Research Paper 34 | **2015**

EMISSIONS TRADING IN THE PRESENCE OF PRICE-REGULATED POLLUTING FIRMS: HOW COSTLY ARE FREE ALLOWANCES?

B. LANZ AND S. RAUSCH

Emissions Trading in the Presence of Price-Regulated Polluting Firms: How Costly Are Free Allowances?*

Bruno Lanz[†] Sebastian Rausch[‡]

January 2015

Abstract

We study whether to auction or to freely distribute emissions allowances when some firms participating in emissions trading are subject to price regulation. We show that free allowances allocated to price-regulated firms effectively act as a subsidy to output, distort consumer choices, and generally induce higher output and emissions by price-regulated firms. This provides a cost-effectiveness argument for an auction-based allocation of allowances (or equivalently an emissions tax). For real-world economies such as the Unites States, in which about 20 percent of total carbon dioxide emissions are generated by price-regulated electricity producers, our quantitative analysis suggests that free allowances increase economy-wide welfare costs of the policy by 40-80 percent relative to an auction. Given large disparities in regional welfare impacts, we show that the inefficiencies are mainly driven by the emissions intensity of electricity producers in regions with a high degree of price regulation.

Keywords: Tradable Pollution Permits; Climate policy; Auctioning; Free Allocation; Price Regulation; Electricity Generation.

JEL Classification numbers: C6, D4, D5, Q4.

---

*We are grateful for comments and suggestions from Christoph Böhringer, Julien Daubanes, Massimo Filippini, Wei Hu, Jean-Philippe Nicolai, John Reilly, Charles Rossmann, Thomas F. Rutherford, and participants at the 2012 Conference on Sustainable Resource Use and Economic Dynamics, 2012 Cowles/Yale Summer Conference on Macroeconomics and Climate, 2012 Annual Meeting of the European Association of Environmental and Resource Economics, 2012 Atlantic Workshop on Energy and Environmental Economics, and 2012 Meeting of the European Economic Association. We thank three anonymous referees for helpful comments. Any remaining errors are ours.

[†]Corresponding author. Department of International Economics and Center for International Environmental Studies, Graduate Institute of International and Development Studies, Chemin Eugène Rigot 2, 1202 Geneva, Switzerland, E-mail: bruno.lanz@graduateinstitute.ch; and Chair for Integrative Risk Management and Economics, ETH Zurich, Switzerland.

[‡]Department of Management, Technology, and Economics, ETH Zurich, Zürichbergstrasse 18, ZUE E, 8032 Zurich, Switzerland, E-mail: srausch@ethz.ch; Center for Economic Research at ETH Zurich (CER-ETH) and Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, USA.

1 Introduction

Emissions trading programs have become a centerpiece of environmental policy in Europe and the United States. In a competitive setting with full information, the creation of a market for emissions permits works to equalize marginal abatement costs across sources and minimizes aggregate compliance costs (Dales, 1968; Montgomery, 1972).[1] An appealing feature of emissions trading is the independence of permit market outcomes and the initial allocation of permits (Hahn and Stavins, 2011). This enables separating efficiency (or cost-effectiveness) from equity considerations, creating the flexibility to secure political support for such policies. Free allowances or the revenue from auctioned permits can be used to relieve participating firms from their compliance costs and offset profit losses (Hepburn et al., 2012; Goulder et al., 2010), or to address unintended distributional outcomes (Stavins, 2008). The independence property also means that the central design question of emissions trading regulation, namely whether to auction or give away emissions permits for free, does not affect the aggregate cost of the policy.

This paper challenges this view by investigating the extent to which the presence of price-regulated firms affects the outcome of alternative permit allocations rules.[2] Our analysis is motivated by two observable features of present-day economies. First, in many industrialized countries a large share of greenhouse gas emissions stems from electricity generation. Second, despite the ongoing liberalization of electricity markets in many countries, the electricity sector remains highly regulated, with electricity prices determined by some form of cost-of-service regulation. For example, in the United States about 30% of economy-wide carbon dioxide (CO_2) emissions in 2011 stemmed from electricity generation (U.S. Environmental Protection Agency, 2014) and around 60% of electricity was generated by producers that were subject to cost-of-service regulation (Energy Information Administration, 2012).

Combining stylized partial equilibrium analysis with numerical general equilibrium simulations, this paper shows that these two observations taken together have important implications

[1] Throughout this paper, we use the terms "permits" and "allowances" interchangeably.

[2] The existence of pre-existing distortions in the fiscal system may also interact with the outcome of environmental policy (Bovenberg and Goulder, 1996; Bovenberg, 1999). From an efficiency perspective, auctioning permits (or equivalently, using emissions taxes) is preferred to a free allocation as revenues can be used to lower pre-existing distortions in the tax system (e.g., taxes on labor and capital income). This paper rather focuses on pre-existing policy that is central in the regulation of utilities.

for the design of emissions trading policies. The reason is that price-regulated firms need to adjust output prices with the value of free allowances. If emissions permits are instead auctioned (or if an emissions tax is used), the cost of buying emissions permits is fully reflected in the output price of price-regulated firms. Because an auction also generates an income effect for households, the impact of free permits relative to an auction depends on income and substitution effects for the good produced by the regulated firm. We show that when income effects associated with revenues from the auction dominate, distributing free permits induces higher output by price-regualted firms as compared to an auction. In turn, with free permits, emissions from price-regulated firms will be higher than under an auction, and abatement will have to shift to other (i.e., non-electric) sectors in the economy, potentially undermining cost-effectiveness.[3]

To get a sense about the likely order of magnitude of efficiency costs and distributional impacts of alternative designs for emissions trading regulation, we develop a numerical general equilibrium model for the U.S. economy. The model is based on standard neoclassical optimizing behavior of firms and households, but it integrates a number of features that are essential for being able to provide an empirical analysis of the likely economic impacts.

First, to characterize abatement opportunities in the electricity sector, we use data on all 16,891 electricity generators active in 2006 published by the Energy Information Agency (EIA) (2007a). Generators are owned by a set of operators, and we identify 319 operators subject to cost-of-service regulation (EIA, 2007b). Regulated operators are treated as cost-minimizers charging average costs, whereas generators owned by non-regulated operators trade on imperfectly competitive regional wholesale markets.[4] By providing a structural "bottom-up" representation of abatement options in the electricity sector, we avoid using overly simplistic aggregate production functions typically employed in aggregated economy-wide general equilibrium models for electricity generation (Paltsev et al., 2005; Goulder et al., 2010). On the one hand,

[3] Alternatively, the authority in charge of price regulation could require any profits from freely distributed emissions permits to be transferred in a lump-sum manner to consumers. Even if the regulatory authority passes on the value of permits independently from the amount of electricity consumed, associating such payment with electricity bills may still induce distortions in consumption behavior to the extent that consumers do not (are not able to) separate the lump-sum payment from the actual price paid per unit of electricity.

[4] Although the degree of competition on wholesale markets is not the primary focus of this paper, it matters in the outcome of market-based environmental policy (e.g. Malueg, 1990). We thus follow Bushnell et al. (2008) and Fowlie (2009) and model wholesale markets as a set of large Cournot players interacting with a competitive fringe.

it enables us to capture some of the complexity of the market structure of the U.S. electricity sector. On the other hand, and relevant for studying the impact of a carbon pricing policy, substitution among different types of electricity technologies is modeled at the generator-level and is based on detailed data for generation costs, fuel switching possibilities, and time-varying (diurnal and seasonal) demand for electricity (see Lanz and Rausch, 2011).

Second, we embed the operator-level representation of electricity generation into a static general equilibrium model of the U.S. economy calibrated based on a set of regional Social Accounting Matrices for 2006. The sub-national detail of the model allows us to capture region-specific detail of energy use and production of various industries and final consumption sectors, and also how electricity demand by private and industrial consumers might change in response to a carbon pricing policy. Moreover, it characterizes abatement possibilities in non-electricity sectors and allows us to evaluate the equilibrium price for tradable emissions permits and economy-wide welfare costs of alternative initial allocations of emissions permits.

Third, to illustrate the distributional impacts of alternative policy design, we build on previous work by Rutherford and Tarr (2008) and Rausch et al. (2011) and integrate "real" households as individual agents in the model. In particular, we include all 15,588 respondents from the Consumer Expenditure Survey (CEX), a representative sample of the U.S. population (Bureau of Labor Statistics (BLS), 2006), as individual households in the model. Using an economy-wide model with heterogeneous consumers allows us to measure impacts both on the uses- and source-side of income, i.e. how do consumers spend and earn their income.[5]

In our quantitative analysis, we consider two alternative bases to determine the quantity of free permits allocated to price-regulated firms, namely historic emissions (i.e., grandfathering) or historic output.[6] As free allowances distributed to price-regulated firms effectively work as a subsidy of electricity rates, allocating permits based on benchmark emissions mitigates electricity price increases of the most CO_2-intensive operators. While this can partially smooth price differentials across operators, it is likely to magnify distortions associated with free allowances.

[5] See, for example, (Musgrave, 1964) for a formal definition of the "uses- and source-side of income" terminology and Fullerton and Heutel (2007) for a discussion of general equilibrium-based incidence measures of environmental taxation.

[6] While the present paper does not explicitly discuss international trade consequences of climate policies, one important motivation for using an output-based allocation is that it mitigates international competitiveness effects for firms subject to the carbon regulation (see Böhringer and Lange, 2005).

In contrast, using benchmark output as a basis for allowance allocation provides an intermediate case, as it equalizes the subsidy rate across regulated operators and thus partially preserves the link between emissions intensity and output prices.

Besides the aforementioned literature that is focused on the choice between auctioning and free allowances, this paper is germane to a number of studies that have investigated the implications of price-regulation for emissions trading policies. Theoretical work by Bohi and Burtraw (1992), Coggins and Smith (1993) and Fullerton et al. (1997) show that cost-of-service regulation can induce inefficient abatement behavior, potentially increasing the welfare costs of a cap-and-trade policy. Paul et al. (2010) and Burtraw et al. (2009) investigate the impacts of various assumptions about allowance allocation in the context of cap-and-trade policies in a numerical simulation model of the U.S. electricity sector that incorporates regional detail about cost-of-service regulation. While they find, in line with this paper, that distributing free allowances to regulated electricity producers substantially increase the equilibrium carbon price, looking at the electricity sector alone prevents addressing the broader policy-design question raised by the present paper. A related paper by Rausch et al. (2010) uses a general equilibrium model to investigate the welfare impacts of subsidized electricity prices, but assumes that the entire electricity produced in the U.S. is subject to price regulation. Our contribution relative to Rausch et al. (2010) is to identify the drivers of welfare impacts in a model capturing heterogeneity in the electricity markets both within and across regions, as well as across households.

The structure of the paper is as follows. Section 2 employs a stylized partial equilibrium model to illustrate the fundamental implications of alternative allowance distribution in the presence of price-regulated firms. Section 3 provides some background about U.S. electricity markets, price regulation, and CO_2 emissions. Section 4 describes the numerical model used to quantify the economic impacts of alternative allowance allocation designs. Section 5 lays out the policy scenarios, reports our quantitative results, and discusses our findings and assumptions. Section 6 concludes.

2 Pricing Behavior of Firms and Allowance Allocation

Our theoretical analysis builds on Fisher (2001) and Böhringer and Lange (2005) who study alternative allowance allocation rules in a one-sector partial equilibrium setting. In the present paper, we extend their analysis to consider the equilibrium outcome when the supply-side of the market is controlled by a price-regulated monopoly, and show how the outcome with free allowances differs from the social optimum. The analytical expressions we obtain provide the basic intuition for the results derived from numerical general equilibrium analysis reported in Section 5.

Consider a good X whose production entails total emissions E and associated environmental damages $G(E) > 0$, with $\partial G(E)/\partial E > 0$ and $\partial^2 G(E)/\partial E^2 \geq 0$. Emissions are the product of output level X and the emissions rate ζ. The marginal cost of production is constant in output but decreasing and convex in the emissions rate, i.e. $c(\zeta) \geq 0, \partial c(\zeta)/\partial\zeta < 0, \partial^2 c(\zeta)/\partial\zeta^2 > 0$. The social optimum on the market for good X is defined as the sum of consumer surplus minus production costs and environmental damages:

$$\max_{X,\zeta} \int_0^X P(s)ds - c(\zeta)X - G(E) \tag{1}$$

where $P(\cdot)$ is the inverse demand function, assumed decreasing and differentiable in X. The first order conditions imply that:

$$P(X) = c(\zeta) + \zeta P_e \tag{2}$$

$$P_e = -\partial c(\zeta)/\partial\zeta = \partial G(E)/\partial E \tag{3}$$

Equation (2) is the condition equating the marginal willingness to pay for good X to marginal private production cost plus marginal damage cost. Equation (3) states that the optimal value of emissions P_e is equal to the marginal cost of abatement and to the marginal damage. These two equations together determine output X^* and emissions rate ζ^*, and thus the socially optimal level of emissions.

In a competitive setting with a continuum of symmetric firms, the social optimum can be decentralized through an emissions trading policy constraining emissions at its efficient level $\overline{E} = X^*\zeta^*$. By requiring each firm to surrender one permit per unit of emissions, the equilibrium allowance price P_e acts as an Pigovian tax on emissions. For the time being, we assume that P_e is exogenous to the firm's decisions (we will return to this issue below). Denoting the amount of free allowances received by firm i by $\phi_i \geq 0$, a representative firm would chose x_i and ζ_i to maximize $\pi_i = x_i[P(X) - c(\zeta_i) - \zeta P_e] + \phi_i P_e$, so that first order conditions reduce to Equations (2) and (3). As well known, in a competitive setting any free allowances received by the firms do not affect the firms' decisions.

We now turn to the case where there is a monopoly for product X and price-setting behavior is regulated. If permits are auctioned, the profit function of the monopolist is given by: $\pi_m = X[P(X) - c(\zeta_m) - \zeta P_e]$. We consider the standard form of price regulation where the price of good X is set so that the monopolist makes no economic profit. Hence the price for good X is obtained by setting $\pi_m = 0$:

$$P(X) = c(\zeta_m) + \zeta P_e \, . \tag{4}$$

Price regulation thus induces the monopoly to chose a socially efficient level of output. Furthermore, if the regulator successfully induces the monopolist to minimize costs $c(\zeta_m) + \zeta P_e$, as we will assume throughout the paper, the opportunity cost of allowances enters the firm's decision, and the first order condition $P_e = -\partial c(\zeta)/\partial \zeta$ coincides with the social optimum (Equation (3)).

When allowances are initially free, profits of the regulated monopolist now include free allowances: $\pi_m = X[P(X) - c(\zeta_m) - \zeta P_e] + \phi_m P_e$. Setting this expression to zero, the equilibrium price of good X is:

$$P(X) = c(\zeta_m) + \zeta P_e - \phi_m P_e / X \, . \tag{5}$$

From Equation (5) it is straightforward to see that the value of free allowances $\phi_m P_e$ lowers the *cum-regulative* price of good X in proportion to the level of output. As a result, consumption of the polluting good will be distorted, and output decisions will typically not be efficient. However, while total emissions will be affected through output decisions, the cost minimizing choice of the emissions rate is not affected and remains at its efficient level.

