technology that equates marginal cost with the tax. If the regulator chooses a cap-and-trade program, by contrast, the firm identifies its abatement options as in the tax case and then must form expectations about the market-clearing price for allowances to guide its abatement technology investment.

Our paper examines this residual uncertainty in allowance prices, which increases the risk that firms may not equate their marginal abatement costs.[1] Firms may err in their allowance price forecasts and make investments that appear optimal ex ante but that are too high or too low ex post. Such forecast errors may reflect different expectations about (1) other firms' abatement costs, (2) economic output, and (3) overlapping public policies that may restrict abatement decisions and influence allowance prices. By increasing the aggregate costs of achieving any given emission goal, this residual uncertainty would increase the welfare costs of implementing a quantity-based instrument. Given the potential trillion-dollar scale of carbon markets, such forecast errors under cap and trade could be economically significant.

To examine the welfare impacts of firm-level uncertainty, we develop a modified version of Weitzman's canonical prices versus quantities framework. In our version,

---

Environmental Economics Workshop, Harvard Kennedy School Energy Policy seminar, American Economic Association annual conference, Berkeley Climate Economics Workshop, National Bureau of Economics Research (NBER) Energy and Environmental Economics spring workshop, and EAERE annual conference. Aldy acknowledges financial support from BP, Resources for the Future, the Taubman Center for State and Local Government, the Belfer Center for Science and International Affairs, and the Roy Family Fund, and Armitage acknowledges financial support from the Vicki Norberg-Bohm Fellowship, the Joseph Crump Fellowship, and the NBER Pre-Doctoral Fellowship in Energy Economics.

1. This analysis generally applies to various quantity-based pollution markets that enable trading of allowances or credits, including cap-and-trade programs and tradable performance standards, and thus we use allowance and credit prices interchangeably.

quantity orders are not imposed directly on individual firms, as in Weitzman (1974) and much of the subsequent literature. Instead, we model an "aggregate quantity order," akin to an emissions cap, that determines the total quantity of emission allowances in a given pollution market. With the allocation and subsequent trading of these allowances, firms face uncertainty over the market-clearing allowance price. We first show how firm-level uncertainty over rivals' cost shocks or overlapping policies gives rise to forecast errors over the market-clearing price for allowances. These firm-level price forecast errors introduce cost inefficiencies for quantity-based regulations. Holding constant other factors, these additional cost inefficiencies would lead a welfare-maximizing regulator to favor price-based instruments over quantity-based instruments, with the relative benefit of price instruments increasing in the variance of the forecast error term. We acknowledge, however, that the cost inefficiencies from these price forecast errors may be augmented or mitigated in cases where individual firms also have residual uncertainty about their own abatement costs when committing to particular compliance strategies.

Since policy makers have revealed their preferences for quantity-based instruments in practice (Santikarn et al. 2021), we also address how to reduce uncertainty-related cost inefficiencies through instrument design, specifically price collars that impose a floor and a ceiling on allowance prices. Price collars may reduce the variance of firm forecast errors and associated welfare losses. We discuss how the information provision role of price collars should be considered alongside standard arguments about the welfare consequences of introducing price ceilings or floors.

Before turning to the model, let us examine the assumptions behind the cost-minimizing result and how real-world experience deviates from those assumptions. Montgomery (1972) formally showed how, under specific assumptions, firms operating under a cap-and-trade program each have an incentive to equate their marginal abatement costs with the allowance price clearing in the market. In this full-information equilibrium, the allowance price reflects a common, cost-minimizing marginal abatement cost among all firms in the pollution market. Many environmental textbooks illustrate how firms equate marginal abatement costs with allowance prices under cap-and-trade programs through either graphical illustrations or stylized numerical examples.[2] Stavins (2003, 361) summarizes a point made in these textbooks: "tradable permits can achieve the same cost-minimizing allocation of the control burden as a charge system." This cost-minimization characteristic of cap-and-trade programs has been described in the Economic Report of the President (Council of Economic Advisers 2000, 2010), Energy Information Administration (2009) modeling analyses of carbon dioxide cap-and-trade programs, and the Environmental Protection Agency (2015) fact sheet on the Clean Power Plan.

2. For example, refer to Kolstad (2000, 163), Callan and Thomas (2013, 145–48), Field and Field (2016, 258–59), Keohane and Olmstead (2016, fig. 9-1), Harris and Roach (2017, fig. 16.5), Tietenberg and Lewis (2018, fig. 15-6), and Hanley et al. (2019, fig. 11.5).

Complementing the static cost-effectiveness across firms, Rubin (1996) and Kling and Rubin (1997) demonstrate the potential for dynamic cost-effectiveness in cap-and-trade programs with intertemporal trading. The resulting Hotelling-style allowance price path in their theoretical set-up is reflected in virtually every integrated assessment modeling analysis of long-term quantitative climate change goals (Clarke et al. 2014). These analyses report carbon prices that minimize the costs of achieving quantitative goals because they assume full-information, well-functioning, liquid allowance markets; that is, every emission source equates its marginal abatement cost with the observed carbon price.

In practice, pollution markets with allowance and credit trading reveal evidence of frequent violations of these assumptions. The Carlson et al. (2000) analysis of SO_2 emission trading in the United States found that more than half of electricity-generating units failed to minimize compliance costs at some point during phase I of the program. Bilateral credit transactions under the California Low Carbon Fuel Standard (LCFS) reveal remarkable price dispersion: on days with multiple transactions reported to the regulator, the within-day credit price spread averages about 20% (see fig. 1). On only 2% of days with multiple transactions is there a single, common price for LCFS credits. In 2020, transportation sector firms paid four times as much, on average, for a CO_2 emission offset as telecommunications sector firms (Ecosystem Marketplace 2021). This variation reflects the dispersion in internal carbon prices used by major corporations, including those operating in the same sector (Aldy and Gianfrate 2019).

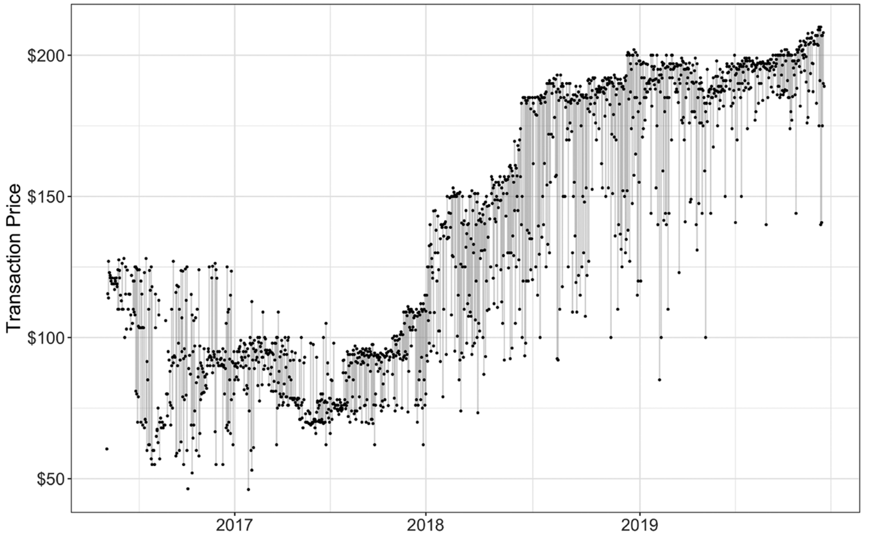

Figure 1. Minimum and maximum daily transaction prices under California's Low Carbon Fuel Standard (LCFS). Constructed by authors using LCFS transaction prices posted online by the California Air Resources Board. Refer to appendix section A.8 for details on construction and data sourcing.

Economic and policy shocks have caused unexpected swings in allowance prices. RECLAIM allowance prices increased by two orders of magnitude during the California electricity crisis (Fowlie et al. 2012). SO_2 allowance prices fell an order of magnitude over two years after the courts overturned related regulations. EU ETS CO_2 allowances have been more volatile than crude oil futures prices since their inception (Aldy and Viscusi 2014).

Overlapping economic and environmental regulations likewise undermine attainment of cost-effective abatement. Fowlie (2010) and Cicala (2015) show that deregulated firms underinvest in capital-intensive compliance strategies for NO_x and SO_2 cap-and-trade programs, combined with evidence of overinvestment by regulated firms. In the context of California's CO_2 cap-and-trade program, the uncertainty around nonmarket overlapping policies, combined with uncertainty around business-as-usual emissions, dwarfs uncertainty over price-responsive abatement quantities (Borenstein et al. 2019).

These illustrations may not be that surprising given that they reflect the challenges a given firm faces in making long-lived capital investments in abatement technology that can influence the firm's demand and supply of emission allowances in a market that may not exist, could be illiquid, or lacks risk management financial instruments at the time of the investment decision. Even without such circumstances, there are theoretical and empirical contexts in which markets realize price dispersion—and hence variation in marginal costs of supply—in practice (Salop and Stiglitz 1977; Burdett and Judd 1983; Carlson and McAfee 1983; Dana 1994; Allen 2014).

