The Welfare Implications of Carbon Price Certainty

Joseph Aldy and Sarah Armitage*

Draft: February 15, 2020

Abstract

The functioning of real-world pollution markets suggests that firms face persistent price forecast errors in making abatement decisions. The residual uncertainty in allowance trading means that pollution markets may fail to deliver cost-effective abatement, in contrast to price-based policies where firms set marginal abatement cost equal to an emission tax. We develop a theoretical model of firm behavior in an allowance trading market that accounts for price uncertainty and dynamic investment in abatement. We show how the additional cost of forecast errors under quantity-based programs can be incorporated into a standard Weitzman-style analysis. Finally, we simulate the potential magnitude of forecast errors in cap-and-trade markets using parameters calibrated to historical and modeled climate policies. Future work will examine the interaction between allowance price uncertainty and abatement cost uncertainty.

JEL Codes: Q52, Q54, Q58

*Aldy: Harvard Kennedy School, Resources for the Future, National Bureau of Economic Research, and Center for Strategic and International Studies, joseph_aldy@hks.harvard.edu; Armitage: 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, Joshua Pfeffer, Robert Pindyck, and Jim Stock, as well as participants at the AERE Summer Conference, Northeast Environmental Economics Workshop, Harvard Kennedy School Energy Policy seminar, Allied Social Science Association Annual Conference, and Berkeley Climate Economics Workshop. We thank Alex Barron for providing data from the Stanford Energy Modeling Forum U.S. carbon tax exercise. Aldy acknowledges financial support from BP, RFF, 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 and the Joseph Crump Fellowship.

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

1 Introduction

Policy uncertainty significantly influences firms' investment decisions. Firms may opt to delay investment until resolving uncertainty about its returns – especially uncertainty related to a pending policy decision, such as over monetary policy, tax policy, and other issues that affect the volatility and uncertainty of a firm's economic context (Dixit and Pindyck, 1994; Arrow and Fisher, 1974; Bernanke, 1983; Rodrik, 1991; Hassett and Metcalf, 1999; Baker et al., 2016). A special case of policy uncertainty occurs when the government decides upon a policy framework, but residual uncertainty over returns to firm investment remain due to the inherent characteristics of policy implementation. For example, suppose that a regulator decides to price a technological externality, such as carbon dioxide emissions, through a cap-and-trade program instead of a carbon tax. The inherent uncertainty of tradable allowance prices under cap-and-trade exceeds that of the tax alternative. We examine the impact of this residual policy uncertainty on the relative efficiencies of price- versus quantity-based policies abating carbon dioxide emissions.

Much of the literature comparing cap-and-trade programs and emission taxes has focused on the regulator's information deficit (Weitzman, 1974). Firms typically know their marginal abatement costs with greater precision than the regulator, and firms may not have the incentive to reveal their true marginal abatement costs to the regulator (and to their competitors). However, under cap-and-trade programs, this information asymmetry affects not only the regulator, but also each firm with respect to other firms' marginal abatement costs. Consider the problem a firm faces in complying with an emission tax versus a cap-and-trade program. Under any policy instrument, firms must first resolve uncertainty surrounding policy design. Then, if the regulator chooses a tax, the firm learns the tax rate, identifies its abatement options, and complies with the policy by investing in abatement that equates marginal cost with the tax. If the regulator instead chooses a cap-and-trade program, the firm not only must identify its abatement options as in the tax case, but also must form expectations about the market-clearing price for emissions allowances which will then guide the firm's investment.

Our paper examines this residual uncertainty about allowance prices, inherent to quantity-based instruments such as cap-and-trade programs and tradable performance standards, which increases the risk that firms may not equate their marginal abatement costs and hence increase aggregate costs for any given emission goal. Firms may err in their allowance price forecasts and make investments that would appear to be optimal ex ante given their expectations, but, ex post, are recognized as having been too high or too low. Such forecast errors may reflect different expectations about (1) abatement technology costs; (2) economic output; (3) overlapping public policies that may restrict abatement decisions and influence the clearing price in allowance markets. Given the common role of banking in cap-and-trade programs, these expectations would need to reflect a dynamic assessment of these factors. This residual uncertainty coupled with irreversible investments can increase the welfare costs of choosing a quantity-based instrument, such as cap-and-trade, relative to a carbon tax. Given the potential scale of tradable allowance markets, such forecast errors under cap-and-trade could be economically significant.

We build a theoretical model that starts from the perspective of the firm choosing emission abatement investments in the presence of uncertainty over the price of carbon in a tradable allowance market. In contrast to many papers the literature in which firms can instantaneously adjust emission abatement to move up or down their firm-specific marginal abatement cost function as allowance prices evolve, our model includes irreversible investment, uncertainty, and intertemporal emission allowance trading. Our model first shows that price trajectories still follow a Hotelling rule in expectation, but the overall price process is determined by volatility as well as this expected drift. This volatility creates potentially significant forecast errors, as the realized trajectory of prices does not follow a smooth Hotelling trajectory. Instead, our model is consistent with what is observed in allowance markets in practice – a frequent jumping from one Hotelling trajectory to another as shocks to allowance prices are realized.

The fact that firms often make dynamic decisions to invest in long-lived abatement capital magnifies the impact of these forecast errors. By deriving the necessary optimality conditions for firm investment in abatement, we show that firms not only set their marginal abatement costs equal to the avoided allowance prices, as in standard models of abatement as a variable input, but also seek to smooth their rate of investment over time. Given uncertainty over allowance prices, the value of smoothing also depends on the extent of price volatility. This finding introduces the possibility that firms might err not only in their expectations of the realized allowance price, but also in their beliefs about the *overall price process*, as the volatility parameter depends on many different random processes, only some of which are observable to individual firms.

To examine the welfare impacts of these forecast errors, we develop a modified version of Weitzman's canonical prices versus quantities comparison. In our version, quantity orders are not imposed directly on individual firms, as in Weitzman (1974) and much of the subsequent literature, but are instead transmitted to firms through some market-clearing price, as in modern allowance and credit trading markets. We show that firm-level forecast errors in a given period push the regulator to favor price-based instruments over quantity-based instruments by inefficiently allocating quantity across firms, with the relative benefit of price instruments increasing in the variance of the forecast error term. Forecast errors that affect the overall distribution of quantity across compliance periods introduce further cost inefficiencies for quantity-based regulations with banking and borrowing, though additional considerations also emerge around benefit smoothing for non-stock pollutants.

To understand the potential magnitude of these forecast errors, we turn to simulations calibrated to the performance of actual and modeled climate policies. We first simulate stochastic allowance price trajectories, with drift and volatility parameters estimated from historical prices from the European Union's Emissions Trading System (EU ETS). We model the corresponding abatement investment response by firms using an abatement cost function estimated from models of U.S. carbon pricing policy. We then identify the smoothly increasing "Hotelling" price trajectories that would achieve the same overall emissions reductions as each of the stochastic price simulations, and compare the total resource cost required to achieve a particular level of emissions reductions under the two pricing scenarios. We show that the "Hotelling" scenario, approximately equivalent to an emissions tax, serves as a lower bound on total resource costs for achieving a given level of emissions control, by eliminating forecast errors on the part of regulated firms. We find that the median percentage difference in abatement costs between a stochastic price scenario and the corresponding Hotelling price scenario is approximately 20 percent. In ongoing work, we also illustrate how a hybrid policy instrument – such as a cap-and-trade program with price collars – can reduce the adverse economic impacts of forecast errors.

The structure of the paper follows. Section 2 presents evidence of cost-effectiveness anomalies in cap-and-trade and tradable performance programs in practice that illustrates the economic significance of this residual policy uncertainty inherent to quantity-based approaches. Section 3 develops our theoretical model of firm abatement decisions given investment in long-lived abatement capital and uncertainty over future allowance prices. Section 4 discusses how we can interpret price forecast errors in a Weitzman-style prices versus quantities welfare comparison. Section 5 calibrates a simulation model to illustrate the potential magnitude of forecast errors in these markets, using parameters calibrated to historical and modeled climate policies. Finally, Section 6 concludes and offers directions for future research.

2 Cost-Effectiveness Anomalies in the Implementation of Market-Based Instruments

Since the 1980s, policymakers have employed two major types of quantity-oriented, market-based instruments to address environmental and energy objectives: cap-and-trade programs and tradable performance standards. A cap-and-trade program establishes an emission cap that limits the aggregate quantity of emissions among all sources covered by the program. The cap is subdivided into emission allowances that grant the holder the right to emit a unit of pollution, and the government typically allocates these allowances through an auction and/or freely to sources based on their historic emissions. Firms must hold sufficient allowances to cover their emissions to demonstrate compliance, and a secondary market in emission allowances emerges where firms may buy and sell allowances. Policymakers have designed such markets for sulfur dioxide (SO_2), nitrogen oxides (NO_x), and carbon dioxide (CO_2).

Tradable performance standards establish a quantitative benchmark that firms must meet. If a firm beats the benchmark, its overcompliance generates a credit that may be traded to another firm, such as one that fails to meet the benchmark. Policymakers have designed such markets to reduce lead in gasoline, promote fuel economy among vehicle manufacturers, increase the renewable share of electricity generation, and raise the biofuel share of transportation fuel markets.

The theoretical appeal of cap-and-trade programs and tradable performance standards lies in the potential for the market to allocate effort in a cost-effective manner, just as in any other, efficient market. Montgomery (1972) formally showed how firms operating under a cap-and-trade program each have an incentive to equate their marginal abatement costs with the allowance price and, as a result, marginal abatement costs are equalized among all firms in the market. Complementing this static cost-effectiveness across firms, Rubin (1996) and Kling and Rubin (1997) demonstrate the potential for dynamic cost-effectiveness in cap-and-trade programs that permit intertemporal trading (banking allowances for future compliance purposes, or borrowing future vintage allowances for contemporary compliance purposes), with a Hotelling-style allowance price path over time emerging. In practice, the behavior in a variety of cap-and-trade programs and tradable performance standards deviate from these conditions in the underlying theory. This increases the costs of achieving any emission goal or energy objective.

2.1 Heterogeneity in Allowance Prices

Pollution markets deliver cost-effective abatement when firms equate their marginal abatement costs to the price of an emission allowance or credit (Montgomery, 1972). In order for firms to do this, there must be a single allowance or credit price. With many such markets, trading occurs via brokers, with less transparency about prices than under exchange-based trading. As a result, prices may deviate significantly across transactions.

Consider the California Low Carbon Fuel Standard (LCFS), which requires refineries to satisfy a performance benchmark based on the carbon content of transportation fuels. Since April 2016, the State of California has reported transaction-level data (credit prices and number of credits) on credit trades on 842 days (through December 1, 2019).[1] On 83 percent of these days with trades, California reported at least two completed credit trades – what we refer to as a multi-transaction trading days. On only 15 of these 699 multi-transaction trading days did the transactions include the same credit prices. The within-trading day credit price standard deviation averaged about \$10/tCO2. The maximum price paid for credits exceeded the minimum price for credits by more than 20 percent, on average, within multi-trade trading days. There are as many credit days in which the maximum price paid was double the minimum price paid as there are days with identical credit prices across transactions. If buying firms are equating their marginal costs of compliance with the credit price paid on the date of transaction, then this market is not resulting in the equating of marginal costs of compliance among firms.

2.2 Absence of Hotelling Price Path

The dynamic cost-effectiveness condition for cap-and-trade programs with banking calls for allowance prices to increase with the rate of interest. As Figure 1 illustrates for the SO_2, NO_x, EU ETS CO_2, and LCFS markets, prices are quite volatile, reveal occasional spikes

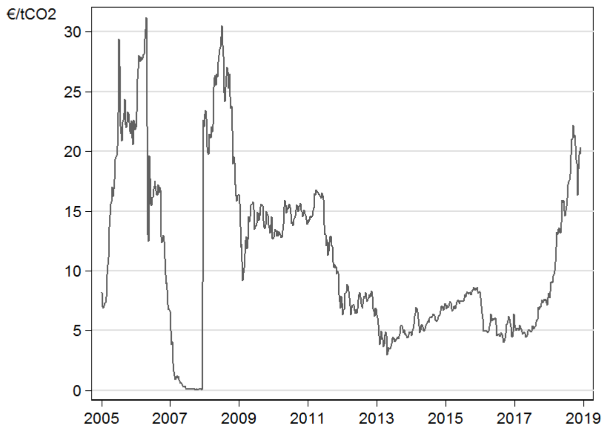
[1] We accessed these data on December 5, 2019 from: `https://ww3.arb.ca.gov/fuels/lcfs/credit/Weekly%20LCFS%20Credit%20Activity%20%28upto%201%20December,%202019%29.xlsx`.

and troughs, and do not follow what would be interpreted, even loosely, as a price path that increases with the rate of interest. This extreme volatility suggests that uncertainty about future prices may be substantial.

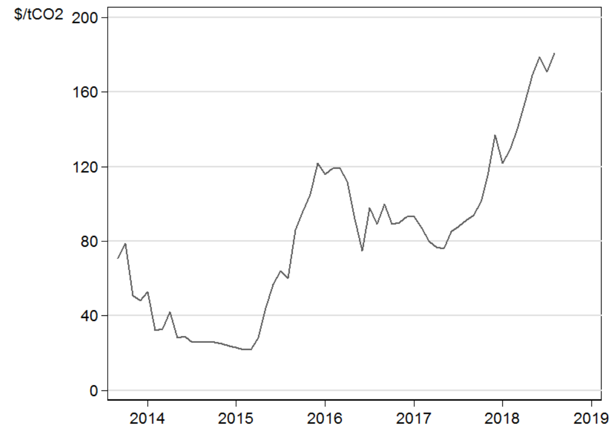
Under the EU ETS, market-clearing prices for carbon have ranged from over €30 per ton of CO_2 to under €5 per ton in the span of five years. Likewise, prices for SO_2 allowances under the U.S. Acid Rain Program have ranged from $1600 per ton to $100 per ton in a five-year period, and prices for NO_x allowances have ranged from $4,500 per ton to $800 per ton under the NO_x Budget Trading Program. Note that the observed volatility in these markets is not simply a function of the volatility of the underlying energy commodities, whether oil, natural gas, or coal. Indeed, the allowance price volatility observed under the EU ETS and the U.S. Acid Rain Program exceeds the volatility of oil or natural gas futures prices over comparable periods (see Figure 2). In all four markets, the volatility of allowance or credit prices exceeds the volatility of S&P 500 index prices over a comparable period.

2.3 Economic Shocks

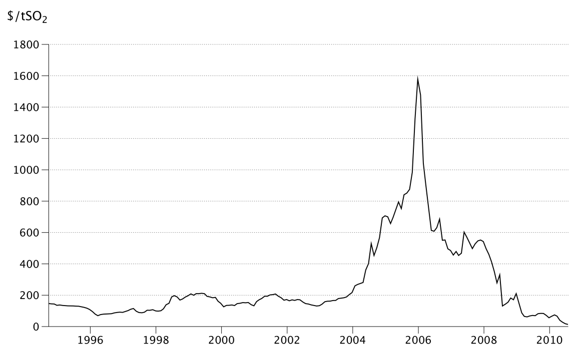
The failure to observe Hotelling-like price paths can reflect shocks to the system. For example, the California RECLAIM program witnessed NO_x allowance prices increase from about $1,000 per ton in 1999, to more than $20,000 per ton in 2000, to more than $120,000 per ton in 2001 (Fowlie et al., 2012). The dramatic run-up in allowance prices over 2000-2001 resulted from the California electricity crisis when insufficient generation existed to meet demand, causing the dirtiest generators, often relied on to meet occasional peak load, to run much more often during the crisis. This kind of output shock translated increased demand for pollution-intensive output into increased demand for emission allowances. As we noted in the introduction, firms need to acquire information about the actions of other firms in the market to form expectations about allowance prices. The start-up of the EU ETS signaled how poorly firms had done this. In April 2006, EU member states released data on the previous year's emissions of facilities covered by the ETS. The emission levels



(a) Emissions Trading System CO_2 Allowance Price (EU)



(b) Low Carbon Fuel Standard Allowance Price (California)



(c) SO_2 Allowance Price (U.S.)

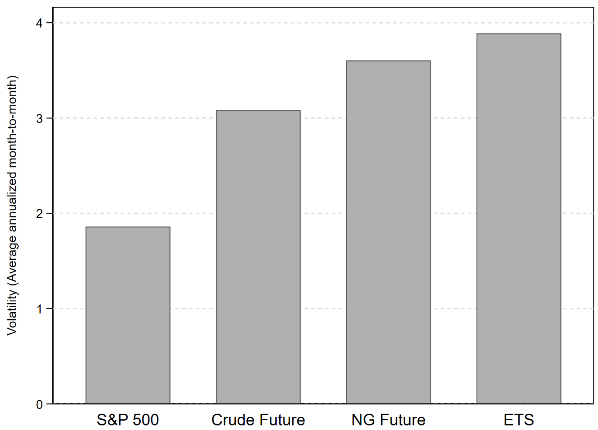
(d) NO_x Allowance Price (U.S.)

