The Welfare Implications of Carbon Price Certainty

Joseph Aldy and Sarah Armitage*

Draft: August 13, 2020

Abstract

Experiences in real-world pollution markets suggest that firms make persistent errors in forecasting allowance and credit prices that inform their investment decisions. The residual uncertainty characterizing allowance and credit trading means that pollution markets may fail to deliver cost-effective abatement. This contrasts with price-based policies under which firms make investments that equate marginal abatement cost to an emission tax. We show how the additional cost of forecast errors under quantity-based programs can be incorporated into a standard Weitzman-style prices versus quantity framework. We distinguish between individual firms’ uncertainty over competitors’ costs and systemic uncertainty over future cost shocks. We show that a welfare-maximizing regulator would favor price instruments in response to the prospect of firm-specific forecast errors under quantity instruments, ceteris paribus, and the relative net benefit of price instruments increases with the variance of the forecast errors.

*Alidy: 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, Derek Lemoine, Robert Pindyck, and Jim Stock, as well as participants at the AERE Summer Conference, Northeast Environmental Economics Workshop, Harvard Kennedy School Energy Policy seminar, American Economic Association Annual Conference, Berkeley Climate Economics Workshop, NBER Energy and Environmental Economics Spring Workshop, and EAERE Annual Conference. 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.
Given strong regulatory revealed preference for quantity-based instruments, we also discuss the role of policy design, including price collars and policy updating, in mitigating the cost-inefficiencies of price forecast errors.

JEL Codes: Q52, Q54, Q58

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

Public policies – and uncertainty over public policies – play critical roles in firms’ investment and strategic decision-making (Arrow and Fisher, 1974; Rodrik, 1991; Hassett and Metcalf, 1999; Baker et al., 2016). A special case of policy uncertainty occurs when the government decides and communicates publicly a particular policy design and stringency, but residual uncertainty over returns to firm investment still remains due to the characteristics of policy implementation. This residual policy uncertainty is more pronounced under certain policy designs than others. 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 for abating carbon dioxide emissions and other environmental externalities.

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). Under cap-and-trade programs, however, this information asymmetry affects not only the regulator, but also each firm with respect to other firms’ marginal abatement costs and other forces driving the market-clearing allowance price. 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 identifies its abatement options as in the tax case, but also must form expectations about the market-clearing price for emissions allowances to guide its investment.
Our paper examines this residual uncertainty about allowance and credit 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 increases the aggregate costs of achieving 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 are recognized as having been too high or too low ex post. Such forecast errors may reflect different expectations about (1) abatement technology costs; (2) economic output; and (3) overlapping public policies that may restrict abatement decisions and influence the clearing price in allowance markets. This residual uncertainty 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.

To examine the welfare impacts of firm-level uncertainty, we develop a modified version of Weitzman’s canonical prices versus quantities framework. In our version, 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 encourage a welfare-maximizing regulator to favor price-based instruments over quantity-based instruments, with the relative benefit of price instruments increasing in the variance of the forecast error term. Forecast errors that affect the overall distribution of quantity across compliance periods introduce further cost inefficiencies for quantity-based regulations with banking and borrowing.

In the real world, policymakers have revealed their preferences for quantity-based instruments. Given this political economy, we also consider how modifying the design of tradable allowance or credit programs could reduce the cost inefficiencies created by firm-level uncertainty. First, we consider price collars – combining a price floor and a price ceiling on

\[1\] We address distinctions between stock and flow pollutants for benefit-smoothing over time in this discussion.
allowance and credit prices—which (weakly) reduce the variance of firm forecast errors and therefore the associated welfare losses. We show that the information provision role of price collars should be considered alongside standard arguments about the welfare consequences of introducing price ceilings or floors. Second, we consider the role of policy updating. While the presence of firm-level uncertainty precludes the first-best outcome, we show that policy updating, given information about realized cost shocks that diminishes uncertainty over time, can offset some of the cost inefficiencies introduced by forecast errors.

To motivate our theoretical model, we provide evidence of cost-effectiveness anomalies in cap-and-trade and tradable performance programs in practice in section 2 of the paper. These empirical examples illustrate the economic significance of this residual policy uncertainty inherent to quantity-based approaches. Section 3 describes how to interpret price forecast errors in a Weitzman-style prices versus quantities welfare framework. Section 4 addresses the role of policy design in ameliorating the inherent uncertainty of allowance and credit trading programs. Finally, section 5 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. To demonstrate compliance, firms must hold sufficient allowances to cover their emissions, 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 the emergence of a Hotelling-style allowance price path over time.

In practice, however, behavior in a variety of cap-and-trade programs and tradable performance standards deviates from these conditions in the underlying theory. Firms may fail to predict the market-clearing price for allowances or credits before making their abatement decisions, or may face non-transparent prices even after determining their compliance strategies. In markets that allow intertemporal banking and borrowing, current prices depend on firms’ expectations of future shocks, which may or may not prove to be accurate. Shocks external to the firm, such as unexpected changes to economic output or impacts of overlapping policies, may alter the effective cap or standard faced by regulated firms and increase firms’ uncertainty about market-clearing prices. These deviations from the cost-effectiveness
predicted by theory all increase the costs of achieving any emission goal or energy objective. We explore each of these issues in detail below.

2.1 Heterogeneity in Allowance and Credit 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).\(^2\) 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. Figure 1 plots the minimum and maximum daily transaction price, illustrating the significant range in daily LCFS prices. The within-trading day credit price standard deviation averaged about $10 per ton of CO\(_2\). The maximum price paid for credits exceeded the minimum price for credits by more than 20 percent, on average, within multi-transaction trading days. There are as many trading 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.

\(^2\)We accessed these data on December 5, 2019 from: https://ww3.arb.ca.gov/fuels/lcfs/credit/Weekly%20LCFS%20Credit%20Activity%20upto%20%20December,%202019%20.xlsx.
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 smoothly with the rate of interest. As Figure 2 illustrates for the U.S. \(SO_2\), U.S. \(NO_x\), EU ETS \(CO_2\), and LCFS markets, prices are quite volatile, reveal occasional spikes 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 3). In all four markets, the volatility of allowance or credit prices exceeds the volatility of S&P 500 index prices over a comparable period.

(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 2: Historical Prices in Allowance and Credit Trading Markets

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 increased demand for pollution-intensive output, which translated into increased demand for emission allowances.

As we have noted, 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 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 per ton of $CO_2$ to about €13 per ton of $CO_2$. 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 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. The presence of cost-of-service regulation 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

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 change the level of emissions so long as the cap under the cap-and-trade program is binding (Goulder and Stavins 2011). 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 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, and 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.

The uncertainty around non-market overlapping policies, combined with uncertainty around business-as-usual emissions, dwarfs uncertainty over price-responsive abatement quantities in the context of California’s CO$_2$ cap-and-trade program [Borenstein et al. 2019]. Since the CO$_2$ market covers power plants and refineries (among other sources), implementing an overlapping policy on power plants, such as a renewable portfolio standard, influences power plant emissions and behavior in the CO$_2$ market in a way that imposes a pecuniary externality on refineries through the CO$_2$ allowance price. Likewise, layering the Low Carbon Fuel Standard on top of the CO$_2$ market for refineries alters their incentives for emission abatement in a way that affects their behavior in the CO$_2$ market and hence imposes a pecuniary externality on power plants. In contrast, under a price-based program, the presence of overlapping policies on one set of firms does not affect the tax rate and hence does not influence the investment decisions of other firms in the market.

2.6 Potential for Anomalies in Future Carbon Markets

This evidence of cost inefficiencies in the implementation of allowance and credit trading markets is particularly important to consider given that policymakers at the supranational, national, and sub-national levels are moving forward with carbon pricing policies. Nearly
20 percent of the world’s carbon dioxide emissions are covered (or will soon be covered) by some carbon pricing policy (see Figure 4). 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 mitigate emissions under the 2015 Paris Agreement signaled an interest in using carbon markets to do so (Ramstein et al., 2018).

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, to the extent that virtually all electricity and gasoline consumed is subject to some sort allowance or credit trading program because of an environmental attribute (Aldy, 2020). 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\textsubscript{x} Budget Trading Program for controlling SO\textsubscript{2} and NO\textsubscript{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
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 per ton of $CO_2$ in 2017, with a standard deviation of $33 per ton of $CO_2$. Of
course, these statistics do not include 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 per ton of $CO_2$, ConocoPhillips uses $43
per ton of $CO_2$, and Devon Energy uses $24 per ton of $CO_2$.

While it is possible that financial instruments would help regulated firms to mitigate un-
certainty 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 set-
ting. 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 al-
lowance 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. U.S. government analysts predicted a wide range of initial annual market values through their modeling scenarios, ranging from $177 billion to $683 billion under the Kyoto Protocol and from $54 billion to $254 billion under the Waxman-Markey Bill.

3 Welfare Implications of Firm-Level Uncertainty

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

---

3 Constructed by the authors based on U.S. Energy Information Administration (U.S. EIA, 1998, 2009) and updated to 2018 $. 

15
an emission cap. This aggregate quantity order is then implemented through a decentralized process – the allowance market – with firm-level uncertainty over the market clearing price. One consequence of the complex price formation process is that all types of shocks affect the equilibrium price, so regulated firms are not only impacted by direct shocks to their own abatement cost functions.

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

To understand the welfare consequences of firm-level forecast errors, we develop a modified version of the Weitzman (1974, 2018) framework for evaluating the welfare consequences of price-based and quantity-based instruments under uncertainty. Before we introduce firm-level forecast errors, our model also closely resembles the decentralized allowance markets in Williams (2002) and Yates (2002). We first introduce firm-level forecast errors around

---

4 Other closely related papers in this recent literature on prices-versus-quantities with banking and borrowing are Heutel (2018), Pizer and Prest (2020), 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 initially consider 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. We later explore how policy updating might be
the cost shocks of competitor firms, but as we demonstrate below, our results are mathematically similar when forecast errors instead stem from uncertainty around the impact of overlapping policies. Finally, while our model captures many key features of multi-period price- and quantity-based regulation in the presence of uncertainty, we adopt the standard assumption that abatement is variable. Future work will examine the impact of price and cost uncertainty on dynamic firm-level investment in abatement.

3.1 Model of tradable Quantities

We begin by defining the benefit function associated with reducing some pollutant and the cost function associated with abatement of that pollutant. Let $B_t(Q_t)$ represent the benefits in period $t$ associated with pollution abatement at level $Q_t$.\(^5\) Likewise, let $C_i(q^i_t, \theta^i_t)$ be the cost to firm $i$ associated with producing quantity $q^i_t$ of the pollutant, where $\theta^i_t$ 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 $\sum_{i=1}^{N} C_i(q^i_t, \theta^i_t)$, where $Q_t = \sum_{i=1}^{N} q^i_t$. We assume uniform mixing of the pollutant in question, 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 carbon dioxide and most other greenhouse gas emissions but could be relaxed to model local pollutants. By contrast, the costs of abatement depend on the pollution level achieved by each individual firm.\(^7\)

We assume for tractability that there are two periods in the current regulatory cycle.\(^8\)

---

\(^5\)Weitzman (2018) provides further discussion of the need for this assumption in footnote 8.

\(^6\)Here we assume no uncertainty in the benefit function, which is consistent with Weitzman (2018).

\(^7\)We also assume that the benefits of abatement can be approximated with a marginal benefit function that is separable over time. This assumption seems reasonable in the case of climate change given evidence that the damage function from greenhouse gas emissions may be approximately linear in temperature (Burke et al., 2015). We acknowledge the work of existing papers that have examined how optimal policies may depend on the characteristics of marginal damages (Williams, 2002; Yates, 2002; Gerlagh and Heijmans, 2018), though this discussion is not directly relevant to our modeling of the cost inefficiencies introduced by firm-level uncertainty. We also note that the key insights from our model remain intact when we explicitly model a stock pollutant where avoided damages (benefits) are intertemporally linked; this derivation is available from the authors upon request.

\(^8\)We also assume here that there is no discounting between periods and that the intertemporal permit
Let us consider the regulator’s problem under two instrument options. Under a tax, the regulator sets an optimal price order in the presence of uncertainty by solving the following maximization problem:

$$\max \ E\left[B_1\left(\sum_{i=1}^{N} q_i^1(p_1, \theta_1^i)\right) - \sum_{i=1}^{N} C_i^1(q_i^1, \theta_1^i, \theta_1^i) + B_2\left(\sum_{i=1}^{N} q_i^2(p_2, \theta_2^i)\right) - \sum_{i=1}^{N} C_i^2(q_i^2, \theta_2^i, \theta_2^i)\right]$$

(1)

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

$$\max \ E\left[B_1\left(\sum_{i=1}^{N} q_i^1(p_1(\hat{Q}, \theta_1, \theta_2), \theta_1^i)\right) - \sum_{i=1}^{N} C_i^1(q_i^1(p_1(\hat{Q}, \theta_1, \theta_2), \theta_1^i, \theta_1^i)\right) + B_2\left(\sum_{i=1}^{N} q_i^2(p_2(\hat{Q}, \theta_1, \theta_2), \theta_2^i)\right) - \sum_{i=1}^{N} C_i^2(q_i^2(p_2(\hat{Q}, \theta_1, \theta_2), \theta_2^i, \theta_2^i)\right]$$

(2)

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

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

$$C_i^t(q_i^t, \theta_i^t) = a_i(\theta_i^t) + (C_i' + \theta_i')(q_i^t - \bar{q}_i^t) + \frac{C_i''}{2} (q_i^t - \bar{q}_i^t)^2$$

(3)

where $C_i'$ represents the expected marginal abatement cost at $\bar{q}_i^t$ and $C_i''$ represents the slope of the marginal abatement cost function. As in Weitzman and much of the subsequent 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.