In order to compare equilibrium behavior of the monopolist when allowances are auctioned and when these are freely allocated, we posit an iso-elastic demand function $P(X) = \alpha M X^{-\frac{1}{\beta}}$, where $M > 0$ is income, β is the price elasticity and $\alpha \in (0, 1)$ is a share parameter measuring expenditures on good X within total income. When allowances are auctioned, we will further account for the fact that the value of allowances $P_e\phi_m$ is part of households' income. Comparing Equations (4) and (5) yields the following proposition:

Proposition 1.

(a) *For a regulated monopoly, the ratio of output under free allowances, denoted by X_{Free}, to that under an auction X_{Auct} is:*

$$\frac{X_{Free}}{X_{Auct}} = \left(\frac{\alpha M + \phi_m P_e X_{Free}^{1/\beta - 1}}{\alpha(M + \phi_m P_e)} \right)^{\beta}$$

(b) *If $\beta = 1$, output by the regulated monopoly is larger under free permits ($X_{Free} > X_{Auct}$), and the difference (i) increases with the value of allowances $\phi_m P_e$, (ii) decreases with α, and (iii) increases with $\phi_m P_e/M$.*

Proof. See Appendix A. □

The difference between X_{Free} and X_{Auct} measures the distortion induced by freely allocating allowances, and effectively subsidizing the price of X, relative to auctioning permits. Proposition 1 exposes two key determinants of the magnitude of this distortion. First, distortions increases with the value of permits, $\phi_m P_e$. This implies that economies for which a large share of emissions stems from regulated producers will potentially suffer relatively large distortions. Second, under free permits output for goods might be higher or lower than under auctioning depending on the demand elasticity and strength of the income effects. For β less than one, which is the relevant case for electricity, a more inelastic demand may in fact give rise to larger distortions.

Proposition 1 also bears out two important implications. On the one hand, if the good produced by the regulated firm represents a relatively large share of income, alternative allocation mechanisms have a similar impact on the quantity produced, and hence the difference between X_{Free} and X_{Auct} will be small. In turn this will mean lower distortions on the market for permits. On the other hand, when the value of free permits represents a large share of income, or when the ratio $\frac{\phi_m P_e}{M}$ is high, the difference between X_{Free} and X_{Auct} will be large, as household may lose a sizable fraction of income. This implies that distortions on the permit market will be large.

There are a number of aspects affecting the size of the distortion that are not included in the simplified partial equilibrium analysis. First, the equilibrium permit price is treated as exogenous. Under a fixed emissions cap, higher output by regulated firms will induce a higher equilibrium permit price. In turn, an increase in P_e will induce a reduction in the emissions rate (Equation (3)), and imply an increase in the marginal cost of production above the social optimum. To see why this is relevant, consider the case in which the emissions cap also applies to firms other than the regulated monopoly. As pricing decisions by regulated and non-regulated firms are linked through the market for emissions permits, an increase in the demand for permits in the regulated sectors will induce higher abatement in non-regulated sectors. The extent to which abatement will be shifted from the regulated to the non-regulated firms depends on the relative marginal abatement costs among sectors.

The remainder of this paper thus provides a quantitative economy-wide assessment of the issues touched upon above by investigating the case of the U.S. economy. We begin with a brief description of the nature of price regulation in the U.S., and how CO_2 intensity of regulated and non-regulated electricity firms vary across regions.

3 Electricity Markets, Price Regulation, and CO_2 Emissions in the U.S.

Historically, the U.S. electricity sector has developed through regional monopolies, where generation, transmission and distribution are vertically integrated (Joskow, 2008). On each market, electricity rates are regulated by a "Public Utility Commission" to protect customers from monopoly pricing. The "rate of return" regulation allows utilities to recover prudently incurred operating costs, so that consumers pay a price comparable to the average accounting cost of service. In the 1970s, a movement of deregulation took place across numerous regulated industries (Winston, 1993), and the 1978 Public Utilities Regulatory Policies Act provided initial legal support for a separation of generation from transmission. In addition, limited economies of scale in modern generation technologies and advances in high-voltage transmission technologies increased opportunities for mutually beneficial trades to take place in a highly balkanized system (Joskow and Schmalensee, 1983).

Through regulatory and technological evolution, traditional regional monopolies were progressively complemented by investor-owned independent power producers that had no network ownership and directly supplied large industrial activities. This situation created a demand from other industrial consumers to be able to purchase electric power from alternative suppliers, particularly in areas with high electricity prices (Joskow, 2005). Through the Energy Policy Act of 1992, the Federal Energy Regulatory Commission (FERC) could order electric utilities to allow current to transit on their network, implicitly inviting market transactions to take place on the network for a fee. In 1999, the FERC called for the creation of Regional Transmission Organizations in order to provide independent supervision of transmission grids.

The trend towards competitive wholesale markets slowed down significantly after the 2000-01 electricity crisis in California (Joskow, 2008). As of 2006, the base year for our analysis, around 60% of electric power is generated by regulated utilities (EIA Form 906-920, 2007b). The electricity sector in continental U.S. can can be divided in 10 regions, which we approximate by state-level borders in Figure 1.[7] While in most of these regions the administration of transmission networks has been transferred to an Independent System Operator, in all regions there remain a number of regulated utilities (Table 1).[8] For example, in the state of Texas, where most electricity producers have joined the ERCOT wholesale market, some 20 regulated monopolies are active within state borders. In regions such as NY and CA, a small number of regulated operators hold large, mainly hydroelectric capacity, while in SEAST, SPP and MOUNT, electric power is almost entirely generated by regulated operators.[9]

Empirical evidence suggests that regional wholesale electricity markets are best described as oligopolies (see for example, Wolak, 2003; Mansur, 2007; Puller, 2007; Bushnell et al., 2008). While conventional market concentration indexes have drawbacks as a measure of imperfect competition for non-storable goods (Borenstein et al., 1999), the Herfindahl-Hirschman Indexes show that wholesale markets in highly regulated regions exhibit the highest concentration. The only exception is the wholesale market in SEAST, which features a large number of relatively

[7] These regions are: California ISO (CA), Northwest Power Pool (NWPP), Mountain Power Area (MOUNT), Electric Reliability Council of Texas (ERCOT), Southwest Power Pool (SPP), Midwest ISO (MISO), Southeast Power Pool (SEAST), Pennsylvania-New Jersey-Maryland Interconnection (PJM), New York ISO (NY), and New England ISO (NENGL).

[8] Because there are regulated electricity producers in each regions, wholesale electricity markets effectively cover a smaller area than regions reported in Figure 1. But for simplicity we refer to geographical areas by the name of associated regional wholesale electricity market.

[9] Regulated operators also sell and buy power through wholesale transactions. For our purposes, the key feature of regulated operators is that their rates reflect generation costs as these are still subject to approval by the Public Utility Commissions.

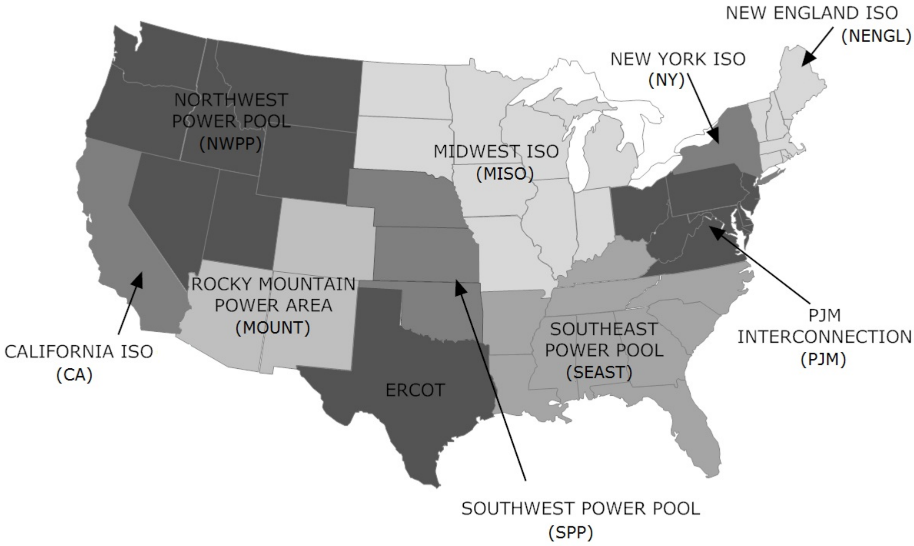


Figure 1: State-level aggregation of national electric power market regions.

small non-regulated producers.

The CO_2 content of electricity from regulated operators is on average about one third higher as compared to non-regulated operators, with large variations at the regional level. For example, electricity produced by regulated operators in NENGL and ERCOT emits almost twice as much CO_2 per MWh as compared to non-regulated operators in these regions, as the latter hold large natural gas capacity. In CA, NWPP, and NY, regulated operators mainly hold hydroelectric resources, and the CO_2 intensity of non-regulated operators is higher. Moreover, the fragmentation of the U.S. electricity sector and differences in generation technologies available on each market implies that the impact of CO_2 pricing policies will be highly heterogeneous.

4 Data, Modeling Framework and Computational Strategy

To quantitatively investigate how price-regulated firms affect the outcome of an emissions trading system under alternative permit allocation rules, we employ a numerical general equilibrium model of the U.S. economy that is calibrated to conditions in the year 2006. Our model features (i) a technology-rich "bottom-up" generator-level model of electricity generation, (ii) market structure detail for regional electricity markets, including regulated markets and imperfectly competitive regional wholesale markets, (iii) a multi-sector and multi-region description

Table 1: Regional electricity generation, market structure and CO_2 intensity in 2006.

Region	Generation	Regulated generation			Non-regulated generation			
	(TWh)	%	N[a]	tCO_2/MWh	%	N[b]	HHI[c]	tCO_2/MWh
SEAST	1,126.6	87.0	87	0.61	13.0	287	310	0.60
SPP	142.4	86.2	133	0.78	13.8	30	1,570	0.42
MOUNT	214.1	85.7	38	0.73	14.3	57	1,160	0.38
NWPP	317.4	79.5	64	0.38	20.5	154	1,130	0.63
MISO	724.4	67.7	305	0.85	32.3	315	1,680	0.47
CA	231.3	49.8	39	0.19	50.2	317	220	0.42
PJM	665.0	35.5	51	0.79	64.5	259	580	0.58
NY	142.9	29.6	14	0.30	70.4	148	550	0.37
ERCOT	348.9	13.2	20	0.84	86.8	157	820	0.52
NENGL	132.8	4.8	28	0.79	95.2	214	510	0.40
US-wide	4045.7	61.2	731	0.65	38.8	1938	–	0.51

Notes: [a] Number of traditional vertically integrated electric utilities. [b] Number of independent electric power producers. [c] Herfindahl-Hirschman index. *Sources*: Authors' own calculations. Data on generation and operator's regulatory status is from EIA Form 906-920 (2007b). CO_2 emissions are based on fuel consumption for each operator (EIA Form 906-920, 2007b) and fuel-specific CO_2 emissions factors (EIA, 2009a).

of economy-wide activities, and (iv) heterogeneous consumers represented as individual economic agents based on micro-household survey data.

The model is formulated as a mixed complementarity problem (Rutherford, 1995) distinguishing two classes of equilibrium conditions: zero economic profits and market clearing. The former condition determines a vector of activity levels and the latter determines a vector of prices.[10] Given the large number of electricity markets and households, however, it is computationally not feasible to operate directly on the system of equations defining the vector of equilibrium prices and quantities. We therefore make use of recent advances in decomposition methods pertaining to the computation of equilibria in numerical general equilibrium models with bottom-up technology representation Böhringer and Rutherford (2009) and many heterogeneous households Rutherford and Tarr (2008). This involves formulating electricity markets and households optimization problems as partial equilibrium problems, and consistently inte-

[10] The mixed complementarity format embodies weak inequalities and complementary slackness, relevant features for problems with corner solutions and bounds on specific variables. Moreover, as it essentially solves the system of Karush-Kuhn-Tucker conditions of the associated optimization problem, the mixed complementarity formulation can accommodate multiple decision-makers engaged in strategic interaction, whose objective are not integrable. Formally, given a function $F\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$, we search for a vector $z \in \mathbb{R}^n$ such that $F(z) \geq 0$, $z \geq 0$, and $z^T F(z) = 0$. A complementary-based approach has been shown to be convenient, robust, and efficient (Rutherford, 1995). We formulate the numerical model in the GAMS software and solve it using the PATH solver (Dirkse and Ferris, 1995).

grate the solution to these problems into an economy-wide framework.

The remainder of this section proceeds as follows. We begin by describing the electricity generation model and then provide a brief overview of the general equilibrium model, followed by a discussion of how we incorporate household heterogeneity. We then elaborate on the numerical techniques we need to employ in order to solve for a general equilibrium of the integrated model. Further details about the modeling framework is provided in Appendix B.

4.1 Electricity Generation

4.1.1 Data

We use 2006 data on all 16,891 generators active in continental U.S., with information on generation technology, capacity (i.e. maximum output), and up to three fuels that can be used (EIA Form 860, 2007a).[11] Each generator is matched to plant level data reported in EIA Form 906-920 (2007b), where a plant can include multiple generators. EIA Form 906-920 provides plant-level monthly output per technology and fuel type, fuel consumption, as well as the operator of each plant, its regulatory status (i.e. whether it is a traditional vertically integrated electric utility or an independent operator), and its region of operation. Our data set therefore comprises information on the portfolio of generation technologies of each operator and its regulatory status.

The marginal cost of generation (in US$/MWh) comprises fuel costs and operation and maintenance (O&M) costs. Fuel costs are based on plant-specific efficiency (in MBTU/MWh), calculated using fuel consumption and electricity output reported in EIA Form 906-920 (2007b) and state-level fuel prices for 2006 (in US$/MBTU) from EIA (2009d).[12] Second, as we do not observe O&M costs at the generator level, we use technology-specific data from EIA (2009b). This includes labor, capital, material and waste disposition costs per MWh.

CO_2 emissions depend on the CO_2 content of the fuel used to generate electricity (in tCO_2/ MBTU), as reported by EIA (2008). Implicitly, the CO_2 intensity of each operator also depends on the efficiency of the plant, as it determines the fuel requirement to generate electricity.

[11] We obtain the dependable capacity by scaling installed capacity figures from EIA Form 860 (2007a) with technology-specific availability data reported by the North American Electric Reliability Council (2007).

[12] Since information on output and fuel consumption at the generator level is not available, generators that belong to the same plant and share the same combination of fuel and technology are assumed to have the same efficiency.