Figure 1: Historical Prices in Allowance and Credit Trading Markets

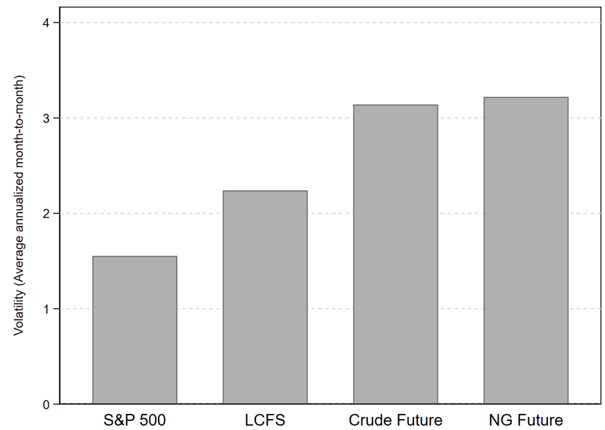
for 2005 suggested less allowance scarcity than the market had been pricing in. In a period of two weeks, the weekly average allowance prices fell from more than €31/tCO2 to about €13/tCO2. This represents a decline in market value of allowances of more than $50 billion.

2.4 Overlapping Regulations: Economic Regulation

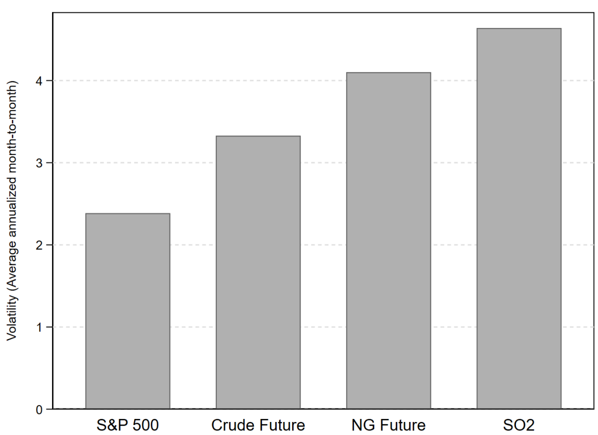
In the U.S. power sector, cap-and-trade programs may cross jurisdictional boundaries that separate power plants operating in competitive markets from those plants subject to state economic regulation (typically some form of cost-of-service rate regulation). This variation in economic regulation may influence the investment decisions and abatement behavior under



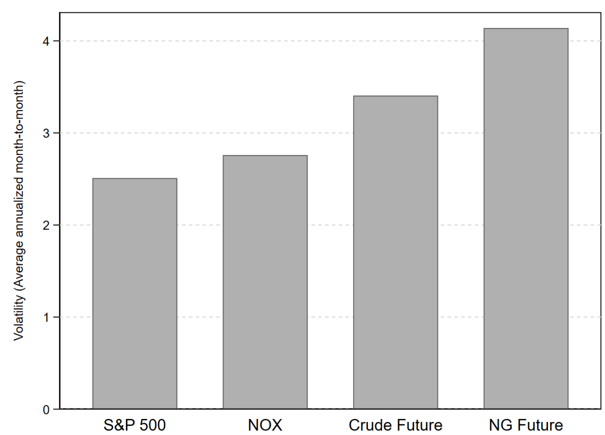
(a) Emissions Trading System CO_2 Allowance Price Volatility (EU)



(b) Low Carbon Fuel Standard Allowance Price Volatility (California)



(c) SO_2 Allowance Price Volatility (U.S.)



(d) NO_x Allowance Price Volatility (U.S.)

Figure 2: Average Annualized Month-to-Month Price Volatility in Allowance and Credit Trading Markets

the cap-and-trade program. For example, both Fowlie (2010) and Cicala (2015) show that deregulated firms may have underinvested in capital-intensive compliance strategies for SO_2 and NO_x cap-and-trade programs, combined with evidence of overinvestment by regulated firms. This may also help explain the finding by Carlson et al. (2000) that more than half of the units operating in Phase I of the SO_2 cap-and-trade program failed to minimize costs during at least part of the study period.

2.5 Overlapping Regulations: Environmental Regulation

As Goulder and Stavins (2011) describe, implementing an environmental mandate on top of an existing cap-and-trade program can cause firms to modify abatement investment to satisfy the mandate, which would increase costs, but would not increase emission reductions so long as the cap under the cap-and-trade program is binding. In effect, the mandate – to the extent it is binding – causes firms to undertake abatement investment that is not cost-effective and reduces the residual effort necessary to comply with the cap-and-trade program. The mandate imposes a pecuniary externality on the balance of the emission sources in the cap-and-trade program through lower allowance prices.

The EU ETS has experienced low allowance prices due to both low economic output – and hence demand for energy – and aggressive renewable power policies in some member states. The latter results in considerable divergence in implicit carbon prices between investment induced only by the cap-and-trade program and investment driven by high feed-in tariffs for wind and solar power. Marcantonini and Ellerman (2015) estimate that the German subsidies for wind and solar cost one and two orders of magnitude more, respectively, than the going EU ETS allowance price over 2007-2010. The State of California likewise employs a wide array of climate-oriented energy policies – a renewable portfolio standard, a solar roof mandate, an energy efficiency resource standard, the LCFS – that all overlap with the CO_2 cap-and-trade program. Any changes in the stringency of these overlapping instruments – or the introduction of new policies – would then affect the allowance prices in the cap-and-trade program.

Frey (2013) finds, for example, that overlapping state regulations limited the cost-effectiveness of Phase I of the Acid Rain Program, by inducing higher-cost units to invest in scrubbers. Under an allowance trading program, the impact of such state policies would not only affect the units directly subject to state regulation, but also all other units in the allowance trading market. Uncertainty about the impact of overlapping policies translates into uncertainty about the net quantity limit in the allowance trading program, which then translates into further uncertainty over the market-clearing price that should guide compliance investments. More recently, the implementation of the Clean Air Interstate Rule and the Cross-State Air Pollution Rule contributed to the collapse in the SO_2 emission allowance prices by requiring additional state-specific SO_2 emission reductions such that the cap under the Acid Rain Program is no longer binding. In contrast, under a price-based program, the presence of overlapping policies does not affect the investment decisions of other firms in the market.

2.6 Prospect for Anomalies in Future Carbon Markets

Policymakers at the supranational, national, and sub-national levels are moving forward with carbon pricing policies. The World Bank recently estimated that nearly 20 percent of the world's carbon dioxide emissions are covered (or will soon be covered) by some carbon pricing policy (see Figure 3). While numerous hybrid policies exist, the majority of existing or planned policies are emissions trading programs rooted in quantity targets, rather than tax policies. Moreover, about one-half of the nations pledging to mitigation emissions under the 2015 Paris Agreement signaled an interest in using carbon markets to do so.

In the United States, regulators have also exhibited strong revealed preference for using allowance or credit trading programs to correct for carbon dioxide as well as other environmental externalities. For example, virtually all energy produced or consumed is subject to some sort allowance or credit trading program because of an environmental attribute. These programs range from state-level trading programs for renewable energy credits as part of Renewable Portfolio Standards, to the nationwide sulfur and benzene credit trading program for gasoline, to the Acid Rain Program and the NO_x Budget Trading Program for controlling SO_2 and NO_x emissions, respectively, to the credit trading under the nationwide renewable fuel standard for gasoline and diesel markets, to the carbon dioxide cap-and-trade programs in California and the Regional Greenhouse Gas Initiative in the northeast and mid-Atlantic states, to the Low Carbon Fuel Standard in California.

The overlapping nature of energy and climate policies at the state and federal level also

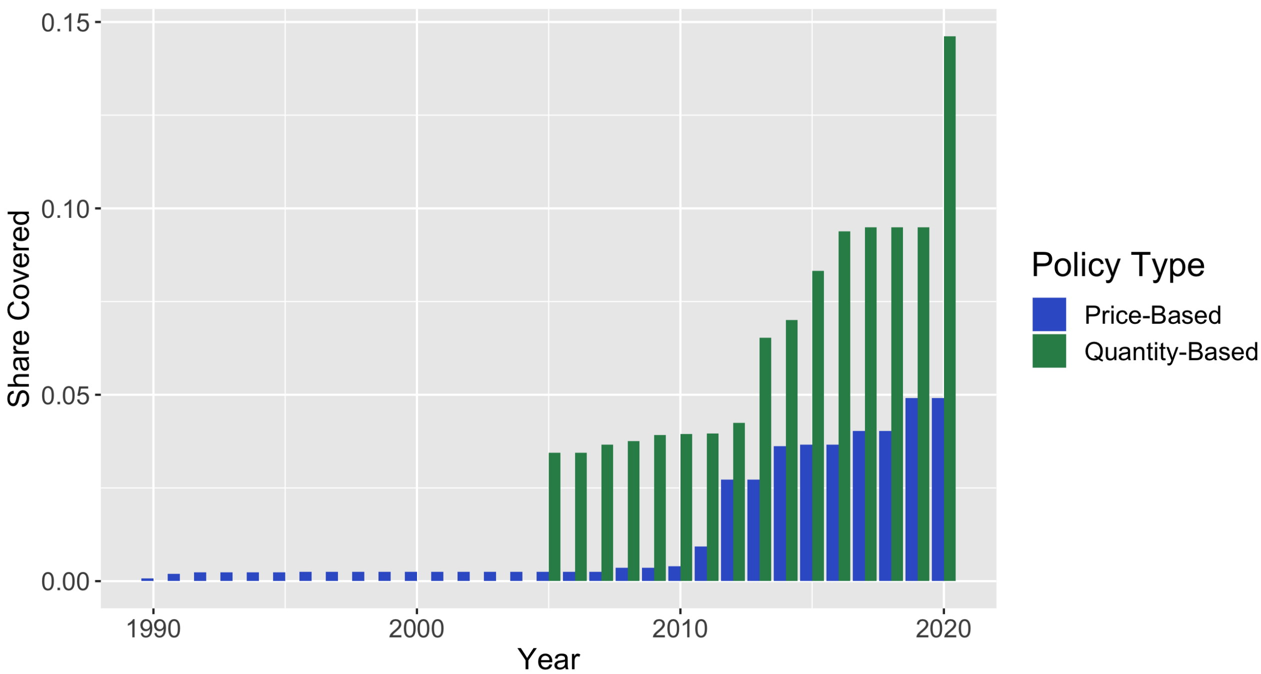


Figure 3: Global Carbon Emissions Covered by Policy (Ramstein et al., 2018).

raise questions about how firms would formulate expectations over carbon allowance prices. Indeed, firms have signaled quite varying expectations over carbon prices in the "internal carbon prices" they have employed in their investment analysis and strategic planning (Aldy and Gianfrate, 2019). According to the Carbon Disclosure Project (CDP), the average carbon price among U.S. firms disclosing use of an internal carbon price in their operations was about \$40/tCO2 in 2017, with a standard deviation of \$33/tCO2. Of course, this omits consideration of the mass of firms that implicitly use a price of zero and do not participate in such disclosure efforts. Even within the same industry and country, there is substantial variation: ExxonMobil uses \$80/tCO2, ConocoPhillips uses \$43/tCO2, and Devon Energy uses \$24/tCO2.

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 U.S. 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. To illustrate the potential magnitude of this start-up uncertainty, consider that proposals for U.S. economy-wide carbon cap-and-trade programs under the Kyoto Protocol and the Waxman-Markey Bill would have created multi-billion dollar markets in their first year of operation, without any historical data on market performance to guide the supply of financial instruments. Even regulators designing these programs predicted a wide range of initial market values through their modeling scenarios, ranging from $177 to $683 billion under the Kyoto Protocol and from $54 to $254 billion under the Waxman-Markey Bill (in 2018 $).

2.7 Building a Theory to Account for Such Anomalies

Our paper seeks to develop existing theories of cap-and-trade markets to account for these observations. Existing studies of allowance trading markets have generally adopted one of two assumptions which are incompatible with the actual functioning of these programs. On the one hand, papers in the spirit of Weitzman's canonical work on the relative advantage of price versus quantity advantage instruments have assumed that quantity orders are imposed directly on individual firms (Hoel and Karp, 2002; Pizer, 2002; Newell and Pizer, 2003). These papers generally focus on uncertainties on the part of the regulator rather than on the part of regulated firms; one exception is Yohe (1978), which 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. By contrast, in real-world allowance and credit trading markets, aggregate quantity orders are transmitted to firms through some (uncertain) 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 not only impacted by direct shocks to their own abatement cost functions.

On the other hand, another body of existing literature that has explicitly modeled the market-clearing price associated with a given quantity limit has instead assumed that prices follow a Hotelling trajectory given banking and borrowing (Rubin, 1996; Cronshaw and Kruse, 1996; Newell et al., 2005). In these models, equilibrium allowance prices are pinned down by an inter-temporal arbitrage condition which requires prices to rise smoothly at the rate of interest. While Rubin (1996) achieves this Hotelling trajectory by assuming perfect certainty in the allowance trading market, Newell et al. (2005) examine how to maintain this smooth trajectory when shocks to abatement cost functions are realized each period. They show that banking and borrowing combined with a commitment on the part of the regulator to fix the last period expected price at a certain level will achieve this price trajectory. However, they do not explore how firm behavior would differ if allowance prices were allowed to remain stochastic, which is the focus of this paper. Real-world allowance prices evidently do not follow this smooth trajectory, perhaps suggesting that allowance prices are continually jumping from one Hotelling path to another as shocks are realized in the market, which we explore further below.

We therefore develop a model of cap-and-trade markets that integrates the two key empirical facts – one, that market-clearing allowance prices are volatile and uncertain and, two, that many abatement options are long-lived capital investments. Our approach combines three modeling strategies: Rubin's work on allowance trading markets with certainty and variable abatement decisions; Anderson et al. (2018) on dynamic investment decisions in natural resource extraction; and Pindyck (1980) on natural resource extraction under uncertainty. Our model recognizes that quantity orders are transmitted to individual firms through an uncertain market-clearing price; that firms may take into account price uncertainty in making abatement decisions; and that abatement decisions are linked over time.

The two papers that relate most closely to this paper in directly accounting for firm investment decisions given allowance price uncertainty are Chao and Wilson (1993) and Zhao (2003). Chao and Wilson estimate the option value component of the SO_2 allowance price in the Acid Rain Program; this option value term reflects the fact that purchasing allowances may provide regulated firms with greater flexibility given uncertainty about future market conditions, as compared to investing in long-lived scrubbers. Then, Zhao (2003) considers how firm investment depends on uncertainty over the abatement cost function, where firms may invest in long-lived abatement capital or "technology" that reduces the cost of achieving a given level of abatement; in his model, actual compliance decisions are modeled as variable inputs. In comparing the impact of abatement cost uncertainty on both tradeable allowance programs and emission taxes, Zhao finds that tradeable allowance programs can be beneficial for responding to these uncertainties, as prices adjust to leave the overall quantity of variable abatement unaffected.

Our paper serves as a useful complement to this earlier work. Zhao's model considers how both firm-level and market-wide shocks affect firm investment in abatement technology but assumes that only market-wide shocks directly affect equilibrium prices or firms' marginal abatement costs. In contrast, our model allows price volatility to stem from any of the myriad random shocks affecting allowance trading markets, only some of which would affect firm abatement costs directly. Zhao's model also assumes a rational expectations equilibrium, whereas our paper explicitly considers the welfare consequences of forecast errors as part of the trade-off between price- and quantity-based policy instruments. Finally, unlike Zhao's model, our approach incorporates banking and borrowing of allowances, thereby allowing us to recover a version of the Hotelling rule for the expected growth in allowance prices that embeds existing models of allowance trading as a special case.

3 Theoretical Model of Abatement Investment

To build intuition for the impact of allowance price uncertainty, we develop a theoretical model of dynamic abatement investment decisions. This model focuses on the firm's cost-minimizing compliance strategy. We abstract from changes to the firm's production levels, though Montgomery (1972) shows how the cost of achieving a certain level of emissions may include both foregone profit due to deviations from unconstrained optimal production levels and the direct costs of installing abatement equipment. Our model shows that firms making abatement decisions will take into account expected price volatility as part of their optimizing behavior. While the expected price trajectory is pinned down by a Hotelling rule, as in previous papers, the price volatility depends on a wide variety of random processes that affect this market. This volatility causes prices to jump from one Hotelling trajectory to another, thereby causing firms to err in their beliefs about the realized allowance price. Firms may also not know all of the random processes that determine price volatility, causing them to err further in their forecasts of the overall price process.

Our model integrates modeling strategies from three related papers. First, we adopt the basic model of an allowance trading market, with banking and borrowing of allowances across periods, developed in Rubin (1996). As with much of the standard literature on multi-period allowance trading markets, Rubin's model assumes no accumulation of abatement from one period to the next. We therefore build on Rubin's work by modeling long-lived investment as the key margin of firm decision-making around compliance. Our approach is similar in the spirit to Anderson et al. (2018), which derives a modified version of the Hotelling rule for natural resource extraction where the key margin of decision-making is investment in well drilling rather than production from already-drilled wells. Lastly, we also recognize that firms face a stochastic price process for allowances. We allow firms to take into account price volatility in their optimization decisions, following methods used in Pindyck (1980) for natural resource extraction under uncertainty.