In the next section, we develop a Weitzman-style prices versus quantities welfare framework to model tradable quantities. Section 2 introduces firm uncertainty over rivals' abatement costs to illustrate the welfare implications of allowance price forecast errors. Section 3 presents the welfare impacts of uncertainty associated with overlapping policies. Section 4 addresses firms' uncertainty over their own abatement costs. Section 5 explores the role of policy design in ameliorating the inherent uncertainty of allowance prices. The final section concludes and offers directions for future research.

1. MODEL OF TRADABLE QUANTITIES

We build on existing theories of cap-and-trade markets to account for cost inefficiencies characterizing firms' behavior in these markets. Previous studies in the spirit of Weitzman's canonical work on the relative advantage of price versus quantity instruments have generally focused on uncertainties on the part of the regulator rather than on the part of regulated firms (Hoel and Karp 2002; Pizer 2002; Williams 2002; Yates 2002; Newell and Pizer 2003). Much of this literature assumes that cap-and-trade programs allocate "quantity orders"—Weitzman's term for assigning a quantity to each firm—directly to individual firms, with the consequence that firms cannot face the types of uncertainties over market-clearing allowance prices that we cataloged above. One exception is Yohe (1978), who considers the possibility that firms may not know their own abatement cost functions with certainty, so quantity orders imposed directly on firms may not be achieved exactly. Williams (2002) and Yates (2002) have also compared decentralized allowance markets to direct quantity targets, but both assume perfect certainty on the part of firms participating in these markets, even over multiple periods. By contrast, in real-world allowance markets, the regulator issues an "aggregate quantity order" by setting an emission cap. This aggregate quantity order is then implemented through a decentralized process—the allowance market—with firm-level uncertainty over the market-clearing price. One consequence of the complex price formation process is that all types of shocks affect the equilibrium price, so regulated firms are affected by more than just direct shocks to their own abatement cost functions.

Given our focus on firm-level uncertainty in allowance trading markets, we distinguish between cross-sectional uncertainty over competitors' cost shocks and other sources of private information and intertemporal uncertainty over as yet unrealized shocks. In the former case, we would expect firm-specific uncertainty over competitors' cost shocks to be resolved simultaneously, as markets clear and firms' compliance decisions become known. These idiosyncratic forecast errors are evident, for example, in the range of transaction prices observed even within a single trading day under California's Low Carbon Fuel Standard. In the latter case, uncertainty is resolved sequentially, as cost shocks are realized over time. Indeed, the former represents epistemological uncertainty over shocks that have been realized but are not perfectly observed, while the latter represents fundamental uncertainty over shocks that have yet to be realized. As we see in the derivation that follows, the intertemporal linkages of equilibrium prices in cap-and-trade markets with banking and borrowing mean that uncertainty in early periods of a policy may continue to affect the policy's cost-effectiveness even after information has been fully revealed to participants in the market.

To understand the welfare consequences of firm-level forecast errors, we develop a modified version of the Weitzman (1974, 2020) framework for evaluating the welfare consequences of price-based and quantity-based instruments under uncertainty. Before we introduce firm-level forecast errors, our model also closely resembles the decentralized allowance markets in Williams (2002) and Yates (2002).[3] We first introduce firm-level forecast errors around the cost shocks of competitor firms, but as we demonstrate below, our results are mathematically similar when forecast errors instead stem from uncertainty around the impact of overlapping policies. We then consider the case where firms determine their compliance strategies based on imperfect information about their own abatement costs, to understand interactions with uncertainty over the market-clearing price in

3. Other closely related papers in this recent literature on prices versus quantities with banking and borrowing are Karp (2019), Heutel (2020), and Pizer and Prest (2020). These papers also consider how the relative advantage of prices, quantities, and quantities with banking and borrowing depend on whether the policy maker is able to update policies over time as information is revealed in the market. Our baseline model considers what amounts to an "open loop" policy, or a policy without updating, in which the regulator sets prices or quantities at the start of the two-period regulatory cycle.

a tradable quantities program versus certainty over the regulated price in a tax regime. Finally, while our model captures many key features of multiperiod price- and quantity-based regulation in the presence of uncertainty, we adopt the standard assumption that abatement is variable.[4] Future work will examine in greater detail the impact of price and cost uncertainty on dynamic firm-level investment in abatement.

We begin by defining the benefit function associated with reducing some pollutant and the cost function associated with abatement of that pollutant. Let $B_t(Q_t)$ represent the benefits in period t associated with pollution abatement at level Q_t.[5] Likewise, let $C_t^i(q_t^i, \theta_t^i)$ be the cost to firm i associated with producing quantity q_t^i of abatement, where θ_t^i represents a firm-specific random shock to the cost function in period t. Therefore, the aggregate costs associated with abatement level Q_t are given by $\Sigma_{i=1}^N C_t^i(q_t^i, \theta_t^i)$, where $Q_t = \Sigma_{i=1}^N q_t^i$. We assume uniform mixing of the pollutant in question, such that only the total level of abatement enters into the benefits function, not the identity of each polluting entity. This assumption reflects the characteristics of carbon dioxide and most other greenhouse gas emissions but could be relaxed to model local pollutants. By contrast, the costs of abatement depend on the level achieved by each individual firm.[6]

We assume for tractability that there are two periods in the current regulatory cycle.[7] Let us consider the regulator's problem under two instrument options. Under a tax, the regulator sets an optimal price order in the presence of uncertainty by solving the following maximization problem:

$$\max_{\tilde{p}_1,\tilde{p}_2} \mathrm{E}\left[B_1\left(\sum_{i=1}^N q_1^i\left(\tilde{p}_1, \theta_1^i\right)\right) - \sum_{i=1}^N C_1^i\left(q_1^i\left(\tilde{p}_1, \theta_1^i\right), \theta_1^i\right) + B_2\left(\sum_{i=1}^N q_2^i\left(\tilde{p}_2, \theta_2^i\right)\right) - \sum_{i=1}^N C_2^i\left(q_2^i\left(\tilde{p}_2, \theta_2^i\right), \theta_2^i\right)\right]. \tag{1}$$

4. Weitzman (2020) provides further discussion of the need for this assumption in n. 8.

5. Here we assume no uncertainty in the benefit function, which is consistent with Weitzman (2020)

6. We also assume that the benefits of abatement can be approximated with a marginal benefit function that is separable over time. This assumption seems reasonable in the case of climate change given evidence that the damage function from greenhouse gas emissions may be approximately linear in temperature (Burke et al. 2015). We acknowledge the work of existing papers that have examined how optimal policies may depend on the characteristics of marginal damages (Williams 2002; Yates 2002; Gerlagh and Heijmans 2018), though this discussion is not directly relevant to our modeling of the cost inefficiencies introduced by firm-level uncertainty. We also note that the key insights from our model remain intact when we explicitly model a stock pollutant where avoided damages (benefits) are intertemporally linked; this derivation is available from the authors upon request.

7. As discussed below, we also assume here that there is no discounting between periods and that the intertemporal permit trading ratio is equal to 1. These assumptions greatly simplify our derivation and allow us to highlight the impact of forecast errors. Karp (2019) shows that the choice of discount factor and permit trading ratio may affect the relative advantages of prices, quantities, and quantities with banking and borrowing.

Under a cap-and-trade regime, the regulator sets the optimal (aggregate) quantity order by solving the following maximization problem:

$$\max_{\hat{Q}} \mathrm{E}\Bigg[B_1\left(\sum_{i=1}^{N} q_1^i\left(p_1\left(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}\right), \theta_1^i\right)\right) - \sum_{i=1}^{N} C_1^i\left(q_1^i\left(p_1\left(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}\right), \theta_1^i\right), \theta_1^i\right) + B_2\left(\sum_{i=1}^{N} q_2^i\left(p_2\left(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}\right), \theta_2^i\right)\right) - \sum_{i=1}^{N} C_2^i\left(q_2^i\left(p_2\left(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}\right), \theta_2^i\right), \theta_2^i\right)\Bigg]. \quad (2)$$

Here $p_t(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2})$ represents the market-clearing price associated with the regulated quantity order $\hat{Q}$ and the vectors of per-period marginal cost shocks $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$.

Following Weitzman and related literature, we expand the cost and benefit functions by taking a second-order Taylor expansion about the quantity $\bar{q}_t^i$. We define each $\bar{q}_t^i$ as the level of abatement that sets expected benefits equal to expected costs for each individual firm. The Taylor expansion of each abatement cost function about $\bar{q}_t^i$ is then given by

$$C_t^i(q_t^i, \theta_t^i) = a_i(\theta_t^i) + (C' + \theta_t^i)(q_t^i - \bar{q}_t^i) + \frac{C_i''}{2}(q_t^i - \bar{q}_t^i)^2, \quad (3)$$

where C' represents the expected marginal abatement cost at $\bar{q}_t^i$, and C_i'' represents the slope of the marginal abatement cost function.[8] As in Weitzman and much of the subsequent literature, we assume for tractability that the abatement cost function is quadratic or can be well approximated by a second-order Taylor expansion. The term θ_t^i represents how the random cost shock affects the slope of the abatement cost function for firm i, and $a_i(\theta_t^i)$ represents how the cost shock affects the level of abatement costs. As in Weitzman's derivation, we assume without loss of generality that $\mathrm{E}[a_i(\theta_t^i)] = 0$ and $\mathrm{E}[\theta_t^i] = 0$. Note that this assumption does not preclude that the conditional expectation of a firm's cost shock differs from 0; in general, $\mathrm{E}[a_i(\theta_t^i)|a_j(\theta_t^j)] \neq 0$ and $\mathrm{E}[\theta_t^i|\theta_t^j] \neq 0$ for some i and j, and $\mathrm{E}[a_i(\theta_t^i)|a_i(\theta_{t'}{}^i)] \neq 0$ and $\mathrm{E}[\theta_t^i|\theta_{t'}{}^i] \neq 0$ for t and t'.