The benchmark demand for electricity (i.e. in the absence of a CO_2 policy) at each regulated operator is given by observed monthly output (EIA Form 906-920, 2007b). We consider only the 319 regulated operators with annual output greater than 10 GWh.[13] To determine the demand on regional wholesale markets, we first map all non-regulated operators to their wholesale market region, and then determine the monthly benchmark demand by summing monthly electricity output for all non-regulated operators within each region. For both regulated and wholesale markets, we capture variations in electricity demand over the year by dividing the year into nine load segments. Specifically, monthly demand on the 319 regulated markets and 10 wholesale markets is aggregated into three seasons (summer, winter and fall/spring), and then seasonal demand is divided into three load blocks (peak, intermediate, and base load) based on region- and season-specific load distribution data (EIA, 2009b).

4.1.2 Regulated Electricity Markets

Regulated operators $f = 1, \ldots, 319$ are assumed to minimize generation costs to meet the demand, and thus implicitly construct a piece-wise linear supply function by ranking available technologies by increasing marginal cost (the "merit order"). In equilibrium, generator g is thus active in load segment $t = 1, \ldots, 9$ if its marginal cost c^g is lower than the marginal cost of the generator used to cover the last unit of demand, denoted C_t^f. This is summarized by the following complementarity condition:

$$c^g + \nu^g \tau + \mu_t^g \geq C_t^f \quad \perp \quad Y_t^g \geq 0 \tag{6}$$

where Y_t^g is the output level, ν^g is the CO_2 intensity, τ denotes the price of emissions and $\perp$ indicates a complementary relationship. μ_t^g represents the shadow value of installed capacity, and it is the complementarity variable of the capacity constraint of each generator:

$$Y_t^g \leq \kappa_t^g \quad \perp \quad \mu_t^g \geq 0\,, \tag{7}$$

[13] This roughly corresponds to the yearly consumption of 1,000 households. Generation from the 412 regulated operators that are not included in the model represents less than 0.1% of electricity generated in each region.

where κ_t^g is the dependable capacity of generator g in load segment t. Generators listed with multiple fuel options endogenously select the least-cost fuel based prevailing fuel prices.

The equilibrium marginal generation cost C_t^f is determined by a market clearing condition for each load segment:

$$\sum_{g \in G_f} Y_t^g \geq d_t^f \quad \perp \quad C_t^f \geq 0\,, \tag{8}$$

where G_f denotes the set of generators owned by regulated operators f and d_t^f is electricity demand in t.

The price of electricity at regulated operator f, P^f, is given by the average generation costs:

$$P^f = \frac{\sum_{g \in G_f} \sum_t Y_t^g c^g + \nu_t^g \tau}{D^f} - s^f\,. \tag{9}$$

where $D^f = \sum_t d_t^f$ is the total demand for generation at operators f over the year, and s^f is a firm-specific subsidy rate that reflects the value of free allowances received, denoted V_f (see Equation 5):

$$s^f = \frac{V_f}{D^f}\,. \tag{10}$$

While this pricing rule is an important simplification of reality, notably because of regional idiosyncrasies in the application of cost of service regulation and the existence of other rules such as block-pricing, it mainly captures the fact that the price signal for many consumers reflects some measure of average production costs and is close to constant throughout the year. Note that capacity rents μ_t^g are not included in the price.

Under an emissions trading policy, generation costs increase proportionally to the emissions rate ν_t^g. Since we assume that regulated operators are minimizing costs, the fact that they must surrender allowances induces fuel switching and a reordering of generators along the supply schedule (merit order effect). The demand response at operator f is a linear approximation of the non-linear economy-wide demand calibrated at benchmark price $\overline{P}^f$ and demand $\overline{D}^f$:

$$D^f = \overline{D}^f \left(1 + \epsilon \left(\frac{P^f}{\overline{P}^f} - 1\right)\right)\,, \tag{11}$$

where $\epsilon < 0$ is the *local* price elasticity of demand. The demand in load segment t is then given by: $d_t^f = D^f\, \overline{d}_t^f / \overline{D}^f$.

4.1.3 Wholesale Electricity Markets

Each region $r = 1, \dots, 10$ is associated with a wholesale market which brings together generators owned by non-regulated operators. In line with Bushnell et al. (2008) and Fowlie (2009), we assume that operators holding more than 3% of wholesale generation capacity behave as Cournot players.[14] Smaller operators act as a price-taking competitive fringe. The Cournot-Nash equilibrium unit profit function for strategic players (denoted by the set G_r^{cournot}) and non-strategic players (denoted by the set G_r^{fringe}) are respectively:

$$\pi_t^g = \begin{cases} p_t^r + \frac{\partial D^r(p_t^r)^{-1}}{\partial Y_t^g} - c^g - \mu_t^g - \nu_t^g \tau & \text{if } g \in G_r^{\text{cournot}} \\ p_t^r - c^g - \mu_t^g - \nu_t^g \tau & \text{if } g \in G_r^{\text{fringe}} \, . \end{cases} \tag{12}$$

Here p_t^r is the wholesale price and $D_t^r(p_t^r)^{-1}$ denotes the inverse demand function. Equilibrium electricity output by each generator is determined by the following zero profit condition:

$$-\pi_t^g \geq 0 \quad \perp \quad Y_t^g \geq 0 \, . \tag{13}$$

Non-regulated operators who own generators with marginal cost below the market clearing price earn capacity rents μ_t^g according to:

$$Y_t^g \leq \kappa_t^g \quad \perp \quad \mu_t^g \geq 0 \, . \tag{14}$$

The wholesale equilibrium price p_t^r is the complementary variable associated with the following market clearing conditions:

$$\sum_{g \in G_r} Y_t^g \geq d_t^r \quad \perp \quad p_t^r \geq 0 \, , \tag{15}$$

where G_r denotes the set of generators in wholesale market r.

The wholesale price signal transmitted to consumers is an output-weighted average of the

[14] As in Bushnell et al. (2008), we find that this simple representation of wholesale markets provides a relatively good fit to observed outcomes (see Appendix B.2). Nevertheless, and mainly because this is not the main focus of the paper, this representation of deregulated electricity markets features several important simplifications. First, strategic firms may withhold capacity to exercise market power without detection, so that capacity data may already reflect strategic behavior. Second, firms interact dynamically in several power market simultaneously, so that the market power of each individual firms will likely be overestimated by our threshold. Third, in our representation firms only receive payments when they sell power on the market, and not for making capacity available on the market. We thank one anonymous reviewer for pointing out these limitations.

prices in each load segment:

$$P^r = \frac{1}{D^r} \sum_t p_t^r d_t^r \, . \tag{16}$$

The annual demand response for wholesale power is locally approximated by a linear demand function:

$$D^r = \overline{D}^r \left(1 + \epsilon \left(\frac{P^r}{\overline{P}^r} - 1 \right) \right) \, , \tag{17}$$

so that demand in load segment t is given by: $d_t^r = D^r \, \overline{d}_t^r / \overline{D}^r$.

4.2 Economy-wide General Equilibrium Model

4.2.1 Data

We use 2006 state-level economy-energy data where each state is described by a social accounting matrix. The IMPLAN data set (IMPLAN, 2008) provides an input-output representation of social accounts for production, consumption and trade for 509 commodities, existing taxes, government revenues and transfers. To expand the characterization of energy markets in the IMPLAN data, we supplement it with data on energy quantities and prices for 2006 (EIA, 2009c). Energy commodities identified in our study include coal (COL), natural gas (GAS), crude oil (CRU), refined oil (OIL), and electricity (ELE); this allows us to account for the substitutability between different energy sources in industrial production and final demand. Our commodity aggregation further comprises five non-energy composites: energy-intensive products (EIS), other manufacturing products (MAN), agriculture (AGR), transportation (TRN), and Services (SRV). Primary production factors included are labor, capital, land, and fossil-fuel resources.[15]

We aggregate state-level data into 10 U.S. regions as identified in Figure 1 in order to approximate wholesale transmission regions by state-level border. Table 2 reports benchmark CO_2 emissions, sectoral shares of total emissions, and emissions intensity by region. CO_2 emissions from regulated electricity generation represent about 30% of total national emissions, transportation (industrial and private) being the other main contributor besides electricity. There is significant variation among regions in terms of emissions intensity of industrial output. For

[15] The aggregation and reconciliation of IMPLAN state-level economic accounts needed to generate a micro-consistent benchmark data set which can be used for model calibration is documented in Rausch and Rutherford (2009).

Table 2: Sectoral CO_2 emissions and regional emissions intensity.

Region	Total emissions	Sectoral share of emissions (%)								Emissions intensity
		Electricity sector		Non-electricity sectors						
	(MtCO_2)	Wholesale	Regulated	AGR	EIS	SRV	TRN	MAN	CONS[a]	(tCO_2/US$)
ERCOT	657.7	27.0	6.3	1.0	25.8	2.5	31.3	3.3	2.7	0.71
SPP	211.8	5.1	43.5	3.1	8.6	3.0	29.0	4.1	3.6	0.65
SEAST	1371.4	5.7	41.9	0.8	10.3	1.3	35.4	1.9	2.6	0.52
MISO	1155.2	10.2	35.4	1.6	9.2	3.7	27.2	4.1	8.5	0.51
MOUNT	261.6	1.2	47.0	0.7	4.3	2.2	38.4	2.4	3.8	0.47
PJM	974.5	25.2	19.1	0.2	10.4	3.5	33.1	1.7	6.7	0.44
NWPP	347.0	12.4	27.1	1.7	5.0	2.0	40.4	1.6	9.8	0.43
NENGL	173.2	26.3	3.1	0.4	3.4	3.2	45.5	0.7	17.4	0.24
CA	387.0	8.9	5.2	0.6	6.4	3.6	62.8	5.3	7.2	0.24
NY	194.4	18.4	11.5	0.3	4.5	11.9	36.5	1.3	15.7	0.22
US-wide	5733.8	13.9	27.4	0.9	10.6	3.0	35.2	2.7	6.3	0.44

Sources: CO_2 emissions from the electricity sector are based on simulated fuel consumption in the benchmark and fuel-specific CO_2 emissions factors (EIA, 2009a). Emissions calculations for non-electricity sectors are based on EIA's State Energy Data System (EIA, 2009c), witch also underlies the calculation of emissions intensity together with economic value flows for industrial output from IMPLAN data 2008. [a]: Emissions from private consumption comprising from natural gas and fuel oil for heating buildings.

example, ERCOT, as the most CO_2 intensive region, shows an emissions intensity that is three times as large as those of CA, NENGL, and NY.[16]

4.2.2 Model Overview

Economy-wide interactions are represented by a static numerical general equilibrium model of the U.S. economy. We here provide a brief description of the model structure, and presents the equilibrium conditions of the model in Appendix B.1.

For each industry but electricity generation ($i = 1, \dots, I$), gross output X_i is produced using inputs of labor (L_i), capital (K_i), natural resources (R_{zi}, $z = 1, \dots, Z$) including coal, natural gas, crude oil, and land, and produced intermediate inputs (x_{ji}, $j = 1, \dots, I$) including electricity.[17] All industries are characterized by constant returns to scale, except for fossil fuels and agriculture, which are produced subject to decreasing returns. Apart from electricity, all commodities are traded in perfectly competitive markets, where firms maximize profits given technology and prices. Labor is assumed to be fully mobile across sectors within a given region,

[16] This can be traced back to large-scale activities in oil refining and energy-intensive industries in Texas.

[17] For ease of notation, we omit the region index when no ambiguity can result.

but immobile across regions. Capital is mobile across sectors and regions.

We distinguish three different representations of *intra*-national trade which depends on the type of commodity and associated regional integration. First, non-energy goods are treated as regionally heterogeneous and the price transmitted to producers and consumers is a CES index of varieties from U.S. regions. Second, domestically traded energy goods (excluding electricity) are assumed to be homogeneous products, so that each region trades with a national pool where all regions supply and demand goods. This reflects the high degree of integration of U.S. market for natural gas, crude and refined oil, and coal. Third, for electricity we approximate the three asynchronous interconnects in the U.S. by defining three regional electricity pools: the Eastern Interconnection, Western Electricity Coordinating Council (WECC), and the Electric Reliability Council of Texas (ERCOT).[18] Each region thus trades directly with its regional pool, within which electricity is homogeneous, and there is no electricity trade between regional pools.

The U.S. economy as a whole is modeled as a large open economy, so that the U.S. can affect world market prices. The international trade closure of the model is determined through a national balance-of-payments constraint. Hence the total value of U.S. exports equals the total value of U.S. imports accounting for an initial balance-of-payments deficit given by 2006 statistics.

In each region, a single government entity approximates government activities at all levels – federal, state, and local. The government raises revenues through taxes, purchases goods and services, and provides lump-sum transfers to households (i.e., social security). Government consumption is represented by a Leontief composite of goods $x_i, \ldots, x_I$ where benchmark value shares are based on social accounting matrix data. Revenues are based on observed ad-valorem output taxes, corporate capital income taxes, and payroll taxes (employers' and employees' contribution).

[18] In terms of the regional aggregation described in Figure 1, the Eastern Interconnection thus comprises SPP, MISO, SEAST, PJM, NY, and NENGL, and the WECC comprises CA, NWPP, and MOUNT.

4.3 Heterogeneous Households

4.3.1 Data

We use data on 15,588 households from the 2006 CEX survey (BLS, 2006), which provides consumption expenditures and income sources for a representative sample of the U.S. population.[19] Since the CEX focuses primarily on recording households' spending, a well-known issue with this survey is quality of income-side data. First, households with income above a certain level are "top-coded" and their income is replaced with the national average. We observe a substantial amount of top-coding for the top 4% of the income distribution (with pre-tax income above US$250k), and our analysis cannot break out the top 4% of the income distribution.

Second, capital income is low as compared to data reported in official National Accounts (e.g. Deaton, 2005; Rutherford and Tarr, 2008). Metcalf et al. (2010) also suggest that capital income may misrepresent capital holdings across income groups. Indeed, if financial assets are disproportionately held by higher income groups then the CEX capital income measure will be biased towards more capital holdings in lower income groups. To supplement capital income data, we use data from the 2007 Survey of Consumer Finances (SCF Federal Reserve Board, 2007), which provides detailed information on different components of wealth holdings. The SCF combines a core representative sample with a high income supplement, which is drawn from the Internal Revenue Service's Statistics of Income data file. This data thus captures both the wealth at the top of the distribution and wealth portfolio of other households. Following Metcalf et al. (2010), we replace capital income reported in the CEX by imputed capital income based on capital income shares by income decile from SCF and total household income from CEX.

An other issue with the CEX data pertains to the implied tax rates reported by households. In particular, imputing personal income tax rates from tax payments in the CEX sample results in tax rates that are significantly lower than observed tax rates. For each households, we thus use data on 2006 average and marginal personal income tax rates by income decile from the

[19] Each household is interviewed every three months over five calendar quarters, and in every quarter 20% of the sample is replaced by new households. We include all households that report expenditures and income for 2006 even if they have only been interviewed for a subset of quarters in this year by following the procedure outlined in BLS (2006, p. 271).

Table 3: Selected expenditure and income shares (%) and median household income (2006 US$).