As we show in the full derivation in appendix section A.1 and A.2, the Taylor expansion is defined such that $C_i'$ 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.
literature, we assume for tractability that the abatement cost function is quadratic or can be well approximated by a second-order Taylor expansion. $\theta_i^t$ represents how the random cost shock affects the slope of the abatement cost function for firm $i$, and $a_i(\theta_i^t)$ represents how the cost shock $\theta_i^t$ affects the level of abatement costs. As in Weitzman’s derivation, we assume without loss of generality that $\mathbb{E}[a_i(\theta_i^t)] = 0$ and $\mathbb{E}[\theta_i^t] = 0$. Note that this assumption does not preclude that the conditional expectation of a firm’s cost shock differs from 0; in general, $\mathbb{E}[a_i(\theta_i^t)|a_j(\theta_j^t)] \neq 0$ and $\mathbb{E}[\theta_i^t|\theta_j^t] \neq 0$ for some $i$ and $j$, and $\mathbb{E}[a_i(\theta_i^t)|a_i(\theta_i^t)'] \neq 0$ and $\mathbb{E}[\theta_i^t|\theta_i^t'] \neq 0$ for $t$ and $t'$.

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

$$B(Q_t) = b + B'(\sum_{i=1}^N q_i^t - \bar{q}_i^t) - \frac{B''}{2}(\sum_{i=1}^N q_i^t - \bar{q}_i^t)^2$$

(4)

Here $B'$ captures the marginal benefit at $Q_t = \sum_{i=1}^N \bar{q}_i^t$, and $B''$ captures the slope of the marginal benefit function (where $B'' \geq 0$).

For the optimal tax, the derivation here closely follows Weitzman’s derivation with multiple production units, except we constrain the regulated price to be the same across all units. We again find that the optimal price order $\tilde{p}_t$ is equal to $B' = C'_i$ for all $i$ (and for all $t$ where these parameters are constant). The full derivation is provided in appendix section A.1. Assuming cost minimization, each firm will set its realized marginal cost function equal to this price, yielding the following firm-level response function:

$$q_i^t(\tilde{p}_t, \theta_i^t) = \tilde{q}_i^t = \bar{q}_i^t - \frac{\theta_i^t}{C_i'}$$

(5)

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

$$\sum_{i=1}^N q_i^t(\tilde{p}_t, \theta_i^t) = \sum_{i=1}^N \tilde{q}_i^t - \frac{\theta_i^t}{C_i'} = \tilde{Q}_t - \sum_{i=1}^N \frac{\theta_i^t}{C_i'}$$

(6)

By contrast, for the optimal cap-and-trade program, 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.\(^\text{10}\) 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:

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

(7)

In the first period, the firm response function dictates that firm \(i\) will produce \(\hat{q}_i^1 = \bar{q}_i^1 + \frac{\sum_{j=1}^{N} \frac{\theta_i^1 + \theta_j^2}{2C_i''}}{\sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\theta_i^1}{C_i''};\) the aggregate first-period quantity produced will be \(\hat{Q}_1 = \bar{Q}_1 + \sum_j \frac{\theta_j^2 - \theta_j^1}{2C_i''}.\)

Likewise, in the second period, firm \(i\) will produce \(\hat{q}_i^2 = \bar{q}_i^2 + \frac{\sum_{j=1}^{N} \frac{\theta_i^1 + \theta_j^2}{2C_i''}}{\sum_{j=1}^{N} \frac{1}{C_j''}} - \frac{\theta_i^2}{C_i''};\) the aggregate second-period quantity is then \(\hat{Q}_2 = \bar{Q}_2 + \sum_j \frac{\theta_j^1 - \theta_j^2}{2C_i''}.\) (The full derivation is given in appendix section A.2.)

Finally, following Weitzman, we define the relative advantage of prices over quantities as the expected difference between net benefits from the optimal tax and net benefits from the

\(^{10}\)A cap-and-trade program indirectly implements an aggregate quantity order in the spirit of Weitzman. While an emissions cap and the subsequent trading of allowances allocates the quantity of allowed emissions, the quantity order in the Weitzman framework instead determines the quantity of emissions abatement. While not explicitly considered in this model, uncertainty over business-as-usual (BAU) emissions – and therefore uncertainty over how the quantity of allowed emissions will translate into a quantity of emissions abatement – is an additional source of firm-level uncertainty. Borenstein et al. (2019) discuss the importance of uncertainty over BAU emissions in driving systematic uncertainty over future allowance prices.
optimal cap-and-trade program with banking and borrowing. That is:

\[
\Delta = E\left[ B_1\left( \sum_{i=1}^{N} \tilde{q}_1^i \right) + B_2\left( \sum_{i=1}^{N} \tilde{q}_2^i \right) - \sum_{i=1}^{N} C_1^i(\tilde{q}_1^i, \theta_1^i) - \sum_{i=1}^{N} C_2^i(\tilde{q}_2^i, \theta_2^i) \right] - E\left[ B_1\left( \sum_{i=1}^{N} \hat{q}_1^i \right) + B_2\left( \sum_{i=1}^{N} \hat{q}_2^i \right) - \sum_{i=1}^{N} C_1^i(\hat{q}_1^i, \theta_1^i) - \sum_{i=1}^{N} C_2^i(\hat{q}_2^i, \theta_2^i) \right]
\]  

(8)

Substituting each firm’s response to the price and quantity orders, respectively, we obtain the following expression for the relative advantage of prices over tradable quantities:

\[
\Delta = E \left[ \frac{1}{4} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} - B'' \right) \left( \sum_{i=1}^{N} \frac{\theta_1^i + \theta_2^i}{C''_i} \right)^2 \right]
\]  

(9)

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

### 3.2 Firm Uncertainty Over Cost Shocks: Idiosyncratic Errors

Because the market-clearing price associated with the regulator’s aggregate quantity order (emissions cap) depends on shocks to the marginal abatement cost functions of all firms, a given firm \(i\) 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

\[\text{We can immediately compare this expression to the result in } \text{Weitzman (2018) for the comparative advantage of fixed prices over time-flexible quantities with perfect information for a single representative firm: } \Delta = E\left[\frac{1}{4}(C'' - C''') \cdot \left( \frac{\theta_1 + \theta_2}{C''} \right)^2 \right]. \text{ We also note that this expression is consistent with the results in Williams (2002) and Yates (2002) in the case of perfect substitutability across polluting entities (i.e., a uniformly mixed pollutant).}\]
abatement cost function equal to the cost-effective price \( \hat{p}_t \) associated with aggregate quantity order \( \hat{Q} \), as derived in the previous section, plus some idiosyncratic expectation error derived from its uncertainty over the aggregate cost shocks realized in the market.\(^{12}\) That is, if the cost-effective price is given by \( \hat{p}_1(\hat{Q}, \theta_1, \theta_2) = C' + \frac{\sum_{i=1}^N \theta_i}{C''} \), as derived in the previous section, then firm \( i \)'s forecast errors in the first period would result in \( \text{E}_i[\hat{p}_1(\hat{Q}, \theta_1, \theta_2)] = \hat{p}_1(\hat{Q}, \theta_1, \theta_2) + \epsilon_i = C' + \frac{\sum_{i=1}^N \theta_i}{C''} + \epsilon_i \). Here \( \epsilon_i \) represents firm \( i \)'s idiosyncratic uncertainty around the impact of aggregate cost shocks on the market-clearing price. In the presence of these forecast errors, the firm’s quantity response becomes:

\[
q_i^1(\hat{p}_1, \theta_1, \epsilon_i) = \frac{\hat{p}_1(\hat{Q}, \theta_1, \theta_2) + \epsilon_i - C' - \theta_i}{C''} + \bar{q}_i
\]

(10)

To distinguish from systematic uncertainty that affects all firms alike, we define these idiosyncratic firm-level forecast errors such that they 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.\(^{13}\) In the next section, we introduce systematic forecast errors, which do affect the overall quantity of abatement in each period. To identify the impact of these idiosyncratic 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.\(^{14}\)

By contrast, under a tax, the price is set by regulation and does not depend on private information about other firms’ marginal abatement costs. We maintain our earlier assumption that the regulated price is known with certainty to all firms under a price-based policy, to maintain the consistent assumption across both price and quantity instruments that un-

\(^{12}\)Throughout this paper, we use the term “cost-effective price” to refer to the market-clearing allowance price that achieves the regulator’s aggregate quantity limit at lowest cost. Firms generally do not have perfect foresight about what this cost-effective price will be.

\(^{13}\)That is, we require \( \sum_{i=1}^N \frac{\epsilon_i}{C''} = 0 \), which ensures that firms’ idiosyncratic expectation errors will collectively cancel with each other in determining the overall quantity.

\(^{14}\)This assumption does not substantively affect our results and could be easily relaxed.
certainty around policy design and stringency has already been resolved. Therefore, firms’ responses to a price order do 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 idiosyncratic forecast errors under quantity-based regulation but holding constant the net benefits of price-based regulation. Our welfare expression now becomes:

\[ \Delta = \mathbb{E} \left[ \frac{1}{4} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C_i''}} - B''' \right) \left( \sum_{i=1}^{N} \frac{\theta_i + \theta_i^2}{C_i''} \right)^2 + \sum_{i=1}^{N} \frac{\epsilon_i^2}{2C_i''} \right] \]  

Equation (11) 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 idiosyncratic 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 aggregate quantity order (the emissions cap), 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 idiosyncratic 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.

---

\[ 15 \]

This derivation (and those in subsequent sections) assumes that the regulator has no foresight about the extent of firm forecast errors. That is, \( \mathbb{E}[\epsilon_i] = 0 \) for all \( i \).

\[ 16 \]

The impact of firm-level forecast errors is also decreasing in the slope of the marginal abatement cost function, \( C_i'' \). The intuition here is that individual firm quantities are less responsive to (expected) price when the slope of the marginal abatement function is large. (Recall that the firm’s quantity response function is \( q_i^t = \bar{q}_i + \frac{p + \theta_i - C_i''}{\epsilon_i} \).)
3.3 Firm Uncertainty Over Cost Shocks: Systematic Errors

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 abatement cost functions 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 cost-effective 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 future shocks. As a result, these systematic forecast errors may also cause firms to collectively over- or under-abate relative to the abatement level that would be intertemporally optimal ex post.[17]

To understand the welfare consequences of this information revelation over time, we relax the earlier restriction that first-period forecast errors do not affect the aggregate quantity produced in this first period. Instead, firms now have price expectations of the form:

$$E_i[\hat{p}_1(\hat{Q}, \theta_1, \theta_2)] = \hat{p}_1(\hat{Q}, \theta_1, \theta_2) + \epsilon_i + \xi_1 = C' + \frac{\sum_{i=1}^{N} \theta_i}{\sum_{i=1}^{N}} + \epsilon_i + \xi_1,$$

where $\xi_1$ represents a systematic first-period expectation error over future cost shocks that is common to all firms. The aggregate quantity in the first period may be higher or lower than what is intertemporally optimal, depending on the sign of $\xi_1$.

To ensure that the regulator’s overall quantity limit is still met by the end of the final reg-

---

[17] 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 allowance price.
ulatory period, the aggregate second-period quantity must also shift upwards or downwards to compensate. The market-clearing price in the second period therefore adjusts accordingly, where \( \hat{p}'_2(\hat{Q}, \theta_1, \theta_2, \xi_1) = C' + \frac{\sum_{i=1}^{N} \frac{\theta_i' + \theta_i}{2C_i}}{\sum_{i=1}^{N} \frac{1}{C_i}} - \xi_1 \). Full details of the changes in expected benefits and costs under quantity orders with systematic forecast errors are provided in appendix section A.5. Because this type of forecast error does not apply to tax policies, the expected benefits and costs of price orders are again unchanged. The relative advantage of prices over quantities with banking and borrowing now becomes:

\[
\Delta = E \left[ \frac{1}{4} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C_i}} - B'' \right) \left( \sum_{i=1}^{N} \frac{\theta_i' + \theta_i}{C_i} \right)^2 + \frac{1}{2} \left( \sum_{i=1}^{N} \frac{\xi_i'^2}{C_i} \right) \right.
\]

\[
+ E \left( \sum_{i=1}^{N} \frac{\xi_i'^2}{C_i} \right) \right]^{12}
\]

Considering that systematic 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 least-cost distribution of quantity across compliance periods, captured by the new term \( E[\sum_{i=1}^{N} \frac{\xi_i'^2}{C_i}] \), 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 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 systematic 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 \( \sum_{i=1}^{N} \frac{\xi_i'}{C_i} \) and \( \sum_{i=1}^{N} \frac{\theta_i' - \theta_i}{2C_i} \) are negatively correlated. The intuition here is that tradable allowance programs ordinarily deviate from perfect benefit smoothing in response to cost shocks in specific periods; to the extent that the systematic forecast errors dampen firm responses to these period-specific cost shocks, benefit smoothing may improve. 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 considera-
tion 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.\footnote{In the case of greenhouse gases, $B''$ may be close to zero given linear damages \cite{Burke2015}. This term also drops out for pure stock pollutants, modeled in the spirit of \cite{Gerlagh2018}; this derivation is available from the authors upon request.}

### 3.4 Firm Uncertainty Over Overlapping Policies

In addition to uncertainty about cost shocks, whether idiosyncratic or systematic, firms
may also make forecast errors around the market-clearing price due to uncertainty about
the impact of overlapping policies. As discussed previously, under many market-based envi-
ronmental policies, a subset of regulated firms are also subject to additional standards from
overlapping policies, which are often more stringent or more prescriptive than the price or
quantity instrument. For firms where the prescribed technology is not the lowest-cost option
for reducing emissions, the existence of this second set of standards raises the total costs for
meeting the requirements of the overlapping price- or quantity-based policy.\footnote{Insofar as there are fixed costs to participating in allowance or credit trading markets, some firms
may elect to treat their allowance allocations as if they represented command-and-control regulation. This
behavior may produce similar impacts as overlapping policies, where some subset of firms do not abate at
the efficient level. The results described in this section would apply to this scenario as well.} Under a price
instrument, the inefficiencies due to overlapping policies are limited to the subset of firms or
business units that are subject to the more prescriptive regulation. Under an allowance or
credit trading program, by contrast, these cost-inefficiencies affect both firms covered by the
more prescriptive regulation and non-covered firms, by affecting the market-clearing price
for allowances or credits. Insofar as firms are uncertain about the impact of these overlap-
ping policies on the allowance price, we again observe expected welfare losses due to forecast
errors for a tradable quantity instrument relative to a price instrument.
To illustrate, consider a one-period market-based policy where a subset $R$ firms are subject to an additional regulatory standard, while the remaining $M$ firms are subject only to the market-based policy (where $R + M = N$, the total number of firms in the market). Assume that the supplementary regulation causes firms to abate at the level $\hat{q}_j^R$, where $\sum_{j=1}^R \hat{q}_j^R - \sum_{j=1}^R \hat{q}_j = \kappa \geq 0$. That is, the supplementary regulation results in a total quantity of abatement from this subset of firms that is higher than the ex ante efficient quantity for those firms.