For the benefits function, we also take the Taylor expansion around $\bar{Q}_t = \Sigma_{i=1}^N \bar{q}_t^i$:

$$B(Q_t) = b + B'\left(\sum_{i=1}^{N} q_t^i - \sum_{i=1}^{N} \bar{q}_t^i\right) - \frac{B''}{2}\left(\sum_{i=1}^{N} q_t^i - \sum_{i=1}^{N} \bar{q}_t^i\right)^2. \quad (4)$$

Here B' captures the marginal benefit at $\bar{Q}_t = \Sigma_{i=1}^N \bar{q}_t^i$, and B'' captures the slope of the marginal benefit function (where $B'' \geq 0$).

For the optimal tax, the derivation here closely follows Weitzman's derivation with multiple production units, except we constrain the regulated price to be the same across all units. We again find that the optimal price order $\tilde{p}_t$ is equal to $B' = C_i'$

8. As we show in the full derivation in appendix secs. A.1 and A.2, the Taylor expansion is defined such that C' is constant for all i. Furthermore, while we could allow for the parameters of the cost and benefit function to differ across periods, we assume for analytic tractability that C_i'' and B'' are constant over time.

for all i (and for all t where these parameters are constant). The full derivation is provided in appendix section A.1 (appendix is available online). Assuming cost minimization, each firm will set its realized marginal cost function equal to this price, yielding the following firm-level response function:

$$\tilde{q}_t^i = q_t^i(\tilde{p}_t, \theta_t^i) = \bar{q}_t^i - \frac{\theta_t^i}{C_i''}. \quad (5)$$

The aggregate quantity produced in each period under the optimal tax will be

$$\sum_{i=1}^{N} \tilde{q}_t^i = \sum_{i=1}^{N} q_t^i(\tilde{p}_t, \theta_t^i) = \sum_{i=1}^{N} \left(\bar{q}_t^i - \frac{\theta_t^i}{C_i''} \right) = \bar{Q}_t - \sum_{i=1}^{N} \frac{\theta_t^i}{C_i''}. \quad (6)$$

For the optimal cap-and-trade program, our derivation differs from the original Weitzman derivation in that the regulator sets an overall quantity target but cannot directly determine how that overall quantity is allocated across firms. We find that the optimal quantity order $\hat{Q}$ is equal to $\Sigma_{i=1}^{N} \bar{q}_1^i + \Sigma_{i=1}^{N} \bar{q}_2^i$; the full derivation is provided in appendix section A.2. To determine how this overall quantity is then allocated across firms, we must solve for the market-clearing price such that the aggregate quantity order is achieved after shocks are realized in the market.[9] Furthermore, because we allow banking and borrowing across the two periods, we must also apply a no-arbitrage condition that relates the first-period price to the expected second-period price. We assume that both the discount factor and the intertemporal trading ratio are equal to 1, so the first-period price and expected second-period price must be equal in equilibrium.

To build intuition for the basic model set-up, we initially follow Weitzman and assume that firms know all shocks in the market before making any compliance decisions and therefore are able to set their marginal abatement costs equal to a known market-clearing price; we then relax this assumption in subsequent sections to show the impact of firm-level uncertainty. Assuming cost minimization, the equilibrium price associated with the overall optimal quantity order $\hat{Q}$ is given by

$$\hat{p}_1(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}) = \hat{p}_2(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2}) = C' + \frac{\sum_{i=1}^{N} \frac{\theta_1^i + \theta_2^i}{2C_i''}}{\sum_{i=1}^{N} \frac{1}{C_i''}}. \quad (7)$$

In the first period, the firm response function dictates that firm i will produce

9. A cap-and-trade program indirectly implements an aggregate quantity order in the spirit of Weitzman. While an emissions cap and the subsequent trading of allowances allocates the quantity of allowed emissions, the quantity order in the Weitzman framework instead determines the quantity of emissions abatement. While not explicitly considered in this model, uncertainty over business-as-usual (BAU) emissions—and therefore uncertainty over how the quantity of allowed emissions will translate into a quantity of emissions abatement—is an additional source of firm-level uncertainty. Borenstein et al. (2019) discuss the importance of uncertainty over BAU emissions in driving systematic uncertainty over future allowance prices.

$$\hat{q}_1^i = \bar{q}_1^i + \frac{\sum_{j=1}^N \dfrac{\theta_1^j + \theta_2^j}{2C_j''}}{C_i''\sum_{j=1}^N \dfrac{1}{C_j''}} - \frac{\theta_1^i}{C_i''};$$

the aggregate first-period quantity produced will be $\hat{Q}_1 = \bar{Q}_1 + \Sigma_i[(\theta_2^i - \theta_1^i)/2C_i'']$. Likewise, in the second period, firm i will produce

$$\hat{q}_2^i = \bar{q}_2^i + \frac{\sum_{j=1}^N \dfrac{\theta_1^j + \theta_2^j}{2C_j''}}{C_i''\sum_{j=1}^N \dfrac{1}{C_j''}} - \frac{\theta_2^i}{C_i''};$$

the aggregate second-period quantity is then $\hat{Q}_2 = \bar{Q}_2 + \Sigma_i[(\theta_1^i - \theta_2^i)/2C_i'']$. (The full derivation is given in appendix sec. A.2.)

Finally, following Weitzman, we define the relative advantage of prices over quantities as the expected difference between net benefits from the optimal tax and net benefits from the optimal cap-and-trade program with banking and borrowing. That is,

$$\begin{aligned}\Delta = {} & \mathrm{E}\left[B_1\left(\sum_{i=1}^N \tilde{q}_1^i\right) + B_2\left(\sum_{i=1}^N \tilde{q}_2^i\right) - \sum_{i=1}^N C_1^i\left(\tilde{q}_1^i, \theta_1^i\right) - \sum_{i=1}^N C_2^i\left(\tilde{q}_2^i, \theta_2^i\right)\right] \\ & -\mathrm{E}\left[B_1\left(\sum_{i=1}^N \hat{q}_1^i\right) + B_2\left(\sum_{i=1}^N \hat{q}_2^i\right) - \sum_{i=1}^N C_1^i\left(\hat{q}_1^i, \theta_1^i\right) - \sum_{i=1}^N C_2^i\left(\hat{q}_2^i, \theta_2^i\right)\right].\end{aligned} \tag{8}$$

Substituting each firm's response to the price and quantity orders, respectively, we obtain the following expression for the relative advantage of prices over tradable quantities given perfect firm-level certainty:[10]

$$\Delta = \mathrm{E}\left[\frac{1}{4}\left(\frac{1}{\sum_{i=1}^N \frac{1}{C_i''}} - B''\right)\left(\sum_{i=1}^N \frac{\theta_1^i + \theta_2^i}{C_i''}\right)^2\right]. \tag{9}$$

The full derivation of this expression is provided in appendix section A.3.

10. We can compare this baseline welfare expression directly to similar derivations in existing literature. For example, we can immediately compare this expression to the result in Weitzman (2020, eq. [51]) for the comparative advantage of fixed prices over time-flexible quantities with perfect information for a single representative firm. Our expression is also consistent with the relative advantage of prices over tradable quantities in Williams (2002, eq. [31]), given our assumption that abatement from different firms is perfectly substitutable in the benefits function and if we were to further assume that abatement cost shocks are independent across firms.

While this expression maintains some of the standard logic about comparing the relative slopes of the marginal cost and marginal benefit functions, note that we cannot evaluate the costs associated with the quantity order for each firm separately, since the quantity produced by each firm depends on the shocks to all other firms' marginal abatement cost functions via the market-clearing price. Instead we must compare the slope of marginal benefits to an expression that combines the slopes of all marginal costs. Given this baseline result, we now proceed with relaxing the strong assumption that firms have perfect information about all shocks before making any compliance decisions.

2. FIRM UNCERTAINTY OVER RIVALS' ABATEMENT COSTS

In introducing firm-level uncertainty, we distinguish between "idiosyncratic uncertainty" facing specific firms and "systemic uncertainty" facing all firms in the industry. Our goal is to differentiate between the impact of uncertainty on the allocation of abatement across firms and the impact on the allocation of aggregate abatement across periods. We first derive the impact of firm-level uncertainty over rivals' abatement cost shocks and then show how results are qualitatively similar for firm-level uncertainty over the impact of overlapping policies.