Income decile	Electricity	Natural Gas	Capital	Labor	Transfers	Capital-labor ratio	Median income
1	4.7	1.8	5.7	35.8	58.5	16.0	13,090
2	3.7	1.3	4.1	33.9	62.1	12.0	22,366
3	3.2	1.1	6.5	55.1	38.4	11.8	31,398
4	2.8	1.0	7.4	68.1	24.5	10.9	40,026
5	2.4	0.9	7.8	79.9	12.2	9.8	49,169
6	2.5	0.8	8.8	83.4	7.8	10.6	59,941
7	2.2	0.8	9.1	86.6	4.3	10.5	72,433
8	1.9	0.7	10.6	86.8	2.6	12.2	87,987
9	1.8	0.7	13.2	84.9	1.9	15.6	114,628
10	1.5	0.6	45.6	53.5	0.9	85.3	187,365
All	2.6	1.0	24.6	69.1	6.3	35.6	55,140

Notes: Population-weighted within-income group averages based on CEX data.

National Bureau of Economic Research's tax simulator (Feenberg and Coutts, 1993).

Finally, to obtain expenditure data that are consistent with the definition of consumption goods in our macroeconomic data, we aggregate expenditures into *Personal Consumption Expenditure* accounts, and mapped these to *North American Industry Classification System* accounts with a bridge matrix from the Bureau of Economic Analysis (2007). As savings are not reported directly in the CEX data, they are imputed as pre-tax household income minus the sum of consumption expenditures and tax payments. This ensures that pre-tax household income is equal to the sum of consumption expenditures, tax payments, and savings.

Table 3 reports expenditure shares for electricity and natural gas, and income shares for capital and labor by annual income decile.

4.3.2 Optimal Household Behavior

Each household is incorporated as a separate agent within the general equilibrium framework, so that aggregate consumption, labor supply, and savings result from the decisions of $h = 1, \ldots, 15,588$ households, each maximizing its utility subject to an income constraint. The preferences of each household is represented by a nested CES function that combines material consumption, savings, and leisure thus making consumption-investment and labor supply decisions endogenous. The nested utility structure is specified to reflect estimates of substitution elasticities among energy and non-energy goods (Paltsev et al., 2005) and reported in Appendix

B.1. Household income is derived from government transfers and from supplying regional markets with capital, labor, and natural resources.

4.4 Model Calibration and Computational Strategy

Model calibration is achieved by first constructing a consistent benchmark data set reconciling electricity generation data, the social accounting matrices, and data on household consumption and income. For example, cost-minimizing labor demand by electricity operators needs to be consistent with economy-wide equilibrium on the labor market, which in turn depends on labor supply decisions by the set of heterogeneous households. This procedure is described in Appendix B.2 and Appendix B.3 for electricity generation and households' demand respectively. Prices and quantities from the constructed benchmark data are then used to calibrate the value share and level parameters in CES production and consumption functions. This procedure ensures that the benchmark data is consistent with the notion of a general equilibrium (see e.g. Robinson, 1991).[20]

The aim of the solution method is to compute the vector of price and quantities that solves the system of simultaneous equations given by the equilibrium conditions of the models. Given the highly non-linear nature and large dimensionality of the numerical problem at hand, an integrated solution approach is not feasible. We make use of recent advances in decomposition methods to numerically compute the general equilibrium of the the integrated model in the presence of a policy shock. Note that given our calibration procedure, solving the model in the absence of a policy replicates the benchmark data.

The electricity sector and economy-wide general equilibrium components are solved based on a decomposition algorithm put forward by Böhringer and Rutherford (2009). As described in detail in Appendix B.2, the algorithm involves sequentially solving the electricity and economy-wide components under the same policy shock, ensuring consistency between general equilibrium prices and quantity of electricity produced and associated demand of inputs determined in the electricity generation model. Hence, consistency is also achieved in terms of prices and

[20] The benchmark situation on each wholesale and regulated electricity market is derived as an endogenous solution to the electricity generation model based on generation costs and benchmark demand. Figures reported in Appendix B.2 suggest that our model provides a good representation of generation costs, and also accurately predicts CO_2 intensity.

demands for fuels, capital, labor, and other commodities and services used to produce electricity.

Endogenous decisions by all households are integrated in the economy-wide framework through a decomposition algorithm based on Rutherford and Tarr (2008). The key idea is to compute a sequence of artificial agent equilibria which replicate choices of the many "real" households. The algorithm employs an iterative procedure which is undertaken after each solution of the electricity sector model, as described in Appendix B.3. This procedure ensures that the general equilibrium prices derived from the economy-wide model with a single representative consumer are consistent with partial equilibrium demands by individual households.[21]

5 Quantitative Results

5.1 Allowance Allocation Scenarios

In a general equilibrium setting, alternative policy designs with either auctioned or freely distributed emissions permits ultimately translate into a statement about how the value of allowances accrues to households. Let T_0 denote economy-wide CO_2 emissions in the benchmark, ξ the emissions reduction target (expressed as a fraction of benchmark emissions), τ the equilibrium allowance price, and ϑ the fraction of allowances or revenue retained by the government to achieve budget neutrality.[22] We can then write the value of allowances received by household h as:

$$A_h = T_0 \xi \tau (1 - \vartheta) \cdot (a_h + b_h + c_h) \,, \tag{18}$$

where a_h, b_h, and c_h denote the share of the allowance's value allocated to regulated electricity producers, non-regulated electricity producers, and non-electricity sectors respectively. Further, define λ_m, the share of allowances allocated to electricity market $m = r \cup f$, as a linear combination of the share of benchmark electricity emissions ($\overline{E}_m$) and benchmark electricity output

[21] Note that this procedure does not alter preferences of the "real" households nor does it rely on any form of aggregation of preferences; the single representative agent is simply used as a computational device to incorporate general equilibrium effects.

[22] ϑ is determined endogenously in each scenario as the amount of allowances required in equilibrium to compensate for changes in non-CO_2 tax revenue. This corresponds to a (non-distortionary) lump-sum tax and ensures that the aggregate budget of the government remains constant across all counterfactual equilibria.

$(\overline{O}_m)$:

$$\lambda_m = \alpha_E \frac{\overline{E}_m}{\overline{E}} + (1-\alpha_E)\frac{\overline{O}_m}{\overline{O}}\,, \tag{19}$$

where $\alpha_E \in [0,1]$, $\overline{E}$ represents benchmark emissions from the electricity sector, and $\overline{O}$ is total electricity generation in the benchmark.

We now define alternative allowance allocation designs in terms of a_h, b_h, c_h, α_E, and the value of free allowances distributed to price-regulated electricity producers, V_f. For the regulated sector, the parameters a_h and V_f reflect whether emissions permits are auctioned ($a_h > 0$ and $V_f = 0$) or freely distributed ($a_h = 0$ and $V_f > 0$). We consider the following three cases:

- Auctioning (AUCT): $a_h = \upsilon_h w \sum_g \lambda_g I_{f,h},\ V_f = 0,\ \alpha_E = 1,$
- Free allocation based on emissions (FREE_E): $a_h = 0,\ V_f = T_0 \xi \tau (1-\vartheta) \cdot w\lambda_f,\ \alpha_E = 1,$
- Free allocation based on output (FREE_O): $a_h = 0,\ V_f = T_0 \xi \tau (1-\vartheta) \cdot w\lambda_f,\ \alpha_E = 0,$

where υ_h denotes the weight of household h in total population, $w = \overline{E}/T_0$ is the benchmark share of emissions from the electricity sector and $I_{m,h}$ is an indicator variable which is equal to one if household h is a consumer in market m, and zero otherwise.

In order to isolate the potential loss in cost-effectiveness due to the presence of regulated firms in the economy, we keep the treatment of the other non-regulated sectors constant across all scenarios. In particular, we assume that all non-regulated firms, (i.e. non-regulated electricity operators and non-electricity sectors) receive free allowances in proportion to their benchmark emissions ($\alpha_E = 1$), or equivalently, a share of the revenue from auctioned permits that is proportional to their CO_2 emissions in the benchmark. We further assume that the value of free allowances distributed to non-regulated firms accrues to households in proportion to their share of capital income in aggregate capital income, denoted by κ_h. Intuitively, free allowances represent windfall profits, increasing revenues from capital ownership. For all scenarios we therefore assume that:

$$b_h = \kappa_h w \sum_r \lambda_r \tag{20}$$

$$c_h = \kappa_h (1-w)\,. \tag{21}$$

Different interpretations for the AUCT scenario are conceivable. It can be viewed as an emissions trading scheme in which permits are fully auctioned and the resulting revenue is recycled to households through per-capita lump-sum transfers. Alternatively, AUCT could be viewed as a situation in which allowances are allocated for free based on benchmark emissions, and where the entity that is regulating output prices of electricity firms (e.g., Public Utility Commissions in the U.S.) require the value of free allowances to be transferred to households through lump-sum transfers.

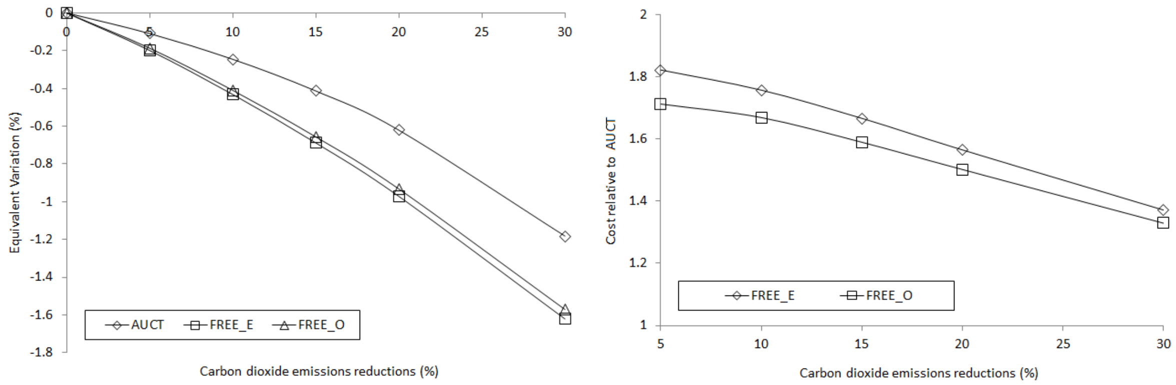
The FREE_E and FREE_O cases represent policies where free allowances are allocated based on benchmark emissions or output, respectively, and price-regulated firms transfer the value of free allowances through electricity rates. The value of allowances allocated to the regulated firm f, V_f, determines firm-specific subsidies s_f according to Equation (10).

5.2 Cost-effectiveness of Alternative Policy Designs

Figure 2 and Table 4 summarize the impacts of alternative cap-and-trade designs on aggregate welfare.[23] emissions permits are freely allocated, the welfare costs of the cap-and-trade policy for the same level of emissions reductions are between 40% and 80% higher relative to a policy design that chooses to auction permits. For a 20% emissions reduction target, the economy-wide efficiency cost is equivalent to an additional burden of around US$50 billion (in 2006 dollars). As shown in Figure 2 panel (b), the efficiency costs become smaller as the cap increases since, for higher targets, welfare costs due to freely allocating permits represent a smaller share of total welfare costs.

If electricity rates of regulated operators do not fully reflect the CO_2 price signal, electricity consumption and in turn CO_2 emissions from electricity production are sub-optimally high. The magnitude of efficiency costs is therefore closely related to the size of electricity price changes. As reported in Table 4, a 20% target under the AUCT scenario induces average electricity price

[23] Aggregate welfare costs are the weighted average of each household's equivalent variation as a percentage of full income, where a household's weight is proportional to its share of the total population.



(a) (b)

Figure 2: (a) Average welfare impacts and CO_2 abatement. (b) Excess welfare costs relative to AUCT.

increases of about 70% for regulated operators.[24] Because demand for electricity is relatively inelastic, this induces only a 12% US-wide decline in electricity output by regulated operators. Subsidizing electricity prices with free allowances substantially reduces the average increase of electricity prices. Under an emissions-based allocation (FREE_E), a 20% target raises electricity prices at regulated operators by 38% on average, and by around 45% under an output-based allowance allocation. Given the lower price increase under free permits, the associated change in output for regulated operators is substantially dampened (-4.5% and -5.3% for FREE_E and FREE_O respectively).

Figure 3 summarizes the distribution of operator-level electricity price changes for a 20% reduction in emissions. For the AUCT scenario, where electricity prices fully reflect CO_2 emissions, price increases range from about zero for producers with low CO_2 intensity to around 250% for operators holding a portfolio composed mainly of coal-fired plants. When free allowances subsidize regulated electricity rates, both the mean and the dispersion of price changes decline. The maximum price increase under an emissions-based subsidy is about 100%, and 181.5% under an output-based subsidy. Under an output-based subsidy, regulated operators with low CO_2

[24] Similarly, wholesale electricity prices increase on average by about 20%. Differences between regulated and non-regulated markets reflect the higher CO_2 intensity of regulated producers, but also the lower substitution possibilities among fuels and technologies, as regulated operators typically hold a much smaller set of generators compared to the set of generators active on regional wholesale markets. Moreover, regulated operators set prices according to the average cost of generation, so that electricity rate reflect the average CO_2 content of electricity. On wholesale markets, the price reflects (a function of) the generation costs of the marginal producer, and hence the CO_2 price is reflected in wholesale electricity prices only through the CO_2 content of the marginal producer.

Table 4: Welfare costs, CO_2 prices, and sectoral CO_2 abatement.

	AUCT			FREE_E			FREE_O		
Reduction target[a] (%)	10	20	30	10	20	30	10	20	30
Welfare cost[b]									
Total (US$billion)	34.4	83.0	155.3	60.4	129.9	213.0	57.4	124.7	206.5
Per avoided ton of CO_2 (US$)	58.8	70.9	88.4	103.8	111.4	121.6	98.4	106.8	117.8
CO_2 price (US$ per ton)	14.1	31.2	51.3	18.9	40.5	63.2	17.3	37.4	60.0
Electricity price changes[c] (%)	20.8	38.7	68.1	13.9	29.7	52.8	15.3	34.1	56.4
Regulated operators (%)	28.8	67.7	115.5	15.7	38.0	66.6	18.8	46.3	78.5
Non-regulated operators (%)	8.2	20.2	38.0	11.1	24.5	44.0	9.8	26.4	42.4
Sectoral abatement									
Economy-wide (million tons)	585	1,170	1,756	585	1,170	1,756	585	1,170	1,756
Sectoral contribution (%)									
Regulated electricity (%)	38.1	38.9	38.8	19.3	23.7	28.8	25.3	27.9	31.4
Wholesale electricity (%)	11.0	14.0	16.5	17.8	20.5	20.8	15.8	19.1	19.7
Non-electricity sectors (%)	50.9	47.1	44.7	63.0	55.8	50.4	58.9	53.0	48.9

Notes: [a]Emissions reductions relative to benchmark ($100(1-\xi)$). [b]Negative of the weighted sum of equivalent variations of each household. [c]Weighted average across electricity markets, net of transmission and distribution costs.

emissions can be overcompensated by the subsidy, which can induce a reduction in prices.