Under a price regime, the $M$ non-covered firms behave no differently, and continue to set $q_i = \hat{q}_i - \frac{\theta_i}{C''_i}$. However, under tradable quantities, the aggregate quantity target still determines overall abatement in the market, so the market-clearing price must adjust to account for the excess $\kappa$ compliance produced by the $R$ firms subject to the supplementary regulation. The market-clearing price will therefore adjust to $p^\kappa = C' + \frac{\sum_{i=1}^M \theta_i}{\sum_{i=1}^M \frac{1}{C''_i}} - \frac{M \kappa}{\sum_{i=1}^M \frac{1}{C''_i}}$, and the non-covered firms will respond with quantity $q_i^\kappa = \frac{p^\kappa - C' - \theta_i}{C''_i} + \hat{q}_i$. When we compare the expected welfare of prices versus tradable quantities under this scenario, we obtain the following relation:

$$
\Delta = E \left[ \frac{1}{2} \left( \frac{1}{\sum_{i=1}^M \frac{1}{C''_i}} - B'' \right) \left( \sum_{i=1}^M \frac{\theta_i}{C''_i} \right)^2 + \frac{1}{2} \left( \frac{1}{\sum_{i=1}^M \frac{1}{C''_i}} - B'' \right) \kappa^2 \right]
$$

As we see here, the impact of the overlapping policy on the expected welfare of a price instrument relative to a tradable quantity instrument reduces to the standard “relative slopes” comparison. The intuition is straightforward: the overlapping policy causes a subset of firms to over-comply relative to what would be most cost effective; the overcompliance for this subset of firms occurs under either a tax or a cap-and-trade program. Under a tradable

---

20In this extension of the model, we abstract away from the distinction between idiosyncratic and systematic forecast errors discussed in the previous section and focus simply on a one-period model. We note, however, that forecast errors around the impact of overlapping policies could be either specific to each firm or common to all affected firms.
quantity regime, the market-clearing price adjusts downward such that the aggregate target is still met, resulting in lower abatement than would be cost effective among the remaining firms. Under a price instrument, the remaining firms still abate at the cost-effective level, but there is excess compliance across the full set of firms. The tradeoff between these two sources of inefficiency depends on the relative slopes of the marginal benefit and marginal cost functions.

This comparison changes when we introduce the prospect of firm-level uncertainty. Some firms not subject to the supplementary regulation may not know the full extent of over-compliance among doubly regulated firms. Instead, each firm $i$ might make some forecast error $\eta_i$ about how the overlapping policy will affect the market-clearing price for tradable allowances (credits). Under the tradable quantity regime, firm $i$ will therefore choose an abatement strategy given its price expectation $E_i[p^\kappa] = C' + \sum_{j=1}^{M} \frac{\theta_j}{C'_j} - \sum_{j=1}^{M} \frac{\kappa}{C'_j} + \eta_i$, where $\eta_i$ represents firm $i$’s forecast error around the impact of the overlapping policy. In this scenario, the expected welfare of a price-based policy relative to tradable allowances or credits is given by:

$$\Delta = E \left[ \frac{1}{2} \left( \frac{1}{\sum_{i=1}^{M} \frac{1}{C'_i}} - B'' \right) \left( \sum_{i=1}^{M} \frac{\theta_i}{C'_i} \right)^2 + \frac{1}{2} \left( \frac{1}{\sum_{i=1}^{M} \frac{1}{C'_i}} - B'' \right) \kappa^2 + \left( \sum_{i=1}^{M} \frac{\eta_i^2}{2C''_i} \right) \right] \quad (13)$$

Mathematically, the impact of uncertainty around overlapping policies is identical to the results derived above for uncertainty around costs shocks; the full derivation is presented in appendix section A.6. Once again, we recover our initial prices versus tradable quantities result, here for the overlapping policies scenario, as well as an additional term that captures the cost inefficiencies created by firm-level forecast errors. All else equal, the presence of these forecast errors again pushes the regulator to prefer prices over tradable quantities.
3.5 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 encourage the welfare-maximizing 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.
4 Policy Responses to Firm-Level Uncertainty

We have established that firm-level uncertainty may have significant implications for the cost-effectiveness of allowance and credit trading programs, by modeling these additional uncertainties in a Weitzman-style framework that has traditionally focused on the regulator’s uncertainty. However, political economy considerations may continue to push policymakers to favor quantity instruments irrespective of any welfare advantage of price instruments. In this section, we consider that certain policy design tools may be available to reduce – even if not eliminate – the impacts of this firm-level uncertainty. We discuss the potential role for price ceilings and price floors (“price collars”) and for policy updating.

4.1 Price Collars

Under a tradable quantities program, a price ceiling serves as a maximum price above which the market-clearing price for allowances or credits is not allowed to rise\textsuperscript{21} The regulator commits to selling additional allowances whenever the price ceiling is reached, charging regulated entities this maximum price but no more. The effect is to relax the overall quantity constraint to avoid large increases in firms’ compliance costs. A price floor works analogously, where the regulator commits to buying allowances at a specified minimum price whenever needed, or specifies a reserve price in allowance auctions. Taken together, the price ceiling and price floor represent a hybrid policy between a pure quantity-based instrument and a pure price-based instrument\textsuperscript{22} We discuss here the role of price collars in reducing the variance of firm-level forecast errors, by providing firms with additional information about the range of allowance prices that they may face. This information directly reduces the expected cost of firm forecast errors. Just as the cost inefficiencies of firm-level uncertainty should be considered alongside standard comparisons of relative slopes, reductions in these cost

\textsuperscript{21}In the model that follows, we focus on a “hard” price ceiling where the regulator commits to maintaining a certain maximum price, as opposed to a “soft” price, where the regulator commits only to releasing a finite allowance reserve if the market-clearing price rises above some level.

\textsuperscript{22}Roberts and Spence (1976) and Weitzman (1978) have demonstrated that this type of hybrid policy may be welfare-enhancing relative to a pure quantity or price instrument.
inefficiencies should be included in standard evaluations of price collar mechanisms.

We begin by discussing how to evaluate price collars in a Weitzman-style framework. In general, the impact of the price collar falls into one of three cases, depending on the relationship between the price ceiling or floor, the cost-effective price given the aggregate quantity order (emission cap), and the ex post optimal price (first-best price) given the realization of shocks in the market. These scenarios are depicted in Figure 5 for a price ceiling in a market with a representative firm.

(a) Price Ceiling: \(|MB| < |MC|\)

(b) Price Ceiling: \(|MB| > |MC|\) and \(\bar{p} > p^*\)

(c) Price Ceiling: \(|MB| > |MC|\) and \(\bar{p} < p^*\)

Figure 5: Impact of Price Ceiling on Deadweight Loss, Relative to Pure Quantities Policy. The lighter triangles represent the deadweight loss associated with the cost-effective price from allowance trading \((p^{C\&T})\), while the darker triangles represent the deadweight loss associated with the price ceiling \(\bar{p}\). (There is no deadweight loss associated with the ex post optimal price \(p^*\).)
First, whenever the slope of the marginal cost function is steeper than the slope of the marginal benefit function, the price ceiling is welfare-improving relative to an unconstrained cap-and-trade program. In this case, a price-based policy is preferred to a quantity-based policy, but may not be possible for political economy or other reasons; a price collar recreates some characteristics of a fixed price policy and therefore increases welfare under the tradable quantities regime. Second, whenever the price ceiling is higher than the ex post optimal price, the deadweight loss under the price ceiling is (weakly) smaller than if the price is permitted to float under the cap-and-trade program. Finally, when the price ceiling is lower than the ex post optimal price, the impact on deadweight loss is ambiguous and depends on the relative magnitude of the difference between the ex post optimal price and the price ceiling versus the difference between the optimal price and the cost-effective price from allowance trading.

To express these relationships analytically for a market with heterogeneous firms, note that any given price ceiling can be expressed in terms of some threshold for the weighted average of cost shocks across the market. That is, \( \bar{p} = \frac{\bar{t}}{\sum_{i=1}^{N} \frac{\theta_i}{C_i}} + C' \). Therefore, the probability that the equilibrium allowance price reaches the price ceiling is equivalent to the probability that the weighted sum of cost shocks rises above this threshold \( \bar{t} \). If we define \( G(\cdot) \) as the CDF of the weighted sum of cost shocks \( \sum_{i=1}^{N} \frac{\theta_i}{C_i} \), then we can express the probability that the price ceiling is triggered as \( 1 - G(\bar{t}) \); the probability that the price ceiling is not triggered is \( G(\bar{t}) \)\(^{23}\)

\[
\text{Prob}(p^{C&K} > \bar{p}) = \text{Prob}(\sum_{i=1}^{N} \frac{\theta_i}{C_i} > \bar{t}) = 1 - G(\bar{t})
\]

Likewise, the probability of triggering the price floor is equivalent to the probability that the weighted sum of cost shocks falls below some threshold \( \bar{t} \):

\[
\text{Prob}(p^{C&K} < \bar{p}) = \text{Prob}(\sum_{i=1}^{N} \frac{\theta_i}{C_i} < \bar{t}) = G(\bar{t})
\]

The expected welfare of introducing a price ceiling \( \bar{p} = \frac{\bar{t}}{\sum_{i=1}^{N} \frac{\theta_i}{C_i}} + C' \) and a price floor

\(^{23}\)Here we consider a tradable quantities policy over one period.
\[ p = \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} + C' \]

in a tradable quantities program is given by:

\[
\Delta = (1 - G(\bar{t})) \cdot E \left[ \frac{1}{2} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} - B'' \right) \left( \bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \right)^2 \right] + G(\bar{t}) \cdot E \left[ \frac{1}{2} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} - B'' \right) \left( \bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \right)^2 \right]
\]

(14)

To interpret this expression, we note first that the expected welfare of tradable quantities with and without a price collar is exactly the same with probability \( G(\bar{t}) - G(\bar{t}) \), so the expected difference in welfare depends only on the scenarios where the price ceiling or price floor is triggered. In this expression, the second term in each set of square brackets is (weakly) positive, since \( \bar{t} \geq 0 \) for a price ceiling and \( \sum_{i=1}^{N} \frac{\theta_i}{C''_i} > \bar{t} \) whenever the price ceiling is binding, and \( \bar{t} \leq 0 \) for a price floor and \( \sum_{i=1}^{N} \frac{\theta_i}{C''_i} < \bar{t} \) whenever the price floor is binding.\(^{24}\) The first term in each set of square brackets depends on the relative slopes of the marginal benefit and marginal cost functions; if the marginal cost function is steeper than the marginal benefit function, then the introduction of the price collar is unambiguously welfare enhancing, as discussed with our graphical analysis above. If instead the marginal benefit function is steeper, then whether the price collar is welfare enhancing depends on the relative magnitudes of the first versus second terms, which reflects the other scenarios discussed as part of our graphical analysis.