Because the market-clearing price associated with the regulator's aggregate quantity order (emissions cap) depends on shocks to the marginal abatement cost functions of all firms, a given firm may not know this price with certainty when making its abatement decisions. Instead, we assume that before the first compliance period, each firm i observes its current abatement cost shock θ_1^i but none of the other abatement cost shocks in the market.[11] Given its own realized shock and industry-wide common priors over the joint distribution of all cost shocks, each firm i forms a posterior expectation over its rivals'

11. While it is possible that financial instruments would help regulated firms to mitigate uncertainty associated with volatile allowance and credit prices, evidence on hedging decisions more generally suggests that firms are likely to hedge incompletely, if at all. In studying hedging of input fuel prices by US airlines, Rampini et al. (2014) find that the airlines in their sample hedge only 20% of expected next-year jet fuel expenses—despite the fact that financial instruments are widely available in this market and jet fuel represents a substantial and highly volatile operating expense for these firms. The authors attribute this imperfect hedging partly to firm financial constraints, which would certainly be relevant in our setting. Moreover, firms may also be unable to hedge their full exposure to uncertain allowance prices since the total quantity of allowances demanded depends on both the uncertain future price and potential additional uncertainties around future abatement cost. Finally, since allowance and credit trading markets are created virtually overnight through regulation, there is considerable uncertainty associated with the start-up of these markets which may reduce the availability of financial instruments in their early phases.

current cost shocks θ_1^{-i} and all future cost shocks $\boldsymbol{\theta}_2$.[12] Based on its posterior expectations, each firm i then formulates a forecast for the cost-effective price:[13]

$$\begin{aligned} \mathrm{E}[\hat{p}_1(\hat{Q}, \boldsymbol{\theta_1}, \boldsymbol{\theta_2})|\theta_1^i] &= C' + \frac{\mathrm{E}\left[\sum_{j=1}^{N} \frac{\theta_1^j + \theta_2^j}{2C_j''} \,\middle|\, \theta_1^i\right]}{\sum_{j=1}^{N} \frac{1}{C_j''}} \\ &= C' + \frac{\frac{\theta_1^i}{2C_i} + \mathrm{E}\left[\sum_{j\neq i}^{N-1} \frac{\theta_1^j}{2C_j''} + \sum_{j=1}^{N} \frac{\theta_2^j}{2C_j''} \,\middle|\, \theta_1^i\right]}{\sum_{j=1}^{N} \frac{1}{C_j''}}. \end{aligned} \tag{10}$$

Firm i will then set its marginal abatement cost equal to this forecast of the cost-effective price. In real-world emissions trading markets, most compliance options available to firms require time-to-build or other up-front (and often irreversible) investments, motivating our assumption that firms cannot adjust their abatement quantities after the realized market-clearing price is observed.[14]

To simplify notation going forward, we suppress the dependence of $\hat{p}_t$ on $\hat{Q}$, $\boldsymbol{\theta}_1$, and $\boldsymbol{\theta}_2$ and define the difference between firm i's forecasted price and the cost-effective price as follows:

---

12. Note that this set-up is compatible with a Bayesian Nash equilibrium concept, in which firms choose their optimal actions taking the (expected) actions of their competitors as given, based on their respective information sets and common priors over the distribution of shocks in the market. Because we have not placed any restrictions on potential correlation between firm i's cost shock and those of specific rivals, this set-up still permits the firm's observation of θ_t^i to be more informative about certain firms' cost shocks than others (or to be more informative about current period shocks than shocks in future periods). For further discussion of firm behavior in emissions trading markets with a Bayesian Nash equilibrium concept, see Malueg and Yates (2009), building on Vives (2002).

13. Throughout this paper, we use the term "cost-effective price" to refer to the market-clearing allowance price that achieves the regulator's aggregate quantity limit at lowest cost. Firms generally do not have perfect foresight about what this cost-effective price will be.

14. In contrast to our modeling approach, Malueg and Yates (2009) assume that firms submit their entire abatement cost schedule to the regulator, who then settles an efficient market-clearing price, which is then communicated to all firms before they choose their actual abatement levels. Consequently, Malueg and Yates find no difference in firm behavior when rivals' cost shocks are known versus held as private information, as the regulator effectively resolves all firm-level uncertainty before (competitive) firms choose their optimal compliance strategies. Real-world compliance behavior more closely matches the model developed here in which firms enter the allowance market with a particular emissions quantity rather than a perfectly flexible emissions schedule.

$\mathrm{E}[\hat{p}_1|\theta_1^i] - \hat{p}_1 =$ forecast error

$$= \frac{\mathrm{E}\left[\sum_{j\neq i}^{N-1} \frac{\theta_1^j}{2C_j''} + \sum_{j=1}^{N} \frac{\theta_2^j}{2C_j''} \,\middle|\, \theta_1^i\right]}{\sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\sum_{j\neq i}^{N-1} \frac{\theta_1^j}{2C_j''} + \sum_{j=1}^{N} \frac{\theta_2^j}{2C_j''}}{\sum_{j=1}^{N} \frac{1}{C_j''}}. \quad (11)$$

The firm's quantity response becomes

$$q_1^i(\mathrm{E}[\hat{p}_1|\theta_1^i], \theta_1^i) = \frac{\hat{p}_1 + \text{forecast error} - C' - \theta_1^i}{C_i''} + \bar{q}_1^i. \quad (12)$$

Next, we decompose each firm's forecast error into two parts: ξ_1 is the component of the forecast error common to all firms (systemic uncertainty), and ϵ_1^i is the component of the forecast error specific to firm i (idiosyncratic uncertainty). By construction, ξ_1 affects industry-wide abatement levels across periods, while ϵ_1^i affects only the distribution of abatement across firms within the initial compliance period.[15] That is, the aggregate quantity in the first period may be higher or lower than what is intertemporally optimal, depending on the sign of ξ_1; likewise, firm *i*'s share of that industry-wide abatement may be higher or lower than what is cross-sectionally optimal, depending on the sign of ϵ_1^i.

We can rewrite the firm's quantity response as

$$q_1^i(\hat{p}_1 + \xi_1 + \epsilon_1^i, \theta_1^i) = \frac{\hat{p}_1 + \xi_1 + \epsilon_1^i - C' - \theta_1^i}{C_i''} + \bar{q}_1^i. \quad (13)$$

Aggregate first-period abatement is therefore given by

$$Q_1 = \sum_{i=1}^{N} \frac{\theta_2^i - \theta_1^i}{2C_i''} + \sum_{i=1}^{N} \frac{\xi_1}{C_i''} + \bar{Q}_1.$$

Consequently, after accounting for the impact of firm-level uncertainty, aggregate abatement differs from the cost-effective level derived in the previous section. To ensure that the regulator's overall quantity limit is still met by the end of the final regulatory period, the aggregate second-period quantity must also shift upward or downward from the cost-effective level to compensate, and the market-clearing price in the second period adjusts accordingly. Full details of this derivation are provided in appendix section A.4.

By contrast, under a tax, the price is set by regulation and does not depend on private information about other firms' marginal abatement costs or about future marginal

15. We construct ϵ_1^i such that $\Sigma_{i=1}^{N}(\epsilon_1^i/C_i'') = 0$, which ensures that firms' idiosyncratic expectation errors collectively cancel with each other in determining the per-period abatement quantity. By contrast, the net impact of firms' systemic expectation errors on industry-wide abatement in the first period is given by $\Sigma_{i=1}^{N}(\xi_1/C_i'') \neq 0$.

abatement costs. We maintain our earlier assumption that the regulated price is known with certainty to all firms under a price-based policy, to maintain the consistent assumption across both price and quantity instruments that uncertainty around policy design and stringency has already been resolved. Therefore, firms' responses to a price order do not change from the version derived in section 1.

Based on this set-up, we re-derive the relative advantage of prices over quantities, allowing for the presence of both idiosyncratic and systemic forecast errors under quantity-based regulation but holding constant the net benefits of price-based regulation. Our welfare expression now becomes[16]

$$\Delta = \mathrm{E}\Bigg[\underbrace{\frac{1}{4}\left(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}} - B''\right)\left(\sum_{i=1}^{N}\frac{\theta_1^i + \theta_2^i}{C_i''}\right)^2}_{\text{Original Expression}} + \underbrace{\frac{1}{2}\left(\sum_{i=1}^{N}\frac{\epsilon_1^{i2}}{C_i''}\right)}_{\text{Cost Inefficiencies from Idiosyncratic Errors} \geq 0} + \underbrace{\left(\sum_{i=1}^{N}\frac{\xi_1^2}{C_i''}\right)}_{\text{Cost Inefficiencies from Systemic Errors} \geq 0} + \underbrace{B''\left(\sum_{i=1}^{N}\frac{\xi_1}{C_i''}\right)^2 + 2B''\left(\sum_{i=1}^{N}\frac{\xi_1}{C_i''}\right)\left(\sum_{i=1}^{N}\frac{\theta_2^i - \theta_1^i}{2C_i''}\right)}_{\text{Additional Benefit Term} \geq \text{ or } \leq 0}\Bigg]. \tag{14}$$

We parse equation (14) in stages. The first term replicates the "relative slopes" comparison from our baseline derivation when we assumed perfect firm-level certainty. The second and third terms together indicate that firm-level forecast errors under quantity regulation create an additional advantage of price instruments relative to quantity instruments, with the relative advantage increasing in the variance of both the idiosyncratic error terms and the systemic error terms.[17] To interpret this finding, in light of Weitzman's original result, we recognize that the regulator is no longer imposing quantity orders directly on individual firms. Instead, as we have noted, the regulator imposes an aggregate quantity order (the emissions cap), which a market mechanism then translates into individual quantity orders. Given idiosyncratic firm-level expectation errors in a single compliance period, the same relative advantages of price and quantity instruments still exist, but we must also consider the possibility that the aggregate quantity order is not distributed in the least-cost manner across firms. Likewise, given systemic firm-level expectation errors, we must also consider cost-inefficient distributions of aggregate abatement across

16. This derivation (and those in subsequent sections) assumes that the regulator has no foresight about the extent of firm forecast errors. That is, $\mathrm{E}[\xi_1] = 0$ and $\mathrm{E}[\epsilon_1^i] = 0$ for all i. This follows immediately from the law of iterated expectations.