While the magnitude of efficiency costs depends on the characteristics of the electricity sector, it also depends on how costly it is to abate in other sectors. As shown in Table 4, our analysis suggests that the slope of the (implicit) marginal abatement cost schedule for regulated operators is lower than for other activities. This implies that the share of total abatement undertaken by regulated operator is relatively large.[25] Under the AUCT scenario, about 40% of total abatement comes from regulated electricity producers, amounting to a 29% emissions reduction in this sector. Under FREE_E and FREE_O, regulated operators are still required to surrender allowances, and cost minimizing behavior will induce fuel switching and a merit order effect, but consumers do not fully see the CO_2 price signal. In turn, the share of total abatement by regulated operators drops from 40% to 23% under FREE_E and 28% under FREE_O.

Higher emissions by regulated operators increase the demand for allowances, which raises the equilibrium CO_2 price, and incentivizes sub-optimally large levels of abatement in the wholesale electricity and non-electricity sectors. For a 20% target, the contribution of non-regulated

[25] Importantly, marginal abatement cost curves for each regulated operator is endogenously determined given available technologies, relative prices and electricity demand.

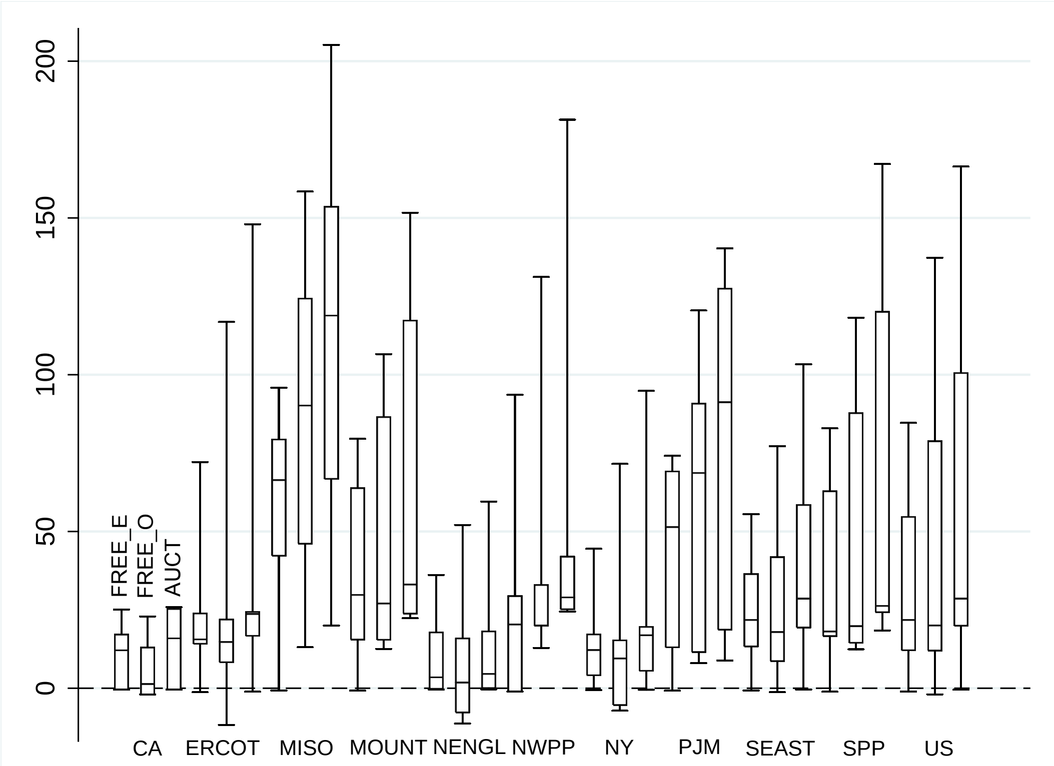


Figure 3: Distribution of (subsidy-inclusive) electricity price changes by region ($\xi = 0.8$). *Notes:* Electricity prices are net of transmission and distribution costs. For each region, the box-whisker plots from left to right refer to the AUCT, FREE_O, and FREE_E cases, respectively. The whiskers show outlier values at the 5^{th} and 95^{th} percentile, respectively.

electricity producers increases from 14% for AUCT to about 20% under both subsidy cases. Similarly, non-electricity sectors contribute about 47% of total abatement in the AUCT case while the corresponding share increases to 56% and 53% under FREE_E and FREE_O, respectively. Across the economy, the equilibrium marginal abatement cost, which is equal to the price of carbon, increases by around 30% for a 20% target.

5.3 Regional Variations and the Size of Efficiency Costs

Regional variations in the share of regulated electricity output and CO_2 intensity can be used to identify the source of the efficiency costs. Table 5 summarizes regional welfare changes relative to the AUCT scenario and provides information on the average level of the electricity subsidy rate for a 20% emissions reductions target. The pattern of regional welfare losses correlates closely with the size of subsidy rates, confirming that the value of free allowances allocated to regulated firms is the main driver of the magnitude of efficiency costs. First, all the regions

Table 5: Efficiency cost, subsidy rate, and CO_2 emissions by region ($\xi = 0.8$).

	Δ Welfare rel. to AUCT (%)[a]		Subsidy rate (cents/kWh)[b]		Δ CO_2 emissions rel. to AUCT (%)[a]	
	FREE_E	FREE_O	FREE_E	FREE_O	FREE_E	FREE_O
MOUNT	-0.99	-0.90	1.2	1.0	3.2	2.2
SPP	-0.90	-0.82	1.3	0.9	4.8	2.7
SEAST	-0.65	-0.63	1.0	0.9	1.0	1.7
MISO	-0.41	-0.37	1.2	0.8	3.7	1.9
NWPP	-0.31	-0.28	0.5	0.8	-0.2	-1.0
PJM	-0.23	-0.21	0.6	0.4	-1.4	-1.4
ERCOT	-0.13	-0.11	0.2	0.1	-4.5	-3.5
NY	-0.11	-0.09	0.3	0.3	-2.9	-1.8
CA	-0.09	-0.07	0.2	0.5	-2.8	-1.2
NENGL	-0.07	-0.05	0.1	0.1	-2.5	-1.8

Notes: [a]Difference in percentage points of percentage mean welfare changes under AUCT relative to FREE_E and FREE_O. [b]Output-weighted average across regulated electricity producers in each region.

are worse off when regulated firms receive free permits. Second, regions with high shares of electricity produced under cost-of-service regulation (SEAST, SPP, MOUNT) suffer relatively large adverse welfare impacts. Conversely, regions with a low degree of regulation (NENGL, ERCOT, NY, PJM) experience almost negligible welfare losses.

When free permits are allocated in proportion to benchmark emissions, benchmark CO_2 intensity of regulated electricity generation is an important driver of regional efficiency costs. For example, CA experiences the second smallest efficiency costs despite the fact that almost half of electricity is produced under regulation. This is because regulated operators in this region mostly hold hydroelectric resources and have the lowest CO_2 intensity among all regional regulated operators (see Table 1). Similarly, SEAST has the largest share of output from regulated operators, but CO_2 intensity is lower than other highly regulated regions (SPP and MOUNT), leading to significantly lower efficiency costs. Conversely, under an output-based allocation, subsidy rates are lower for CO_2 intensive regions, so that the welfare gains relative to FREE_E are largest in regions with CO_2-intensive generation.

Table 5 also reports changes in regional abatement relative to AUCT. Under a subsidy, regions where regulated operators have a large market share and hold CO_2 intensive technologies, reduce their abatement effort, i.e. emit more CO_2. Because aggregate emissions are capped, other regions have to abate more. An output-based subsidy generally leads to smaller changes in abatement, which mitigates redistribution of abatement.

5.4 Household Distributional Impacts

An important dimension for assessing alternative policy designs and instruments are distributional impacts across income. The incidence analysis based on our quantitative model allows us to capture two important aspects. First, the general equilibrium framework means that both the uses- and sources-side of income effects are incorporated. Second, incorporating expenditure and income patterns of real households based on micro-data has the advantage of being able to characterize incidence both across and within income groups rather than measuring mean impacts based on highly aggregated groups of (representative) consumers.

Table 6 thus summarizes the within and across income decile distribution of efficiency costs of the FREE scenarios relative to the AUCT scenario for a 20% emissions reduction target.[26,27] Specifically, we report the difference in equivalent variation between FREE_E and AUCT, and FREE_O and AUCT, respectively. First, the additional efficiency cost borne by an average household expressed in 2006 US$ is 227 for an emission-based allocation and 200 for an output based allocation. Second, looking at the mean welfare impacts by income decile suggest that the efficiency costs from subsidizing electricity rates are regressive. Third, within the three lowest income deciles, there is a substantial number of households that experience large negative welfare impacts, implying that regressivity is more pronounced at the mean of the distribution. For the top 20% of the distribution, however, some households also experience significant losses, making average policy impacts slightly progressive in this part of the income distribution. Fourth, Table 6 shows that focusing on aggregate welfare impacts, even when looking at representative households by income class, masks important variations across individual households. In particular, the variation in impacts across households within a given income decile swamps

[26] It is well-known in the literature that distributional impacts of a revenue-generating carbon pricing policy crucially depend on the ways the carbon revenue is returned back to households, and this dominates the impacts of the carbon price itself. Our paper does not aim to contribute to this aspect of policy but rather focuses on the question of how the efficiency costs stemming from free allocation and (unintended) interaction with existing price regulation in the electricity sector are distributed across household with different incomes.

[27] Previous literature has pointed out the caveat that consumption taxes—including energy or carbon taxes—tend to look less regressive when lifetime income measures are used than when annual income measures are used (Poterba, 1989, 1991; Rausch et al., 2011). One way to adjust for this bias is by proxying lifetime income with "current expenditures" (Grainger and Kolstad, 2010). The lifetime income approach can be an important caveat to distributional findings from annual incidence analyses but it relies on strong assumptions about household consumption decisions. In particular it assumes that households base current consumption decisions knowing their full stream of earnings over their lifetime. Given these alternative approaches, we have performed our analysis using both an annual income measure and a lifetime income measure. As the pattern of welfare impacts is virtually identical for both approaches, we decided to present the distributional impact based on annual income.

Table 6: Distribution of household welfare impacts across income groups (FREE_E relative to AUCT, $\xi = 0.8$).

	FREE_E					FREE_O				
Income deciles	Mean[a]	US$ per hh[b]	25%[a]	50%[a]	75%[a]	Mean[a]	US$ per hh[b]	25%[a]	50%y	75%[a]
1	-0.50	-76	-0.77	-0.38	-0.11	-0.50	-77	-0.74	-0.38	-0.13
2	-0.34	-91	-0.53	-0.26	-0.07	-0.34	-90	-0.50	-0.25	-0.10
3	-0.32	-111	-0.52	-0.27	-0.06	-0.31	-108	-0.48	-0.25	-0.08
4	-0.32	-136	-0.50	-0.27	-0.11	-0.29	-127	-0.44	-0.24	-0.11
5	-0.29	-151	-0.43	-0.27	-0.10	-0.26	-139	-0.39	-0.23	-0.10
6	-0.29	-182	-0.42	-0.27	-0.13	-0.26	-163	-0.37	-0.22	-0.12
7	-0.30	-220	-0.41	-0.28	-0.14	-0.26	-194	-0.34	-0.23	-0.12
8	-0.29	-253	-0.39	-0.27	-0.15	-0.25	-220	-0.32	-0.22	-0.12
9	-0.34	-362	-0.40	-0.25	-0.11	-0.29	-308	-0.31	-0.19	-0.09
10	-0.46	-676	-0.48	-0.20	-0.06	-0.38	-564	-0.35	-0.15	-0. 05
Weighted average	-0.34	-227	-0.46	-0.27	-0.10	-0.31	-200	-0.41	-0.23	-0.09

Notes: [a]Difference in percentage points of population-weighted within-income group percentage welfare changes under FREE_E relative to AUCT. [b]Population-weighted within-income group average of equivalent variation expressed in 2006 US$ relative to AUCT (absolute difference).

the variation in means across income deciles.

Thus while the auctioning scheme with per-capita lump-sum recycling produces a U-shaped outcome in terms of welfare impacts across income (relative to no policy benchmark), freely allocating permits to regulated electricity firms disproportionately hits low and high income households (inverted U-shaped profile) relative to auctioning. This is tantamount to saying that while there are substantial efficiency costs at the aggregate level, the distribution of welfare gains becomes slightly more equal (although at larger negative impacts for any household group). There are two reasons for the inverted U-shaped pattern.[28] First, under FREE_E and FREE_O low-income households no longer receive a per-capita lump-sum transfer from the carbon revenue of regulated firms as was the case under AUCT. As taking away $1 of income creates a disproportionate large welfare loss for low income households, these households are worse off despite the fact that they spend more of their income on electricity whose price is subsidized. Second, high income households bear a disproportionately large burden from free allocation (relative to auctioning) as they are most affected by changes in factor prices. Indeed, as in our model capital is assumed to be more mobile than labor, capital is a better substitute for CO_2

[28] See Parry (2004) for a discussion of the conditions under which free permits (in the absence of pre-existing distortions) would be regressive.

abatement. It follows that a CO_2 policy increases the relative price of capital to labor.[29] Under a subsidy, inefficiencies in economy-wide abatement further depress the demand for capital relative to that for labor, so that the relative price of capital to labor is lower under the FREE cases as compared to the AUCT case. As households in the top income decile derive a relatively large fraction of their income from capital, they are more adversely impacted on the source of income side vis-à-vis low and middle income households.

5.5 Sensitivity Analysis

Our *ex-ante* policy analysis should be best interpreted as an attempt to identify the policy-relevant drivers of distortions introduced by price regulation. A key driver of the quantitative results is the marginal abatement cost of CO_2, which are driven primarily by two sets of parameters in the model. First, in the electricity sector, substitution among generation technologies is based on changes in relative generation costs. Second, for households and non-electricity sectors, the behavioral response is described by calibrated CES functions that are parameterized based on assumptions about elasticities of substitution.

Naturally, empirical estimates about elasticity parameters are fraught with uncertainties. However, by focusing on the impact of free permits relative to an auction, we limit the sensitivity of our results with respect to these parameters as they are kept constant across scenarios. For example, one key parameter that governs the substitutability between value-added (i.e., capital and labor) versus energy (see Appendix B.1), σ_{KLE}, does significantly affect the level of welfare costs. It does not, however, affect much the difference of welfare costs across scenarios. One important exception are elasticity parameters that measure the degree of market integration in the electricity sector, i.e., σ and σ_{xELEr} (see Appendix B.2). As different policy assumptions induce substantial price differentials across electricity markets, the level of market integration can potentially affect the size and regional distribution of efficiency costs unevenly across scenarios.