3.1 Model Set-Up

Assume that a representative firm chooses some optimal rate of abatement investment $A(t)$ and allowance purchases $Y(t)$ at each instant, where positive $Y(t)$ corresponds to net allowances purchases and negative $Y(t)$ corresponds to net allowances sales. The investment cost function is given by $\psi(\cdot)$ and the instantaneous price of allowances is $P(t)$. The firm discounts future costs and benefits at some exogenous discount rate r. The firm accumulates investment in abatement over time, and the total stock of abatement capital is given by $K(t)$; this stock depreciates at rate δ. The firm is also able to bank unused allowances over time, where $B(t)$ gives the total size of the firm's bank at time t. The firm's baseline emissions rate is $\bar{E}$. Therefore, at each instant, the firm's compliance requirement is $\bar{E} - K(t)$, which may result in the firm buying or selling allowances or adding or removing allowances from its bank.[2] Finally, following (Anderson et al., 2018), we define $R(t)$ as the remaining abatement opportunities available to the firm, perhaps due to technological limitations on the firm's ability to reduce its emissions rate beyond a particular amount.

Therefore, at the start of the allowance trading market, the firm's problem is given by:

$$\max_{A,Y} \mathrm{E}_0[\int_0^T e^{-rt}\{-\psi(A(t)) - P(t)Y(t)\}dt] \tag{1}$$

subject to:

$$\dot{K} = A(t) - \delta K(t) \tag{2}$$

$$A(t) \geq 0,\ \ K_0 \text{ given} \tag{3}$$

$$\dot{B} = K(t) + Y(t) - \bar{E} \tag{4}$$

$$B(T) \geq 0,\ \ B_0 = 0 \tag{5}$$

[2]Firms may also receive some allowances for free (grandfathered) as a function of historic emissions, and then buy or sell allowances and abate emissions as necessary for compliance. In that case, the firm's compliance requirement at each instant would instead be $\bar{E} - \bar{G} - K(t)$, where $\bar{G}$ represents the flow of grandfathered allowances. The optimality conditions derived in this section would remain unchanged.

$$\dot{R} = -A(t) + \delta K(t) \tag{6}$$

$$R(t) = \bar{E} - K(t) \geq 0,\ R_0 \text{ given} \tag{7}$$

$$dP = \alpha P dt + \sigma P dz \text{ , where } dz \text{ is an increment of a Weiner process} \tag{8}$$

Let us discuss each of these constraints in turn. Equation 2 describes the law of motion for abatement capital, where $\dot{K}$ represents the time derivative of capital stock; at each instant, investment at rate $A(t)$ adds to the stock, while depreciation at rate $\delta K(t)$ depletes the stock. Equation 3 sets the initial stock of abatement capital and constrains the rate of investment to be weakly positive, meaning that firms cannot reverse investments once they have occurred, besides waiting for depreciation to run its course. Likewise, Equation 4 describes the law of motion for the allowance bank, where $\dot{B}$ gives the time derivative of the allowance bank. Additional allowances are added to the bank at the rate of allowances purchased $Y(t)$, net of allowances used for compliance $\bar{E} - K(t)$. Equation 5 sets the initial level of the bank and constrains the bank to be weakly positive in the final period (a no-Ponzi condition).

Equation 6 gives the law of motion of the stock of abatement opportunities R, where current investment results in the loss of future opportunities, but depreciation of installed abatement capital creates additional opportunities. Here we define the stock of abatement opportunities as total baseline emissions less installed abatement capital, but we could assume technological limitations on the firm's ability to abate some of its baseline emissions without altering any of the major results below. Finally, we assume that prices follow a Geometric Brownian Motion process, with drift parameter α and volatility parameter σ, following Equation 8. We discuss in detail below the implications of assuming an exogenous price process in deriving the firm's optimal behavior.

To solve for the firm's necessary optimality conditions, we follow Pindyck (1980) and Karp and Traeger (2013) and define the flow profit function Π and the optimal value function J:

$$\Pi(t) = e^{-rt}\{-\psi(A(t)) - P(t)Y(t)\}$$

$$J(t,P,K,B,R) = \max_{A,Y} \mathrm{E}_t[\int_t^T e^{-r\tau}\{-\psi(A(\tau)) - P(\tau)Y(\tau)\}d\tau]$$

We then write the firm's problem as a dynamic programming equation in continuous time and expand using Ito's Lemma:[3]

$$0 = \max_{A,Y} \Pi(t) + \frac{1}{dt}\mathrm{E}_t[d(J)]$$

$$\begin{aligned} 0 = \max_{A,Y} \Pi(t) + J_t + J_B(K + Y - \bar{E}) + J_K(A - \delta K) + J_R(-A + \delta K) \\ + J_P \alpha P + \frac{1}{2} J_{PP}\sigma^2 P^2 \end{aligned} \tag{9}$$

For simplicity, we assume the firm chooses an interior solution for both $A(t)^*$ and $Y(t)^*$.[4] In the case of $Y(t)^*$, Rubin (1996) notes that if we wish to consider arbitrary price paths in this market, we would need to apply a technical condition to ensure that firms do not seek to buy or sell an unbounded number of allowances, since the firm's maximization problem is linear in the number of allowances purchased or sold. By applying the first-order condition directly in this derivation, we are assuming that the process of price formation is such that no firm finds it optimal to buy or sell an unbounded number of allowances in equilibrium. In the case of $A(t)^*$, our optimality condition dictates firm investment conditional on some interior solution $A(t)^* > 0$; it is possible that the first-order condition does not hold with equality for any positive value of $A(t)^*$ in which case the firm will be bound by the irreversibility constraint $A(t) \geq 0$ and set $A(t)^* = 0$.

With these assumptions in place, taking the first-order conditions with respect to the choice variables $A(t)$ and $Y(t)$ yields:

$$\Pi_A + J_K - J_R = 0$$

[3] We follow the notation in Pindyck (1980) and use the notation J_P to denote $\frac{\partial J}{\partial P}$ and $\frac{1}{dt}\mathrm{E}_t[d(\cdot)]$ to denote Ito's differential operator. As in Karp and Traeger (2013), we are able to eliminate the expectation operator in this step since we are taking the expectation conditional on information at time t, so all functions are measurable at time t.

[4] For additional discussion of modifying the firm's constraints to allow for banking but not borrowing, see Rubin (1996).

$$\Pi_Y + J_B = 0$$

which can be rewritten as:

$$\psi'(A(t))e^{-rt} = J_K - J_R \tag{10}$$

$$P(t)e^{-rt} = J_B \tag{11}$$

3.2 Firm's Optimal Behavior in the Allowance Market

First, we focus on the optimality conditions governing firm behavior in the allowance market. We evaluate Equation 9 at the optimal values of $Y(t)^*$ and $A(t)^*$ and differentiate with respect to B:

$$\begin{aligned} 0 =& \Pi_B + J_{tB} + J_{BB}(K(t) + Y(t) - \bar{E}) + J_{KB}(A(t) - \delta K(t)) + J_{RB}(-A(t) + \delta K(t)) \\ &+ J_{PB}\alpha P(t) + \frac{1}{2} J_{PPB}\sigma^2 P(t)^2 \end{aligned}$$

Applying Ito's Lemma allows us to rewrite this expression as:

$$0 = \Pi_B + \frac{1}{dt}\mathrm{E}_t[d(J_B)]$$

Since $\Pi_B = 0$, this expression simplifies to:

$$0 = \frac{1}{dt}\mathrm{E}_t[d(J_B)]$$

Substituting in the firm's first-order condition for $Y(t)$ (Equation 11) and expanding the differential operator gives:

$$0 = \frac{1}{dt}\mathrm{E}_t[d(P(t)e^{-rt})] = -rP(t)e^{-rt} + e^{-rt}\frac{1}{dt}\mathrm{E}_t[d(P(t))]$$

Rearranging terms gives:

$$rP(t) = \frac{1}{dt}\mathrm{E}_t[dP(t)] \tag{12}$$

This result recovers the standard Hotelling rule for the expected growth in equilibrium allowance prices – namely, that prices are expected to increase over time at the discount rate. It is useful to compare this result to the analogous finding in Rubin (1996), which assumes both perfect certainty about the future allowance market and no accumulation of abatement investment. When a no-borrowing constraint does not bind in his model, Rubin recovers the following rule for price dynamics:

$$rP(t) = \dot{P(t)} \tag{13}$$

In our model, Rubin's result holds in expectation. This finding then allows us to endogenize the drift parameter in our Geometric Brownian Motion process for P, as the Hotelling rule pins down that drift parameter α is equal to discount rate r in equilibrium.

3.3 Firm's Optimal Abatement Investment

Next, we examine optimal firm investment in abatement. We follow the same general steps described above; we first differentiate Equation 9 with respect to $K(t)$ and then apply Ito's Lemma, substitute the first-order conditions for $A(t)$ and $Y(t)$ in the resulting expression, and expand the remaining differential operator. This procedure yields the following condition:

$$(r+\delta)\psi'(A(t)^*) = P(t) + \frac{1}{dt}\mathrm{E}_t[d(\psi'(A(t)^*))] \tag{14}$$

Note that $A(t)^*$ is a function of the firm's state variables B, K, and R, as well as price P. Therefore, we can expand the differential operator on the right-hand side of Equation 14 using the following identities:

$$\begin{aligned} \frac{1}{dt}\mathrm{E}_t[dA(t)^*] &= \frac{1}{dt}\mathrm{E}_t[A_K dK + A_B dB + A_R dR + A_P dP + \frac{1}{2}A_{PP}dP^2] \\ &= A_K\dot{K} + A_B\dot{B} + A_R\dot{R} + A_P\alpha P + \frac{1}{2}A_{PP}\sigma^2 P^2 \end{aligned} \tag{15}$$

$$\frac{1}{dt}\mathrm{E}_t[(dA(t)^*)^2] = \frac{1}{dt}\mathrm{E}_t[(A_K dK + A_B dB + A_R dR + A_P dP + \frac{1}{2}A_{PP} dP^2)^2] \\ = A_P \sigma^2 P^2 \tag{16}$$

where the second set of equalities follows from eliminating all higher-order expressions of dt that vanish as $dt \to 0$. (We suppress the explicit time notation in these expressions for notational simplicity.)

Our optimality condition then becomes:

$$(r+\delta)\psi'(A(t)^*) = \underbrace{P(t)}_{\text{Avoided allowance purchases}} + \underbrace{\psi''(A(t)^*)\frac{1}{dt}\mathrm{E}_t[dA^*] + \frac{1}{2}\psi''(A^*)\frac{1}{dt}\mathrm{E}_t[(dA(t)^*)^2]}_{\text{Value of smoothing investment over time}} \\ = P + \psi''(A^*)[A_K\dot{K} + A_B\dot{B} + A_R\dot{R} + A_P\alpha P + \frac{1}{2}A_{PP}\sigma^2 P^2] + \frac{1}{2}\psi'''(A^*)A_P^2\sigma^2 P^2 \tag{17}$$

Given an interior solution for the optimal rate of abatement investment, we find that the firm not only sets its amortized marginal abatement cost equal to the avoided allowance payment, but also takes into account additional terms that capture the value of smoothing its rate of investment over time. This value of smoothing depends, in turn, on the extent of price volatility in the market, captured by the parameter σ. The value of smoothing also depends on both the convexity of the abatement cost function and of the marginal abatement cost function. The latter relationship may be understood as analogous to "prudence" terms in lifetime consumption-savings models in macroeconomics. In these models, convexity of marginal utility of consumption induces individuals to engage in precautionary savings in the presence of uncertainty. In our context, we can understand increased abatement investment as analogous to precautionary savings in the face of uncertain allowance prices, which occurs whenever the marginal abatement cost function is concave (such that $-\psi'(\cdot)$ is convex).[5]

We can lend further interpretation to this optimality condition by considering what happens when we set price volatility to zero or instead model abatement decisions as variable. First, if we set price volatility to zero, we effectively assume that allowance prices will follow

[5] Provided that optimal abatement is decreasing in allowance price, or $A_P \leq 0$.

the same Hotelling trajectory indefinitely. In that case, we recover the following variant of Equation 17:

$$(r+\delta)\psi'(A(t)^*) = P(t) + \psi''(A(t)^*)\dot{A(t)} \tag{18}$$

In this modified optimality condition, the firm still wishes to smooth abatement investment over time, but now the time derivative of A^* is known with certainty.[6] The "prudence" term, which depends linearly on price volatility, also vanishes.

Likewise, it is also instructive to compare our result in Equation 17 to a model where abatement is a variable decision, rather than a dynamic investment. We could recover this scenario as a special case of the model presented above by assuming full depreciation ($\delta = 1$). In this case, the firm's first-order conditions become:

$$\Pi_Y + J_B = 0$$

$$\Pi_A + J_B = 0$$

which can be rewritten as:

$$e^{-rt}\psi'(A(t)) + J_B = 0$$

$$e^{-rt}P(t) + J_B = 0$$

Rearranging terms and combining these expressions immediately yields:

$$\psi'(A(t)) = P(t) \tag{19}$$

Here the firm sets the rate of abatement such that marginal cost is equal to the instantaneous price of allowances. This result matches the abatement optimality condition in (Rubin, 1996).

[6]This result has parallels to the modified Hotelling rule for oil drilling in Anderson et al. (2018), which models firms making investment decisions as if future prices were certain.

3.4 Discussion

Our optimality conditions show that allowance trading activity, combined with banking and borrowing, pins down expected price growth in equilibrium. We can then interpret the volatility parameter as causing the price process to jump continually from one expected Hotelling path to another. Where does this volatility come from? While much existing literature discusses the Hotelling rule and allowance price trajectories over time, considerably fewer papers have endogenized the overall price *level*. One exception is Stock et al. (Forthcoming), which determines that the volatility of RIN prices in the Renewable Fuel Standard credit trading market is largely determined by economic fundamentals such as the global prices of oil, corn, and soybeans or the impact of overlapping policies – which are, at least to an extent, exogenous to decisions in the abatement market. From this perspective, assuming an exogenous volatility parameter may be a reasonable approximation. Additional sources of price volatility in these markets may also be, to first order, exogenous to the abatement behavior of market participants, including changes in economic output, changing expectations about future policy reforms, or unpredictable impacts of overlapping policies.

In a supplemental analysis (see Appendix), we also solved for the firm's necessary optimality conditions given an endogenous price process. In this extended version of the model, we model the market-clearing price as a flexible function of the overall emissions limit set by the regulator, the baseline level of emissions, total installed abatement capital, total accumulated allowance bank, and, critically, random shocks to both baseline emissions (e.g., due to changes in economic output or overlapping policies) and the abatement cost function (e.g., due to changes in the relative cost of high- versus low-emission fuels). This extended model therefore allows us to endogenize price volatility as a function of market fundamentals. Assuming a competitive market, in which firms do not internalize the impact of their decisions on the market-clearing price or behave strategically, we find that the firm's optimality conditions bear close resemblance to those derived above. We again recover the Hotelling rule for the expected path of allowance prices, and the optimal level of abatement investment again comes from setting the amortized marginal cost of abatement equal to avoided allowance payments plus terms that capture the value of smoothing investment over time. However, now the smoothing terms depend on the volatility of the underlying random processes, instead of the volatility of the exogenous price process. Likewise, the “prudence” term is expanded to include cross-partial derivatives of the investment cost function between abatement level and the random shock to the cost function; this source of volatility – in contrast to shocks to baseline emissions – affects both the market-clearing price and, conditional on that price, the optimal level of abatement investment. However, our substantive conclusions remain unchanged after we endogenize price volatility in this manner, which lends additional support to our use of an exogenous price process in our baseline model.

Given the significant price volatility observed in real-world allowance and credit trading markets, it seems likely that actual prices do indeed continually jump from one Hotelling path to another, as we have modeled here. As a consequence, firms are likely to experience significant forecast errors between the expected allowance price, against which firms make their abatement decisions, and the actual market-clearing price realized ex post. Moreover, firms may also err in their forecasts of the price *process*, in addition to specific realized prices. To the extent that the volatility parameter depends on many random variables, only some of which are known to an individual firm, each firm may not know the true value of σ when making its abatement decisions. Furthermore, firms may not be able to learn this parameter simply by observing the past history of prices, given that the extent of price volatility has varied significantly across different allowance and credit trading markets and over time.

4 Welfare

To understand the welfare consequences of firm-level forecast errors over allowance prices, we turn to a modified version of the Weitzman (1974, 2018) derivations for the relative advantage of price-based instruments over quantity-based instruments. This analysis allows us to examine the respective roles of information revelation over time versus firm-specific forecast errors about market-clearing prices. The former creates uncertainty in the market that is fundamentally unresolvable; firms can never know what abatement cost shocks will be several periods in the future, for example. 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 this uncertainty continues to affect the policy's cost-effectiveness even after information has been fully revealed in the market. Firm-specific forecast errors, on the other hand, may result from the presence of private information in the market even if all current shocks have been realized; 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.

To perform this analysis, we first develop Weitzman's 1974 derivation with multiple production units to account for the fact that aggregate quantity orders are transmitted through an equilibrium price in modern quantity-based regulations. We also integrate Weitzman's 2018 model of quantity regulation with banking and borrowing over two compliance periods.[7] To be consistent with Weitzman's set-up in both of these papers, all randomness in the market-clearing price stems from shocks to each firm's marginal abatement costs; however, as we have discussed earlier, real-world price volatility stems from myriad other factors including changes in overall economic output or the impact of overlapping policies. The key feature that links the uncertainty modeled here and these other sources of randomness in the allowance price is that no individual regulated firm has perfect information about all shocks affecting the market, which leads to firm-level forecast errors as firms make compliance decisions before the market-clearing price is realized and all information is revealed.