17. The impact of firm-level forecast errors is also decreasing in the slope of the marginal abatement cost function, C_i''. The intuition here is that individual firm quantities are less responsive to (expected) price when the slope of the marginal abatement function is large. (Recall that the firm's quantity response function is $q_t^i = \bar{q}_t^i + [(p_t - \theta_t^i - C')/C_i'']$.)

compliance periods.[18] Both types of abatement reallocation push the regulator to prefer the price-based instrument over the quantity-based instrument with banking and borrowing.

The final term in equation (14) captures the idea that systemic forecast errors influence the extent to which the marginal benefits of abatement are equated across time periods. Tradable allowance programs ordinarily deviate from perfect benefit smoothing in response to cost shocks in specific periods. To the extent that the systemic forecast errors dampen firm responses to these period-specific cost shocks, quantity variability may decrease and therefore benefit smoothing may improve; to the extent that these forecast errors amplify responses to cost shocks, however, benefit smoothing may be further reduced. Consequently, this second set of additional terms has an ambiguous impact on the regulator's preference for prices over quantities with banking and borrowing, depending on the overall impact on benefit smoothing. The overall impact, in turn, depends on the sign and magnitude of the correlation between $\Sigma_{i=1}^{N}(\xi_1/C_i'')$ and $\Sigma_{i=1}^{N}[(\theta_2^i - \theta_1^i)/2C_i'']$. We note, however, that this additional consideration would disappear given a stock pollutant, such as greenhouse gases, where the timing of production is not a first-order concern over short-time horizons—leaving only the additional cost inefficiencies due to forecast errors under quantity-based policies.[19]

Finally, we also note that in this derivation, as in subsequent derivations featuring firm-level uncertainty, we assume that firms incur forecast errors in the first compliance period but not the second compliance period.[20] While stylized, this assumption is necessary to ensure that the regulator's overall quantity order is still met exactly at the

---

18. Both Karp (2019) and Feng and Zhao (2006) also relax the assumption in Weitzman (2020) that the representative firm has perfect information about second-period cost shocks. Karp (2019) assumes instead rational expectations over future shocks which evolve according to an AR-1 process. In the derivation presented here, we instead model heterogeneous firms with different expectations about market-clearing prices. This derivation is still compatible with rational expectations if we require each firm to have rational expectations over the cost shocks of other firms in the market, given the revelation of its own cost shock. However, our derivation depends on heterogeneous expectations about the allowance price. Feng and Zhao (2006) show how the welfare gains from intertemporal trading decrease as firms' knowledge of current shocks becomes less informative about future shocks, thereby decreasing the extent of information asymmetry between firms and the regulator.

19. In the case of greenhouse gases, B'' may be close to zero given linear damages (Burke et al. 2015). This term also drops out for pure stock pollutants, modeled in the spirit of Gerlagh and Heijmans (2018); this derivation is available from the authors upon request.

20. Here we emphasize, however, that although firm-level uncertainty is restricted to the first period, the systemic forecast errors continue to create cost inefficiencies in later periods due to the intertemporal linkages of cap and trade. By contrast, the cost inefficiencies of idiosyncratic errors are limited to the period in which the uncertainty occurs. This distinction illuminates why we multiply the idiosyncratic forecast error term in eq. (14) by 1/2, but not the systemic error term.

end of the specified banking and borrowing horizon, which in turn allows us to apply the consistent assumption across both price and quantity instruments that the regulator's policy target is met exactly. In real-world emissions trading markets, of course, the firm-level uncertainty that we describe here is likely to persist across all compliance periods, and we rarely observe a terminal date of a credit trading program with banking or borrowing; instead credits are generally allowed to cross over into the next phase of the program, often at an adjusted trading ratio.

3. FIRM UNCERTAINTY OVER THE IMPACT OF OVERLAPPING POLICIES

In addition to uncertainty about cost shocks, firms may also make forecast errors around the market-clearing price due to uncertainty about the impact of overlapping policies. As discussed previously, under many market-based environmental policies, a subset of regulated firms are also subject to additional standards from overlapping policies, which are often more stringent or more prescriptive than the price or quantity instrument. For firms where the prescribed technology is not the lowest-cost option for reducing emissions, the existence of this second set of standards raises the total costs for meeting the requirements of the overlapping price- or quantity-based policy.[21] Under a price instrument, the inefficiencies due to overlapping policies are limited to the subset of firms or business units that are subject to the more prescriptive regulation. Under an allowance or credit trading program, by contrast, these cost-inefficiencies affect both firms covered by the more prescriptive regulation and noncovered firms, by affecting the market-clearing price for allowances or credits. Insofar as firms are uncertain about the impact of these overlapping policies on the allowance price, we again observe expected welfare losses due to forecast errors for a tradable quantity instrument relative to a price instrument.

To illustrate, consider the same two-period market-based policy modeled above, but now assume that a subset R firms are subject to an additional regulatory standard in each compliance period. The remaining M firms are subject only to the market-based policy (where $R + M = N$, the total number of firms in the market). Assume that the supplementary regulation causes firms to abate at the level $(q_R)_t^i$, where $\Sigma_{i=1}^R (q_R)_t^i - \Sigma_{i=1}^R \bar{q}_t^i = \kappa \geq 0$. That is, the supplementary regulation results in a total quantity of abatement from this subset of firms that is higher than their ex ante efficient quantity.

Under a price regime, the M noncovered firms behave no differently and continue to set $\tilde{q}_1^i = \bar{q}_1^i - (\theta_1^i / C_i'')$ (and analogously for $\tilde{q}_2^i$). However, under tradable quantities, the aggregate quantity target still determines overall abatement in the market, so the

21. Insofar as there are fixed costs to participating in allowance or credit trading markets, some firms may elect to treat their allowance allocations as if they represented command-and-control regulation. This behavior may produce similar impacts as overlapping policies, where some subset of firms do not abate at the efficient level. The results described in this section would apply to this scenario as well.

market-clearing price must adjust to account for the excess κ compliance produced by the firms subject to the supplementary regulation. The market-clearing price will therefore adjust to

$$\hat{p}_1^{\kappa} = \hat{p}_2^{\kappa} = C' + \frac{\sum_{i=1}^{M} \frac{\theta_1^i + \theta_2^i}{2C_i''}}{\sum_{i=1}^{M} \frac{1}{C_i''}} - \frac{\kappa}{\sum_{i=1}^{M} \frac{1}{C_i''}},$$

and under perfect certainty the noncovered firms will each respond with quantity $q_1^i(\hat{p}_1^{\kappa}, \theta_1^i) = [(\hat{p}^{\kappa} - C' - \theta_1^i)/C_i''] + \bar{q}_1^i$ (and analogously for $q_2^i(\hat{p}_2^{\kappa}, \theta_2^i)$). When we compare the expected welfare of prices versus tradable quantities under this scenario, we obtain the following relation:

$$\Delta = E\left[\underbrace{\frac{1}{4}\left(\frac{1}{\sum_{i=1}^{M} \frac{1}{C_i''}} - B''\right)\left(\sum_{i=1}^{M} \frac{\theta_1^i + \theta_2^i}{C_i''}\right)^2}_{\text{Standard Comparison}} + \underbrace{\left(\frac{1}{\sum_{i=1}^{M} \frac{1}{C_i''}} - B''\right)\kappa^2}_{\text{Impact of Overlapping Policy}}\right].$$

As we see here, the impact of the overlapping policy on the expected welfare of a price instrument relative to a tradable quantity instrument reduces to the standard "relative slopes" comparison. The intuition is straightforward: the overlapping policy causes a subset of firms to overcomply relative to what would be most cost effective; the overcompliance for this subset of firms occurs under either a tax or a cap-and-trade program. Under a tradable quantity regime, the market-clearing price adjusts downward such that the aggregate target is still met, resulting in lower abatement among the remaining firms, relative to their cost-effective level. Under a price instrument, the remaining firms still abate at the cost-effective level, but there is excess compliance across the full set of firms. The trade-off between these two sources of inefficiency depends on the relative slopes of the marginal benefit and marginal cost functions.