From this set-up, it is straightforward to introduce forecast errors. With probability \( G(\bar{t}) - G(\bar{t}) \), neither the price ceiling nor the price floor is binding, but the greater information

\(^{24}\)We assume that the regulator does not set the price ceiling below the expected optimal price, \( E[p^*] = C' \), since this policy could never be optimal, so \( \bar{t} \geq 0 \). Likewise, the regulator does not set the price floor above the expected optimal price, so \( \bar{t} \leq 0 \).
provided to firms about the range of possible prices reduces the variance of forecast errors in the regime with the price collar: 
\[ E[\sum_{i=1}^{N} \hat{\epsilon}_i^2] \geq E[\sum_{i=1}^{N} \bar{\epsilon}_i^2], \]
where \( \hat{\epsilon}_i \) represents the forecast error of firm \( i \) given the price collar.\(^{25}\) With probability \( 1 - G(\bar{t}) + G(t) \), either the price ceiling or the price floor binds, and the change in expected welfare is again given by the expression derived above, plus the difference in forecast errors across the two regimes. This impact of the price collar on firm-level expectations of price comports with the intuition provided in \textit{Salant et al. (2020)}, that price collars may affect firms’ expectations of allowance prices in a dynamic emissions trading market with banking, even when those price collars are not binding.

\[
\Delta = (1 - G(\bar{t})) \cdot E \left[ \frac{1}{2} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C_i''}} - B'' \right) \left( \bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_i''} \right)^2 + \bar{t} \left( \sum_{i=1}^{N} \frac{\theta_i}{C_i''} - \bar{t} \right) \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C_i''}} \right) \mid \sum_{i=1}^{N} \frac{\theta_i}{C_i''} > \bar{t} \right] \\
+ G(t) \cdot E \left[ \frac{1}{2} \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C_i''}} - B'' \right) \left( \bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_i''} \right)^2 + \bar{t} \left( \sum_{i=1}^{N} \frac{\theta_i}{C_i''} - \bar{t} \right) \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C_i''}} \right) \mid \sum_{i=1}^{N} \frac{\theta_i}{C_i''} < \bar{t} \right] \\
+ E \left[ \frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i''} - \frac{1}{2} \sum_{i=1}^{N} \frac{\bar{\epsilon}_i^2}{C_i''} \right] \quad \text{Additional Term} \geq 0
\]

Relative to the baseline model of the price collar without forecast errors, we see from this expression that we must also consider the information provision role of price collars in addition to the standard analysis described above. Indeed, it is possible that the expected change in welfare from introducing the price collar would be negative in the absence of firm-level uncertainty around prices, but becomes positive given the reduction in variances of price forecast errors. The full derivation of these welfare expressions is provided in appendix section A.7. In general, our welfare expression becomes more complicated if we relax the assumption that forecast errors do not affect the overall quantity produced in a given period.\(^{25}\)

\(^{25}\)Here we return to our earlier assumption that the forecast errors collectively have no impact on the overall quantity produced in a given period, or \( \sum_{i=1}^{N} \frac{\hat{\epsilon}_i}{C_i''} = \sum_{i=1}^{N} \frac{\bar{\epsilon}_i}{C_i''} = 0 \). This assumption is especially reasonable for price ceilings and floors that are symmetric around the ex ante expected allowance price.
that is, if we include systematic as well as idiosyncratic forecast errors. In that case, our welfare results then depend on how the presence of the price ceiling alters the magnitude and direction of deviations from the optimal aggregate quantity, for which our theory does not provide clear predictions.

4.2 Policy Updating

Next we consider the potential role of policy updating in mitigating the cost inefficiencies created by firm-level forecast errors. Policy updating has received attention in recent literature, with [Pizer and Prest (2020)] suggesting that tradable quantity instruments can recover the first-best outcome by updating the quantity target after information about cost shocks is revealed. [Heutel (2018)] shows that this result depends on the assumption that cost shocks are identical across periods, and the expected change in welfare is ambiguous when cost shocks are correlated but not identical over time. However, these findings both depend on the assumption that a representative firm has perfect foresight over all cost shocks affecting the market.

We begin by introducing policy updating in our model, generalizing the framework in [Pizer and Prest (2020)] and [Heutel (2018)] to allow for heterogeneous firms. This set-up allows us to model the impact of individual firms’ uncertainty over their competitors’ cost shocks on opportunities for policy updating. To highlight the core insights of this model, we consider the case where cost shocks are identical over time for a given firm, following [Pizer and Prest (2020)]. That is, in the first compliance period, each firm learns its own cost shock, which it knows will persist into all subsequent periods, but does not know the value of other firms’ cost shocks before making its initial compliance decision. After the market clears at the end of the first compliance period, each firm observes the other shocks in the market, as does the regulator. Since both the firms and the regulator know that these shocks will persist in future periods, all uncertainty has been resolved by this point, but the impact of firms’ initial forecast errors persists.
Unlike in Pizer and Prest (2020) or Heutel (2018) with time-invariant shocks, we can no longer recover the first-best outcome in the presence of firm-level uncertainty. Because firms’ generally incur forecast errors in predicting the impact of cost shocks in the market, and therefore in predicting market-clearing price, they also generally have incorrect expectations over the extent of policy updating. However, we demonstrate that policy updating, following the updating rule considered in these earlier papers, still reduces the impact of these forecast errors on expected welfare. The intuition is that ex post policy updating dampens the impact of firm-level cost shocks, thereby dampening the impact of the forecast errors.

To see this result, note that the first-best quantity target for a two-period regulation is given by: $Q^* = \bar{Q} - \frac{2 \sum_{i=1}^{N} \theta_i}{1 - B'' \sum_{i=1}^{N} \frac{\theta_i}{C''_i}}$, where $\bar{Q}$ is the expected optimal quantity target. (The full derivation is provided in appendix section A.8.) After the first compliance period, the regulator is now able to observe realized cost shocks, determine the exact value of $Q^*$, and update the aggregate quantity target accordingly. Assuming that the regulator sets the optimal ex ante quantity target, this first-best result requires updating the initial quantity target by $\delta^* = Q^* - \bar{Q} = \frac{-2 \sum_{i=1}^{N} \theta_i}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}}$. For consistency with Pizer and Prest (2020) and Heutel (2018), we assume that firms anticipate that the regulator will update its quantity target after additional information has been revealed. For example, the regulator may precommit to a policy rule that requires updating the quantity target after the first compliance period. Consequently, firms will adjust their own behavior in the first compliance period based on whatever signals they have received about the magnitude of the regulator’s eventual quantity adjustment.

However, in our model with heterogeneous firms, the regulated entities may not have perfect information about all other firms’ cost shocks prior to making their compliance decisions, even after their own shocks have been realized. Consequently, just as we might expect firms to face forecast errors around the impact of competitors’ shocks on the market-clearing allowance price, so too would we expect firms to make forecast errors around the

\[26\] Here we assume that $\theta_1 = \theta_2 = \theta'$. As we showed earlier, $\bar{Q}$ is equal to the aggregate quantity target that the regulator sets before the first compliance period, $\bar{Q}$.
regulator’s quantity updating. These two channels through which uncertainty over other firms’ cost shocks enters into each firm’s optimization function are partially offsetting.

The cost-effective market-clearing price changes in response to this policy updating and is now given by \( p^\delta = C' + \sum_{i=1}^{N} \frac{\theta_i}{C''_i} + \frac{\delta^*}{2} \sum_{i=1}^{N} \frac{1}{C''_i} \). Each firm \( i \) forms some expectation of this market-clearing price. Since the extent of policy updating is a function of the aggregate cost shocks realized in the market, we can also write firms’ expectation errors around the extent of policy updating in terms of their expectation errors around the aggregate cost shocks, \( \epsilon_1^i \) and \( \xi_1 \). Therefore, each firm’s price expectation is given by:

\[
E_i[p^\delta] = C' + \sum_{i=1}^{N} \frac{\theta_i}{C''_i} + \frac{\delta^*}{2} \sum_{i=1}^{N} \frac{1}{C''_i} + (\epsilon_1^i + \xi_1)(1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}})
\]

In this expression, note that the magnitude of the idiosyncratic and systematic forecast errors is adjusted downwards, as the multiplier in parentheses lies between 0 and 1. The magnitude of the forecast error affecting the firm’s quantity response is smaller given this policy updating rule. If we again use \( \tilde{\epsilon}_1^i + \tilde{\xi}_1 \) to denote the reduced forecast errors under this modified policy design, the expected change in welfare from introducing policy updating in a two-period tradable quantities program is given by:

\[
\Delta = \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i} + B''} \right) \left( \frac{\delta^*}{2} \right) + \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i} + B''} \right) \left[ \sum_{i=1}^{N} \frac{\xi_1}{C'''_i} \right] - \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i} + B''} \right) \left[ \sum_{i=1}^{N} \frac{\tilde{\epsilon}_1^i}{C'''_i} \right] \]

\[
+ \frac{1}{2} \left( \sum_{i=1}^{N} \frac{(\epsilon_1^i)^2}{C'''_i} - \sum_{i=1}^{N} \frac{(\tilde{\epsilon}_1^i)^2}{C'''_i} \right)
\]

This expression compares expected welfare under tradable quantities with policy updating to tradable quantities without policy updating, in the presence of firm forecast errors. As

\[27\] We previously defined the idiosyncratic and systematic expectation errors around cost shocks in the market as \( \epsilon_1^i + \xi_1 = E[\sum_{i=1}^{N} \frac{\theta_i}{C''_i}] - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \). Therefore, we can also write \( E_i[\delta^*] - \delta^* = \frac{-2(\sum_{i=1}^{N} \frac{\theta_i}{C''_i})(\epsilon_1^i + \xi_1)}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}} \).

\[28\] In this context, \( \tilde{\epsilon}_1^i = \epsilon_1^i (1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}}) \) and \( \tilde{\xi}_1 = \xi_1 (1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}}) \).
we see from this expression, there are three types of welfare improvements from ex-post policy updating: first, the ability to more closely reach the first-best quantity level, which occurs irrespective of firm expectation errors; second, the reduced impact of systematic forecast errors on firm behavior; and third, the reduced impact of idiosyncratic forecast errors. While policy updating does not entirely resolve the cost inefficiencies we have highlighted for allowance and credit trading programs with firm-level uncertainty, it does help to mitigate these losses.

5 Conclusion

In this paper, we examine the impact of firm-level uncertainty over allowance prices in cap-and-trade and tradable credit markets – a form of residual uncertainty which is inherent to this type of policy instrument. Motivated by numerous empirical illustrations of cap-and-trade and tradable quantity markets deviating from cost-effectiveness, we develop a theory model that elucidates the welfare implications of instrument choice with allowance and credit price forecast errors. Building on the Weitzman prices-versus-quantities framework, we highlight the additional costs associated with imperfect information about future market-clearing prices when modeling cap-and-trade as an “aggregate quantity order” that allocates quantities in a decentralized manner to firms. All else equal, the cost inefficiencies created by firm forecast errors would encourage the welfare-maximizing regulator to favor price instruments over tradable quantity instruments.

We focus our model on a uniformly-mixed pollutant. While a special case, it addresses carbon dioxide and other greenhouse gas emissions, which represent the most pressing environmental policy challenge of the 21st century – and one that drew considerable attention from Marty Weitzman in his scholarship over the last several decades of his career. Given the scale of investment necessary to combat climate change – on the order of trillions of dollars in the coming decades – and the growing policy interest around the world in market-based
instruments in abating greenhouse gas emissions, our work highlights another dimension of the instrument choice problem in promoting cost-effective and economically efficient emission reductions. We also show how variations in instrument design – such as price collars and policy updating of cap-and-trade programs – can mitigate some, but not all, of the welfare costs associated with the residual uncertainty in tradable quantity instruments. We also discuss how intertemporal benefit smoothing for flow pollutants (or non-uniformly mixed pollutants) may also affect the regulator’s choice of instrument.

Examining the interactions between firm-level forecast errors, uncertainty over future abatement costs, and other shocks to pollution markets represents a fruitful direction for research. Many economists responded to the insights of the standard prices-versus-quantities framework by exploring its empirical implications, such as the estimation of the damage and abatement cost functions for a variety of pollutants, including carbon dioxide (e.g., Pizer [2002]). Further examination of the economic fundamentals driving allowance and credit prices, and rigorous evaluations of cost-effectiveness anomalies in these markets could further enhance understanding of instrument choice and design. In future research, we plan to build on the static abatement decisions analyzed here, to understand the impact of these various shocks on long-lived firm investment, including R&D.