While our model captures many key features of multi-period price- and quantity-based

[7] Other closely related papers in this recent literature on prices-versus-quantities with banking and borrowing are Heutel (2018), Pizer and Prest (Forthcoming), and Karp (2019). These papers also consider how the relative advantage of prices, quantities, and quantities with banking and borrowing depend on whether the policymaker is able to update policies over time as information is revealed in the market. We consider here 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.

regulation in the presence of uncertainty, we return to the standard assumption that abatement is variable in this section.[8] As a result, we consider intertemporal smoothing of the quantity produced only at the aggregate level in this model; firm-level smoothing occurs simply by setting marginal cost equal to the market-clearing price, which then equalizes firm-level marginal costs over time (at least in expectation) through the no-arbitrage condition created by banking and borrowing. Future work will examine the welfare consequences of smoothing investment over time in the presence of price uncertainty.

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 associated with producing an aggregate quantity Q_t of the pollutant.[9] Likewise, let $C_t^i(q_t^i, \theta_t^i)$ be the costs to firm i associated with producing quantity q_t^i of the pollutant, where θ_t^i represents a firm-specific random shock to the cost function in period t. Therefore, the aggregate costs associated with pollution level Q_t are given by $\sum_{i=1}^{N} C_t^i(q_t^i, \theta_t^i)$, where $Q_t = \sum_{i=1}^{N} q_t^i$. Here we are assuming that the pollutant in question is uniformly mixed, such that only the total level of pollutant enters into the benefits function, not the identity of each polluting entity; this assumption reflects the characteristics of most greenhouse gas emissions but could be relaxed to model local pollutants. By contrast, the costs of abatement depend on the pollution level achieved by each individual firm.

We assume for tractability that there are two periods in the current regulatory cycle.[10] In each period, the regulator sets an optimal price order in the presence of uncertainty by

[8] Weitzman (2018) provides further discussion of the need for this assumption in footnote 8.

[9] Here we assume no uncertainty in the benefit function, which is consistent with Weitzman (2018).

[10] 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 intra-firm and inter-temporal 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.

solving the following maximization problem:

$$\max_{\tilde{p}_1,\tilde{p}_2} \mathrm{E}[B_1(\sum_{i=1}^{N} q_1^i(\tilde{p}_1,\theta_1^i)) - \sum_{i=1}^{N} C_1^i(q_1^i(\tilde{p}_1,\theta_1^i),\theta_1^i) + B_2(\sum_{i=1}^{N} q_2^i(\tilde{p}_2,\theta_2^i)) - \sum_{i=1}^{N} C_2^i(q_2^i(\tilde{p}_2,\theta_2^i),\theta_2^i)] \tag{20}$$

Likewise, the regulator sets the optimal (aggregate) quantity order by solving the following maximization problem:

$$\begin{aligned}\max_{\hat{Q}} \mathrm{E}[B_1(\sum_{i=1}^{N} q_1^i(p_1(\hat{Q},\theta_1,\theta_2),\theta_1^i)) - \sum_{i=1}^{N} C_1^i(q_1^i(p_1(\hat{Q},\theta_1,\theta_2),\theta_1^i),\theta_1^i) \\ + B_2(\sum_{i=1}^{N} q_2^i(p_2(\hat{Q},\theta_1,\theta_2),\theta_2^i)) - \sum_{i=1}^{N} C_2^i(q_2^i(p_2(\hat{Q},\theta_1,\theta_2),\theta_2^i),\theta_2^i)]\end{aligned} \tag{21}$$

Here $p_t(\hat{Q},\theta_1,\theta_2)$ represents the market-clearing price associated with the regulated quantity order $\hat{Q}$ and the marginal cost shocks θ_1 and θ_2.

Following Weitzman, 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 the pollutant that sets expected benefits equal to expected costs for each individual firm. However, $\bar{q}_t^i$ no longer represents the regulator's optimal quantity limit imposed on each firm, since the regulator now chooses the aggregate quantity only. The Taylor expansion of each abatement cost function about $\bar{q}_t^i$ is then given by:

$$C_t^i(q_t^i,\theta_t^i) = a_i(\theta_t^i) + (C' + \theta_t^i)(q_t^i - \bar{q}_t^i) + \frac{C_i''}{2}(q_t^i - \bar{q}_t^i)^2 \tag{22}$$

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.[11] 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. θ_t^i represents how the random

[11]As we show in the full derivation in the Appendix, 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.

cost shock affects the slope of the abatement cost function for firm i, and $a_i(\theta_t^i)$ represents how the cost shock θ_t^i 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 $\bar{Q}_t = \sum_{i=1}^N \bar{q}_t^i$:

$$B(Q_t) = b + B'(\sum_{i=1}^N q_t^i - \bar{q}_t^i) - \frac{B''}{2}(\sum_{i=1}^N q_t^i - \bar{q}_t^i)^2 \tag{23}$$

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

For the optimal price order, 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 $B' = C_i'$ for all i (and for all t where these parameters are constant). The full derivation is provided in the Appendix. Assuming cost minimization, each firm will set its realized marginal cost function equal to this price, yielding the following firm-level response function:

$$q_t^i(\tilde{p}_t, \theta_t^i) = \tilde{q}_t^i = \bar{q}_t^i - \frac{\theta_t^i}{C_i''} \tag{24}$$

The aggregate quantity produced in each period t will then be:

$$\sum_{i=1}^N q_t^i(\tilde{p}_t, \theta_t^i) = \sum_{i=1}^N \bar{q}_t^i - \frac{\theta_t^i}{C_i''} = \bar{Q}_t - \sum_{i=1}^N \frac{\theta_t^i}{C_i''} \tag{25}$$

By contrast, for the optimal quantity order, we must solve for the market-clearing price such that the aggregate quantity order is achieved after the realization of all shocks in the market. Furthermore, because we allow banking and borrowing across the two periods, we must also apply a no-arbitrage condition that requires that the first-period price is equal to the expected second-period price. To build intuition for the basic model set-up, we initially follow Weitzman and assume that firms know both first- and second-period cost shocks before making any compliance decisions; we then relax this assumption in subsequent sections. Assuming cost minimization, the equilibrium price associated with the overall optimal quantity order $\hat{Q}$ is given by:[12]

$$\hat{p}_1(\hat{Q}, \theta_1, \theta_2) = \hat{p}_2(\hat{Q}, \theta_1, \theta_2) = C' + \frac{\sum_{i=1}^{N} \frac{\theta_1^i + \theta_2^i}{2C_i''}}{\sum_{i=1}^{N} \frac{1}{C_i''}} \tag{26}$$

In the first period, firm i will produce $\hat{q}_1^i = \bar{q}_1^i + \frac{\sum_{j=1}^{N} \frac{\theta_1^i + \theta_2^i}{2C_j''}}{C_i'' \sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\theta_1^i}{C_i''}$; the aggregate first-period quantity produced will be $\hat{Q}_1 = \bar{Q}_1 + \sum_j \frac{\theta_2^i - \theta_1^i}{2C_i''}$. Likewise, in the second period, firm i will produce $\hat{q}_2^i = \bar{q}_2^i + \frac{\sum_{j=1}^{N} \frac{\theta_1^i + \theta_2^i}{2C_j''}}{C_i'' \sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\theta_2^i}{C_i''}$; the aggregate second-period quantity is then $\hat{Q}_2 = \bar{Q}_2 + \sum_j \frac{\theta_1^j - \theta_2^j}{2C_i''}$. (The full derivation is again given in the Appendix.)

Finally, following Weitzman, we define the relative advantage of prices over quantities as the expected difference between net benefits from the optimal price order and net benefits from the optimal quantity order with banking and borrowing. That is:

$$\begin{aligned} \Delta = & \mathrm{E}[B_1(\sum_{i=1}^{N} \tilde{q}_1^i) + B_2(\sum_{i=1}^{N} \tilde{q}_2^i) - \sum_{i=1}^{N} C_1^i(\tilde{q}_1^i, \theta_1^i) - \sum_{i=1}^{N} C_2^i(\tilde{q}_2^i, \theta_2^i)] \\ & - \mathrm{E}[B_1(\sum_{i=1}^{N} \hat{q}_1^i) + B_2(\sum_{i=1}^{N} \hat{q}_2^i) - \sum_{i=1}^{N} C_1^i(\hat{q}_1^i, \theta_1^i) - \sum_{i=1}^{N} C_2^i(\hat{q}_2^i, \theta_2^i)] \end{aligned} \tag{27}$$

Substituting each firm's response to the price and quantity orders, respectively, we obtain

[12] The Appendix also derives the market-clearing price for a single regulatory period (i.e., without banking and borrowing and the resultant no-arbitrage condition).

the following variant on Weitzman's original derivation:[13]

$$\Delta = \mathrm{E}[\frac{1}{4}(\frac{1}{\sum_{i=1}\frac{1}{C_i''}} - B'')(\sum_{i=1}\frac{\theta_1^i + \theta_2^i}{C_i''})^2] \quad (28)$$

Here the relative advantage of prices over quantities with banking and borrowing may be interpreted analogously to the original Weitzman derivation with multiple production units, where the regulator compares the net benefits of fixing aggregate quantity over two periods versus fixing marginal cost over two periods. Note that in this modified version of Weitzman's derivation, 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. Consequently, we must compare the slope of marginal benefits to an expression that combines the slopes of all marginal costs.

Given this baseline result for the relative advantage of prices versus quantities with banking and borrowing and multiple production units, we now proceed with relaxing the strong assumption that firms have perfect information about all shocks before making any compliance decisions.

4.1 Firm Forecast Errors

Because the market-clearing price associated with the regulator's quantity order depends on shocks to the marginal abatement cost functions of all firms, a given firm i may not know this price with certainty even in making abatement decisions for the current compliance period. As a result, we assume that firm i will set its marginal abatement cost function equal to the efficient price $\hat{p}_t$ associated with aggregate quantity order $\hat{Q}$, as derived in the previous section, plus some expectation error ϵ_t^i. For example, firm-level forecast errors in

[13] We can immediately compare this expression to the result in Weitzman (2018) for the comparative advantage of fixed prices over time-flexible quantities with perfect information for a single representative firm: $\Delta = \mathrm{E}[\frac{1}{4}(C'' - B'') \cdot (\frac{\theta_2+\theta_1}{C''})^2]$.

the first period would result in $\mathrm{E}_i[\hat{p}_1(\hat{Q}, \theta_1, \theta_2)] = \hat{p}_1(\hat{Q}, \theta_1, \theta_2) + \epsilon_1^i$. In the presence of forecast errors, the firm's quantity response becomes:

$$q_1^i(\hat{p}_1, \theta_1^i, \epsilon_1^i) = \frac{\hat{p}_1(\hat{Q}, \theta_1, \theta_2) + \epsilon_1^i - C' - \theta_1^i}{C_i''} + \bar{q}_1^i \tag{29}$$

Let us assume initially that these firm-level forecast errors do not affect the overall quantity produced in a given period. That is, some firms may have higher price expectations and other firms may have lower price expectations in a given period, but the aggregate quantity response is unchanged within that period.[14] To identify the impact of these forecast errors more clearly, we will assume that they only appear in the first period, whereas firms have accurate information about the market-clearing price in the second period.

By contrast, under a price order, the price is set by regulation and does not depend on private information about other firm's marginal abatement costs. We maintain our earlier assumption that the price is known with certainty to all firms under price-based regulation. Therefore, firms' response to a price order does not change from the version derived above.

We therefore re-derive the relative advantage of prices over quantities with banking and borrowing, allowing for the presence of forecast errors under quantity-based regulation but holding constant the net benefits of price-based regulation. Our welfare expression now becomes:

$$\Delta = \mathrm{E}[\frac{1}{4}(\frac{1}{\sum_{i=1} \frac{1}{C_i''}} - B'')(\sum_{i=1} \frac{\theta_1^i + \theta_2^i}{C_i''})^2 + \underbrace{\sum_i \frac{{\epsilon_1^i}^2}{2C_i''}}_{\text{Additional Term}}] \tag{30}$$

Equation 30 indicates 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 the error terms. One way to interpret this finding, in light of Weitzman's original result, stems from the fact that the regulator is no longer imposing quantity orders directly on individual firms. Instead, the regulator imposes an

[14] This initial assumption requires $\sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''} = 0$, which ensures that firms' expectation errors will collectively cancel with each other in determining the overall quantity.

aggregate quantity order, which a market mechanism then translates into individual quantity orders through the market-clearing price and the magnitude of marginal cost shocks realized for each firm. Given firm-level expectation errors in a given 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* in that period.

4.2 Information Revelation over Time

Beyond inefficiencies in the distribution of quantity across firms within a given period, uncertainty over market-clearing prices may also affect the overall distribution of the regulated quantity across periods. From this perspective, forecast errors may reflect not only idiosyncratic firm-level uncertainty over the marginal cost curves of other market participants, but also fundamental uncertainty over future marginal cost shocks. No-arbitrage conditions under quantity-based policies with banking and borrowing mean that the first-period market-clearing price must incorporate information about second-period shocks as well as first-period shocks, but these second-period shocks are generally not realized until after the first compliance period. Therefore, systematic forecast errors around the ex post efficient market-clearing price (that is, the price that incorporates accurate information about both first- and second-period shocks) may reflect this fundamental uncertainty about the realization of these future shocks. As a result, forecast errors may also cause firms to collectively over- or under-abate relative to the abatement level that would be intertemporally optimal ex post.[15]

To understand the welfare consequences of this information revelation over time, we relax

[15]Karp (2019) also relaxes the assumption in Weitzman (2018) that the representative firm has perfect information about second-period cost shocks, assuming 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 price.

the earlier restriction that first-period forecast errors do not affect the aggregate quantity produced in this first period.[16] Instead, the aggregate quantity in the first period may be higher or lower than what is intertemporally optimal, depending on the net impact of firms' forecast errors around the ex post efficient price. 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 upwards or downwards to compensate. To accomplish this adjustment, the second-period price must also adjust accordingly:

$$\hat{p}_2'(\hat{Q}, \theta_1, \theta_2, \epsilon_1) = C' + \frac{\sum_{i=1}^N \frac{\theta_1^i + \theta_2^i}{2C_i''} - \sum_{i=1}^N \frac{\epsilon_1^i}{C_i''}}{\sum_{i=1}^N \frac{1}{C_i''}}$$

Full details of the changes in expected benefits and costs under quantity orders with this type of forecast errors are provided in the Appendix; because this type of forecast error does not apply to regulated prices, the expected benefits and costs of price orders are again unchanged. The relative advantage of prices over quantities with banking and borrowing now becomes:

$$\begin{aligned} \Delta =& \mathrm{E}[\frac{1}{4}(\frac{1}{\sum_{i=1} \frac{1}{C_i''}} - B'')(\sum_{i=1} \frac{\theta_1^i + \theta_2^i}{C_i''})^2 + \sum_i \frac{{\epsilon_1^i}^2}{2C_i''} \\ & + \underbrace{\frac{1}{2}(\frac{1}{\sum_{i=1} \frac{1}{C_i''}})(\sum_{i=1}^N \frac{\epsilon_1^i}{C_i''})^2}_{\text{Additional Cost Term}} + \underbrace{B''(\sum_{i=1}^N \frac{\epsilon_1^i}{C_i''})^2 + 2B''(\sum_{i=1}^N \frac{\epsilon_1^i}{C_i''})(\sum_{i=1}^N \frac{\theta_2^i - \theta_1^i}{2C_i''})]}_{\text{Additional Benefit Term}} \end{aligned} \tag{31}$$

Considering that forecast errors may result in overall under- or over-abatement in a given compliance period creates several additional considerations for the regulator. First, the additional cost imposed by failing to achieve the optimal distribution of quantity across compliance periods, captured by the new term $\frac{1}{2}(\frac{1}{\sum_{i=1}^N \frac{1}{C_i''}})(\frac{\epsilon_1^i}{C_i''})^2$, pushes the regulator to prefer the price-based instrument over the quantity-based instrument with cost-ineffective banking and borrowing. Here we emphasize that although we have restricted forecast errors

[16] We now allow $\sum_{i=1} \frac{\epsilon_1^i}{C_i''} \neq 0$, where this sum reflects market-wide uncertainty about the ex post efficient price.

to occur in the first period only, this initial uncertainty continues to create cost-inefficiencies in later periods due to the intertemporal linkages of cap-and-trade.

On the other hand, while the effect of forecast errors on the intertemporal distribution of quantity creates another opportunity for the policy to deviate from perfect benefit smoothing over time, this concern is mitigated if the forecast errors reduce the net reallocation of quantity across periods (i.e., $\sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''}$ and $\sum_{i=1}^{N} \frac{\theta_1^i+\theta_2^i}{2C_i''}$ have opposite signs). Consequently, this second set of additional terms, which depend on the slope of the marginal benefit function, have an ambiguous impact on the regulator's preference for prices over quantities with banking and borrowing, depending on the overall impact on benefit smoothing. We note, however, that this second 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.

4.3 Discussion

In this comparison of expected welfare under prices versus quantities with banking and borrowing, we find that the presence of firm-specific forecast errors creates cost inefficiencies, both across firms and over time, that push the regulator to prefer price-based instruments over quantity-based instruments. These cost inefficiencies asymmetrically affect quantity instruments, which are transmitted to regulated entities through the equilibrium market-clearing price, but not price instruments, where firms know the regulated price ex ante. Moreover, we show that these cost inefficiencies should be considered alongside the standard comparison of the relative slopes of the marginal cost and marginal benefit functions.