This comparison changes when we introduce the prospect of firm-level uncertainty. Some firms not subject to the supplementary regulation may not know the full extent of overcompliance among doubly regulated firms. Under a price-based policy, this uncertainty about the impact of overlapping policies does not affect the firm's own compliance decision, so abatement quantities do not change. Under tradable quantities, however, each firm i forms some posterior expectation over the extent of excess compliance, which then informs the firm's forecast of the market-clearing price.[22] The firm then chooses its optimal abatement level given its forecasted market-clearing

22. Here we return to assuming perfect certainty over cost shocks to isolate the impact of uncertainty over overlapping policies. It is also possible, of course, to use our model set-up to explore interactions between these two sources of uncertainty.

price. Again let $\xi_1 + \epsilon_1^i$ represent the difference between the expected price and the cost-effective price, to capture systemic and idiosyncratic forecast errors, respectively. The expected welfare of a price-based policy relative to tradable allowances or credits is then given by

$$\Delta = \mathrm{E}\Bigg[\underbrace{\left(\frac{1}{\sum_{i=1}^{M}\frac{1}{C_i''}} - B''\right)\left(\sum_{i=1}^{M}\frac{\theta_1^i + \theta_2^i}{2C_i''}\right)^2}_{\text{Standard Comparison}} + \underbrace{\left(\frac{1}{\sum_{i=1}^{M}\frac{1}{C_i''}} - B''\right)\kappa^2}_{\text{Impact of Overlapping Policy}}$$

$$+ \underbrace{\frac{1}{2}\left(\sum_{i=1}^{M}\frac{(\epsilon_1^i)^2}{C_i''}\right) + \left(\sum_{i=1}^{M}\frac{\xi_1^2}{C_i''}\right)}_{\text{Cost Inefficiencies from Forecast Errors} \geq 0} + \underbrace{B''\left(\sum_{i=1}^{M}\frac{\xi_1}{C_i''}\right)^2 + 2B''\left(\sum_{i=1}^{M}\frac{\xi_1}{C_i''}\right)\left(\sum_{i=1}^{M}\frac{\theta_1^i - \theta_2^i}{2C_i''}\right)}_{\text{Additional Benefit Term} \geq \text{ or } \leq 0}\Bigg]. \tag{15}$$

Mathematically, the impact of uncertainty around overlapping policies is almost identical to the results derived above for uncertainty around cost shocks; the full derivation is presented in appendix section A.5. Once again, we recover our initial prices versus tradable quantities result, here for the overlapping policies scenario, as well as additional terms that capture the cost inefficiencies associated with both idiosyncratic and systemic firm-level forecast errors. All else equal, these cost inefficiencies again push the regulator to prefer prices over tradable quantities. However, as before, we also have a term with ambiguous sign that depends on whether the forecast errors contribute to more or less benefit smoothing across periods, which in turn depends on the correlation between aggregate cost shocks and systemic forecast errors.

4. UNCERTAINTY OVER OWN ABATEMENT COSTS

Thus far, we have focused on several ways in which firms may face uncertainty over the market-clearing price in allowance trading markets. As we have emphasized, these uncertainties introduce asymmetric cost inefficiencies for quantity-based instruments relative to price-based instruments, since for the latter, regulators set and announce the price of pollution in advance. Yet we have ignored other potential sources of uncertainty that may interact with uncertainty over allowance prices. In practice, these other sources of uncertainty may magnify or mitigate the baseline cost inefficiencies modeled above.

To illustrate, we relax our earlier assumption that firms know their own abatement cost shocks with certainty. Under this version of the model, firms must choose their compliance strategies before they have full information about either their own abatement costs during the compliance period or the market-clearing price for allowances. For example, output prices in the energy market often influence both own abatement costs and the market-clearing price for allowances, and firms may not have full information about energy prices when determining abatement strategies (and often fail to hedge completely their exposure to these prices).

Consequently, under a quantity regime, each firm i chooses its optimal quantity of abatement given its information set $\mathcal{I}_t^i$ and resulting expectations over own costs and the market-clearing price for allowances:

$$q_1^i(\mathrm{E}[\hat{p}_1|\mathcal{I}_1^i], \mathrm{E}[\theta_1^i|\mathcal{I}_1^i]) = \bar{q}_1^i + \frac{\mathrm{E}\left[\sum_{j=1}^{N} \frac{\theta_1^j + \theta_2^j}{2C_j''} \,\middle|\, \mathcal{I}_1^i\right]}{C_i'' \sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\mathrm{E}[\theta_1^i|\mathcal{I}_1^i]}{C_i''}.$$

Likewise, under a price regime, firm i's optimal abatement quantity is now a function of the regulated price and its expectation of its own abatement costs, given its current information set:

$$q_1^i(\tilde{p}_1, \mathrm{E}[\theta_1^i|\mathcal{I}_1^i]) = \bar{q}_1^i - \frac{\mathrm{E}[\theta_1^i|\mathcal{I}_1^i]}{C_i''}.$$

For this more general model, we define the following firm-level forecast errors over own cost shocks and the market-clearing price, respectively:

$$\text{own cost forecast errors:}\ \ \epsilon_1^i(\theta) + \xi_1(\theta) = \mathrm{E}[\theta_1^i|\mathcal{I}_1^i] - \theta_1^i$$

$$\text{price forecast errors:}\ \ \epsilon_1^i(\hat{p}) + \xi_1(\hat{p}) = \frac{\mathrm{E}\left[\sum_{j=1}^{N} \frac{\theta_1^j + \theta_2^j}{2C_j''} \,\middle|\, \mathcal{I}_1^i\right]}{\sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\sum_{j=1}^{N} \frac{\theta_1^j + \theta_2^j}{2C_j''}}{\sum_{j=1}^{N} \frac{1}{C_j''}}.$$

Therefore, we can rewrite the above quantity responses in terms of these two sets of forecast errors. Under a quantity instrument, firm i's quantity response becomes

$$q_1^i(\mathrm{E}[\hat{p}_1 \mid \mathcal{I}_1^i], \mathrm{E}[\theta_1^i \mid \mathcal{I}_1^i]) = \bar{q}_1^i + \frac{\sum_{j=1}^{N} \frac{\theta_1^j + \theta_2^j}{2C_j''}}{C_i''} + \frac{\epsilon_1^i(\hat{p}) + \xi_1(\hat{p})}{C_i''} - \frac{\theta_1^i}{C_i''} - \frac{\epsilon_1^i(\theta) + \xi_1(\theta)}{C_i''}.$$

Under a price instrument, firm i's quantity response becomes

$$q_1^i(\tilde{p}_1, \mathrm{E}[\theta_1^i|\mathcal{I}_1^i]) = \bar{q}_1^i - \frac{\theta_1^i}{C_i''} - \frac{\epsilon_1^i(\theta) + \xi_1(\theta)}{C_i''}.$$

From these expressions, it is straightforward to see that the overall impact of firm-level uncertainty on each firm's quantity response—which then directly translates into the market's overall cost-effectiveness—may be higher or lower under quantity instruments compared to price instruments. The overall effect depends on the relative magnitude of firms' uncertainty over own cost shocks versus their uncertainty over the market-clearing allowance price and whether these two sources of firm-level uncertainty tend to move in the same or opposite directions. To formalize this intuition, we present the full expression for the relative advantage of prices over quantities when firms face uncertainty over own abatement costs as well as over allowance prices, with the full derivation in appendix section A.6:

$$
\begin{aligned}
\Delta = & \underbrace{\frac{1}{4}\left(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}} - B''\right)\left(\sum_{i=1}^{N}\frac{\theta_1^i + \theta_2^i}{C_i''}\right)^2}_{\text{Original Expression}} + \underbrace{\frac{1}{2}\left(\sum_{i=1}^{N}\frac{\epsilon_1^i(\hat{p})^2}{C_i''}\right) + \left(\sum_{i=1}^{N}\frac{\xi_1(\hat{p})^2}{C_i''}\right) + \frac{1}{2}\left(\sum_{i=1}^{N}\frac{\xi_1(\theta)^2}{C_i''}\right)}_{\text{Cost Inefficiencies from Idiosyncratic and Systemic Errors} \geq 0} \\
& - \underbrace{\left(\sum_{i=1}^{N}\frac{\epsilon_1^i(\hat{p})\epsilon_1^i(\theta)}{C_i''}\right)}_{\text{Correlation in Idiosyncratic Errors} \geq \text{ or } \leq 0} - \underbrace{2\left(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}} + B''\right)\left(\sum_{i=1}^{N}\frac{1}{C_i''}\right)^2(\xi_1(\hat{p})\xi_1(\theta))}_{\text{Correlation in Systemic Errors} \geq \text{ or } \leq 0} \\
& + \underbrace{B''\left(\sum_{i=1}^{N}\frac{\xi_1(\hat{p})^2}{C_i''}\right) + 2B''\left(\sum_{i=1}^{N}\frac{\xi_1(\hat{p})}{C_i''}\right)\left(\sum_{i=1}^{N}\frac{\theta_2^i - \theta_1^i}{2C_i''}\right)}_{\text{Benefit Smoothing for Price Errors} \geq \text{ or } \leq 0} \\
& + \underbrace{\frac{1}{2}B''\sum_{i=1}^{N}\left(\frac{\xi_1(\theta)^2}{C_i''}\right) - B''\left(\sum_{i=1}^{N}\frac{\theta_2^i}{C_i''}\right)\left(\sum_{i=1}^{N}\frac{\xi_1(\theta)}{C_i''}\right)}_{\text{Benefit Smoothing for Own Cost Errors} \geq \text{ or } \leq 0}.
\end{aligned}
\tag{16}
$$

We tackle each of these terms in succession. As in all previous derivations, we have a standard comparison of the relative slopes of the benefit function and firms' cost functions. Also as before, we have terms capturing the cost inefficiencies associated with firm uncertainty, whether over prices or own abatement costs; these cost inefficiencies push the regulator to prefer prices over tradable quantities, all else equal. Note that inefficiencies arising from own cost forecast errors drop out in the first period, as they appear across both price and quantity regimes; however, the impact of these errors persists into the second period under quantities but not prices, due to the intertemporal linkages of cap and trade with banking and borrowing.