References


A Derivation of Welfare Result

A.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:

$$\max_{p_1, p_2} E[B_1\left(\sum_{i=1}^{N} q^i_1(\tilde{p}_1, \theta^i_1)\right) - \sum_{i=1}^{N} C^i_1(q^i_1(\tilde{p}_1, \theta^i_1), \theta^i_1) + B_2\left(\sum_{i=1}^{N} q^i_2(\tilde{p}_2, \theta^i_2)\right) - \sum_{i=1}^{N} C^i_2(q^i_2(\tilde{p}_2, \theta^i_2), \theta^i_2)]$$

which yields the following first-order conditions:

$$E[\sum_{i=1}^{N} \frac{\partial B_1}{\partial Q_1} \cdot \frac{\partial Q_1}{\partial q^i_1} \cdot \frac{dq^i_1}{dp_1}] = E[\sum_{i=1}^{N} \frac{\partial C^i_1}{\partial q^i_1} \cdot \frac{dq^i_1}{dp_1}]$$

$$E[\sum_{i=1}^{N} \frac{\partial B_2}{\partial Q_2} \cdot \frac{\partial Q_2}{\partial q^i_2} \cdot \frac{dq^i_2}{dp_2}] = E[\sum_{i=1}^{N} \frac{\partial C^i_2}{\partial q^i_2} \cdot \frac{dq^i_2}{dp_2}]$$

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

$$\frac{\partial C^i_t}{\partial q^i_t} = p_t = C' + \theta^i_t + C''(q^i_t - \bar{q}_t)$$

Rearranging gives:

$$q^i_t(p_t, \theta^i_t) = q^i_t = \frac{p_t - \theta^i_t}{C''} + \bar{q}_t$$

Differentiating with respect to $p_t$ gives:

$$\frac{dq^i_t}{dp_t} = \frac{1}{C''}$$

Substituting this result into the regulator’s first-order condition for the optimal price...
order and recognizing that all firms set \( \frac{\partial C_i}{\partial q_i} = \tilde{p}_t \), we have:

\[
\frac{\partial B_1}{\partial Q_1} \cdot \sum_{i=1}^{N} 1 \frac{1}{C''_i} = \tilde{p}_1 \sum_{i=1}^{N} \frac{1}{C''_i}
\]

\[
\frac{\partial B_2}{\partial Q_2} \cdot \sum_{i=1}^{N} 1 \frac{1}{C''_i} = \tilde{p}_2 \sum_{i=1}^{N} \frac{1}{C''_i}
\]

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:

\[
q_i^t(\tilde{p}_t, \theta_i^t) = \tilde{q}_i = -\frac{\theta_i^t}{C''_i} + \bar{q}_i
\]

### A.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:

\[
\max_{\hat{Q}} E[B_1(\sum_{i=1}^{N} q_i^1(p_1(\hat{Q}, \theta_1), \theta_1))) - \sum_{i=1}^{N} C_i^1(q_i^1(p_1(\hat{Q}, \theta_1), \theta_1)), \theta_1) + B_2(\sum_{i=1}^{N} q_i^2(p_2(\hat{Q}, \theta_2), \theta_2))) - \sum_{i=1}^{N} C_i^2(q_i^2(p_2(\hat{Q}, \theta_2), \theta_2), \theta_2)]
\]

which yields the first-order condition:

\[
E\left[\sum_{i=1}^{N} \frac{\partial B_1}{\partial Q_1} \cdot \frac{\partial Q_1}{\partial q_i} \cdot \frac{d\tilde{p}_1}{dQ} + \sum_{i=1}^{N} \frac{\partial B_2}{\partial Q_2} \cdot \frac{\partial Q_2}{\partial q_i} \cdot \frac{d\tilde{p}_2}{dQ}\right] = E\left[\sum_{i=1}^{N} \frac{\partial C_i^1}{\partial q_i} \cdot \frac{d\tilde{p}_1}{dQ} + \sum_{i=1}^{N} \frac{\partial C_i^2}{\partial q_i} \cdot \frac{d\tilde{p}_2}{dQ}\right]
\]

First, note that we still have \( \frac{\partial q_i}{\partial p_i} = \frac{1}{C''_i} \), a constant. We can also use the price response
equation from above to derive \( \frac{dp}{dQ} \)

\[
\sum_{i=1}^{N} q_i^t(p_t, \theta_t^i) = \sum_{i=1}^{N} \frac{p_t - C_i' - \theta_t^i}{C_i''} + \bar{q}_t
\]

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

Therefore, aggregate quantity is given by:

\[
Q = \sum_{i=1}^{N} \frac{p_1 - C_i' - \theta_1^i}{C_i''} + Q_1 + \sum_{i=1}^{N} \frac{p_2 - C_i' - \theta_2^i}{C_i''} + Q_2
\]

Differentiating with respect to \( Q \) gives:

\[
1 = \sum_{i} \frac{1}{C_i''} \frac{dp_1}{dQ} + \sum_{i} \frac{1}{C_i''} \frac{dp_2}{dQ} \Rightarrow \frac{dp_1}{dQ} + \frac{dp_2}{dQ} = 1/(\sum_{i} \frac{1}{C_i''})
\]

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:

\[
\sum_{i=1}^{N} B' \cdot \frac{1}{C_i''} \cdot \frac{1}{2 \sum_{j} \frac{1}{C_j''}} + \sum_{i=1}^{N} B' \cdot \frac{1}{C_i''} \cdot \frac{1}{2 \sum_{j} \frac{1}{C_j''}}
\]

\[
= \sum_{i=1}^{N} E[\frac{\partial C_i^{i*}}{\partial q_t^i}] \cdot \frac{1}{C_i''} \cdot \frac{1}{2 \sum_{j} \frac{1}{C_j''}} + \sum_{i=1}^{N} E[\frac{\partial C_2^{i*}}{\partial q_t^i}] \cdot \frac{1}{C_i''} \cdot \frac{1}{2 \sum_{j} \frac{1}{C_j''}}
\]

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

\[
\sum_{i=1}^{N} B' \cdot \frac{1}{C_i''} \cdot \frac{1}{\sum_{j} \frac{1}{C_j''}} = \sum_{i=1}^{N} E[\hat{p}(\hat{Q}, \theta_1, \theta_2)] \cdot \frac{1}{C_i''} \cdot \frac{1}{\sum_{j} \frac{1}{C_j''}}
\]
Which ultimately yields (for the regulator’s optimal aggregate quantity order \( \hat{Q} \)):

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

Note that we have defined \( \bar{q}_t \) such that \( \mathbb{E}\left[\frac{\partial B}{\partial q_t}\right] = \mathbb{E}\left[\frac{\partial C_i}{\partial q_t}\right] \). 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 cost-effective price resulting from the optimal quantity instrument is not necessarily equal to \( B' \) or \( C' \), but instead depends on the realizations of cost shocks \( \theta \). 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:

\[
Q_1 + Q_2 = \hat{Q}_1 + \sum_{i=1}^{N} \hat{p}_1 - C'_i - \theta_1^i + \hat{Q}_2 + \sum_{i=1}^{N} \hat{p}_2 - C'_i - \theta_2^i
\]

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:

\[
\hat{p}_1(\hat{Q}, \theta_1, \theta_2) = \hat{p}_2(\hat{Q}, \theta_1, \theta_2) = C' + \sum_{i} \frac{\theta_1^i + \theta_2^i}{2C''_i} \sum_{j} \frac{1}{C''_j}
\]

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

\[
q_1^i(\hat{p}_1, \theta_1^i) = \frac{\hat{p}_1 - C'_i - \theta_1^i}{C''_i} + \bar{q}_1^i = \frac{C' + \sum_j \frac{\theta_1^j + \theta_2^j}{2C''_j} - C'_i - \theta_1^i}{\sum_j \frac{1}{C''_j}} + \bar{q}_1^i = \frac{\sum_j \frac{\theta_1^j + \theta_2^j}{2C''_j} - \theta_1^i}{\sum_j \frac{1}{C''_j}} + \bar{q}_1^i
\]
Likewise, in the second period:

\[
q^i_t(\hat{p}_2, \theta^i_t) = \frac{\hat{p}_2 - C' - \theta^i_2}{C''_i} + \tilde{q}^i_2 = \frac{C'' + \sum_j \frac{\theta_j^i + \theta_j^i}{2C_J} - C' - \theta^i_2}{C''_i} + \tilde{q}^i_2 = \frac{\sum_j \frac{\theta_j^i + \theta_j^i}{2C_J} - \theta^i_2}{C''_i} + \tilde{q}^i_2
\]

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

### A.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)](https://www.jstor.org/stable/2685937) that firms have perfect information about both first- and second-period cost shocks before making any compliance decisions. We substitute our expressions for $q^i_t(\hat{p}_t, \theta^i_t)$ and $q^i_t(\tilde{p}_t, \theta^i_t)$ into the Taylor expansions
for (expected) benefits and costs. For the expected benefits of the quantity order, we obtain:

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

Expected costs from the quantity order:

$$\sum_{i=1}^{N} E[C_i(q_1^i(\hat{p}_1, \theta_1^i), \theta_1^i)] + \sum_{i=1}^{N} E[C_i(q_2^i(\hat{p}_2, \theta_2^i), \theta_2^i)]
= \sum_{i=1}^{N} E[\alpha_i(\theta_1^i) + (C' + \theta_1^i)(q_1^i(\hat{p}_1, \theta_1^i) - \bar{q}_1^i) + \frac{C''}{2} (q_1^i(\hat{p}_1, \theta_1^i) - \bar{q}_1^i)^2]
+ \sum_{i=1}^{N} E[\alpha_i(\theta_2^i) + (C' + \theta_2^i)(q_2^i(\hat{p}_2, \theta_2^i) - \bar{q}_2^i) + \frac{C''}{2} (q_2^i(\hat{p}_2, \theta_2^i) - \bar{q}_2^i)^2]
= \sum_{i=1}^{N} E[\theta_1^i (\frac{\theta_1^i}{2C_i} - \theta_1^i) + \frac{C''}{2} (\frac{\theta_1^i}{2C_i})^2]
+ \sum_{i=1}^{N} E[\theta_2^i (\frac{\theta_2^i}{2C_i} - \theta_2^i) + \frac{C''}{2} (\frac{\theta_2^i}{2C_i})^2]
= -\frac{1}{2} (\sum_{i=1}^{N} \frac{\theta_1^2}{C_i}) - \frac{1}{2} (\sum_{i=1}^{N} \frac{\theta_2^2}{C_i}) + (\frac{1}{4} \sum_{i=1}^{N} \frac{1}{C_i}) (\sum_{i=1}^{N} \frac{\theta_1^2}{C_i} - \sum_{i=1}^{N} \frac{\theta_2^2}{C_i})^2$$
Expected benefits from the price order:

\[
E[B_1(\sum_{i=1}^{N} q_i^1(\tilde{p}_1, \theta_i^1)) + B_2(\sum_{i=1}^{N} q_i^2(\tilde{p}_2, \theta_i^2))]
\]

\[
= E[b + B'(\sum_{i=1}^{N} q_i^1(\tilde{p}_1, \theta_i^1) - \bar{q}_i^1) + \frac{-B''}{2}(\sum_{i=1}^{N} q_i^1(\tilde{p}_1, \theta_i^1) - \bar{q}_i^1)^2]
\]

\[
+ E[b + B'(\sum_{i=1}^{N} q_i^2(\tilde{p}_2, \theta_i^2) - \bar{q}_i^2) + \frac{-B''}{2}(\sum_{i=1}^{N} q_i^2(\tilde{p}_2, \theta_i^2) - \bar{q}_i^2)^2]
\]

\[
= \frac{-B''}{2} \left( \sum_{i=1}^{N} -\theta_i^1 \right)^2 + \frac{-B''}{2} \left( \sum_{i=1}^{N} -\theta_i^2 \right)^2
\]

Expected costs from the price order:

\[
\sum_{i=1}^{N} E[C_i(\tilde{q}_1^1(\tilde{p}_1, \theta_i^1), \theta_i^1)] + \sum_{i=1}^{N} E[C_i(\tilde{q}_1^2(\tilde{p}_2, \theta_i^2), \theta_i^2)]
\]

\[
= E[a_i(\theta_i^1) + (C' + \theta_i^1)(\tilde{q}_1^1(\tilde{p}_1, \theta_i^1) - \bar{q}_i^1) + \frac{C''}{2}(\tilde{q}_1^1(\tilde{p}_1, \theta_i^1) - \bar{q}_i^1)^2]
\]

\[
+ E[a_i(\theta_i^2) + (C' + \theta_i^2)(\tilde{q}_1^2(\tilde{p}_2, \theta_i^2) - \bar{q}_i^2) + \frac{C''}{2}(\tilde{q}_1^2(\tilde{p}_2, \theta_i^2) - \bar{q}_i^2)^2]
\]

\[
= \sum_{i=1}^{N} E[\theta_i^1 \frac{-\theta_i^1}{C_i} + \frac{C''}{2}(\frac{-\theta_i^1}{C_i})^2] + \sum_{i=1}^{N} E[\theta_i^2 \frac{-\theta_i^2}{C_i} + \frac{C''}{2}(\frac{-\theta_i^2}{C_i})^2]
\]

\[
= E[\sum_{i=1}^{N} \frac{-\theta_i^1}{2C_i} + \sum_{i=1}^{N} \frac{-\theta_i^2}{2C_i}]
\]

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

\[
\Delta = E[(\frac{1}{4} \sum_{i=1}^{N} \frac{\theta_i^1}{C_i^m} - B'')(\sum_{i=1}^{N} \frac{\theta_i^1}{C_i^m} - \sum_{i=1}^{N} \frac{\theta_i^2}{C_i^m})^2]
\]

### A.4 Relative Advantage of Prices Over Quantities with Idiosyncratic 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 \( \hat{p}_i(\hat{Q}, \theta_t, \theta_{\sigma}) \) when making its compliance decision, because this quantity depends on \( \theta_i^{-1} \) whereas the production unit \( i \) may only observe \( \theta_i^i \). 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, \( \text{E}[\hat{p}_i] = \hat{p}_i(\hat{Q}, \theta_t, \theta_{\sigma}) + \epsilon_i^t \), where \( \epsilon_i^t \) reflects the firm’s idiosyncratic forecast error in period \( t \). The corresponding quantity response is given by:

\[
q_i^t(p_t, \theta_i^t) = \frac{\hat{p}_i(\hat{Q}, \theta_t, \theta_{\sigma}) + \epsilon_i^t - C' - \theta_i^i}{C_i''} + \bar{q}_i^t
\]

where \( \hat{p}_i(\hat{Q}, \theta_t, \theta_{\sigma}) \) is the cost-effective price defined above.

We construct the idiosyncratic forecast errors to have no impact on the aggregate quantity produced in period \( t \), which requires that the following condition is met:

\[
\hat{Q} = \sum_i C'' + \frac{\sum_j \frac{\theta_j^2 + \theta_{\sigma j}^2}{2C_j''}}{\sum_j \frac{1}{C_j''}} + \epsilon_i^t - C' - \theta_1^i \quad + \quad \bar{Q}_1 + \sum_i C'' + \frac{\sum_j \frac{\theta_j^2 + \theta_{\sigma j}^2}{2C_j''}}{\sum_j \frac{1}{C_j''}} - C' - \theta_2^i \quad + \quad \bar{Q}_2
\]

\[
0 = 2 \sum_{i=1}^N \frac{\theta_1^i + \theta_2^i}{2C_i''} - \frac{\theta_1^i}{C_i''} - \frac{\theta_2^i}{C_i''} + \epsilon_i^t
\]

Here we assume that the firm experiences forecast errors only in the first period, though this assumption could be easily relaxed.