Table 7 reports results for alternative assumptions about electricity market integration. In the first panel, we summarize the impacts of the FREE_E and FREE_O scenarios relative to AUCT

[29] A similar result is obtained in the analytical general equilibrium model by Fullerton and Heutel (2007) who show that the relative price of capital (to labor) falls when capital is a better substitute for pollution. Note also that the distribution of ownership of polluting capital equipment will also have distributional implications, although this effect is difficult to capture empirically.

Table 7: Impacts of different cap-and-trade designs under alternative parameter assumptions.

Parametrization *Scenario*	Mean EV[a]	Standard deviation of EV[b]	Mean EV by region[c]	Electricity price change[d]	Mean electricity price change by region[e]
Central case ($\sigma = 1$, $\sigma_{xELE} = 0.5$)					
FREE_E	-0.34	0.47	0.34	0.29	0.12
FREE_O	-0.31	0.42	0.31	0.46	0.15
Low market integration ($\sigma = 0$)					
FREE_E	-0.35	0.46	0.33	0.29	0.13
FREE_O	-0.32	0.42	0.32	0.47	0.16
High market integration ($\sigma = 10$)					
FREE_E	-0.31	0.50	0.37	0.27	0.10
FREE_O	-0.28	0.46	0.35	0.44	0.12
High market integration ($\sigma = 10$) and high electricity trade elasticity ($\sigma_{xELE} = 5$)					
FREE_E	-0.30	0.51	0.38	0.25	0.09
FREE_O	-0.26	0.49	0.37	0.36	0.11

Notes: Results shown for $\xi = 0.8$. [a]Percentage point difference relative to AUCT, weighted average across households. [b]Percentage change of welfare by household relative to AUCT. [c]Percentage change of mean welfare by region relative to AUCT. [d]Percentage change of price change by market relative to AUCT. [e]Percentage change of mean price change by region relative to AUCT.

for our central assumptions. The second and third panels show results for low and high market integration cases, respectively. For the low σ case, efficiency costs of both subsidy scenarios increase slightly, but it has almost no impact on the dispersion of welfare measures and electricity prices. For the high σ case, efficiency costs decline, as abatement is cheaper in the electricity sector, and the dispersion in electricity price impacts declines. CO_2 prices are lower compared to the central case, translating into lower subsidy rates, and this induces a modest increase of the dispersion of welfare measures. Finally, increasing both σ and σ_{xELEr} further reduces efficiency costs and the dispersion of price changes, but increase the dispersion of welfare impacts.

These results suggest that the magnitude of efficiency costs does not substantially depend on our representation of electricity markets integration. Intuitively, these elasticity parameters affect the tails of the price change distribution across scenarios, but leave the average price impacts almost unaffected. As households in our model do not directly observe electricity prices on each market – but rather trade-off an aggregate electricity commodity with the consumption of other aggregate goods – changes in the dispersion of electricity price are not directly reflected in the distribution of welfare impacts.

A final limitation of our framework of importance for the interpretation of our results is the absence of dynamics. On the one hand, the allocation of allowances is typically repeated each year, providing the regulator with additional flexibility over time. In particular, while allowances may be initially granted for free, the regulation may well gradually move towards an auction-based allocation. Our quantitative results should thus be interpreted as bounds on the welfare cost, since we do not consider the case where the allocation of allowances varies over time. On the other hand, our static representation does not consider investments in low-CO_2 electric technologies or retirement decisions. While higher capacity in low-CO_2 technologies would decrease aggregate welfare costs under all scenarios, thereby not affecting the main insights of our analysis, a subsidy would lower incentives to invest in clean technologies relative to lump-sum transfers. While answering this question in any details is beyond the scope of this paper, one may expect that efficiency costs associated with free allowances in the presence of price regulation will be even larger in the mid- to long-run.

6 Conclusions

This paper has studied the efficiency and distributional implications of alternative designs for emissions trading systems in the presence of price-regulated firms. An emissions trading policy that is designed to distribute emissions permits for free is likely to effectively subsidize output prices of polluting firms that are subject to price-regulation. While this may explain why regulated electricity companies may also have an interest in lobbying for free permits, the failure to pass through the carbon price signal can impede cost-effectiveness and lead to substantial additional welfare cost. To shed light on the empirical relevance of this issue, we focused on the case of U.S. economy where about one fifth of economy-wide CO_2 emissions are produced by price-regulated electricity suppliers. Our quantitative analysis suggests that, for an emissions reduction target of 20%, efficiency costs of freely allocating permits are about 60% higher relative to auctioning of allowances.

We have shown these large welfare costs to be driven by two main factors. First, given the large share of emissions stemming from price-regulated firms, the value of free permits used to subsidize electricity rate is quantitatively important, and has a significant impact on electricity output. In turn, U.S. regions with a large share of electricity produced under cost-of-service regulation suffer from relatively large distortions. Second, as free permits induce a higher output by regulated electricity producers, the economy forgoes low-cost abatement opportunities in the electricity sector associated with fossil-based, in particular coal-fired, electricity generation. The marginal abatement cost schedule in non-electricity sectors is relatively steep compared to regulated electricity producers, so that shifting abatement to other sectors induces a substantial increase in the equilibrium marginal abatement cost.

Freely allocating permits to regulated electricity firms disproportionately hits low and high income households (inverted U-shaped profile) relative to auctioning. Thus, while there are substantial efficiency costs at the aggregate level, the distribution of welfare gains becomes slightly more equal (although at larger negative impacts for any household group). This results is based on two effects. First, under free allocation low-income households bear a disproportionately large burden as they no longer receive the lump-sum transfer from the carbon revenue of regulated firms as was the case under auctioning. We find that this overcompensates the expected progressive effect stemming from a reduction in the subsidy-inclusive electricity price. Second, high-income households bear a disproportionately large burden from free allocation (relative to auctioning) as they are most affected by changes in factor prices.

In light of still ongoing attempts in many countries to introduce market-based instruments to control pollutants,[30] this paper highlights the fact that a price on carbon does not automatically guarantee cost-effectiveness. In fact, if the policy is poorly designed, the market-based instrument may even lose its superiority over command-and-control-type instruments. While the fundamental design aspect of emissions trading systems, namely whether to auction or freely distribute permits, has already been investigated from a variety of angles (for example, to provide compensation of profit losses as in Goulder et al., 2010, or to lower pre-existing fiscal distortions as in Goulder et al., 1999), this paper points to the importance of pre-existing regulatory interventions affecting price-adjustment mechanisms.

[30] Examples of successfully implemented cap-and-trade programs include the Emissions Trading Scheme in Europe and the NO_x budget trading program in the United States. Recently, California has implemented a cap-and-trade program to curb GHG emissions and China has launched over the past years a number of pilot emissions trading programs at the provincial level.

References

Armington, P. (1969) "A theory of demand for products distinguished by place of production," *International Monetary Fund Staff Papers*, 16, pp. 159–76.

Bohi, D. and D. Burtraw (1992) "Utility investment behavior and emission trading market," *Resources and Energy*, 14, pp. 129–153.

Böhringer, C. and A. Lange (2005) "Economic implications of alternative allocation schemes for emission allowances," *Scandinavian Journal of Economics*, 107 (3), pp. 563 – 581.

Böhringer, C. and T. F. Rutherford (2009) "Integrated assessment of energy policies: Decomposing top-down and bottom-up," *Journal of Economic Dynamics and Control*, 33, pp. 1648–1661.

Borenstein, S., J. Bushnell, and C. Knittel (1999) "Market power in electricity markets: Beyond concentration measures," *The Energy Journal*, 20 (4), pp. 65–88.

Bovenberg, A. L. (1999) "Green tax reforms and the double dividend: an updated reader's guide," *International Tax and Public Finance*, 6, 3, pp. 421–443.

Bovenberg, A. L. and L. H. Goulder (1996) "Optimal environmental taxation in the presence of other taxes: General equilibrium analyses," *American Economic Review*, 86, pp. 985–1000.

Bureau of Economic Analysis (2007) "Annual input-output personal consumption expenditures bridge matrix." Washington D.C.

Bureau of Labor Statistics (2006) "2006 consumer expenditure interview survey public use microdata documentation." Washington D.C.

Burtraw, D., M. Walls, and J. Blonz (2009) "Distributional impacts of carbon pricing policies in the electricity sector." American Tax Policy Institute, Washington, D.C.

Bushnell, J. B., E. T. Mansur, and C. Saravia (2008) "Vertical arrangements, market structure and competition: An analysis of restructured US electricity markets," *American Economic Review*, 98(1), pp. 237–266.

Coggins, J. and V. Smith (1993) "Some welfare effects of emission allowance trading in a twice-regulated industry," *Journal of Environmental Economics and Management*, 25, pp. 275–297.

Dales, J. H. (1968) *Pollution, Property and Prices: an Essay in Policy-Making and Economics*: Toronto: University of Toronto Press.

Deaton, A. (2005) "Measuring poverty in a growing world (or measuring growth in a poor world)," *Review of Economics and Statistics*, 87 (1), pp. 1 – 19.

Dirkse, S. P. and M. C. Ferris (1995) "The PATH solver: a non-mononontone stabilization scheme for mixed complementarity problems," *Optimization Methods and Software*, 5, pp. 123–156.

Energy Information Administration (2007a) "Form EIA-860 annual electric generator data." Washington, DC.

——— (2007b) "Form EIA-906-920 power plant operations data." Washington, DC.

——— (2008) "Carbon dioxide uncontrolled emission factors, electric power annual data for 2000–2007." Washington, DC.

——— (2009a) "Electric power annual 2007." Office of Coal, Nuclear, Electric and Alternate Fuels, US Department of Energy, Washington DC.

——— (2009b) "The electricity market module of the national energy modeling system." Washington, DC.

——— (2009c) "State energy data system." Washington, DC.

——— (2009d) "State energy price and expenditure report." Washington, DC.

——— (2012) "Electric power monthly july 2012." Washington, DC.

Federal Reserve Board (2007) "Survey of consumer finances." Board of Governors of the Federal Reserve System, Washington, DC.

Feenberg, D. and E. Coutts (1993) "An introduction to the TAXSIM model," *Journal of Policy Analysis and Management*, 12 (1), pp. 187 – 194.

Fisher, C. (2001) "Rebating environmental policy revenues: Output-based allocation and tradable performance standards." Resource for the Future Discussion Paper no. 01-22, Washington D.C.

Fowlie, M. L. (2009) "Incomplete environmental regulation, imperfect competition, and emissions leakage," *American Economic Journal: Economic Policy*, 1 (2), pp. 72–112.

Fullerton, D., S. P. McDermott, and J. P. Caulkins (1997) "Sulfur dioxide compliance of a regulated utility," *Journal of Environmental Economics and Management*, 34, pp. 34–53.

Fullerton, D. and G. Heutel (2007) "The general equilibrium incidence of environmental taxes," *Journal of Public Economics*, 91, 3-4, pp. 571–591.

Goulder, L. H., I. W. H. Parry, R. Williams III, and D. Burtraw (1999) "The cost-effectiveness of alternative instruments for environmental protection in a second-best setting," *Journal of Public Economics*, 72 (3), pp. 329–360.

Goulder, L. H., M. A. Hafstead, and M. Dworsky (2010) "Impacts of alternative emissions allowance allocation methods under a federal cap-and-trade program," *Journal of Environmental Economics and Management*, 60(3), pp. 161–181.

Grainger, C. and C. Kolstad (2010) "Who pays a price on carbon?" *Environmental and Resource Economics*, 46, 3, pp. 359–376.

Hahn, R. W. and R. N. Stavins (2011) "The effect of allowance allocations on cap-and-trade system performance," *Journal of Law and Economics*, 54, pp. 267–294.

Hepburn, C. J., J. K.-H. Quah, and R. A. Ritz (2012) "Emissions trading with profit-neutral permit allocations," *Journal of Public Economics*, 98, pp. 85 – 99.

IMPLAN (2008) "State-level U.S. data for 2006." Stillwater, MN: Minnesota IMPLAN Group.

Joskow, P. L. (2005) "Transmission policy in the united states," *Utilities Policy*, 13 (2), pp. 95–115.

——— (2008) "Challenges for creating a comprehensive national electricity policy." MIT Technology Policy Institute.

Joskow, P. L. and R. Schmalensee (1983) *Markets for Power: An analysis of Electric Utilities Deregulation*: MIT Press, Cambridge USA.

Lanz, B. and S. Rausch (2011) "General equilibrium, electricity generation technologies and the cost of carbon abatement: A structural sensitivity analysis," *Energy Economics*, 33 (5), pp. 1035–1047.

Malueg, D. A. (1990) "Welfare consequences of emission credit trading programs," *Journal of Environmental Economics and Management*, 18 (1), pp. 66–77.

Mansur, E. T. (2007) "Upstream competition and vertical integration in electricity markets," *Journal of Law and Economics,* 50(1), pp. 126–156.

Metcalf, G. E., A. Mathur, and K. A. Hassett (2010) "Distributional impacts in a comprehensive climate policy package." NBER Working Paper w16101, Cambridge MA.

Montgomery, D. W. (1972) "Markets in licenses and efficient pollution control programs," *Journal of Economic Theory*, 5.

Musgrave, R. A. (1964) "Estimating the distribution of the tax burden," *The review of income and wealth*, 10 (1), pp. 186 – 219.

North American Electric Reliability Council (2007) "2002 – 2006 generating availability report." North American Electric Reliability Council, Washington D.C.

Paltsev, S., J. M. Reilly, H. Jacoby, R. Eckhaus, J. McFarland, M. Sarofim, M. Asadoorian, and M. Babiker (2005) "The MIT emissions prediction and policy analysis (EPPA) model: Version 4." MIT JPSPGC, *Report 125*, August, 72 p.

Parry, I. W. H. (2004) "Are emissions permits regressive?" *Journal of Environmental Economics and Management*, 47 (2), pp. 364 – 387.

Paul, A., D. Burtraw, and K. Palmer (2010) "Compensation for electricity consumers under a US CO_2 emissions cap," in V. Ghosal ed. *Reforming Rules and Regulations: Laws, Institutions and Implementation*: Cambridge, MA: MIT Press.

Poterba, J. (1989) "Lifetime incidence and the distributional burden of excise taxes," *American Economic Review Papers and Proceedings*, 32, pp. 325–330.

——— (1991) "Is the gasoline tax regressive," *Tax Policy and the Economy*, 5, pp. 145–164.

Puller, S. L. (2007) “Pricing and firm conduct in California’s deregulated electricity market,” *Review of Economics and Statistics*, 89, pp. 75–87.

Rausch, S., G. E. Metcalf, J. M. Reilly, and S. Paltsev (2010) “Distributional impacts of a U.S. greenhouse gas policy: a general equilibrium analysis of carbon pricing,” in G. E. Metcalf ed. *U.S. Energy Tax Policy*: Cambridge University Press, MA.

Rausch, S., G. E. Metcalf, and J. M. Reilly (2011) “Distributional impacts of carbon pricing: A general equilibrium approach with micro-data for households,” *Energy Economics*, 33, pp. S20–S33.