Next, given the multiple sources of firm-level uncertainty affecting tradable quantities markets, we have new terms capturing the correlation between uncertainty over prices and uncertainty over own abatement costs. Whenever these two sources of uncertainty are positively correlated across firms, the regulator is pushed to prefer tradable quantities over prices, as positively correlated errors mitigate the cost inefficiencies associated with each error individually. That is, forecast errors that cause firms to overestimate the cost-effective market-clearing price induce firms to overabate relative to the efficient level, while forecast errors that cause firms to overestimate their own abatement cost shocks induce them to underabate. Therefore, these two sources of uncertainty may be, to varying degrees, offsetting under a tradable quantities regime. By contrast, under a price regime, firms only face uncertainty over own abatement costs, so there is no opportunity for mitigating their forecast errors.

Finally, we must now consider the impact of both types of uncertainty on benefit smoothing. For both price- and quantity-based policies, forecast errors could increase or decrease quantity variability across periods and thereby increase or decrease benefit smoothing, depending on the correlation between firms' errors and their cost shocks. Once again, however, the total effect on benefit smoothing is partially offset for forecast errors around own cost shocks, because they appear under both price and quantity regimes.

Considering all terms together, we see that interactions between firm-level uncertainty over the market-clearing price and over own abatement cost shocks add further nuance to our earlier result that firm-level uncertainty pushes regulators to prefer prices over tradable quantities on cost effectiveness grounds. In the case of greenhouse gas abatement, where the slope of the marginal benefit function is believed to be close to zero and therefore considerations around benefit smoothing are less germane, the final impact of these two sources of firm-level uncertainty depends on the extent of positive or negative correlation across firms' various forecast errors. Where forecast errors over the market-clearing price stem from uncertainty over rivals' cost shocks, it seems reasonable that these errors would be positively correlated with forecast errors around own cost shocks, which would mitigate (but not necessarily eliminate) some of the cost-effectiveness concerns in these markets. However, it is less obvious whether forecast errors around the impact of overlapping policies would be positively or negatively correlated with forecast errors around own costs.

5. POLICY RESPONSES TO FIRM-LEVEL UNCERTAINTY

We have established that firm-level uncertainty may have important implications for the cost-effectiveness of allowance and credit trading programs, by modeling these additional uncertainties in a Weitzman-style framework that has traditionally focused on the regulator's uncertainty. However, political economy considerations may continue to push policy makers to favor quantity instruments irrespective of any welfare advantage of price instruments (Stavins 2020). For this reason, it is worthwhile to consider what policy design tools may be available to reduce—even if not eliminate—the impacts of this firm-level uncertainty. We discuss here the potential role for price ceilings and price floors ("price collars"), as one example.

Under a tradable quantities program, a price ceiling serves as a maximum price above which the market-clearing price for allowances or credits is not allowed to rise.[23] The

23. In the model that follows, we focus on a "hard" price ceiling where the regulator commits to maintaining a certain maximum price, as opposed to a "soft" price, where the regulator commits only to releasing a finite allowance reserve if the market-clearing price rises above some level.

regulator commits to selling additional allowances whenever the price ceiling is reached, charging regulated entities this maximum price but no more. The effect is to relax the overall quantity constraint to avoid large increases in firms' compliance costs. A price floor works analogously, where the regulator commits to buying allowances at a specified minimum price whenever needed, or specifies a reserve price in allowance auctions. Taken together, the price ceiling and price floor represent a hybrid policy between a pure quantity-based instrument and a pure price-based instrument.[24]

In appendix section A.7, we derive a Weitzman-style expression for evaluating the welfare consequences of introducing a price collar in a cap-and-trade market. To establish a baseline, we first evaluate the impact of price collars in the absence of firm-level uncertainty. In general, the impact of the price collar falls into one of three cases, depending on the relationship between the price ceiling or floor, the cost-effective price given the aggregate quantity order (emission cap), and the ex post optimal price (first-best price) given the realization of shocks in the market. First, whenever the slope of the marginal cost function is steeper than the slope of the marginal benefit function, the price collar is welfare-improving relative to an unconstrained cap-and-trade program. In this case, a price-based policy is preferred to a quantity-based policy but may not be possible for political economy or other reasons; a price collar recreates some characteristics of a fixed price policy and therefore increases welfare under the tradable quantities regime. Second, whenever the price ceiling (floor) is higher (lower) than the ex post optimal price, the deadweight loss under the price ceiling (floor) is (weakly) smaller than if the price is permitted to float under the cap-and-trade program. Finally, when the price ceiling (floor) is lower (higher) than the ex post optimal price, the impact on deadweight loss is ambiguous and depends on the relative magnitude of the difference between the ex post optimal price and the price ceiling (floor) versus the difference between the optimal price and the cost-effective price from allowance trading. We illustrate these three scenarios for a market with a representative firm in appendix figure 1.

When we account for firm-level uncertainty, the welfare implications of introducing a price collar also depend on how this policy design affects firms' information sets as they make their compliance decisions. A price collar announced in advance means that firms know ex ante that the market-clearing price will not rise above the price ceiling or fall below the price floor; this additional information then affects the magnitude of firms' forecast errors. In the appendix, we extend our Weitzman-style expression for the welfare consequences of introducing price collars by also incorporating firm-level uncertainty over competitors' cost shocks. In this more general welfare expression, we obtain additional terms capturing the difference in the variance of firms' forecast errors with and without the price collar, as well as additional terms capturing changes in benefit smoothing across periods with and without the price collar. Without placing additional structure

24. Roberts and Spence (1976) and Weitzman (1978) have demonstrated that this type of hybrid policy may be welfare-enhancing relative to a pure quantity or price instrument.

on the correlation of firms' cost shocks, we cannot determine exactly how the variance of firm forecast errors changes under a price collar. However, we might expect that the variance would be lower under a price collar, as firms have greater certainty over the market-clearing price when the extremes of the price distribution are truncated. As a consequence of reducing the variance of forecast errors, the cost inefficiencies from firm-level uncertainty would also diminish. These potential reductions in cost inefficiencies should therefore be included in standard evaluations of price collar mechanisms.

6. CONCLUSION

In this paper, we examine the impact of firm-level uncertainty over allowance prices in cap-and-trade and tradable credit markets—a form of residual uncertainty which is inherent to this type of policy instrument. Motivated by numerous empirical illustrations of cap-and-trade and tradable quantity markets deviating from cost-effectiveness, we develop a theory model that elucidates the welfare implications of instrument choice with allowance and credit price forecast errors. Building on the Weitzman prices-versus-quantities framework, we highlight the additional costs associated with imperfect information about future market-clearing prices when modeling cap and trade as an "aggregate quantity order" that allocates quantities in a decentralized manner to firms. All else equal, the cost inefficiencies created by firm forecast errors over the cost-effective price would encourage the welfare-maximizing regulator to favor price instruments over tradable quantity instruments. These cost inefficiencies may be partly offset, however, when firms incur additional forecast errors around their own abatement costs when determining compliance strategies, and those errors around own costs are positively correlated with errors around the allowance price.

We focus our model on a uniformly mixed pollutant. While a special case, it addresses carbon dioxide and other greenhouse gas emissions, which represent the most pressing environmental policy challenge of the twenty-first century—and one that drew considerable attention from Marty Weitzman in his scholarship over the last several decades of his career. Given the scale of investment necessary to combat climate change—on the order of trillions of dollars in the coming decades—and the growing policy interest around the world in market-based instruments in abating greenhouse gas emissions, our work highlights another dimension of the instrument choice problem in promoting cost-effective and economically efficient emission reductions. We also show how variations in instrument design, such as price collars, may mitigate some, but not all, of the welfare costs associated with the residual uncertainty in tradable quantity instruments. We also discuss how intertemporal benefit smoothing for flow pollutants (or nonuniformly mixed pollutants) may also affect the regulator's choice of instrument.