To determine the modified welfare expression in the presence of these expectation errors, note first that expected benefits are a function of the total quantity produced in a given period, which does not change in the presence of idiosyncratic expectation errors. Therefore, the Taylor expansion for \( B_1(\sum_{i=1}^N \bar{q}_i^t) \) does not change from the version derived above. To
determine how costs change under a quantity order, we evaluate the following expression:

\[
\sum_{i=1}^{N} E[C_i(q_1(\hat{p}_1, \theta_i^1), \theta_i^1)] + \sum_{i=1}^{N} E[C_i(q_2(\hat{p}_2, \theta_i^2), \theta_i^2)]
\]

\[
= \sum_{i=1}^{N} E[\theta_i^1(\frac{\sum_{j=1}^{N} \theta_j^1 + \theta_j^2}{2C_i} + \epsilon_i^1 - \theta_1^i)] + C_i^m \left(\frac{\sum_{j=1}^{N} \theta_j^1 + \theta_j^2}{2C_i} + \epsilon_i^1 - \theta_1^i\right)^2]
\]

\[
+ \sum_{i=1}^{N} E[\theta_i^2(\frac{\sum_{j=1}^{N} \theta_j^1 + \theta_j^2}{2C_i} - \theta_2^i)] + C_i^m \left(\frac{\sum_{j=1}^{N} \theta_j^1 + \theta_j^2}{2C_i} - \theta_2^i\right)^2]
\]

\[
= \text{original terms from above + new terms}
\]

\[
= \text{original terms from above + } \sum_{i=1}^{N} \frac{\epsilon_i^1 + \theta_i^1}{2C_i^m} + \sum_{i=1}^{N} \frac{\theta_i^1 \epsilon_i^1}{C_i^m} - \sum_{i=1}^{N} \frac{\theta_i^1 \epsilon_i^1}{C_i^m} + \frac{N}{C_i} \left(\sum_{j=1}^{N} \frac{\theta_i^1 + \theta_j^2}{2C_j^m}\right)
\]

\[
= -\frac{1}{2} \left(\sum_{i=1}^{N} \frac{\theta_i^1}{C_i^m}\right) - \frac{1}{2} \sum_{i=1}^{N} \frac{\theta_i^2}{C_i^m} + \frac{1}{4} \sum_{i=1}^{N} \frac{1}{C_i^m} \left(\sum_{i=1}^{N} \frac{\theta_i^1}{C_i^m} - \sum_{i=1}^{N} \frac{\theta_i^2}{C_i^m}\right)^2 + \sum_{i=1}^{N} \frac{\epsilon_i^1}{C_i^m}
\]

The last equality follows from applying the constraint that \(\sum_{i=1}^{N} \frac{\epsilon_i^1}{C_i^m} = 0\).

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

\[
\Delta = E[(\frac{1}{4} \sum_{i=1}^{N} \frac{1}{C_i^m}) - B^m)(\sum_{i=1}^{N} \frac{\theta_i^1}{C_i^m} - \sum_{i=1}^{N} \frac{\theta_i^2}{C_i^m})^2 + \sum_{i=1}^{N} \frac{\epsilon_i^1}{2C_i^m}]
\]

A.5 Relative Advantage of Prices Over Quantities with Systematic Forecast Errors

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, firm \(i\) chooses its quantity given the following
expectation of market-clearing price: \( E_i[\hat{p}_t] = \hat{p}_t(\hat{Q}, \theta_t, \theta_t') + \epsilon_t^i + \xi_t \), where \( \xi_t \) is common to all firms in period \( t \). Consequently, the overall quantity produced in the first period is now given by:

\[
Q_1 = \frac{\hat{Q}}{2} + \sum_{i=1}^{N} \frac{\theta_i^j - \theta_i^j}{2C''_i} + \sum_{i=1}^{N} \xi_i
\]

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:

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

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

\[
Q_2 = \frac{\hat{Q}}{2} + \sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C''_i} - \sum_{i=1}^{N} \frac{\xi_i}{C''_i}
\]

Consequently, expected benefits (over both compliance periods) from the quantity order are now given by:\[20\]

\[
E[-B''(\sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C''_i} - \sum_{i=1}^{N} \frac{\xi_i}{C''_i})^2]
\]

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

\[
-\frac{1}{2} \left( \sum_{i=1}^{N} \frac{\xi_i^2}{C''_i} \right) + \frac{1}{2} \left( \frac{1}{4 \sum_{i=1}^{N} \frac{1}{C''_i}} \right) \left( \sum_{i=1}^{N} \frac{\theta_1^i}{C''_i} - \sum_{i=1}^{N} \frac{\theta_2^i}{C''_i} \right)^2 + \sum_{i=1}^{N} \frac{\xi_i^2}{2C''_i} + \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} \right) \left( \sum_{i=1}^{N} \frac{\xi_i}{C''_i} \right) \left( \sum_{i=1}^{N} \frac{\theta_1^i}{2C''_i} \right) + \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} \right) \left( \sum_{i=1}^{N} \frac{\xi_i}{C''_i} \right) \left( \sum_{i=1}^{N} \frac{\theta_2^i}{2C''_i} \right)
\]

\[20\text{We could equivalently write this expression as: } E[-B''(\sum_{i=1}^{N} \frac{\theta_1^i - \theta_2^i}{2C''_i} + \sum_{i=1}^{N} \frac{\xi_i}{C''_i})^2].\]
Second-period expected costs are given by:

\[
-\frac{1}{2} \left( \sum_{i=1}^{N} \frac{\theta_i^2}{C''_i} \right) + \frac{1}{2} \left( \frac{1}{4} \sum_{i=1}^{N} \frac{1}{C''_i} \right) \left( \sum_{i=1}^{N} \frac{\theta_i}{C''_i} - \sum_{i=1}^{N} \frac{\theta_i^2}{2C''_i} \right)^2
+ \sum_{i=1}^{N} \frac{\xi_i^2}{2C''_i} - \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} \right) \left( \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \right) - \left( \frac{1}{\sum_{i=1}^{N} \frac{1}{C''_i}} \right) \left( \sum_{i=1}^{N} \frac{\xi_i}{C''_i} \right) \left( \sum_{i=1}^{N} \frac{\theta_i}{2C''_i} \right)
\]

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:

\[
\Delta = E\left[ \frac{1}{4} \sum_{i=1}^{N} \frac{1}{C''_i} - B'' \left( \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \right)^2 + \frac{1}{2} \sum_{i=1}^{N} \frac{\xi_i^2}{C''_i} \right] + \sum_{i=1}^{N} \frac{\xi_i^2}{C''_i} + 2B'' \left( \sum_{i=1}^{N} \frac{\xi_i}{C''_i} \right) \left( \sum_{i=1}^{N} \frac{\theta_i}{2C''_i} \right) + B'' \left( \frac{\xi_1}{C''_i} \right)^2
\]

**A.6 Relative Advantage of Prices Over Quantities with Overlapping Policies**

We first derive the welfare comparison between prices and tradable quantities with overlapping policies in a setting without forecast errors. As described in the main text, some subset of \(R\) firms are subject to both the overlapping regulation and the market-based policy, abating at level \(\hat{q}_j^R\); without loss of generality, we denote the total costs of this subset of firms by \(C^R\).

For the \(M\) firms that are only subject to the market-based policy, firm behavior is unchanged in the presence of overlapping policies under a price regime, with each firm setting \(q_i = \frac{\theta_i}{C''_i}\). The total expected cost associated with this subset of firms is therefore given by:

\[
E\left[ \sum_{i=1}^{M} a_i(\theta_i) + \sum_{i=1}^{M} (C' + \theta_i)(-\theta_i) + \sum_{i=1}^{M} \frac{C''_i}{2} (-\theta_i)^2 \right]
\]

By contrast, the total quantity produced by these \(M\) firms does change in the presence of
the overlapping policy under tradable quantities. In this case, the market-clearing price will fall in response to the over-compliance among the doubly regulated firms, and the quantity produced by the $M$ firms will decrease correspondingly, thereby ensuring that the aggregate quantity target is still met. The market clearing price becomes:

$$p^* = C' + \frac{\sum_{i=1}^M \theta_i}{\sum_{i=1}^M C'_i} - \frac{\kappa}{\sum_{i=1}^M \frac{1}{C'_i}}$$

The corresponding quantity response by firm $i$ is therefore $q_i = \hat{q}_i + \frac{p^* - C' - \theta_i}{C'_i}$. The total expected costs for the $M$ firms subject to only the market-based regulation is given by:

$$\mathbb{E} \left[ \sum_{i=1}^M a_i(\theta_i) + \sum_{i=1}^M (C' + \theta_i)(\frac{\sum_{j=1}^M \theta_j}{C''_{ij}}) - \frac{\kappa}{C''_{ij}} \frac{1}{C''_{ij}} - \frac{\theta_i}{C''_{ij}} \right] + \sum_{i=1}^M \frac{C''_{ij}}{2} \left( \frac{\sum_{j=1}^M \theta_j}{C''_{ij}} - \frac{\theta_i}{C''_{ij}} \right)^2$$

Total expected benefits depend on the quantity produced in the market as a whole, including those firms subject to both regulations and those subject to only the market-based policy. In the case of a price-based policy, these expected benefits are given by

$$\mathbb{E} \left[ b + B' \left( \sum_{i=1}^M \frac{-\theta_i}{C''_{ij}} + \sum_{j=1}^R (\hat{q}_j^R - \hat{q}_j) - \frac{B''}{2} \left( \sum_{i=1}^M \frac{-\theta_i}{C''_{ij}} \right) + \sum_{j=1}^R (\hat{q}_j^R - \hat{q}_j)^2 \right] \right] = \mathbb{E} \left[ b + B' \left( \sum_{i=1}^M \frac{-\theta_i}{C''_{ij}} + \kappa \right) - \frac{B''}{2} \left( \sum_{i=1}^M \frac{-\theta_i}{C''_{ij}} + \kappa \right)^2 \right]$$

Under the tradable quantities policy, expected benefits are simply given by $\mathbb{E}[b]$.

Combining all of these terms yields the expression presented in the text for the relative
advantage of prices over quantities in the presence of overlapping policies:

\[
\Delta = E[b + B'(\sum_{i=1}^{M} -\theta_i C''_i) + \kappa] - \frac{B''}{2} (\sum_{i=1}^{M} -\theta_i C''_i) + \kappa)^2] \\
- E[\sum_{i=1}^{M} a_i(\theta_i) + \sum_{i=1}^{M} (C' + \theta_i)(-\theta_i C''_i) + \sum_{i=1}^{M} \frac{C''_i}{2} (-\theta_i C''_i)^2] - CR - E[b] \\
+ E[\sum_{i=1}^{M} a_i(\theta_i) + \sum_{i=1}^{M} (C' + \theta_i)(\frac{\sum_{j=1}^{M} \theta_j C''_j}{C''_i} - \frac{\kappa}{C''_i} - \frac{\theta_i}{C''_i})] + CR \\
- E[\frac{1}{2}(\sum_{j=1}^{M} \frac{1}{C''_j} - B'')(\frac{\theta_i}{C''_i})^2 + \frac{1}{2}(\sum_{j=1}^{M} \frac{1}{C''_j} - B'')\kappa^2] \\
\]

Next we introduce forecast errors around the impact of the overlapping policy in the market for allowances. The \( M \) firms subject to only the market-based regulation expect the impact of over-compliance on the equilibrium price to be \( E_i(\frac{\kappa}{\sum_{i=1}^{M} 1/C''_i}) = \frac{\kappa}{\sum_{i=1}^{M} 1/C''_i} + \eta_i \). Consequently, their expectation of the market-price for allowances is given by: \( E_i[p^k] = C' + \frac{\sum_{i=1}^{M} \theta_i C''_i}{\sum_{i=1}^{M} 1/C''_i} - \frac{\kappa}{\sum_{i=1}^{M} 1/C''_i} - \eta_i \).

As in the baseline derivation for firm uncertainty over competitors’ abatement costs, we assume that there is no impact on the aggregate quantity produced by all regulated firms; therefore, the expected benefits of tradable quantities are unchanged relative to the scenario without forecast errors. Expected costs are now given by:

\[
E[\sum_{i=1}^{M} a_i(\theta_i) + \sum_{i=1}^{M} (C' + \theta_i)(\frac{\sum_{j=1}^{M} \theta_j C''_j}{C''_i} - \frac{\kappa}{C''_i} - \frac{\theta_i}{C''_i}) + \frac{\sum_{i=1}^{M} \theta_i C''_i}{2} (\frac{\sum_{j=1}^{M} \theta_j C''_j}{C''_i} - \frac{\kappa}{C''_i} - \frac{\theta_i}{C''_i})^2] \\
\]

The expected welfare of prices relative to tradable quantities in the presence of overlap-
ping policies is therefore given by:

\[
\Delta = E\left[ \frac{1}{2} \left( \frac{1}{\sum_{j=1}^{M} C_j} - B'' \right)^2 + \frac{1}{2} \left( \frac{1}{\sum_{j=1}^{M} C''_j} - B'' \right)^2 + \sum_{i=1}^{M} \eta_i^2 \right]
\]

which matches the expression given in the main text.