Rausch, S. and T. F. Rutherford (2009) “Tools for building national economic models using state-level implan social accounts.” Cambridge, Massachusetts, MIT. Unpublished manuscript. Available online at: http://www.cepe.ethz.ch/people/profs/srausch/IMPLAN2006inGAMS.pdf.

Robinson, S. (1991) “Macroeconomics, financial variables, and computable general equilibrium models,” *World Development*, 19, 11, pp. 1509–1525.

Rutherford, T. (2002) “Lecture notes on constant elasticity functions.” Mimeo, University of Colorado.

Rutherford, T. F. (1995) “Extension of GAMS for complementarity problems arising in applied economics,” *Journal of Economic Dynamics and Control*, 19 (8), pp. 1299–1324.

Rutherford, T. F. and D. Tarr (2008) “Poverty effects of russia’s wto accession: Modeling “real” households with endogenous productivity effects,” *Journal of International Economics*, 75 (1), pp. 131–150.

Stavins, R. N. (2008) “A meaningful u.s. cap-and-trade system to address climate change,” *Harvard Environmental Law Review*, 32, pp. 293–371.

U.S. Environmental Protection Agency (2014) “Inventory of U.S. greenhouse gas emissions and sinks: 1990–2012.” Washington, DC.

Winston, C. (1993) “Economic deregulation: Days of reckoning for microeconomists,” *Journal of Economic Literature*, 31, pp. 1263–1289.

Wolak, F. A. (2003) “Measuring unilateral market power in wholesale electricity markets: The California market, 1998–2000,” *American Economic Review*, 93, pp. 425–430.

Appendix A Proof of Proposition 1

Appendix A.1 Proof of (a).

Given the assumed demand function, we can rewrite Equation (4) as:

$$\alpha(M + \phi_m P_e)X^{-\frac{1}{\beta}} = c(\zeta_m) + \zeta P_e \,,$$

and Equation (5) as:

$$\alpha M X^{-\frac{1}{\beta}} = c(\zeta_m) + \zeta P_e - \phi_m P_e / X \,,$$

so that:

$$\alpha(M + \phi_m P_e)X_{\mathit{Auct}}^{-\frac{1}{\beta}} = \alpha M X_{\mathit{Free}}^{-\frac{1}{\beta}} + \phi_m P_e / X_{\mathit{Free}}$$

and hence

$$\frac{\alpha M + \phi_m P_e X_{\mathit{Free}}^{\frac{1}{\beta}-1}}{\alpha(M + \phi_m P_e)} X_{\mathit{Free}}^{-\frac{1}{\beta}} = X_{\mathit{Auct}}^{-\frac{1}{\beta}}$$

or

$$\left(\frac{\alpha M + \phi_m P_e X_{\mathit{Free}}^{\frac{1}{\beta}-1}}{\alpha(M + \phi_m P_e)}\right)^{\beta} = \frac{X_{\mathit{Free}}}{X_{\mathit{Auct}}} \,.$$

Appendix A.2 Proof of (b).

Define

$$Z = \left(\frac{\alpha M + \phi_m P_e X_{\mathit{Free}}^{\frac{1}{\beta}-1}}{\alpha(M + \phi_m P_e)}\right)^{\beta} \,.$$

Setting $\beta = 1$, we obtain:

$$Z = \frac{\alpha M + \phi_m P_e}{\alpha(M + \phi_m P_e)}$$

which is greater than one for $\alpha < 1$. Thus when demand is unit-elastic, we have that $X_{\mathit{Free}} > X_{\mathit{Auct}}$. The remaining of the argument requires taking partial derivatives of Z with respect to $\phi_m P_e$, α, $\phi_m P_e / M$, and β, and is thus omitted.

Appendix B Supplementary details on the quantitative model

This appendix provides more information about the modeling framework. In the first subsection we lay out the equilibrium conditions for the economy-wide model. We then discuss the integration of electricity markets and heterogeneous households in turn.

Appendix B.1 Equilibrium Conditions for Economy-wide Model

Our complementarity-based formulation of the economy-wide model distinguishes two classes of conditions that characterize the competitive equilibrium: zero-profit conditions and market clearance conditions.[31] The zero-profit conditions determine a vector of activity levels (X) and the market clearance conditions determine a vector of prices (P).

Zero profit. Let $\Pi^X_{ir}(p)$ denote the unit profit function of industry i in region r which is calculated as the difference between unit revenue (R_{ir}) and unit costs (C_{ir}) where:

$$C_{ir}(p) = \mathit{min}\{p^l_r L_i + p^k K_i + p^z_r R_{zi} + \sum_j p_{jr} x_{ji} \mid F_{ir}(L_{ir}, K_{ir}, R_{zir}; x_{1ir}, \ldots, x_{10ir}) = 1\} \tag{B1}$$

$$R_{ir}(p) = \mathit{max}\{\sum_j p^X_{ir} X_{ir} \mid X_{ir} = 1\}\,. \tag{B2}$$

where p^X_{ir} is the price of X_{ir}. Zero profits implies that no production activity makes positive profits, i.e.:

$$-\,\Pi^X_{ir}(p) = C_{ir} - R_{ir} \geq 0 \qquad \perp \qquad X_{ir}\,. \tag{B3}$$

Similar conditions hold for Armington aggregation ($\Pi^x_i r$).

Market clearance. The second class of equilibrium conditions is that at equilibrium prices and activity levels, the supply of any commodity must balance or exceed demand. For regional output markets we can express this condition as:

$$X_{ir} \geq \sum_j x_{jr} \frac{\partial \Pi^x_{jr}(p)}{\partial p^X_{ir}} \qquad \perp \qquad p^X_{ir} \tag{B4}$$

[31] An income balance accounting condition is usually specified to simplify the implementation of the problem, but can be substituted out of the model without altering the basic logic. In the present context, this condition is given by aggregating Equation (B6) across households.

The market for Armington good i is in balance if:

$$x_{ir} \geq \sum_j X_{jr} \frac{\partial \Pi^X_{jr}(p)}{\partial p^x_{ir}} + \frac{\partial p^I_r}{\partial p^x_{ir}} I_r + \frac{\partial p^x_{GP}}{\partial p^x_{ir}} GP + d_{ir}(p, M_r) \qquad \perp \qquad p^x_{ir} \tag{B5}$$

where by Shephard's Lemma the first three summands on the right-hand side represent the demand of good i by the constant returns to scale production, investment, and government sectors, respectively. Household income is given by:

$$M_h = p^k \omega^k_h + p^l_r \omega^l_h + \sum_z p^z_r \omega^z_h + T_h \tag{B6}$$

where p^k, p^l_r, and p^z_r are prices for capital, labor, and resources, ω's denote the initial endowment of capital, labor (including leisure time), and resources, and T_h is benchmark transfer income. Final demands $d_{ir}(p, M_r)$ are derived from the budget-constrained maximization:

$$d_{ir}(p, M_r) = \textit{argmax}\{U(x_{1r}, \ldots, x_{10r}, q, w, l) \mid \sum_i p^x_{ir} x_{ir} + p^q_r q + p^w_r w + p^l_r l = M_r\} \tag{B7}$$

where $U(\cdot)$ is a CES utility index. Market clearance conditions for labor, capital, and natural resources are given by:

$$\sum_j Y_j \frac{\partial \Pi^Y_{jr}(p)}{\partial p^f_r} + d_{fr}(p, M_r) \geq \sum \omega^f_r \qquad \perp \qquad p^f_r \tag{B8}$$

where $f = \{k, l; 1, \ldots, Z\}$ denotes the set for primary production factors (labor, capital, and natural resources). Market clearance conditions requiring balanced intra-national trade for non-energy goods that are traded on a bilateral basis are omitted here for simplicity.

Foreign closure of the model is warranted through a national balance-of-payments constraint which determines the price of foreign exchange:

$$\sum_i \textit{EX}_i + B = \sum_i \textit{IM}_i \frac{\partial \textit{pdfm}_i}{\partial \textit{pfx}} \qquad \perp \qquad \textit{pfx} \tag{B9}$$

where *EX* and *IM* denote the level of foreign exports and imports, respectively.

For all activities but electricity generation, we characterize production technology by distinguishing three types of production activities: primary energy sectors (indexed by $\textit{pe} = \{\textit{coal}, \textit{gas}, \textit{oil}\}$),

Table B1: Nested production structure and elasticity parameters.

Function	Description	Elasticity (by sector[a])		
		pe	nr	agr
σ_f	Output	0.6	0	0.7
$KLE = KLE(g, E; \sigma_{gE})$	Capital/labor-energy composite	-	0.5	-
$KLM = KLM(g, M; \sigma_{gM})$	Capital/labor-materials composite	0	-	-
$REM = REM(R, EM; \sigma_{REM})$	Resource-Energy/materials composite	-	-	0.6
$EM = EM(E, M; \sigma_{EM})$	Energy-materials composite	-	-	0.3
$M = M(x_1, \ldots, x_I; \sigma_{xM})$	Materials composite	0	0	0
$g = g(K, L; \sigma_{KL})$	Capital-labor composite	1	1	1
$E = E(x_{ELE}, h; \sigma_{ELEh})$	Energy composite	-	0.5	0.5
$h = h(x_{COL}, x_{GAS}, x_{OIL}; \sigma_{xE})$	Coal-gas-oil composite	-	1	1
$x_i = x_i(xD_i, xT_i; \sigma_{xjr})$	Domestic-imported inputs composite	5	5	5
$xT_i = x_i(xDT_i, xFT_i; \sigma_{Tjr})$	Imported inputs composite	5	5	5

Notes: All functions are CES in form. [a] Primary energy (pe): COL, GAS, CRU; Non-resource using (nr): OIL, EIS, MAN, TRN, SRV; Agricultural (agr): AGR.

non-resource based industries (indexed by *nr*), and agriculture (indexed by *agr*):

$$X_i = \begin{cases} f_i[KLM_i(g_i, M_i), R_{zi}; \sigma_f] & \text{if } i \in \{pe\} \\ f_i[KLE_i(g_i, E_i), M_i(x_{1i}, \ldots, x_{Ii}); \sigma_f] & \text{if } i \in \{nr\} \\ f_i[REM_i(R_i, EM_i), g_i(K_i, L_i); \sigma_f] & \text{if } i \in \{agr\}\,. \end{cases} \tag{B10}$$

where σ_f is the elasticity of substitution among composite inputs. We employ nested constant-elasticity-of-substitution (CES) functions with nesting structures reported in Table B1. Elements in the E and M nests are Armington (1969) composites of local and traded products (σ_{xjr}), where traded products are themselves a composite of intra-and inter-national imports (σ_{Tjr}).

The nested utility structure is summarized in Table B2. We assume that utility from government spending is additively separable with utility derived from private consumption, so that it is left out of the optimization problem.

Household income is given by:

$$M_h = p^k \omega_h^k + p_r^l \omega_h^l + \sum_z p_r^z \omega_h^z + T_h \tag{B11}$$

where p^k, p_r^l, and p_r^z are prices for capital, labor, and resources, ω's denote the initial endowment of capital, labor (including leisure time), and resources, and T_h is benchmark transfer income.

Table B2: Nested Utility Structure and elasticity parameters.

Function	Description	Elasticities
$U = U(CI, l)$	Household utility	$\sigma_c{}^a$
$CI = CI(C, q)$	Consumption-savings composite	0
$C = C(E, NE)$	Composite material consumption	0.25
$E = E(x_1, \ldots, x_I)\ , i \in \{e\}$	Energy consumption	0.4
$NE = NE(x_1, \ldots, x_I)\ , i \in \{ne\}$	Non-energy consumption	0.65

Notes: All functions are CES in form. [a] Calibrated to match an uncompensated (compensated) labor supply elasticity of 0.1 (0.3).

Appendix B.2 Integration of Electricity Generation into Economy-wide Transactions

The electricity sector and economy-wide models are consistently solved based on an algorithm by Böhringer and Rutherford (2009). As a first step, we generate a consistent benchmark data set where electricity sector outputs and inputs are consistent with the aggregate representation of the economy. For each regulated and wholesale electricity market, we simulate utilization of technologies, fuel use, and hence benchmark CO_2 emissions by calibrating the electricity generation model to observed demand for output on each market and fuel/input prices. Formally, for regulated markets, given the benchmark demand at each operator $\overline{d}_t^f$, we simulate benchmark output of each generator $\overline{Y}_t^g$ and associated demand for inputs and fuels, as well as benchmark price $\overline{P}^f$, by solving expressions (6) to (8) as a mixed complementarity problem. Similarly, for wholesale markets, we solve Equation (14) to (15).

To evaluate the fit of the electricity model against observed historic data, we first compare operator-level data on electricity generation for each combination of fuel-type and technology simulated with the model with observed values for 2006 are reported in EIA Form 906-920 (2007b). For regulated operators, the R^2 of the model is 90.2%, and 84.1% for wholesale markets.[32] Second, we compare observed average wholesale prices and emissions-intensity with those simulated from the model. Table B3 reports observed average wholesale prices and emissions intensity for wholesale producers, suggesting that our model provides a good representation of generation costs and CO_2 intensity. Figure B1 provides evidence that the model also

[32] Formally, we compute: $R^2 = 1 - \frac{\sum_{\text{tech,fuel}}(y_{\text{tech,fuel}} - \hat{y}_{\text{tech,fuel}})^2}{\sum_{\text{tech,fuel}}(y_{\text{tech,fuel}} - \bar{y})^2}$ where $y_{\text{tech,fuel}}$ is observed output for each technology-fuel combination, $\hat{y}_{\text{tech,fuel}}$ is the model prediction, and $\bar{y}$ is the average observed outcome.

Table B3: Wholesale electricity markets: Prices and emissions intensity.

Region	Regional wholesale price (US$/MWh)		CO_2 intensity (tCO_2/MWh)	
	Observed[a]	Simulated[b]	Observed[c]	Simulated[b]
NWPP	50.2	48.6	0.63	0.62
SEAST	58.1	53.5	0.60	0.61
PJM	55.1	52.2	0.58	0.58
ERCOT	52.9	57.5	0.52	0.50
MISO	44.0	47.7	0.47	0.50
SPP	55.4	63.6	0.42	0.43
CA	48.9	48.7	0.42	0.34
NENGL	60.8	61.5	0.40	0.36
MOUNT	57.4	44.9	0.38	0.35
NY	70.2	71.2	0.37	0.36

Notes: [a]Load-weighted average reported by FERC (2006); [b]Simulated from the electricity sector model; [c]Computed based on fuel consumption (EIA Form 906-920, 2007b) and fuel-specific CO_2 emissions factors (EIA, 2009a).

performs well in matching the observed CO_2 emissions intensity for regulated operators.