Examining the empirical interactions between firm-level forecast errors, uncertainty over future abatement costs, and other shocks to pollution markets represents a fruitful direction for research. Many economists responded to the insights of the standard prices-versus-quantities framework by exploring its empirical implications, such as the estimation of the damage and abatement cost functions for a variety of pollutants, including carbon dioxide (e.g., Pizer 2002). In our previous work, we simulated the impacts of forecast errors on a hypothetical US cap-and-trade market and estimated that it would increase compliance costs 21% relative to an otherwise equivalent tax (Aldy and Armitage 2020). Further examination of the economic fundamentals driving allowance and credit prices, and rigorous evaluations of cost-effectiveness anomalies in these markets could further enhance understanding of instrument choice and design. In future research, we plan to build on the static abatement decisions analyzed here, to understand the impact of these various shocks on long-lived firm investment, including R&D.

REFERENCES

Aldy, Joseph E., and Sarah Armitage. 2020. The cost-effectiveness implications of carbon price certainty. *AEA Papers and Proceedings* 110:113–18.

Aldy, Joseph E., and Gianfranco Gianfrate. 2019. Future-proof your climate strategy: Smart companies are putting their own price on carbon. *Harvard Business Review* 97 (3): 86–97.

Aldy, Joseph E., and W. Kip Viscusi. 2014. Environmental risk and uncertainty. In *Handbook of the economics of risk and uncertainty*, ed. Mark J. Machina and W. Kip Viscusi, 1:601–49. Amsterdam: Elsevier.

Allen, Treb. 2014. Information frictions in trade. *Econometrica* 82 (6): 2041–83.

Arrow, Kenneth J., and Anthony C. Fisher. 1974. Environmental preservation, uncertainty, and irreversibility. *Quarterly Journal of Economics* 88 (2): 312–19.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis. 2016. Measuring economic policy uncertainty. *Quarterly Journal of Economics* 131 (4): 1593–1636.

Borenstein, Severin, James Bushnell, Frank A. Wolak, and Matthew Zaragoza-Watkins. 2019. Expecting the unexpected: Emissions uncertainty and environmental market design. *American Economic Review* 109 (11): 3953–77.

Burdett, Kenneth, and Kenneth L. Judd. 1983. Equilibrium price dispersion. *Econometrica* 51 (4): 955–69.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015. Global non-linear effect of temperature on economic production. *Nature* 527 (7577): 235–39.

Callan, Scott J., and Janet M. Thomas. 2013. *Environmental economics and management: Theory, policy, and applications*. Fort Worth, TX: Cengage Learning.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L. Palmer. 2000. Sulfur dioxide control by electric utilities: What are the gains from trade? *Journal of Political Economy* 108 (6): 1292–1326.

Carlson, John A., and R. Preston McAfee. 1983. Discrete equilibrium price dispersion. *Journal of Political Economy* 91 (3): 480–93.

Cicala, Steve. 2015. When does regulation distort costs? Lessons from fuel procurement in US electricity generation. *American Economic Review* 105 (1): 411–44.

Clarke, Leon, Kejun Jiang, Keigo Akimoto, Mustafa Babiker, Geoffrey Blanford, Karen Fisher-Vanden, J.-C. Hourcade, et al. 2014. Assessing transformation pathways. In *AR5 climate change 2014: Mitigation of climate change*, 413–510. New York: Cambridge University Press.

Council of Economic Advisers. 2000. *Economic report of the president transmitted to the Congress*. Washington, DC: US Government Printing Office.

———. 2010. *Economic report of the president transmitted to the Congress*. Washington, DC: US Government Printing Office.

Dana, James D., Jr. 1994. Learning in an equilibrium search model. *International Economic Review* 35 (3): 745–71.

Ecosystem Marketplace. 2021. Markets in motion: State of the voluntary carbon markets, installment 1. Technical report. Ecosystem Marketplace.

Energy Information Administration. 2009. Energy market and economic impacts of H.R. 2454, the American Clean Energy and Security Act of 2009. Department of Energy, Washington, DC.

Environmental Protection Agency. 2015. *Fact sheet: Overview of the Clean Power Plan*. Environmental Protection Agency, Washington, DC.

Feng, Hongli, and Jinhua Zhao. 2006. Alternative intertemporal permit trading regimes with stochastic abatement costs. *Resource and Energy Economics* 28 (1): 24–40.

Field, Barry C., and Martha K. Field. 2016. *Environmental economics: An introduction*. 7th ed. New York: McGraw-Hill Education.

Fowlie, Meredith. 2010. Emissions trading, electricity restructuring, and investment in pollution abatement. *American Economic Review* 100 (3): 837–69.

Fowlie, Meredith, Stephen P. Holland, and Erin T. Mansur. 2012. What do emissions markets deliver and to whom? Evidence from Southern California's NOx trading program. *American Economic Review* 102 (2): 965–93.

Gerlagh, Reyer, and Roweno J. R. K. Heijmans. 2018. Regulating stock externalities. Working paper.

Hanley, Nick, Jason Shogren, and Ben White. 2019. *Introduction to environmental economics*. Oxford: Oxford University Press.

Harris, Jonathan M., and Brian Roach. 2017. *Environmental and natural resource economics: A contemporary approach*. Armonk, NY: Routledge.

Hassett, Kevin A., and Gilbert E. Metcalf. 1999. Investment with uncertain tax policy: Does random tax policy discourage investment? *Economic Journal* 109:372–93.

Heutel, Garth. 2020. Bankability and information in pollution policy. *Journal of the Association of Environmental and Resource Economists* 7 (4): 779–99.

Hoel, Michael, and Larry Karp. 2002. Taxes versus quotas for a stock pollutant. *Resource and Energy Economics* 24:367–84.

Karp, Larry S. 2019. Welfare ranking of emissions taxes, cap and trade, and banking and borrowing. Working paper.

Keohane, Nathaniel O., and Sheila M. Olmstead. 2016. *Markets and the environment*. Washington, DC: Island Press.

Kling, Catherine, and Jonathan Rubin. 1997. Bankable permits for the control of environmental pollution. *Journal of Public Economics* 64 (1): 101–15.

Kolstad, Ch. D. 2000. Environmental economics. Technical report, Oxford University.

Malueg, David, and Andrew Yates. 2009. Bilateral oligopoly, private information, and pollution permit markets. *Environmental and Resource Economics* 43 (4): 553–72.

Montgomery, W. David. 1972. Markets in licenses and efficient pollution control programs. *Journal of Economic Theory* 5 (3): 395–418.

Newell, Richard G., and William A. Pizer. 2003. Regulating stock externalities under uncertainty. *Journal of Environmental Economics and Management* 45:416–32.

Pizer, William A. 2002. Combining price and quantity controls to mitigate global climate change. *Journal of Public Economics* 85 (3): 409–34.

Pizer, William A., and Brian C. Prest. 2020. Prices versus quantities with policy updating. *Journal of the Association of Environmental and Resource Economists* 7 (3): 483–518.

Rampini, Adriano A., Amir Sufi, and S. Viswanathan. 2014. Dynamic risk management. *Journal of Financial Economics* 111:271–96.

Roberts, Marc J., and Michael Spence. 1976. Effluent charges and licenses under uncertainty. *Journal of Public Economics* 5 (3–4): 193–208.

Rodrik, Dani. 1991. Policy uncertainty and private investment in developing countries. *Journal of Development Economics* 36 (2): 229–42.

Rubin, Jonathan. 1996. A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics and Management* 31 (3): 269–86.

Salop, Steven, and Joseph Stiglitz. 1977. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *Review of Economic Studies* 44 (3): 493–510.

Santikarn, Marissa, Angela Naneu Churie Kallhauge, Mustafa Ozgur Bozcaga, Leyla Sattler, Michael Stephen Mccormick, Ana Ferran Torres, Darragh Conway, et al. 2021. State and trends of carbon pricing 2021. Technical report, World Bank, Washington, DC.

Stavins, Robert N. 2003. Experience with market-based environmental policy instruments. In *Handbook of environmental economics*, 1:355–435. Amsterdam: Elsevier.

———. 2020. The future of US carbon-pricing policy. *Environmental and energy policy and the economy* 1 (1): 8–64.

Tietenberg, Tom, and Lynne Lewis. 2018. *Environmental and natural resource economics*. Boston: Routledge.

Vives, Xavier. 2002. Private information, strategic behavior, and efficiency in Cournot markets. *RAND Journal of Economics* 33 (3): 361–76.

Weitzman, Martin L. 1974. Prices vs. quantities. *Review of Economic Studies* 41 (4): 477–91.

———. 1978. Optimal rewards for economic regulation. *American Economic Review* 68 (4): 683–91.

———. 2020. Prices or quantities can dominate banking and borrowing. *Scandinavian Journal of Economics* 122 (2): 437–63.

Williams, Roberton C. 2002. Prices vs. quantities vs. tradeable quantities. NBER Working Paper Series, no. 9283, National Bureau of Economic Research, Cambridge, MA.

Yates, Andrew. 2002. Decentralization in pollution permit markets. *Journal of Public Economic Theory* 4 (4): 641–60.

Yohe, Gary W. 1978. Towards a general comparison of price controls and quantity controls under uncertainty. *Review of Economic Studies* 45 (2): 229–38.