A.7 Price Collars

Here we provide more detailed derivations of the expected welfare impacts of introducing a price collar under a tradable quantities policy. We start with the baseline scenario without forecast errors. As mentioned in the main text, with probability \( G(\bar{t}) - G(t) \), neither the price ceiling nor the price collar is binding and the expected benefits and costs are identical under the two policy regimes, at least in the absence of forecast errors. With probability \( 1 - G(\bar{t}) \), the price ceiling is triggered and \( p = \bar{p} = \frac{\bar{t}}{\sum_{i=1}^{N} \frac{\theta_i}{C''_i}} + C' \). Finally, with probability \( G(t) \), the price floor is triggered and \( p = \bar{p} = \frac{\bar{t}}{\sum_{i=1}^{N} \frac{\theta_i}{C''_i}} + C' \).

When the price ceiling is triggered, expected benefits are given by:

\[
E\left[ B\left( \sum_{i=1}^{N} q_i(\bar{p}) \right) \right] = E\left[ b + B'( \frac{\bar{t}}{\sum_{j=1}^{M} \frac{\theta_j}{C''_j}} - \frac{\theta_i}{C''_i} ) + \frac{B''}{2} \left( \frac{1}{\sum_{j=1}^{M} \frac{\theta_j}{C''_j}} - \frac{\theta_i}{C''_i} \right)^2 \right] = E\left[ b + B'( \frac{\bar{t}}{\sum_{i=1}^{N} \frac{\theta_i}{C''_i}} - \frac{B''}{2} \left( \frac{1}{\sum_{i=1}^{N} \frac{\theta_i}{C''_i}} \right)^2 \right] = E\left[ b + B'( \bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} ) - \frac{B''}{2} \left( \bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \right)^2 | \sum_{i=1}^{N} \frac{\theta_i}{C''_i} > \bar{t} \right]
\]

Likewise, when the price floor is triggered, expected benefits are given by:

\[
E\left[ B\left( \sum_{i=1}^{N} q_i(p) \right) \right] = E\left[ b + B'( t - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} ) - \frac{B''}{2} \left( t - \sum_{i=1}^{N} \frac{\theta_i}{C''_i} \right)^2 | \sum_{i=1}^{N} \frac{\theta_i}{C''_i} < t \right]
\]
When the price ceiling is triggered, expected costs are given by:

\[
E\left[\sum_{i=1}^{N} C_i(q_i(\bar{p}))\right] = \sum_{i=1}^{N} E[a_i(\theta_i) + (C' + \theta_i)\left(\frac{\bar{t}}{\sum_{j=1}^{N} \frac{1}{C_{ij}^{'}} } - \frac{\theta_i}{C_{ij}^{'}}\right) + \frac{C''}{2}\left(\frac{\bar{t}}{\sum_{j=1}^{N} \frac{1}{C_{ij}^{'}} } - \theta_i\right)^2\right] \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} > \bar{t}]
\]

When the price floor is triggered, expected costs are defined analogously.

To obtain the expected change in welfare from introducing a price collar to a tradable quantities program, we weight these benefit and cost expressions by the probability of triggering the price ceiling or price floor, respectively. We then subtract expected welfare for tradable quantities without a price collar. We obtain the following expression:

\[
\Delta = \{E[-\frac{B''}{2}(\bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 - \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}}\left(\frac{\bar{t}}{\sum_{j=1}^{N} \frac{1}{C_{ij}^{'}} } - \frac{\theta_i}{C_{ij}^{'}}\right) - \sum_{i=1}^{N} \frac{\theta_i^2}{2C_{ij}^{'}} - \sum_{i=1}^{N} \frac{1}{2C_{ij}^{'}}\left(\frac{\bar{t}}{\sum_{j=1}^{N} \frac{1}{C_{ij}^{'}} } - \theta_i\right)^2\right] \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} > \bar{t}] \}
\]

\[
- E[-\sum_{i=1}^{N} \frac{\theta_i^2}{2C_{ij}^{'}} \sum_{j=1}^{N} \frac{1}{C_{ij}^{'}}\left(\frac{1}{\sum_{j=1}^{N} \frac{1}{C_{ij}^{'}} } \right)] \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} > \bar{t} \} \cdot (1 - G(\bar{t}))
\]

\[
- E[-\sum_{i=1}^{N} \frac{\theta_i^2}{2C_{ij}^{'}} \sum_{j=1}^{N} \frac{1}{C_{ij}^{'}}\left(\frac{1}{\sum_{j=1}^{N} \frac{1}{C_{ij}^{'}} } \right)] \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} < \bar{t} \} \cdot G(\bar{t})
\]

\[
= (1 - G(\bar{t})) \cdot \{E[-\frac{B''}{2}(\bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 + \frac{1}{2}\left(\sum_{i=1}^{N} \frac{1}{C_{ij}^{'}}\right)(\sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 - \bar{t}^2\right)\sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} > \bar{t}] \}
\]

\[
+ G(\bar{t}) \cdot \{E[-\frac{B''}{2}(\bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 + \frac{1}{2}\left(\sum_{i=1}^{N} \frac{1}{C_{ij}^{'}}\right)(\sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 - \bar{t}^2\right)\sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} < \bar{t}] \}
\]

\[
= (1 - G(\bar{t})) \cdot \left\{\frac{1}{2}\left(\sum_{i=1}^{N} \frac{1}{C_{ij}^{'}}\right) - B''(\bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 - \bar{t}\left(\sum_{i=1}^{N} \frac{1}{C_{ij}^{'}}\right)(\sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} - \bar{t})\right\} \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} > \bar{t}
\]

\[
+ G(\bar{t}) \cdot \left\{\frac{1}{2}\left(\sum_{i=1}^{N} \frac{1}{C_{ij}^{'}}\right) - B''(\bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}})^2 - \bar{t}\left(\sum_{i=1}^{N} \frac{1}{C_{ij}^{'}}\right)(\sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} - \bar{t})\right\} \sum_{i=1}^{N} \frac{\theta_i}{C_{ij}^{'}} < \bar{t}
\]

which is the result given in the text.

Introducing forecast errors is now straightforward. First, in the case that the price ceiling is not binding, the only change in expected net benefits under the two policy regimes will be the difference in forecast error variances with and without the price ceiling: 

\[
\frac{1}{2}E[\sum_{i=1}^{N} \frac{\theta_i^2}{C_{ij}^{'}}] - \]
\( \frac{1}{2} \mathbb{E}[\sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m}] \). Then, where the price ceiling is binding, expected benefits with and with the price ceiling do not change from the previous derivation without forecast errors, given our assumption that \( \sum_{i=1}^{N} \frac{\epsilon_i}{C_i^m} = \sum_{i=1}^{N} \frac{\epsilon_i}{C_i^m} = 0 \). Expected costs with the price ceiling now include the following additional terms: \( \mathbb{E}[\sum_{i=1}^{N} \frac{\theta_i \epsilon_i}{C_i^m} + \sum_{i=1}^{N} \frac{\epsilon_i^2}{2C_i^m} - \sum_{i=1}^{N} \frac{\theta_i \epsilon_i}{C_i^m}] = \mathbb{E}[\sum_{i=1}^{N} \frac{\epsilon_i^2}{2C_i^m}] \), with an analogous expression for the price floor. Expected costs without the price ceiling now include: \( \mathbb{E}[\sum_{i=1}^{N} \frac{\theta_i \epsilon_i}{C_i^m} + \sum_{i=1}^{N} \frac{\epsilon_i^2}{2C_i^m} - \sum_{i=1}^{N} \frac{\theta_i \epsilon_i}{C_i^m}] = \mathbb{E}[\sum_{i=1}^{N} \frac{\epsilon_i^2}{2C_i^m}] \).

Combining these results yields:

\[
\Delta = (1 - G(\bar{t})) \cdot \mathbb{E}[\frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m} - \frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m} 
+ \frac{1}{2} (\sum_{i=1}^{N} \frac{1}{C_i^m} - B^o)(\bar{t} - \sum_{i=1}^{N} \frac{\theta_i}{C_i^m})^2 - \bar{t}(\sum_{i=1}^{N} \frac{1}{C_i^m})(\sum_{i=1}^{N} \frac{\theta_i}{C_i^m} - \bar{t}) | \bar{t} < \sum_{i=1}^{N} \frac{\theta_i}{C_i^m}]
+ G(\bar{t}) \cdot \mathbb{E}[\frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m} - \frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m} 
+ \frac{1}{2} (\sum_{i=1}^{N} \frac{1}{C_i^m} - B^o)(t - \sum_{i=1}^{N} \frac{\theta_i}{C_i^m})^2 - t(\sum_{i=1}^{N} \frac{1}{C_i^m})(\sum_{i=1}^{N} \frac{\theta_i}{C_i^m} - t) | t > \sum_{i=1}^{N} \frac{\theta_i}{C_i^m}]
+ (G(\bar{t}) - G(t)) \cdot \mathbb{E}[\frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m} - \frac{1}{2} \sum_{i=1}^{N} \frac{\epsilon_i^2}{C_i^m} | t \leq \sum_{i=1}^{N} \frac{\theta_i}{C_i^m}] \]

which is again is equal to the result given in the main text.

### A.8 Policy Updating

We consider the regulator’s optimal rule for updating the aggregate quantity target based on the information about cost shocks revealed during the first period. We assume that firms anticipate the regulator’s actions given their information set in the first compliance period (namely, the realization of cost shocks). To start, we assume that firms have perfect information about all cost shocks in the market, not just their own. For the case where
\( \theta_1^i = \theta_2^i \), the regulator’s optimization problem is given by:

\[
\max_{\delta} b + B' \left( \sum_{i=1}^{N} \frac{p_1(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''} \right) - \frac{B''}{2} \left( \sum_{i=1}^{N} \frac{p_1(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''} \right)^2
\]

\[
+ b + B' \left( \sum_{i=1}^{N} \frac{p_2(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''} \right) - \frac{B''}{2} \left( \sum_{i=1}^{N} \frac{p_2(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''} \right)^2
\]

\[
- \sum_{i=1}^{N} \alpha(\theta_i^i) - \sum_{i=1}^{N} (C' + \theta_1^i)(\frac{p_1(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''}) - \sum_{i=1}^{N} \frac{C_i''}{2} (\frac{p_1(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''})^2
\]

\[
- \sum_{i=1}^{N} \alpha(\theta_i^i) - \sum_{i=1}^{N} (C' + \theta_2^i)(\frac{p_2(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''}) - \sum_{i=1}^{N} \frac{C_i''}{2} (\frac{p_2(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''})^2
\]

where \( \frac{p_1(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''} = q_1^i - \bar{q}_i \), \( \frac{p_2(\bar{Q}, \delta) - C' - \theta_i^i}{C_i''} = \bar{q}_2^i - \bar{q}_i \), and \( p_1 = p_2 = C' + \frac{\delta}{2 \sum_{i=1}^{N} \frac{\theta_i^i}{C_i''}} + \sum_{i=1}^{N} \frac{\theta_i^i}{C_i''} \).

Taking the regulator’s FOC yields:

\[
B' \sum_{i=1}^{N} \frac{1}{C_i''} \frac{\partial p_1}{\partial \delta} - B'' \left( \sum_{i=1}^{N} \frac{p_1 - C' - \theta_i^i}{C_i''} \right) \sum_{i=1}^{N} \frac{1}{C_i''} \frac{\partial p_1}{\partial \delta}
\]

\[
+ B' \sum_{i=1}^{N} \frac{1}{C_i''} \frac{\partial p_2}{\partial \delta} - B'' \left( \sum_{i=1}^{N} \frac{p_2 - C' - \theta_i^i}{C_i''} \right) \sum_{i=1}^{N} \frac{1}{C_i''} \frac{\partial p_2}{\partial \delta}
\]

\[
- \sum_{i=1}^{N} (C' + \theta_i^i) \frac{1}{C_i''} \frac{\partial p_1}{\partial \delta} - \sum_{i=1}^{N} C_i'' \left( \frac{p_1 - C' - \theta_i^i}{C_i''} \right) \frac{1}{C_i''} \frac{\partial p_1}{\partial \theta_i^i}
\]

\[
- \sum_{i=1}^{N} (C' + \theta_i^i) \frac{1}{C_i''} \frac{\partial p_2}{\partial \delta} - \sum_{i=1}^{N} C_i'' \left( \frac{p_2 - C' - \theta_i^i}{C_i''} \right) \frac{1}{C_i''} \frac{\partial p_1}{\partial \theta_i^i} = 0
\]

Substituting \( \frac{\partial p_1}{\partial \delta} = \frac{\partial p_2}{\partial \delta} = \frac{1}{2 \sum_{i=1}^{N} \frac{1}{C_i''}} \) and simplifying yields:

\[
\delta = \frac{-2 \sum_{i=1}^{N} \frac{\theta_i^i}{C_i''}}{1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''}}
\]

The corresponding market-clearing price is given by:

\[
p_1 = p_2 = C' + \frac{\delta}{2 \sum_{i=1}^{N} \frac{1}{C_i''}} + \sum_{i=1}^{N} \frac{\theta_i^i}{\sum_{i=1}^{N} \frac{1}{C_i''}} = C' - \frac{\sum_{i=1}^{N} \frac{\theta_i^i}{C_i''}}{\sum_{i=1}^{N} \frac{1}{C_i''}} \left( 1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''} \right) + \sum_{i=1}^{N} \frac{\theta_i^i}{C_i''}
\]
which results in individual quantities:

\[ q^i_t = q^i_t + \frac{-\sum_{j=1}^{N} \frac{\theta_j^i}{C''_j}}{C''_i \left( \sum_{j=1}^{N} \frac{1}{C_j''} \right) (1 + B'' \sum_{j=1}^{N} \frac{1}{C_j''})} + \frac{\sum_{j=1}^{N} \frac{\theta_j^i}{C''_j}}{C''_i \sum_{j=1}^{N} \frac{1}{C_j''}} - \theta_i 
\]

and aggregate quantities:

\[ \sum_{i=1}^{N} q^i_1 + \sum_{i=1}^{N} q^i_1 = \sum_{i=1}^{N} \hat{q}^i_1 + \sum_{i=1}^{N} \hat{q}^i_1 - \frac{2 \sum_{i=1}^{N} \frac{\theta_i}{C''_i}}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}} \]

which matches the first-best quantity for a two-period policy. We have therefore reproduced the result in Pizer and Prest (2020) and Heutel (2018) that it is possible to achieve the first-best aggregate quantity when firms have perfect information about the regulator’s policy updating rule and all shocks realized in the market, in this case for a market with heterogeneous firms.