It is important to note that the asymmetry in cost effectiveness between price and quantity instruments is a result of the inherent characteristics of quantity-based instruments, which create residual uncertainty for regulated firms even after questions about policy design or stringency have been resolved. We do not model here uncertainty around policy design – that is, what price level or what quantity target the regulator will set in the future, or whether previously announced policies will be altered – because this type of uncertainty may affect any policy instrument. Of course, the characteristics of quantity-based instruments may amplify the impact of this policy design uncertainty, given that a particular regulated entity is affected not only by its own uncertainty over future policies, but also the uncertainty of all other firms in the market, insofar as their beliefs about future policies impact their compliance decisions and ultimately the market-clearing price. By contrast, under a price-based instrument, a given firm's compliance decisions are affected only by its own marginal abatement costs and its own beliefs about future policies.

5 Simulations

To illustrate the potential magnitude of forecast errors arising from price volatility in cap-and-trade markets, we have calibrated a simulation model based on the recent Stanford Energy Modeling Forum (EMF-32) study focused on U.S. carbon tax policies (Barron et al., 2018a; McFarland et al., 2018). In developing this model, we have also incorporated the volatility of EU ETS prices as a real-world illustration of how uncertainties play out in tradable allowance markets.

We first simulate 100,000 different stochastic price trajectories using drift and volatility parameters estimated from the EU ETS. We then model the optimal abatement investment decisions of a representative firm faced with this stochastic price trajectory over ten periods, with prices leveling off indefinitely after the final period. Each period is calibrated to represent one year. We assume that the firm only knows the previous period's realized price when making its investment decisions, thereby allowing us to examine the impact of forecast errors around the realized price. We do not consider the impact of forecast errors related to the overall price process in these simulations.

Based on the total emissions reductions achieved from a given stochastic price trajectory, summed over all periods, we then calculate the "Hotelling" price trajectory that would achieve the same level of overall emissions reductions – that is, a single price that rises each period at the discount rate, before leveling off indefinitely after the final period. This price trajectory is approximately equivalent to an emissions tax in eliminating uncertainty over the regulated price, although in this thought experiment the quantity of emissions reductions is still equalized across analogous stochastic price and tax trajectories. Nonetheless, this step allows us to compare the total resource costs required to achieve a given level of emissions reduction with and without firm-level uncertainty. We can also illustrate how price collars such as price floors and ceilings can reduce the cost inefficiencies resulting from this residual uncertainty.

5.1 Model Calibration

We first assume that allowance prices follow geometric Brownian motion and estimate the drift and volatility parameters associated with historical prices in the EU ETS. We focus on the periods since the end of the Phase I pilot program (2008-2018), as rules for intertemporal banking and borrowing changed between Phases I and II. We estimate these parameters using maximum likelihood; full details of the estimation procedure are provided in the appendix. We estimate an annual (real) drift parameter of 0.0508 and an annual volatility parameter of 0.3925; our drift parameter corresponds to 5.22% expected annual price growth.

We then use results from the Stanford Energy Modeling Forum to calibrate the abatement cost function. This dataset includes projected U.S. emissions reductions, relative to a baseline scenario, resulting from an emissions price set at either $25 or $50 in 2020 and increasing at an annual rate of either 1% or 5% to 2050. The dataset includes results from 10 models analyzing each of these four price scenarios.

We adopt the simplifying assumption that all abatement is long-lived and thus abatement in the current period persists into the next period, adjusting for depreciation; this assumption matches our theoretical modeling of abatement as durable capital stock. Likewise, we also assume that abatement investment in a given year becomes available for compliance in the following year. Based on these assumptions, we calculate the discounted value of the tax payment avoided through abatement investment in each period. By assuming that the investment cost function takes the form $\psi(A) = \phi A^2$, we use the firm's first-order conditions to set the marginal investment cost equal to the discounted stream of avoided tax payments, assuming depreciation rate of 10% and discount factor of 0.95.[17] We calculate the depreciated sum of abatement and compare that to average emissions reductions observed in the modeling scenarios (relative to the baseline scenario). Setting these two values equal then allows use to estimate the abatement cost parameter ϕ.

We obtain an estimated parameter $\hat{\phi} = 8.30 \cdot 10^{-7}$ for abatement measured in metric tons of CO_2 reduced annually. Figure 4 shows the annual emissions reduction predicted by our calibrated model versus each of the ten modeling scenarios from the Stanford EMF-32 exercise.

5.2 Simulation Results

Given this calibration, Figure 5 illustrates the results of our simulations. The horizontal axis shows the different levels of total emissions reduction (in million metric tons of CO_2) resulting from different price simulations and ten years of firm investment in abatement. The vertical axis shows the total investment cost required to achieve those reductions. As is evident from the figures, the "tax" trajectories serve as a lower bound on the total resource costs required to achieve a range of abatement levels. Additionally, the expected increase in costs from stochastic prices increases in the magnitude of realized abatement. We find that the median percentage difference in abatement costs between a stochastic price scenario and

[17] This functional form assumption is consistent with standard assumptions in the literature about abatement cost functions, including Weitzman (1974). However, it does not fully capture the range of potential dynamics derived in the previous section as, for example, $\psi'''(A) = 0$.

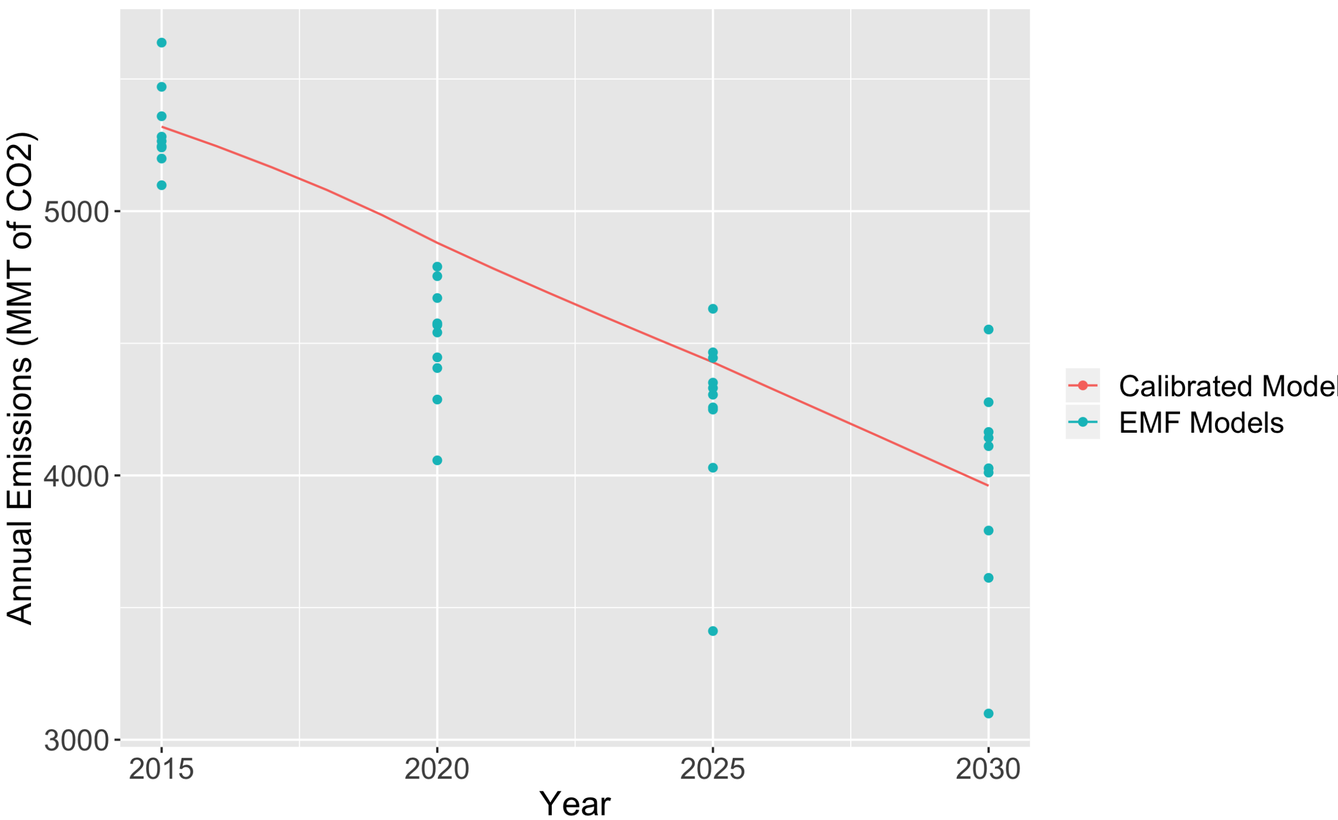


Figure 4: Annual Emissions from Calibrated Model versus EMF 32 Results

The points refer to the projected U.S. CO_2 emissions levels from each of the ten models in EMF 32, assuming an initial carbon price of \$25/ton in 2020, rising at 5% annually until 2050 and then leveling off. The line reflects implied emissions reduction from our calibrated abatement investment cost function, when applying the same carbon price trajectory.

the corresponding tax scenario is approximately 20 percent.[18] Because a given "stochastic price" and "tax" trajectory are constructed to achieve the same total emissions reduction, both scenarios reflect the total investment costs of achieving a range of quantity orders. Under each "tax" trajectory, various shocks to the price process are smoothed over time and firms know the full price trajectory with certainty, as in the Weitzman-style derivation in the previous section that incorporated all firm-specific shocks into a single efficient price. Under the stochastic price trajectories, firms instead face persistent forecast errors in making their investment decisions.

These results therefore illustrate why it is critical to account for the magnitude of forecast errors in considering the relative welfare gains from price or quantity instruments. The presence of firm-level uncertainty over the market-clearing allowance price effectively shifts upwards the expected abatement cost function – that is, the expected total resource cost associated with achieving a given quantity of emissions reduction. Therefore, it is not sufficient to compare the abatement costs of a quantity order imposed directly on firms, on the one hand, with the abatement costs from setting marginal abatement cost equal to the regulated price, on the other hand (as in Pizer (2002), for example). Instead, the effective abatement cost function is itself a function of the type of policy instrument. Persistent forecast errors create an additional welfare cost in the implementation of quantity orders relative to price orders.

6 Conclusion

This paper examines the impact of firm-level uncertainty over allowance prices in cap-and-trade markets – a form of residual uncertainty which is inherent to this type of policy instrument. Our theory model elucidates forecast errors that are not emphasized in the standard literature, both the difference between expected price and realized price and im-

[18]Note that we set an effective price ceiling at \$1000 per ton when discretizing the state space to perform backwards induction; however, given our drift and volatility parameters and the number of periods considered, this upper bound affects fewer than 0.1% of simulated price trajectories.

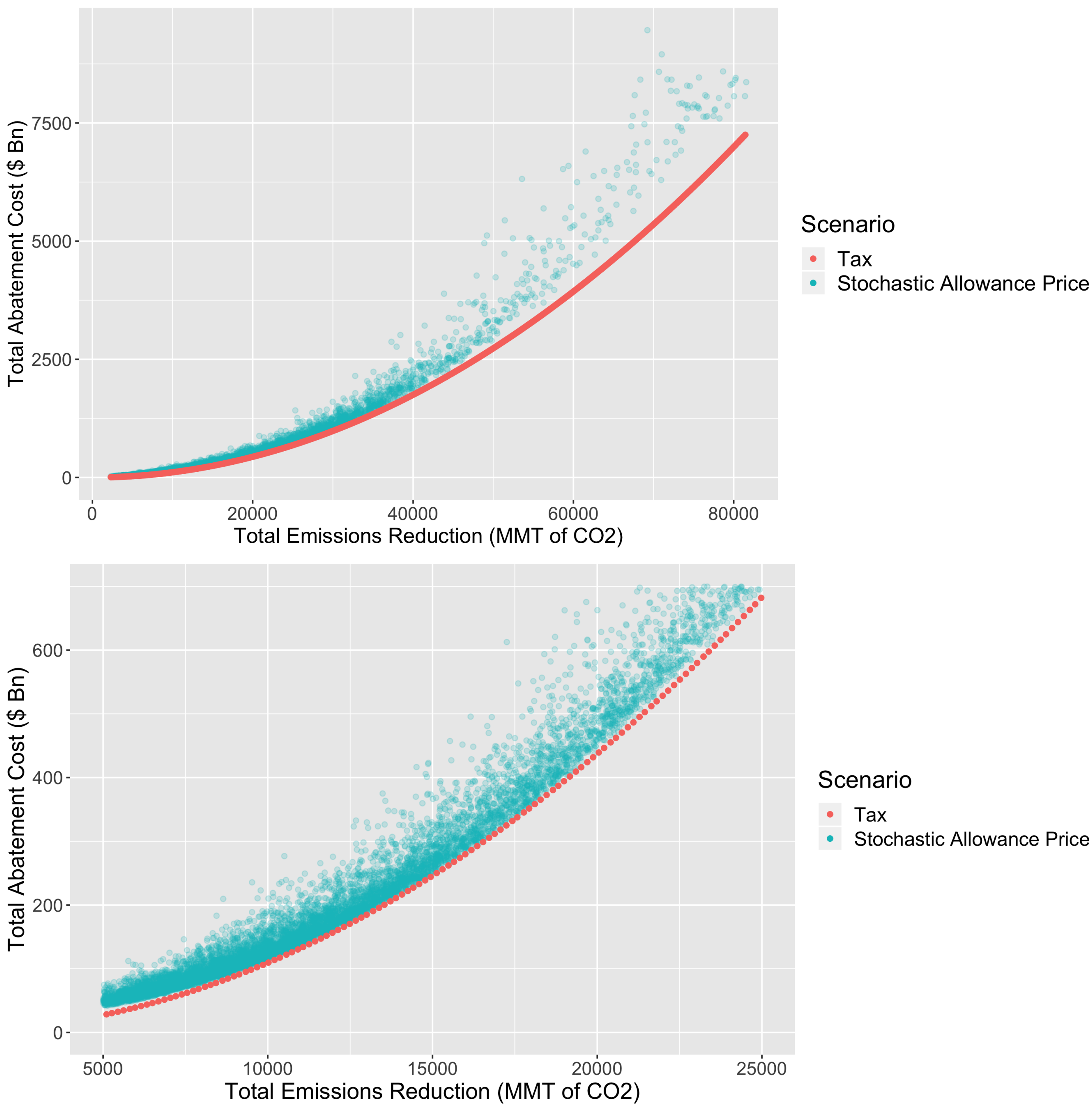


Figure 5: Total Emissions Reduction from 10 Years of Abatement Investment

The "stochastic" price scenario reflects total emissions reductions and abatement investment costs associated with simulated price trajectories, where drift and volatility parameters are calibrated to historic EU ETS prices; the baseline price is $25/ton. The "tax" scenarios achieve the same total emissions reduction given a baseline price that smoothly increases at the calibrated EU ETS drift. In both cases, the abatement cost function is calibrated to the results of Barron et al. (2018a). The lower panel is a close-up of the upper panel.

perfect information about the overall price process. Our welfare analysis then focuses on the first type of forecast error and shows how the additional cost associated with imperfect information about future market-clearing prices can be analyzed in a standard prices-versus-quantities framework. Finally, our simulations suggest that the magnitude of these forecast errors may be substantial in the context of climate policy, creating a wedge between the effective abatement cost function under price certainty versus in the presence of these forecast errors.

In future research, we seek to decompose the extent of price volatility into that resulting from own abatement cost shocks – and therefore correlated with the firm's optimal abatement decisions, conditional on price – and that resulting from other shocks to these markets. These other shocks may include shocks to the abatement costs of other firms, the impacts of overlapping policies, or changes to economic output. This exercise will also enable us to extend our Weitzman-style welfare analysis into a fully dynamic version in which optimal quantity responses are linked over time. We also have ongoing work to incorporate an assessment of overlapping policies in our simulation model.

References

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

S. T. Anderson, R. Kellogg, and S. W. Salant. Hotelling under pressure. *Journal of Political Economy*, 126(3):984–1026, 2018.

K. J. Arrow and A. C. Fisher. Environmental preservation, uncertainty, and irreversibility. *The Quarterly Journal of Economics*, 88(2):312–319, 1974.

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

A. R. Barron, A. A. Fawcett, M. A. C. Hafstead, J. R. McFarland, and A. C. Morris. Policy insights from the emf 32 study on u.s. carbon tax scenarios. *Climate Change Economics*, 9(1):1840003–1 – 1840003–47, 2018a.

A. R. Barron, A. A. Fawcett, M. A. C. Hafstead, J. R. McFarland, and A. C. Morris. Data from: Policy insights from the EMF 32 study on U.S. carbon tax scenarios, 2018b. Data obtained from personal communication with Alex Barron, December 17, 2018.

B. S. Bernanke. Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics*, 98(1):85–106, 1983.

C. Carlson, D. Burtraw, M. Cropper, and K. Palmer. Sulfur dioxide control by electric utilities: What are the gains from trade? *Resources for the Future Discussion Paper*, (98-44-REV), 2000.

H.-P. Chao and R. Wilson. Option value of emission allowances. *Journal of Regulatory Economics*, 5(3):233–249, 1993.

S. Cicala. When does regulation distort costs? lessons from fuel procurement in us electricity generation. *American Economic Review*, 105(1):411–444, 2015.

M. B. Cronshaw and J. B. Kruse. Regulated firms in pollution permit markets with banking. *Journal of Regulatory Economics*, 9:179–189, 1996.

A. K. Dixit and R. S. Pindyck. *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ, 1994.

European Central Bank: Statistical Data Warehouse. Inflation Rate (HICP), 2020. Data retrieved on January 9, 2020, from `https://sdw.ecb.europa.eu/`.

M. Fowlie. Emissions trading, electricity restructuring, and investment in pollution abatement. *American Economic Review*, 100(3):837–869, 2010.

M. Fowlie, S. P. Holland, and E. T. Mansur. What do emissions markets deliver and to whom? evidence from southern california's nox trading program. *American Economic Review*, 102(2):965–993, 2012.

E. F. Frey. Technology diffusion and environmental regulation: The adoption of scrubbers by coal-fired power plants. *The Energy Journal*, 34(1):177–205, 2013.

L. H. Goulder and R. N. Stavins. Challenges from state-federal interactions in us climate change policy. *American Economic Review*, 101(3):253–57, 2011.

K. A. Hassett and G. E. Metcalf. Investment with uncertain tax policy: Does random tax policy discourage investment? *The Economic Journal*, 109:372–393, 1999.

G. Heutel. Bankable prices. *NBER Working Paper Series*, (25235), 2018.

M. Hoel and L. Karp. Taxes versus quotas for a stock pollutant. *Resource and Energy Economics*, 24:367–384, 2002.

L. S. Karp. Welfare ranking of emissions taxes, cap & trade, and banking & borrowing. *Working Paper*, 2019.

L. S. Karp and C. P. Traeger. Dynamic methods in environmental and resource economics. Unpublished, 2013.

C. Kling and J. Rubin. Bankable permits for the control of environmental pollution. *Journal of Public Economics*, 64(1):101–115, 1997.

C. Marcantonini and A. D. Ellerman. The implicit carbon price of renewable energy incentives in germany. *The Energy Journal*, pages 205–239, 2015.

J. R. McFarland, A. A. Fawcett, A. C. Morris, J. M. Reilly, and P. J. Wilcoxen. Overview of the emf 32 study on u.s. carbon tax scenarios. *Climate Change Economics*, 9(1):1840002–1 – 1840002–37, 2018.

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

R. G. Newell and W. A. Pizer. Regulating stock externalities under uncertainty. *Journal of Environmental Economics and Management*, 45:416–432, 2003.

R. G. Newell, W. A. Pizer, and J. Zhang. Managing permit markets to stabilize prices. *Environmental & Resource Economists*, 31:133–157, 2005.

R. S. Pindyck. Uncertainty and exhaustible resource markets. *Journal of Political Economy*, 88(6):1203–1225, 1980.

B. Pizer and B. C. Prest. Prices versus quantities with policy updating. *Journal of the Association of Environmental and Resource Economists*, Forthcoming.

W. A. Pizer. Combining price and quantity controls to mitigate global climate change. *Journal of Public Economics*, 85(3):409–434, 2002.

A. A. Rampini, A. Sufi, and S. Viswanathan. Dynamic risk management. *Journal of Financial Economics*, 111:271–296, 2014.

C. Ramstein, R. Goyal, S. Gray, and A. C. Kallhauge. State and trends of carbon pricing 2018. Technical report, World Bank, Washington, DC, May 2018. URL `https://openknowledge.worldbank.org/bitstream/handle/10986/29687/9781464812927.pdf?sequence=5&isAllowed=y`.

D. Rodrik. Policy uncertainty and private investment in developing countries. *Journal of Development Economics*, 36(2):229–242, 1991.

J. Rubin. A model of intertemporal emission trading, banking, and borrowing. *Journal of Environmental Economics and Management*, 31(3):269–286, 1996. URL `https://EconPapers.repec.org/RePEc:eee:jeeman:v:31:y:1996:i:3:p:269-286`.

Sandbag - Smarter Climate Policy. Carbon Price Viewer, 2018. Data retrieved on December 6, 2018, from `https://sandbag.org.uk/carbon-price-viewer`.

J. H. Stock, S. H. Irwin, and K. McCormack. The price of biodiesel rins and economic fundamentals. *American Journal of Agricultural Economics*, Forthcoming.

M. L. Weitzman. Prices vs. quantities. *The Review of Economic Studies*, 41(4):477–491, 1974.

M. L. Weitzman. Prices or quantities dominate banking and borrowing. *Working Paper*, 2018.

G. W. Yohe. Towards a general comparison of price controls and quantity controls under uncertainty. *The Review of Economic Studies*, 45(2):229–238, 1978.

J. Zhao. Irreversible abatement investment under cost uncertainties: tradable emission permits and emissions charges. *Journal of Public Economics*, 87:2765–2789, 2003.

A Derivation of Theory Model

A.1 Necessary Condition for Optimal Investment

To obtain the dynamics of optimal abatement investment, differentiate the fundamental equation of optimality with respect to K:

$$\Pi_K + J_{tK} + J_{BK}(K + Y + \bar{G} - \bar{E}) + J_B + J_{KK}(A - \delta K) \\ - \delta J_K + J_{RK}(-A + \delta K) + \delta J_R + J_{PK}\alpha P + \frac{1}{2}J_{PPK}\sigma^2 P^2 = 0$$

After applying Ito's Lemma and eliminating $\Pi_K = 0$, this expression becomes:

$$\frac{1}{dt}\mathrm{E}_t[d(J_K)] + J_B - \delta J_K + \delta J_R = 0$$

Substituting the first-order conditions derived in the main text ($J_B = e^{-rt}P$ and $J_K - J_R = \psi'(A)e^{-rt}$), we obtain:

$$\frac{1}{dt}\mathrm{E}_t[d(J_R + \psi'(A)e^{-rt})] + e^{-rt}P - \delta\psi'(A)e^{-rt} = 0$$

Note that this expression is evaluated at the optimal values of A and Y. Next, noting that $\frac{1}{dt}\mathrm{E}_t[d(J_R)] = 0$[19] and expanding differential operator to eliminate e^{-rt} terms, we obtain:

$$(\delta + r)\psi'(A) = P + \frac{1}{dt}\mathrm{E}_t[d(\psi'(A))]$$

The main text provides the expansion and interpretation of this result.

A.2 Endogenous Prices

Now assume that the market-clearing price is some function of disturbances to baseline emissions η, disturbances to the marginal cost function θ, the installed stock of abatement equipment K, and the regulatory target $\bar{Q}$. (However, we assume that the market is competitive, so firms do not internalize the impact of their choice of K on the market-clearing price.)

Now the firm's optimization problem can be written as:

$$\max_{A,Y} \mathrm{E}_0[\int_0^T e^{-rt}\{-\psi(A(t);\theta(t)) - P(t)Y(t)\}dt]$$

subject to:

$$dK = \{A(t) - \delta K(t)\}dt$$

$$A(t) \geq 0,\ \ K_0 \text{ given}$$

[19] This result comes from differentiating the fundamental equation of optimality with respect to R, which yields: $\Pi_R + J_{tR} + J_{BR}(K + Y + \bar{G} - \bar{E}) + J_{KR}(A - \delta K) + J_{RR}(-A + \delta K) + J_{PR}\alpha P + \frac{1}{2}J_{PPR}\sigma^2 P^2 = 0$. Applying Ito's Lemma allows us to rewrite this expression as $\Pi_R + \frac{1}{dt}\mathrm{E}_t[d(J_r)] = 0$. Noting that $\Pi_R = 0$, we therefore have $\frac{1}{dt}\mathrm{E}_t[d(J_R)] = 0$.

$$dB = \{K(t) + Y(t) - \bar{E}\}dt + d\eta(t)$$

$$B(T) \geq 0, \; B_0 = 0$$

$$dR = \{-A(t) + \delta K(t)\}dt$$

$$R(t) = \bar{E} - K(t) \geq 0, \; R_0 \text{ given}$$

$$P(t) = f(\eta(t), \theta(t), K(t), B(t), R(t), \bar{Q})$$

$$d\eta(t) = \sigma_\eta dz_\eta, \text{ where } dz_\eta \text{ is an increment of a Wiener process}$$

$$d\theta(t) = \sigma_\theta dz_\theta, \text{ where } dz_\theta \text{ is an increment of a Wiener process}$$

$$\mathrm{E}_t[dz_\eta dz_\theta] = \rho$$

Now the fundamental equation of optimality can be written as:

$$0 = \max_{A,Y} \Pi(t) + \frac{1}{dt}\mathrm{E}_t[d(J)]$$

Which simplifies to:

$$0 = \max_{A,Y} \Pi_d(t) + J_t + J_B(K + Y - \bar{E}) + J_K(A - \delta K) + J_R(-A + \delta K) + \frac{1}{2}J_{\eta\eta}\sigma_\eta^2 + \frac{1}{2}J_{\theta\theta}\sigma_\theta^2 + J_{\theta\eta}\sigma_\theta\sigma_\eta\rho]$$

The firm's first-order conditions are again given by:

$$\Pi_A + J_K - J_R = 0$$

$$\Pi_Y + J_B = 0$$

To obtain price dynamics, we again differentiate the fundamental equation of optimality with respect to B:

$$
\begin{aligned}
0 =& \Pi_B + J_{tB} + J_{BB}(K + Y + \bar{G} - \bar{E}) + J_{KB}(A - \delta K) + J_{RB}(-A + \delta K) \\
&+ \frac{1}{2} J_{\eta\eta B}\sigma_\eta^2 + \frac{1}{2} J_{\theta\theta B}\sigma_\theta^2 + J_{\eta\theta B}\sigma_\eta\sigma_\theta\rho
\end{aligned}
$$

which can be again written as:

$$
0 = \Pi_B + \frac{1}{dt}\mathrm{E}_t[d(J_B)]
$$

In the competitive case where firms do not internalize the impact of their decisions on $P(t)$ through the state variables $B(t)$, $K(t)$, and $R(t)$, we have again have $\frac{\partial \Pi_d}{\partial B} = 0$.[20] Substituting the first-order condition $J_B = e^{-rt}P$ and expanding using Ito's differential operator again gives:

$$
0 = \frac{1}{dt}\mathrm{E}_t[d(e^{-rt}P(t))] = -re^{-rt}P + e^{-rt}\frac{1}{dt}\mathrm{E}_t[d(P(t))]
$$

We again recover the Hotelling rule for price dynamics:

$$
\frac{1}{dt}\mathrm{E}_t[d(P(t))] = rP(t)
$$

Next we differentiate the fundamental equation of optimality with respect to K, which yields:

$$
\begin{aligned}
&\Pi_K + J_{tK} + J_{BK}(K + Y - \bar{E}) + J_B + J_{KK}(A - \delta K) \\
&\quad - \delta J_K + J_{RK}(-A + \delta K) + \frac{1}{2} J_{\eta\eta K}\sigma_\eta^2 + \frac{1}{2} J_{\theta\theta K}\sigma_\theta^2 + J_{\eta\theta K}\sigma_\eta\sigma_\theta\rho = 0
\end{aligned}
$$

As before, this expression can be rewritten as:

$$
\Pi_K + \frac{1}{dt}\mathrm{E}_t[d(J_K)] + J_B - \delta J_K + \delta J_R = 0
$$

[20] For an analogous derivation, see equation 13 in (Pindyck, 1980).

Under a competitive equilibrium, we have $\Pi_K = 0$. Substituting first-order conditions again yields:

$$(\delta + r)\frac{\partial \psi}{\partial A} = P + \frac{1}{dt}\mathrm{E}_t[d(\frac{\partial \psi}{\partial A})] \tag{32}$$

In this case, we now have $\frac{\partial \psi}{\partial A}$ is a function both of A and θ. Note that we again evaluate the entire expression at the optimal values of A^* and Y^*. Therefore, expanding the differential operator on the right-hand side now yields:

$$\begin{aligned}\frac{1}{dt}\mathrm{E}_t[d(\frac{\partial \psi}{\partial A})] =& \frac{\partial^2 \psi}{\partial A^2}\frac{1}{dt}\mathrm{E}_t[dA] + \frac{\partial^2 \psi}{\partial A \partial \theta}\frac{1}{dt}\mathrm{E}_t[d\theta] \\ &+ \frac{1}{2}\frac{\partial^3 \psi}{\partial A^3}\frac{1}{dt}\mathrm{E}_t[(dA)^2] + \frac{\partial^3 \psi}{\partial A^2 \partial \theta}\frac{1}{dt}\mathrm{E}_t[(d\theta)(dA)] + \frac{1}{2}\frac{\partial^3 \psi}{\partial A \partial \theta^2}\frac{1}{dt}\mathrm{E}_t[(d\theta)^2]\end{aligned}$$

First note that $\mathrm{E}_t[d\theta] = 0$. Then, expanding with Ito's Lemma, note that $A^* = A(K, R, B, P, \theta)$, so we have $dA = A_K dK + A_R dR + A_B dB + A_P dP + A_\theta d\theta + \frac{1}{2}A_{\theta\theta}(d\theta)^2$.

Therefore we can expand the above terms as follows:

$$\frac{1}{dt}\mathrm{E}_t[(dA)^2] = A_P^2 f_\theta^2 \sigma_\theta^2 + A_P^2 f_\eta^2 \sigma_\eta^2 + 2A_P^2 f_\theta f_\eta \sigma_\theta \sigma_\eta \rho \tag{33}$$

$$\frac{1}{dt}\mathrm{E}_t[(dA)(d\theta)] = A_P f_\theta \sigma_\theta^2 + A_P f_\eta \sigma_\theta \sigma_\eta \rho \tag{34}$$

$$\frac{1}{dt}\mathrm{E}_t[(d\theta)^2] = \sigma_\theta^2 \tag{35}$$

Substituting into the above expression yields:

$$\begin{aligned}(r + \delta)\frac{\partial \psi}{\partial A} = P &+ \frac{\partial \psi^2}{\partial A^2}\frac{1}{dt}\mathrm{E}_t[dA^*] \\ &+ \frac{1}{2}\frac{\partial^3 \psi}{\partial A^3}[A_P^2 f_\theta^2 \sigma_\theta^2 + A_P^2 f_\eta^2 \sigma_\eta^2 + 2A_P^2 f_\theta f_\eta \sigma_\theta \sigma_\eta \rho] \\ &+ \frac{\partial^3 \psi}{\partial A^2 \partial \theta}[A_P f_\theta \sigma_\theta^2 + A_P f_\eta \sigma_\eta \sigma_\theta \rho] \\ &+ \frac{1}{2}\frac{\partial^3 \psi}{\partial A \partial \theta^2}\sigma_\theta^2\end{aligned} \tag{36}$$

B Derivation of Welfare Result

B.1 Single Price Order

First, we derive a variant of Weitzman's 1974 and 2018 results with multiple production units and two compliance periods, when the regulator sets a single price order. The problem set-up is given in the main text. The regulator's optimization problem is given by:

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

which yields the following first-order conditions:

$$\mathrm{E}[\sum_{i=1}^{N} \frac{\partial B_1}{\partial Q_1} \cdot \frac{\partial Q_1}{\partial q_1^i} \cdot \frac{dq_1^i}{d\tilde{p}_1}] = \mathrm{E}[\sum_{i=1}^{N} \frac{\partial C_1^i}{\partial q_1^i} \cdot \frac{dq_1^i}{d\tilde{p}_1}]$$

$$\mathrm{E}[\sum_{i=1}^{N} \frac{\partial B_2}{\partial Q_2} \cdot \frac{\partial Q_2}{\partial q_2^i} \cdot \frac{dq_2^i}{d\tilde{p}_2}] = \mathrm{E}[\sum_{i=1}^{N} \frac{\partial C_2^i}{\partial q_2^i} \cdot \frac{dq_2^i}{d\tilde{p}_2}]$$

As in Weitzman, we assume that firms set marginal cost equal to price, which yields the following response function for firm i facing price p_t:

$$\frac{\partial C_t^i}{\partial q_t^i} = p_t = C' + \theta_t^i + C_i''(q_t^i - \bar{q}_t^i)$$

Rearranging gives:

$$q_t^i(p_t,\theta_t^i) = q_t^i = \frac{p_t - C' - \theta_t^i}{C_i''} + \bar{q}_t^i$$

Differentiating with respect to p_t gives:

$$\frac{dq_t^i}{dp_t} = \frac{1}{C_i''}$$

Substituting this result into the regulator's first-order condition for the optimal price order and recognizing that all firms set $\frac{\partial C_t^i}{\partial q_t^i} = \tilde{p}_t$, we have:

$$\frac{\partial B_1}{\partial Q_1} \cdot \sum_{i=1}^{N} \frac{1}{C_i''} = \tilde{p}_1 \sum_{i=1}^{N} \frac{1}{C_i''}$$

$$\frac{\partial B_2}{\partial Q_2} \cdot \sum_{i=1}^{N} \frac{1}{C_i''} = \tilde{p}_2 \sum_{i=1}^{N} \frac{1}{C_i''}$$

Therefore, the optimal price order is given by $\tilde{p}_1 = B'$ and $\tilde{p}_2 = B'$. Using the same steps as Weitzman (1974), we can then show that $C' = \tilde{p}_1 = \tilde{p}_2$. Therefore, plugging this optimal price order into our price response function, we arrive at the same quantity response as the original Weitzman derivation:

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

B.2 Single Quantity Order

To obtain a single quantity order, we retain the set-up given in the main text. Now the optimal aggregate quantity order is given by:

$$\begin{aligned}\max_{\hat{Q}} \mathrm{E}[B_1(\sum_{i=1}^{N} q_1^i(p_1(\hat{Q}, \theta_1), \theta_1)) - \sum_{i=1}^{N} C_1^i(q_1^i(p_1(\hat{Q}, \theta_1), \theta_1^i), \theta_1^i) \\ + B_2(\sum_{i=1}^{N} q_2^i(p_2(\hat{Q}, \theta_2), \theta_2)) - \sum_{i=1}^{N} C_2^i(q_2^i(p_2(\hat{Q}, \theta_2), \theta_2^i), \theta_2^i)]\end{aligned}$$

which yields the first-order condition:

$$\begin{aligned}\mathrm{E}[\sum_{i=1}^{N} \frac{\partial B_1}{\partial Q_1} \cdot \frac{\partial Q_1}{\partial q_1^i} \cdot \frac{\partial q_1^i}{\partial \hat{p}_1} \cdot \frac{d\hat{p}_1}{d\hat{Q}} + \sum_{i=1}^{N} \frac{\partial B_2}{\partial Q_2} \cdot \frac{\partial Q_2}{\partial q_2^i} \cdot \frac{\partial q_2^i}{\partial \hat{p}_2} \cdot \frac{d\hat{p}_2}{d\hat{Q}}] \\ = \mathrm{E}[\sum_{i=1}^{N} \frac{\partial C_1^i}{\partial q_1^i} \cdot \frac{\partial q_1^i}{\partial \hat{p}_1} \cdot \frac{d\hat{p}_1}{d\hat{Q}} + \sum_{i=1}^{N} \frac{\partial C_2^i}{\partial q_2^i} \cdot \frac{\partial q_2^i}{\partial \hat{p}_2} \cdot \frac{d\hat{p}_2}{d\hat{Q}}]\end{aligned}$$

First, note that we still have $\frac{\partial q_t^i}{\partial p_t} = \frac{1}{C_i''}$, a constant. We can also use the price response equation from above to derive $\frac{d\hat{p}_t}{d\hat{Q}}$:

$$\sum_{i=1}^{N} q_t^i(p_t, \theta_t^i) = \sum_{i=1}^{N} \frac{p_t - C' - \theta_t^i}{C_i''} + \bar{q}_t^i$$

$$Q_t = \sum_{i=1}^{N} \frac{p_t - C' - \theta_t^i}{C_i''} + \bar{Q}_t$$

Therefore, aggregate quantity is given by:

$$Q = \sum_{i=1}^{N} \frac{p_1 - C' - \theta_1^i}{C_i''} + \bar{Q}_1 + \sum_{i=1}^{N} \frac{p_2 - C' - \theta_2^i}{C_i''} + \bar{Q}_2$$

Differentiating with respect to Q gives:

$$1 = \sum_i \frac{1}{C_i''} \frac{dp_1}{dQ} + \frac{1}{C_i''} \frac{dp_2}{dQ} \Rightarrow \frac{dp_1}{dQ} + \frac{dp_2}{dQ} = 1/(\sum_i \frac{1}{C_i''})$$

From the regulator's perspective at the time of setting the aggregate quantity limit, the no-arbitrage condition requires $p_1 = p_2$. Therefore, we have $\frac{dp_1}{dQ} = \frac{dp_2}{dQ} = 1/(2\sum_{i=1}^{N} \frac{1}{C_i''})$ Note that this expression is also a constant that can be pulled outside the expectation error. Therefore, we can set $Q = \hat{Q})$ and invoke the linearity of expectation to obtain:

$$\begin{aligned} &\sum_{i=1}^{N} B' \cdot \frac{1}{C_i''} \cdot \frac{1}{2\sum_j \frac{1}{C_j''}} + \sum_{i=1}^{N} B' \cdot \frac{1}{C_i''} \cdot \frac{1}{2\sum_j \frac{1}{C_j''}} \\ &= \sum_{i=1}^{N} \mathrm{E}[\frac{\partial C_1^i}{\partial q_1^i}] \cdot \frac{1}{C_i''} \cdot \frac{1}{2\sum_j \frac{1}{C_j''}} + \sum_{i=1}^{N} \mathrm{E}[\frac{\partial C_2^i}{\partial q_1^i}] \cdot \frac{1}{C_i''} \cdot \frac{1}{2\sum_j \frac{1}{C_j''}} \end{aligned}$$

Recognizing that each $\frac{\partial C_t^i}{\partial q_t^i}$ is set equal to $\hat{p}_t(\hat{Q}, \theta_t, \theta_{t'})$ by cost minimization and that $\mathrm{E}[\hat{p}_1(\hat{Q}, \theta_1, \theta_2)] = \mathrm{E}[\hat{p}_2(\hat{Q}, \theta_1, \theta_2)]$ from the perspective of the regulator deciding on optimal policy, we can rewrite this expression as:

$$\sum_{i=1}^{N} B' \cdot \frac{1}{C_i''} \cdot \frac{1}{\sum_j \frac{1}{C_j''}} = \sum_{i=1}^{N} \mathrm{E}[\hat{p}(\hat{Q}, \theta_1, \theta_2)] \cdot \frac{1}{C_i''} \cdot \frac{1}{\sum_j \frac{1}{C_j''}}$$

Which ultimately yields (for the regulator's optimal aggregate quantity order $\hat{Q}$):

$$B' = \mathrm{E}[\hat{p}(\hat{Q}, \theta)]$$

Note that we have defined $\bar{q}_t^i$ such that $\mathrm{E}[\frac{\partial B_t}{\partial q_t^i}] = \mathrm{E}[\frac{\partial C_t^i}{\partial q_t^i}]$. Therefore, the regulator's first-order condition for Q is satisfied when $\hat{Q} = \sum_{i=1}^{N} \bar{q}_1^i + \sum_{i=1}^{N} \bar{q}_2^i$.

Nonetheless, the realized efficient price resulting from the optimal quantity instrument is not necessarily equal to B' or C', but instead depends of the realizations of cost shocks θ. As in the initial derivation in the main body of the paper, we assume that firms initially have perfect information about first- and second-period cost shocks. Therefore, the first-period market-clearing price is given by:

$$Q_1 + Q_2 = \hat{Q}_1 + \sum_{i=1}^{N} \frac{\hat{p}_1 - C' - \theta_1^i}{C_i''} + \hat{Q}_2 + \sum_{i=1}^{N} \frac{\hat{p}_2 - C' - \theta_2^i}{C_i''}$$

We then apply the two key conditions governing this market: one the aggregate quantity limit must be met ($Q_1 + Q_2 = \hat{Q}_1 + \hat{Q}_2 = \hat{Q}$), and two, the no-arbitrage condition requires that the first-period market-clearing price is equal to the (expected) second-period market-clearing price ($\hat{p}_1 = \hat{p}_2$). Imposing these conditions and rearranging terms then yields the efficient market-clearing price under a multi-period quantity instrument with banking and borrowing:

$$\hat{p}_1(\hat{Q}, \theta_1, \theta_2) = \hat{p}_2(\hat{Q}, \theta_1, \theta_2) = C' + \frac{\sum_i \frac{\theta_1^i + \theta_2^i}{2C_i''}}{\sum_i \frac{1}{C_i''}}$$

Plugging this expression for $\hat{p}_t(\hat{Q}, \theta)$ into each price response function, we see that the individual realized quantities q_t^i will generally not be equal to $\hat{q}_t^i$. In the first period:

$$q_1^i(\hat{p}_1, \theta_1^i) = \frac{\hat{p}_1 - C' - \theta_1^i}{C_i''} + \bar{q}_1^i = \frac{C' + \frac{\sum_j \frac{\theta_1^j + \theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - C' - \theta_1^i}{C_i''} + \bar{q}_1^i = \frac{\frac{\sum_j \frac{\theta_1^j + \theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_1^i}{C_i''} + \bar{q}_1^i$$

Likewise, in the second period:

$$q_2^i(\hat{p}_2, \theta_2^i) = \frac{\hat{p}_2 - C' - \theta_2^i}{C_i''} + \bar{q}_2^i = \frac{C' + \frac{\sum_j \frac{\theta_1^j + \theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - C' - \theta_2^i}{C_i''} + \bar{q}_2^i = \frac{\frac{\sum_j \frac{\theta_1^j + \theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_2^i}{C_i''} + \bar{q}_2^i$$

By our definition of $\bar{q}_t^i$, we have imposed that C' is constant for all i and all t. Since we generally do not have $\theta_t^i = \theta_t^j$ or $\theta_t^i = \theta_{t'}^i$, this expression does not reduce to $\bar{q}_t^i$ except under very special conditions.

B.3 Relative Advantage of Prices Over Quantities with Banking and Borrowing

Here we derive the relative advantage of prices over quantities with banking and borrowing, relying on the baseline assumption from Weitzman (2018) that firms have perfect information about both first- and second-period cost shocks before making any compliance decisions. We substitute our expressions for $q_t^i(\hat{p}_t, \theta_t^i)$ and $q_t^i(\tilde{p}_t, \theta_t^i)$ into the Taylor expansions for (expected) benefits and costs. For the expected benefits of the quantity order, we obtain:

$$
\begin{aligned}
&\mathrm{E}[B_1(\sum_{i=1}^{N} q_1^i(\hat{p}_1, \theta_1^i)) + B_2(\sum_{i=1}^{N} q_2^i(\hat{p}_2, \theta_2^i))] \\
&\quad = \mathrm{E}[b + B'(\sum_{i=1}^{N} q_1^i(\hat{p}_1, \theta_1^i) - \bar{q}_1^i) + \frac{-B''}{2}(\sum_{i=1^N} q_1^i(\hat{p}_1, \theta_1^i) - \bar{q}_1^i)^2] \\
&\quad + \mathrm{E}[b + B'(\sum_{i=1}^{N} q_2^i(\hat{p}_2, \theta_2^i) - \bar{q}_2^i) + \frac{-B''}{2}(\sum_{i=1^N} q_2^i(\hat{p}_2, \theta_2^i) - \bar{q}_2^i)^2] \\
&\quad = \mathrm{E}[B'(\sum_{i=1}^{N} \frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_1^i}{C_i''}) + \frac{-B''}{2}(\sum_{i=1}^{N} \frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_2^i}{C_i''})^2] \\
&\quad + \mathrm{E}[B'(\sum_{i=1}^{N} \frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_1^i}{C_i''}) + \frac{-B''}{2}(\sum_{i=1}^{N} \frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_2^i}{C_i''})] \\
&\quad = -B''(\sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C_i''})^2
\end{aligned}
$$

Expected costs from the quantity order:

$$
\begin{aligned}
&\sum_{i=1}^{N} \mathrm{E}[C_1^i(q_1^i(\hat{p}_1, \theta_1^i), \theta_1^i)] + \sum_{i=1}^{N} \mathrm{E}[C_2^i(q_2^i(\hat{p}_2, \theta_2^i), \theta_2^i)] \\
&\quad = \sum_{i=1}^{N} \mathrm{E}[a_i(\theta_1^i) + (C' + \theta_1^i)(q_1^i(\hat{p}_1, \theta_1^i) - \bar{q}_1^i) + \frac{C_i''}{2}(q_1^i(\hat{p}_1, \theta_1^i) - \bar{q}_1^i)^2] \\
&\quad + \sum_{i=1}^{N} \mathrm{E}[a_i(\theta_2^i) + (C' + \theta_2^i)(q_2^i(\hat{p}_2, \theta_2^i) - \bar{q}_2^i) + \frac{C_i''}{2}(q_2^i(\hat{p}_2, \theta_2^i) - \bar{q}_2^i)^2] \\
&\quad = \sum_{i=1}^{N} \mathrm{E}[\theta_1^i(\frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_1^i}{C_i''}) + \frac{C_i''}{2}(\frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_1^i}{C_i''})^2] \\
&\quad + \sum_{i=1}^{N} \mathrm{E}[\theta_2^i(\frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_2^i}{C_i''}) + \frac{C_i''}{2}(\frac{\frac{\sum_j \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_j \frac{1}{C_j''}} - \theta_2^i}{C_i''})^2] \\
&\quad = -\frac{1}{2}(\sum_{i=1}^{N} \frac{{\theta_1^i}^2}{C_i''}) - \frac{1}{2}(\sum_{i=1}^{N} \frac{{\theta_2^i}^2}{C_i''}) + (\frac{1}{4\sum_{i=1}^{N} \frac{1}{C_i''}})(\sum_{i=1}^{N} \frac{\theta_1^i}{C_i''} - \sum_{i=1}^{N} \frac{\theta_2^i}{C_i''})^2
\end{aligned}
$$

Expected benefits from the price order:

$$\begin{aligned}
&\mathrm{E}[B_1(\sum_{i=1}^{N} q_1^i(\tilde{p}_1, \theta_1^i)) + B_2(\sum_{i=1}^{N} q_2^i(\tilde{p}_2, \theta_2^i))] \\
&\quad = \mathrm{E}[b + B'(\sum_{i=1}^{N} q_1^i(\tilde{p}_1, \theta_1^i) - \bar{q}_1^i) + \frac{-B''}{2}(\sum_{i=1}^{N} q_1^i(\tilde{p}_1, \theta_1^i) - \bar{q}_1^i)^2] \\
&\quad + \mathrm{E}[b + B'(\sum_{i=1}^{N} q_2^i(\tilde{p}_2, \theta_2^i) - \bar{q}_2^i) + \frac{-B''}{2}(\sum_{i=1}^{N} q_2^i(\tilde{p}_2, \theta_2^i) - \bar{q}_2^i)^2] \\
&\quad = \frac{-B''}{2}(\sum_{i=1}^{N} \frac{-\theta_1^i}{C_i''})^2 + \frac{-B''}{2}(\sum_{i=1}^{N} \frac{-\theta_2^i}{C_i''})^2
\end{aligned}$$

Expected costs from the price order:

$$\begin{aligned}
&\sum_{i=1}^{N} \mathrm{E}[C_1^i(q_1^i(\tilde{p}_1, \theta_1^i), \theta_1^i)] + \sum_{i=1}^{N} \mathrm{E}[C_2^i(q_2^i(\tilde{p}_2, \theta_2^i), \theta_2^i)] \\
&\quad = \mathrm{E}[a_i(\theta_1^i) + (C' + \theta_1^i)(q_1^i(\tilde{p}_1, \theta_1^i) - \bar{q}_1^i) + \frac{C_i''}{2}(q_1^i(\tilde{p}_1, \theta_1^i) - \bar{q}_1^i)^2] \\
&\quad + \mathrm{E}[a_i(\theta_2^i) + (C' + \theta_2^i)(q_2^i(\tilde{p}_2, \theta_2^i) - \bar{q}_2^i) + \frac{C_i''}{2}(q_2^i(\tilde{p}_2, \theta_2^i) - \bar{q}_2^i)^2] \\
&\quad = \sum_{i=1}^{N} \mathrm{E}[\theta_1^i(\frac{-\theta_1^i}{C_i''}) + \frac{C_i''}{2}(\frac{-\theta_1^i}{C_i''})^2 + \sum_{i=1}^{N} \mathrm{E}[\theta_2^i(\frac{-\theta_2^i}{C_i''}) + \frac{C_i''}{2}(\frac{-\theta_2^i}{C_i''})^2] \\
&\quad = \mathrm{E}[\sum_{i=1}^{N} -\frac{{\theta_1^i}^2}{2C_i''} + \sum_{i=1}^{N} -\frac{{\theta_2^i}^2}{2C_i''}]
\end{aligned}$$

Combining all terms to form the relative advantage of prices over quantities yields:

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

B.4 Relative Advantage of Prices Over Quantities with Forecast Errors

We now incrementally relax the assumption that firms have perfect certainty over all marginal cost shocks before making any abatement decisions. For one, each production unit may not know the realized $\hat{p}_t(\hat{Q}, \theta_t, \theta_{t'})$ when making its compliance decision, because this quantity depends on θ_t^{-i} whereas the production unit i may only observe θ_t^i. We note here that it is possible to leave the aggregate quantity distribution unchanged across the two compliance periods while mis-allocating quantity across production units within a given period, which produces additional welfare loss relative to a price instrument.

Suppose that firm i optimizes with respect to its signal of the market-clearing price, $\mathrm{E}[\hat{p}_t] = \hat{p}_t(\hat{Q}, \theta_t, \theta_{t'}) + \epsilon_t^i$, where ϵ_t^i reflects the firm's idiosyncratic forecast error in period t. The corresponding quantity response is given by:

$$q_t^i(p_t, \theta_t^i) = \frac{\hat{p}_t(\hat{Q}, \theta_t, \theta_{t'}) + \epsilon_t^i - C' - \theta_t^i}{C_i''} + \bar{q}_t^i$$

where $\hat{p}_t(\hat{Q}, \theta_t, \theta_{t'})$ is the efficient price defined above.

In this case, the aggregate quantity produced in period t will remain unchanged whenever the following condition is met:

$$\hat{Q} = \sum_i \frac{C' + \frac{\sum_j \frac{\theta_1^j + \theta_2^j)}{2C_j''}}{\sum_j \frac{1}{C_j''}} + \epsilon_1^i - C' - \theta_1^i}{C_i''} + \bar{Q}_1 + \sum_i \frac{C' + \frac{\sum_j \frac{\theta_1^j + \theta_2^j)}{2C_j''}}{\sum_j \frac{1}{C_j''}} - C' - \theta_2^i}{C_i''} + \bar{Q}_2$$

$$0 = 2\sum_{i=1}^{N} \frac{\theta_1^i + \theta_2^i}{2C_i''} - \frac{\theta_1^i}{C_i''} - \frac{\theta_2^i}{C_i''} + \frac{\epsilon_1^i}{C_i''}$$

$$0 = \sum_i \frac{\epsilon_1^i}{C_i''}$$

Here we assume that the firm experiences forecast errors only in the first period to cleanly distinguish this result from the results in the subsequent section; this assumption could be easily relaxed.