When instead firms do not perfectly observe their competitors’ shocks when making their compliance decisions in the first period, it is no longer possible to achieve the first best outcome. However, as we noted in the main text, policy updating may still serve to mitigate the cost-inefficiencies introduced by firm-level price uncertainty. To demonstrate this result, we first note that it is theoretically possible to model several different information structures, ranging from the regulator failing to observe past forecast errors to the regulator having a detailed understanding of the process by which firms make these forecast errors. We adopt the assumption that seems to reflect what we observe in the world: regulators do not have any prior knowledge of the magnitude of firm forecast errors or how they depend on policy stringency, and therefore cannot set their updating rule to match the expected forecast errors. The key point is that optimal policy updating given realized cost shocks is sufficient to dampen the negative impact of these forecast errors on overall welfare.
Given forecast errors, the firm response function in the first period is now given by:

\[ q_i^1 = \hat{q}_i^1 + \frac{E_i[-\sum_{i=1}^{N} \theta_i^i]}{C_i^i(\sum_{i=1}^{N} \frac{1}{C_i^i}) (1 + B'' \sum_{i=1}^{N} \frac{1}{C_i^i})} + \frac{E_i[\sum_{i=1}^{N} \theta_i^i]}{C_i'' \sum_{i=1}^{N} \frac{1}{C_i''}} - \theta_i \\

\]

Following our earlier notation that \( \epsilon_i^1 + \xi_1 \) represents firm \( i \)'s overall forecast error around the impact of cost shocks on the market-clearing price, or \( \epsilon_i^1 + \xi_1 = E_i[p_1] - p_1 = E_i[\frac{\sum_{i=1}^{N} \theta_i^i}{\sum_{i=1}^{N} \frac{1}{C_i^i}}] - \frac{\sum_{i=1}^{N} \theta_i^i}{\sum_{i=1}^{N} \frac{1}{C_i^i}} \), we can also write the response function as:

\[ q_i^1 = \hat{q}_i^1 - \theta_i C_i^i + \frac{\sum_{i=1}^{N} \theta_i^i}{C_i'' \sum_{i=1}^{N} \frac{1}{C_i''}} + \frac{-\sum_{i=1}^{N} \theta_i^i}{C_i'' (\sum_{i=1}^{N} \frac{1}{C_i''}) (1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''})} \\
+ \frac{\epsilon_i^1}{C_i''} + \frac{\xi_1}{C_i''} + \frac{-\epsilon_i^1}{C_i'' (1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''})} + \frac{-\xi_1}{C_i'' (1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''})} \]

The aggregate quantity resulting from these individual response functions is given by:

\[ Q_1 = \bar{Q}_1 - \sum_{i=1}^{N} \frac{\theta_i^i}{C_i''} \left( \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''}} \right) - \sum_{i=1}^{N} \frac{\xi_1}{C_i''} \left( \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''}} \right) + \sum_{i=1}^{N} \frac{\xi_1}{C_i''} \]

which can also be expressed as:

\[ Q_1 = \bar{Q}_1 + \frac{\delta}{2} + (1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''}}) \sum_{i=1}^{N} \frac{\xi_1}{C_i''} \]

We can compare this aggregate first-period quantity to the aggregate first-period quantity without policy updating:

\[ Q_1 = \bar{Q}_1 + \sum_{i=1}^{N} \frac{\xi_1}{C_i''} \]

Note that the expression \( B'' \sum_{i=1}^{N} \frac{1}{C_i''} \) is strictly positive (given that \( B'' \geq 0 \) and \( C_i'' \geq 0 \) for all \( i \)), which means that \( (1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C_i''}}) \mid \sum_{i=1}^{N} \xi_i \mid \leq \mid \sum_{i=1}^{N} \xi_i \mid \), with a strict inequality whenever \( B'' > 0 \) and \( C_i'' > 0 \). That is, policy updating reduces the magnitude of systematic forecast errors, which as we show below reduces costs. The intuition here is that
policy updating reduces the impact of aggregate cost shocks on the quantity produced by a proportionality factor \(\frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}}\). Because systematic forecast errors are added to the aggregate cost shocks in determining each firm’s quantity response, these forecast errors are also proportionally reduced in the presence of policy updating.

Given the total quantity produced in the first period, the market-clearing price in the second period must adjust to ensure that the overall quantity target (net of policy updating) is met. For now we assume that the regulator adheres to some pre-announced updating rule and does not revise \(\delta\) based on information about the magnitude of the first-period forecast errors, though we also relax this assumption below for comparison. The second-period price is now given by:

\[
p_2 = C' + \frac{\delta}{2 \sum_{i=1}^{N} \frac{1}{C'_i}} + \frac{\sum_{i=1}^{N} \frac{\hat{q}^i}{C''_i}}{\sum_{i=1}^{N} \frac{1}{C''_i}} - \xi_1 \cdot (1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}})
\]

The second period response function is therefore given by:

\[
q'_2 = \hat{q}'_2 + \frac{\delta}{2C'' \sum_{j=1}^{N} \frac{1}{C''_j}} + \frac{\sum_{i=1}^{N} \frac{\hat{q}^i}{C''_j}}{\sum_{i=1}^{N} \frac{1}{C''_j}} - \xi_1 \frac{1}{1 + B'' \sum_{j=1}^{N} \frac{1}{C''_j}} - \frac{\theta^i}{C''_i}
\]

which results in the following aggregate second-period quantity:

\[
Q_2 = Q_2 + \frac{\delta}{2} - (1 - \frac{1}{1 + B'' \sum_{i=1}^{N} \frac{1}{C''_i}}) \sum_{i=1}^{N} \frac{\xi_1}{C''_i}
\]

Again we can compare this aggregate quantity to a tradable quantities regime without updating:

\[
Q_2 = \bar{Q}_2 - \sum_{i=1}^{N} \frac{\xi_1}{C''_i}
\]

To show the welfare impacts of policy updating in the presence of first-period forecast errors, we substitute these quantity response functions into the expression for the ex post difference in net benefits for quantity trading with and without policy updating. We use the
ex post welfare expression since the regulator determines policy updating based on realized (not expected) cost shocks.

\[\Delta = b + B'(\frac{\delta}{2} + (1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C''i}}) \sum_{i=1}^{N} \frac{\xi_i}{C''i}) - \frac{B''}{2} (\frac{\delta}{2} + (1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C''i}}) \sum_{i=1}^{N} \frac{\xi_i}{C''i})^2 \]

\[b + B'(\frac{\delta}{2} - (1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C''i}}) \sum_{i=1}^{N} \frac{\xi_i}{C''i}) - \frac{B''}{2} \frac{\delta}{2} - (1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C''i}}) \sum_{i=1}^{N} \frac{\xi_i}{C''i})^2 \]

\[-b - B'(\sum_{i=1}^{N} \frac{\xi_i}{C''i}) + \frac{B''}{2} (\sum_{i=1}^{N} \frac{\xi_i}{C''i})^2 \]

\[b - B'(\sum_{i=1}^{N} \frac{\xi_i}{C''i}) + \frac{B''}{2} (-\sum_{i=1}^{N} \frac{\xi_i}{C''i})^2 \]

\[-\sum_{i=1}^{N} a(\theta^i) - \sum_{i=1}^{N} (C' + \theta^i)(\frac{-\delta}{2C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{-\epsilon_i^1 - \xi_1}{C''i(1 + B''\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(1 + B''\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i) \]

\[-\sum_{i=1}^{N} C''i \frac{-\delta}{2C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{-\epsilon_i^1 - \xi_1}{C''i(1 + B''\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(1 + B''\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i)^2 \]

\[-\sum_{i=1}^{N} C''i \frac{-\delta}{2C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{j=1}^{N} \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i)^2 \]

\[+ \sum_{i=1}^{N} a(\theta^i) + \sum_{i=1}^{N} (C' + \theta^i)(\frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i) + \sum_{i=1}^{N} \frac{C''i}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} \frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i)^2 \]

\[+ \sum_{i=1}^{N} a(\theta^i) + \sum_{i=1}^{N} (C' + \theta^i)(\frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{j=1}^{N} \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\sum_{i=1}^{N} \frac{\theta_i}{C''i} \sum_{i=1}^{N} \frac{1}{C''i}}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} + \frac{\epsilon_i^1 + \xi_1}{C''i} - \theta_i)^2 \]

\[+ \sum_{i=1}^{N} \frac{C''i}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} - \sum_{j=1}^{N} \frac{\epsilon_i^1 + \xi_1}{C''i(\sum_{i=1}^{N} \frac{1}{C''i})} - \theta_i)^2 \]
which simplifies to:

\[
\Delta = -B''\frac{(\delta)^2}{2} - B''(1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C_i'}})\left(\sum_{i=1}^{N} \frac{\xi_i}{C_i''}\right)^2 + B''\left(\sum_{i=1}^{N} \frac{\xi_i}{C_i''}\right)^2
\]

\[
-2\sum_{i=1}^{N} \theta^i\left(\frac{\delta}{C_i'}\sum_{j=1}^{N} \frac{1}{C_j'}\right) - \left(\sum_{i=1}^{N} \frac{1}{C_i'}\right)\left(\frac{(\delta)^2}{2} - \frac{1}{2}\left(\sum_{i=1}^{N} \frac{\epsilon_i^1 + \xi_i^1}{C_i''}\right)^2(1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C_i'}})\right)
\]

\[
+ \delta\left(\sum_{i=1}^{N} \frac{\theta^i}{C_i''}\right) + \left(\sum_{i=1}^{N} \frac{\theta^i}{C_i''}\right)^2\left(\sum_{i=1}^{N} \frac{1}{C_i'}\right)\left(1 + B''\sum_{i=1}^{N} \frac{1}{C_i'}\right) - \left(\frac{\epsilon_i^1 + \xi_i^1}{2C_i''}\right)(1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C_i'}})
\]

\[
+ \sum_{i=1}^{N} \frac{(\epsilon_i^1 + \xi_i^1)^2}{2C_i''} + \frac{1}{2}\left(\sum_{i=1}^{N} \frac{\epsilon_i^1 + \xi_i^1}{C_i''}\right)^2\left(\sum_{i=1}^{N} \frac{1}{C_i'}\right)
\]

Let \(\sum_{i=1}^{N} \frac{\xi_i}{C_i'} = (1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C_i'}})\left(\sum_{i=1}^{N} \frac{\xi_i}{C_i'}\right)\), where \(\sum_{i=1}^{N} \frac{\xi_i}{C_i'} \leq \sum_{i=1}^{N} \frac{\xi_i}{C_i''}\). Likewise, \(\sum_{i=1}^{N} \frac{\xi_i}{C_i'} = (1 - \frac{1}{1 + B''\sum_{i=1}^{N} \frac{1}{C_i'}})\left(\sum_{i=1}^{N} \frac{\xi_i}{C_i'}\right)\), where \(\sum_{i=1}^{N} \frac{\xi_i}{C_i'} \leq \sum_{i=1}^{N} \frac{\xi_i}{C_i''}\). This expression finally becomes:

\[
\Delta = \left[\frac{1}{\sum_{i=1}^{N} \frac{1}{C_i'}} + B''\right] + \left[\frac{1}{\sum_{i=1}^{N} \frac{1}{C_i''}} + B''\right] + \left[\frac{N}{\sum_{i=1}^{N} \frac{\xi_i}{C_i'}} - \left(\sum_{i=1}^{N} \frac{\xi_i}{C_i''}\right)^2\right]
\]

\[
\text{Ex Post Updating: } \geq 0 \quad \text{Reduced Systematic Forecast Errors: } \geq 0
\]

\[
+ \frac{1}{2}\left(\sum_{i=1}^{N} \frac{(\epsilon_i^1)^2}{C_i''} - \sum_{i=1}^{N} \frac{(\epsilon_i^1)^2}{C_i''}\right)
\]

\[
\text{Reduced Idiosyncratic Forecast Errors: } \geq 0
\]

As we see here, the welfare impact of introducing policy updating is unambiguously positive even with firm-level forecast errors. The intuition is that the policy updating scenario targets the ex post first-best outcome and introduces a smaller deviation due to (partially diminished) forecast errors, compared to a regime without updating that targets the ex ante optimal quantity and then introduces the full set of forecast errors.