To determine the modified welfare expression in the presence of these expectation errors, note first that the benefits under a quantity order are a function of the total quantity produced in a given period; because we assume in this derivation that the expectation errors do not have an impact on the overall quantity in this compliance period, the Taylor expansion for $B_1(\sum_{i=1}^N \hat{q}_1^i)$ does not change from the version derived above. To determine how costs change under a quantity order, we evaluate the following expression:

$$\begin{aligned}
&\sum_{i=1}^N \mathrm{E}[C_1^i(q_1^i(\hat{p}_1,\theta_1^i),\theta_1^i)] + \sum_{i=1}^N \mathrm{E}[C_2^i(q_2^i(\hat{p}_2,\theta_2^i),\theta_2^i)] \\
&\quad = \sum_{i=1}^N \mathrm{E}[\theta_1^i(\frac{\frac{\sum_{j=1}^N \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_{j=1}^N \frac{1}{C_j''}} + \epsilon_1^i - \theta_1^i}{C_i''}) + \frac{C_i''}{2}(\frac{\frac{\sum_{j=1}^N \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_{j=1}^N \frac{1}{C_j''}} + \epsilon_1^i - \theta_1^i}{C_i''})^2] \\
&\quad + \sum_{i=1}^N \mathrm{E}[\theta_2^i(\frac{\frac{\sum_{j=1}^N \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_{j=1}^N \frac{1}{C_j''}} - \theta_2^i}{C_i''}) + \frac{C_i''}{2}(\frac{\frac{\sum_{j=1}^N \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_{j=1}^N \frac{1}{C_j''}} - \theta_2^i}{C_i''})^2] \\
&\quad = \text{ original terms from above } + \text{ new terms} \\
&\quad = \text{ original terms from above } + \sum_{i=1}^N \frac{{\epsilon_1^i}^2}{2C_i''} + \sum_{i=1}^N \frac{\theta_1^i\epsilon_1^i}{C_i''} - \sum_{i=1}^N \frac{\theta_1^i\epsilon_1^i}{C_i''} + (\sum_{i=1}^N \frac{\epsilon_1^i}{C_i''})(\frac{\sum_{j=1}^N \frac{\theta_1^j+\theta_2^j}{2C_j''}}{\sum_{j=1}^N \frac{1}{C_j''}}) \\
&= -\frac{1}{2}(\sum_{i=1}^N \frac{{\theta_1^i}^2}{C_i''}) - \frac{1}{2}(\sum_{i=1}^N \frac{{\theta_2^i}^2}{C_i''}) + (\frac{1}{4\sum_{i=1}^N \frac{1}{C_i''}})(\sum_{i=1}^N \frac{\theta_1^i}{C_i''} - \sum_{i=1}^N \frac{\theta_2^i}{C_i''})^2 + \underbrace{\sum_{i=1}^N \frac{{\epsilon_1^i}^2}{2C_i''}}_{\text{Additional Term}}
\end{aligned}$$

The last equality follows from applying the constraint that $\sum_{i=1}^N \frac{\epsilon_1^i}{C_i''} = 0$ in order for the aggregate quantity to be unchanged in the first compliance period.

By plugging in this expectation of the cost function under a quantity order with expectation errors, we obtain the following modified expression for the relative advantage of prices over quantities:

$$\Delta = \mathrm{E}[(\frac{1}{4\sum_{i=1}^N \frac{1}{C_i''}} - B'')(\sum_{i=1}^N \frac{\theta_1^i}{C_i''} - \sum_{i=1}^N \frac{\theta_2^i}{C_i''})^2 + \sum_{i=1}^N \frac{{\epsilon_1^i}^2}{2C_i''}]$$

B.5 Relative Advantage of Prices Over Quantities with Information Revelation Over Time

In this variant of the model, we continue to allow firms to make forecast errors with regard to the first-period price, but we now allow those errors to influence the overall distribution of quantity across compliance periods. That is, we now allow $\sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''}$. Consequently, the overall quantity produced in the first period is now given by:

$$Q_1 = \frac{\hat{Q}}{2} + \sum_{i=1}^{N} \frac{\theta_2^i - \theta_1^i}{2C_i''} + \sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''}$$

As discussed in the main text, the second-period price must now adjust to ensure that the aggregate quantity limit is still met, given this adjustment to first-period production. The new market-clearing price in the second period is now given by:

$$\hat{p}_2'(\hat{Q}, \theta_1, \theta_2, \epsilon_1) = C' + \frac{\sum_{i=1}^{N} \frac{\theta_1^i + \theta_2^i}{2C_i''} - \sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''}}{\sum_{i=1}^{N} \frac{1}{C_i''}}$$

This market-clearing price then yields the following overall quantity in the second compliance period:

$$Q_2 = \frac{\hat{Q}}{2} + \sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C_i''} - \sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''}$$

Consequently, expected benefits (over both compliance periods) from the quantity order are now given by:[21]

$$\mathrm{E}[-B''(\sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C_i''} - \sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''})^2]$$

[21] We could equivalently write this expression as: $\mathrm{E}[-B''(\sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C_i'} - \sum_{i=1}^{N} \frac{\epsilon_1^i}{C_i''})^2]$.

First-period expected costs from the quantity order are now given by:

$$-\frac{1}{2}(\sum_{i=1}^{N}\frac{\theta_1^{i\,2}}{C_i''})+\frac{1}{2}(\frac{1}{4\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\theta_1^i}{C_i''}-\sum_{i=1}^{N}\frac{\theta_2^i}{C_i''})^2+\sum_{i=1}^{N}\frac{\epsilon_1^{i\,2}}{2C_i''}$$
$$+(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\epsilon_1^i}{C_i''})(\sum_{i=1}^{N}\frac{\theta_1^i}{2C_i''})+(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\epsilon_1^i}{C_i''})(\sum_{i=1}^{N}\frac{\theta_2^i}{2C_i''})$$

Second-period expected costs are given by:

$$-\frac{1}{2}(\sum_{i=1}^{N}\frac{\theta_2^{i\,2}}{C_i''})+\frac{1}{2}(\frac{1}{4\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\theta_1^i}{C_i''}-\sum_{i=1}^{N}\frac{\theta_2^i}{C_i''})^2$$
$$+\frac{1}{2}(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\epsilon_1^i}{C_i''})^2-(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\epsilon_1^i}{C_i''})(\sum_{i=1}^{N}\frac{\theta_1^i}{2C_i''})-(\frac{1}{\sum_{i=1}^{N}\frac{1}{C_i''}})(\sum_{i=1}^{N}\frac{\epsilon_1^i}{C_i''})(\sum_{i=1}^{N}\frac{\theta_2^i}{2C_i i})$$

Combining terms and rearranging yields the following modified expression for the relative advantage of prices over quantities with banking and borrowing, where firms are subject to forecast errors and market-level information is revealed over time:

$$\Delta = \mathrm{E}[\frac{1}{4}(\frac{1}{\sum_{i=1}\frac{1}{C_i''}}-B'')(\sum_{i=1}\frac{\theta_1^i+\theta_2^i}{C_i''})^2+\sum_{i}\frac{\epsilon_1^{i\,2}}{2C_i''}$$
$$+\frac{1}{2}(\frac{1}{\sum_{i=1}\frac{1}{C_i''}})(\frac{\epsilon_1^i}{C_i''})^2+2B''(\sum_{i=1}^{N}\frac{\epsilon_1^i}{C_i''})(\sum_{i=1}^{N}\frac{\theta_2^i-\theta_1^i}{2C_i''})+B''(\frac{\epsilon_1^i}{C_i''})^2]$$

C Details about Model Calibration

C.1 Details about Price Calibration

We assume prices follow Geometric Brownian Motion and estimate the corresponding drift and volatility parameters by maximum likelihood estimation with data on historical EU ETS prices.

Prices following GBM will evolve according to the following (stochastic) law of motion:

$$P_t = P_0\exp\left((\alpha-\frac{\sigma^2}{2})t+\sigma\sqrt{t}W_t\right)$$

. where $W_t \sim N(0,1)$.

Note that this set-up can also be written as:

$$\ln(\frac{P_t}{P_0}) \sim N((\alpha - \frac{\sigma^2}{2}), \sigma^2 t)$$

We estimate the drift and volatility coefficients using maximum likelihood estimation. Recall that the maximum likelihood estimator for the mean of a normal random variable is $\hat{a} = \frac{1}{n}\sum_{j=1}^{n} x_j$ and the maximum likelihood estimator for the variance of a normal random variable is $\hat{s}^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2$.

Therefore, we have:

$$\hat{a} = (\hat{\alpha} - \frac{\hat{\sigma}^2}{2})t = \frac{1}{n}\sum_{i=1}^{T-1}\ln(\frac{P_{i+1}}{P_i})$$

$$\hat{s}^2 = \hat{\sigma}^2 t = \frac{1}{n}\sum_{j=1}^{n}(\ln(\frac{P_{i+1}}{P_i}) - \hat{a})^2$$

In this case, we have weekly price data, but it is more reasonable to assume that the relevant decision period is quarterly or annually. Therefore, we set $t = \frac{1}{52}$ (to reflect 52 weeks/year) when estimating $\hat{\alpha}$ and $\hat{\sigma}$. Using this procedure with EU ETS prices from 2008 through 2018 yields $\hat{\alpha} = 0.0508$ and $\hat{\sigma} = 0.3925$.

From the set-up above, we have:

$$\begin{aligned}
\mathrm{E}[P_{t+1}] &= \mathrm{E}[P_0 \exp((\alpha - \frac{\sigma^2}{2}) \cdot 1 + \sigma\sqrt{1}W_t)] \\
&= P_t \exp(\alpha - \frac{\sigma^2}{2})\mathrm{E}[\exp(\sigma W_t)] = P_t \exp(\alpha - \frac{\sigma^2}{2})\exp(\frac{\sigma^2}{2}) \\
&= P_t \exp(\alpha)
\end{aligned}$$

For estimated drift parameters around 0.0508, this yields expected price increases of $\exp(0.0508) = 1.0522$, or 5.22%. Data on historical EU ETS allowance prices is taken from Sandbag - Smarter Climate Policy (2018); we convert to real allowance prices using inflation data data from European Central Bank: Statistical Data Warehouse (2020).

C.2 Details about Abatement Function Calibration

We obtained data for the abatement function calibration from Barron et al. (2018b). From reviewing the results of this modeling exercise, it seems reasonable to assume that the estimated emissions reductions in each period relative to the baseline scenario depends on a) expectations of future allowance prices; b) the existing stock of abatement, insofar as the "low hanging fruit" is addressed first; and c) technology improvements over time. With the exception of expected allowance prices, we do not observe these components of the underlying model. Furthermore, the observed emissions reductions likely include both variable abatement (e.g., behavioral responses to reduce energy consumption, carbon capture, etc.) and fixed abatement investment (e.g., retrofitting plant to reduce energy consumption, installing carbon capture equipment, etc.), yet we do not observe the relative contribution of these two types of emissions reductions. As a consequence, we make certain assumptions about the abatement cost function, which we describe below.

As a first step, note that the Stanford EMF data includes estimated emissions at different years over the period 2010 to 2020 for the 10 different models. We extract emissions data for 2015, 2020, 2025, and 2030 specifically, since these years are included for (almost) all models; then we average over all 10 models for each price scenario, to obtain the values underlying the red lines from Figure 1 in Barron et al. (2018a). (See McFarland et al. (2018) for a technical discussion of the models underlying this data.) We calculate net emissions reductions over these periods for each price scenario, averaged across all models and adjusted for any changes in baseline emissions.

We adopt the simplifying assumption that all abatement is long-lived and thus abatement in the current period persists into the next period, adjusting for depreciation; this assumption matches our theoretical modeling of abatement as durable capital stock. Likewise, we also assume that abatement investment in a given year becomes available for compliance in the following year. Based on these assumptions, we calculate the discounted value of the tax payment avoided through abatement investment in each period. Following the Stanford EMF price scenarios, we assume that the emissions price increases at either 1% or 5% annually from 2020 to 2050, after which the price levels off (indefinitely); we also assume that firms correctly anticipate this price trajectory beginning in 2015. We set depreciation $\delta = 0.10$ and the firm's discount factor $\beta = 0.95$.

As illustration, a firm reducing emissions by A_{2020} in the year 2020 avoids the following tax payment:

$$\begin{aligned}\text{Avoided Tax} = \beta \cdot A_{2020} \cdot P_{2021} \cdot \frac{1-[(1+g)(1-\delta)\beta]^{2050-2021+1}}{1-(1+g)(1-\delta)\beta} \\ + \beta^{2051-2020} \cdot (1-\delta)^{2051-2021-1} \cdot A_{2020} \cdot P_{2050} \cdot \frac{1}{1-(1-\delta)\beta}\end{aligned}$$

where the first term refers to the avoided tax payment up to 2050, while the price is growing at rate g, and the second term refers to the avoided tax payment for all periods thereafter.

After computing the avoided tax payment from abatement investments in each year 2015 to 2030, we then rewrite each of these expressions to solve for A_t explicitly:

$$\begin{aligned}\text{Avoided Tax}_{2020}/\{\beta \cdot P_{2021} \cdot \frac{1-[(1+g)(1-\delta)\beta]^{2050-2021+1}}{1-(1+g)(1-\delta)\beta} \\ + \beta^{2051-2020} \cdot (1-\delta)^{2051-2021-1} \cdot P_{2050} \cdot \frac{1}{1-(1-\delta)\beta}\} = A_{2020}\end{aligned}$$

By assuming that the investment cost function takes the form $\psi(A) = \phi A^2$, we use the firm's first-order conditions to set the marginal investment cost equal to the discounted stream of avoided tax payments. We calculate the depreciated sum of abatement and compare that to average emissions reductions observed in the modeling scenarios (relative to the baseline scenario). Setting these two values equal then allows use to estimate the abatement cost parameter ϕ.

To illustrate, the total accumulated abatement stock in 2030 is given by:

$$K_{2030} = A_{2015} \cdot (1-\delta)^{14} + A_{2016} \cdot (1-\delta)^{13} + ... + A_{2028} \cdot (1-\delta) + A_{2029}$$

Emissions reductions relative to baseline are then given by:

$$\Delta E_{2030} = \bar{E}_{2030} - K_{2030}$$

Substituting my expressions for each A_t into this equation then allows me to solve for ϕ. For A denoted in metric tons of CO_2, estimated $\hat{\phi}$ values are provided in the table below:

	Modeling Scenario			
Years	**\$25, 5%**	**\$25, 1%**	**\$50, 5%**	**\$50, 1%**
2015-2030	$8.30 \cdot 10^{-07}$	$6.74 \cdot 10^{-07}$	$1.19 \cdot 10^{-06}$	$8.15 \cdot 10^{-07}$

Table 1: Estimated Abatement Cost Function Parameter from Stanford EMF-32 Modeling Scenarios

We use the parameter associated with a $25 tax growing at 5% annually.

D Details about Model Simulations

To model the representative firm's response to simulated price trajectories, we first performed backward induction to determine the firm's optimal abatement policy as a function of the accumulated abatement cost stock, the realized allowance price in the previous compliance period, and the number of elapsed compliance periods. Given computational limitations and the need to discretize the state space, the representative firm is able to accumulate abatement capital stock in multiples of 1 million metric tons of avoided annual CO_2 emissions; the upper bound on permitted abatement capital stock is total annual U.S. emissions in 2020, as modeled in EMF 32 baseline scenarios. We construct the price transition matrix by simulating 10 million evolutions of a stochastic process with our calibrated drift and volatility parameters and then calculating the probability that the next period allowance price will fall into each "price bin," conditional on the current period price. Each price bin is defined as a particular integer dollar value. Note that we set an effective price ceiling at $1000 per ton when discretizing the state space to perform backwards induction; however, given our drift and volatility parameters and the number of periods considered, this upper bound affects fewer than 0.1% of simulated price trajectories.

After constructing the representative firm's optimal policy matrix, we perform forward simulation to model abatement investment paths for 100,000 simulated stochastic price trajectories. To be consistent with the Stanford EMF modeling exercise, we assume that the price levels off indefinitely after period T. We then sum the total avoided emissions from each year of abatement investment and the firm's total current value cost of that investment. To compare the representative firm's response under each of these stochastic trajectories to responses under "tax trajectories," we calculate the initial price P_0 that would yield the same total emissions reduction if that initial price were to increase smoothly each period at the rate of interest. In this scenario, we assume that firms have perfect information about the price path. We then calculate the difference in total abatement investment costs between the "stochastic" and "tax" scenarios, having constrained total emissions reductions to be the same in both cases.[22]

[22] Because of the discretization of the state space, we cannot always achieve a given level of emissions reduction *exactly* following this approach. In practice, therefore, we calculate the total emissions reduction and total abatement investment cost from a sequence of smoothly increasing initial prices and then plot a curve from these emissions-cost pairs.