To integrate the resulting input demand into markets represented in the economy-wide model, we map fuel categories and input from the electricity sector to commodities in the economy-wide model, and adjust the input-output data with least-square optimization techniques in order to minimize the required adjustments.[33] Second, we calibrate the value share and level parameters of the CES functions in the economy-wide model using benchmark prices and quantities of the integrated electricity-economy data set.[34]

In the economy-wide model the demand for electricity by households and firms is based on a regional "retail" price, P^r_{retail}, for two reasons. First, social accounts that are used to calibrate the economy-wide model only report annual electricity consumption by region. Thus P^r_{retail} links electricity generation to the rest of the economy by aggregating information from multiple electricity markets within each region and across time (load segments). Second, it allows us to incorporate assumptions about the degree of electricity markets' integration within a region

[33] Given our operator-level representation of electricity markets, we are able to precisely match each regulator to its region. Operators that hold generators across regional borders defined in the model lead to small discrepancies in the benchmark data. For non-regulated operators, all the generators are mapped to their appropriate region of operation, so that discrepancies between state-level borders and wholesale markets geography do not affect our analysis.

[34] Nested CES function that characterize technology are formulated in calibrated share form (Rutherford, 2002), which considerably eases anchoring of a CES functions to the calibration point.

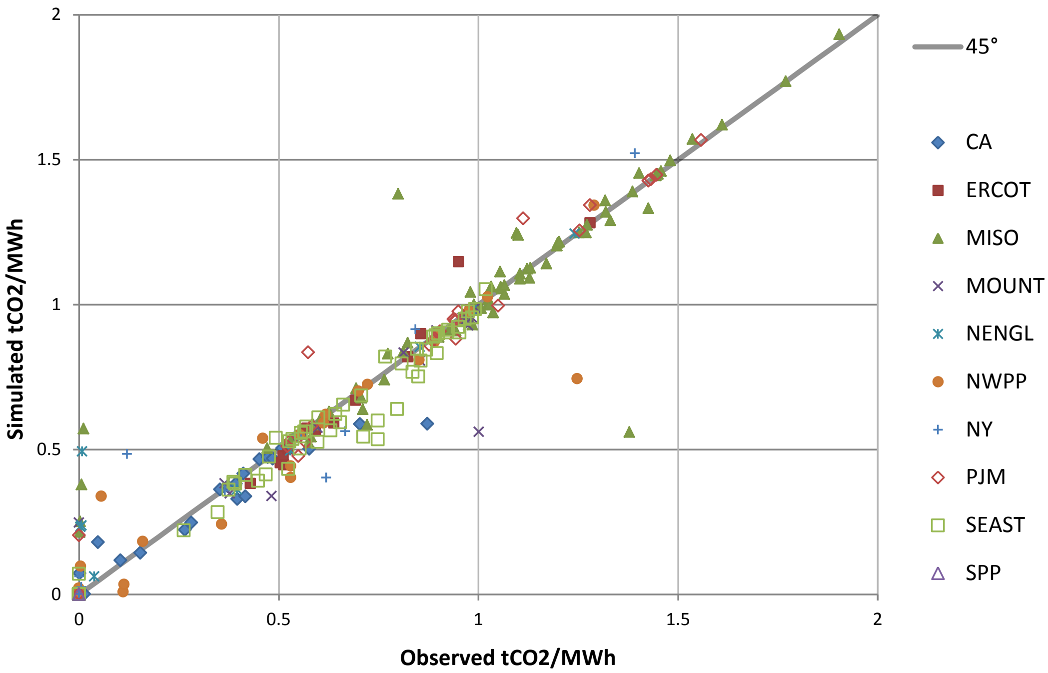


Figure B1: Regulated electricity markets: Emissions intensity. *Notes*: Observed emissions intensity estimates are based on fuel consumption (EIA Form 906-920, 2007b) and fuel-specific CO_2 emissions factors (EIA, 2009a).

without an explicit representation of a transmission network.[35]

In the benchmark, we thus define $\overline{P}^r_{\text{retail}} = \overline{P}^r_{\text{ele}} + TD^r$, where $\overline{P}^r_{\text{ele}}$ is an output-weighted average of generation costs across electricity markets in each region, and TD^r are regional transmission and distribution costs.[36] Away from the benchmark, we represent barriers to market integration by monopolistic competition between regulated and non-regulated operators (with

[35] Observed differences in prices across markets (and operators) suggest that trade opportunities among operators are limited, in turn reflecting both the existing regulatory structure and transmission constraints. It is, however, far from clear how such barriers to market integration will evolve in the future and, in particular, under a CO_2 emissions control policy. For example, given differences in the technology portfolios of operators, establishing a uniform price on CO_2 will generate heterogeneity in generation cost increase, favoring opportunities for mutually beneficial trades. Public Utility Commissions may thus require regulated operators to shut down highly CO_2-intensive plants and purchase power from other sources, weakening monopoly power of regulated operators.

[36] Formally, $\overline{P}^r_{\text{ele}} = (\overline{D}^r + \sum_{f \in r} \overline{D}^f)^{-1}(\overline{D}^r\overline{P}^r + \sum_{f \in r} \overline{D}^f \overline{P}^f)$. As benchmark prices $\overline{P}^r$ and $\overline{P}^f$ only include generation costs, we impute transmission and distribution costs as $\overline{TD}^r = \overline{P}^r_{\text{retail}} - \overline{P}^r_{\text{ele}}$, where $\overline{P}^r_{\text{retail}}$ denotes observed retail prices in the IMPLAN data (2008). For all counterfactual simulations, we assume that these costs remain constant at their benchmark level.

a fixed number of firms), i.e. each market produces a variety of electricity with a distinct price:

$$P^r_{\text{ele}} = \left[\theta^r \left(\frac{P^r}{\overline{P}^r} \right)^{(1-\sigma)} + \sum_{f \in r} \theta^f \left(\frac{P^f}{\overline{P}^f} \right)^{(1-\sigma)} \right]^{\frac{1}{1-\sigma}}, \tag{B12}$$

where θ^r and θ^f denote the observed market shares for wholesale market r and for regulated operator f in region r, respectively. σ capture the degree of market integration, including regulatory and network constraints.[37]

By calibrating Equation (B12) to observed market shares and prices, this model rationalizes existing and observed price differentials in the base year data. Our approach can thus be viewed of as implicitly representing *existing* barriers to price equalization across markets that are independent of σ. Away from the benchmark, the model response is governed by the second- and higher-order properties of Equation (B12), as represented by the elasticity of substitution σ. Our base case assumption is "low integration" ($\sigma = 1$), and our analysis explores the implications of independents markets ($\sigma = 0$) and a "high integration" case (i.e., large σ's).

The solution method by Böhringer and Rutherford (2009) involves sequentially solving the electricity and economy-wide components under the same policy shock. Changes in general equilibrium prices are passed to the electricity generation model, and changes in the quantity of electricity produced and associated demand of inputs determined in the electricity generation model are transmitted back to the economy-wide model. The link between the two models is achieved by the linear demand functions for electricity (Equations (11) and (17)) which are sequentially updated using candidate general equilibrium solutions for electricity price and demand.[38]

We now provide an algebraic description of the integrated electricity-economy model. Let $n = 1, \ldots, N$ denote an iteration index and consider first the economy-wide component. The least-cost input requirements obtained from solving the electricity generation model in iteration $(n-1)$ are used to parametrize the general equilibrium model in (n). This is accomplished by

[37] This structure assumes that trade opportunities among regulated operators and between each regulated operator and the wholesale market are symmetric. We have experimented with more complicated substitution patterns but have found them to yield similar results.

[38] As pointed out by Böhringer and Rutherford (2009), the choice of the local elasticity value in the linear demand approximation can influence convergence speed but does not influence the equilibrium solution.

defining the market clearing condition for electricity (B4) as:

$$\sum_{g,t} Y_t^{g(n-1)} \geq \sum_j x_{jr}^{(n)} \frac{\partial \Pi_{jr}^{x^{(n)}}(p)}{\partial p_{ir}^{Y^{(n)}}} \qquad \perp \qquad p_{ir}^{Y^{(n)}} \qquad i = ele \tag{B13}$$

where the left-hand side represents electricity supply as defined in (B20). Demand for input i comprising fuels and other materials by the electricity sector is accommodated through:

$$Y_{ir}^{(n)} \geq \sum_j x_{jr}^{(n)} \frac{\partial \Pi_{jr}^{x^{(n)}}(p)}{\partial p_{ir}^{Y^{(n)}}} + \sum_{g,t} \phi_i^c \, c^g \, Y_t^{g^{(n-1)}} \qquad \perp \qquad p_{ir}^{Y^{(n)}} \tag{B14}$$

where ϕ_i^f and ϕ_i^c represent the benchmark value share of good i in variable generation costs. Factor market Equation (B8) for capital and labor are modified according to:

$$\sum_j Y_j^{(n)} \frac{\partial \Pi_{jr}^{(n)}(p)}{\partial p_r^{f^{(n)}}} + d_{fr}^{(n)}(p, M_r) + \sum_{g,t} \phi_f^c \, c^g \, Y_t^{g^{(n-1)}} \geq \sum \omega_r^f \qquad \perp \qquad p_r^f \tag{B15}$$

A consistent solution also requires capturing profits earned by non-regulated electricity operators. There are two types of profits. First, generators with marginal costs below the equilibrium price for electricity earn sub-marginal profits that reflect the shadow value of installed capacity (μ^g). Second, profits for Cournot players are due to markups on marginal generation costs. Total profits are implicitly given by the difference between the wholesale market price in each load segment and total generation costs. We assume that profits generated in a given region are distributed nationally in proportion to capital income.[39]

To account for these profits, we modify the income balance (B6) to account for technology-specific rents and profits (Π_r^{ELE}):

$$M_r^{(n)} = p^{k^{(n)}} \omega_r^k + p_r^{l^{(n)}} \omega_r^l + \sum_z p_r^{z^{(n)}} \omega_r^z + T_r^{(n)} + \Pi_r^{ELE^{(n)}} . \tag{B16}$$

Electricity-sector output and inputs are valued implicitly at market prices, and hence we do not

[39] Due to data constraints on the ownership patterns of electric-sector capital, we use base-year capital income as a proxy. We find that alternative assumptions regarding the distribution of electric-sector capital do not materially affect our conclusions.

need to include capacity rents and profits explicitly in the economy-wide model:

$$\Pi_r^{ELE^{(n)}} = \sum_{g \in r} \sum_t Y_t^{g^{(n-1)}} \left(p_{ele,r}^{Y^{(n)}} \, P^{r^{(n-1)}} - P_r^{\mathrm{c}^{(n)}} c^g \right) \tag{B17}$$

where the price indexes for variable generation costs are updated according to:

$$P_r^{\mathrm{f}^{(n)}} = \sum_f \phi_f^f \, p_r^{f^{(n)}} \,,$$

$$P_r^{c^{(n)}} = \sum_i \phi_i^c \, p_{ir}^{Y^{(n)}} + \sum_f \phi_f^c \, p_r^{f^{(n)}} \,,$$

respectively.

In the electricity generation model, the demand schedules are parametrized to locally approximate the response of the top-down model. In each iteration step, the linear function is re-calibrated to price and quantities derived from the top-down solution. Hence the demand function for a regulated operator f in iteration n (compare with Equation (11)) is updated according to:

$$D^{f^{(n)}} = \overline{D}^{f^{(n)}} \zeta^{f^{(n)}} \left(1 + \epsilon \left(\frac{P^{f^{(n)}}}{\overline{P}^{f^{(n)}} \xi^{f^{(n)}}} - 1 \right) \right) \tag{B18}$$

where

$$\zeta^{f^{(n)}} = \sum_j x_{jr}^{(n)} \frac{\partial \Pi_{jr}^{x^{(n)}}(p)}{\partial p_{ir}^{Y^{(n)}}} \overline{D}^{f^{(0)}} \qquad , \; f \in r$$

$$\xi^{f^{(n)}} = p_{ele,r}^{Y^{(n)}} \, \overline{P}^{f^{(0)}} \qquad , \; f \in r$$

are scale factors that are based on the n^{th} solution of the economy-wide model, and reference demand ($D^{f^{(0)}}$) and price ($\overline{P}^{f^{(0)}}$). A similar updating rule applies to wholesale electricity demand in Equation (17).

Finally, using the updated variable cost indexes, the revised unit profit functions for Cournot players and for price takers in iteration (n) are given by:

$$\pi_t^{g^{(n)}} = \begin{cases} p_t^{r^{(n)}} + \dfrac{\partial D^{r^{(n)}}(p_t^r)}{\partial Y_t^{g^{(n)}}}^{-1} - P_r^{c^{(n)}} c^g - \mu_t^{g^{(n)}} & \text{if } g \in r \text{ is a Cournot player} \\ p_t^{r^{(n)}} - P_r^{c^{(n)}} c^g - \mu_t^{g^{(n)}} & \text{if } g \in r \text{ is a price taker} \,. \end{cases} \tag{B19}$$

Non-negative profits and average cost pricing conditions for regulated operators in iteration (n)

are given by:

$$P^{c^{(n)}} c^g \geq \overline{C}_t^{f^{(n)}} \quad \perp \quad Y_t^{g^{(n)}} \geq 0\,, \tag{B20}$$

$$P^{f^{(n)}} = \frac{\sum_{g \in G_f} \sum_t P^{c^{(n)}}\, Y_t^{g^{(n)}}\, c^g}{D^{f^{(n)}}}\,. \tag{B21}$$

Appendix B.3 Integration of Heterogeneous Households into Economy-wide Transactions

The key idea is to compute a sequence of artificial agent equilibria which replicate choices of the many "real" households. First, a candidate equilibrium is computed in the economy-wide model where households in each region are replaced by a single artificial agent. Second, we solve a partial equilibrium relaxation of the utility maximization problem for each 15,588 households given candidate general equilibrium prices from the artificial agent problem. Iterating between both sub-problems involves re-calibrating preferences of the artificial agent in each region based on partial equilibrium quantity choices by "real" households.

To illustrate the key idea of the algorithm, we develop the following notation for nested utility functions. Let the quantity choices be denoted by q_i, for $i = 1, \ldots, I$, corresponding to commodities with prices p_i, respectively. The utility tree consists of $N+1$ levels, $n = 0, 1, \ldots, N$; on each level we distinguish several utility components. At the highest level (indicated by $n = N$) of the utility tree there is only one component, which corresponds to overall utility; this component is a function of utility components at the next-lower level $n = N - 1$. These utility components at $N - 1$ are in turn each a function of disjoint groups of utility components at the next lower level $N - 2$, and so on. Finally, the utility components at level $n = 1$ are functions of the the elementary utility components.

We specify the utility function for household h by assuming that all the utility components are linear homogeneous CES-type functions of the associated components at the next lower level:

$$q_{n,i}^h = \left[\sum_{j \in i} \theta_{n-1,j}^h \left(\frac{q_{n-1,j}^h}{\bar{q}_{n-1,j}^h} \right)^{\rho_{n,i}^h} \right]^{\frac{1}{\rho_{n,i}^h}} \qquad , \qquad \rho_{n,i}^h = \frac{\sigma_{n,i}^h - 1}{\sigma_{n,i}^h}\,, \tag{B22}$$

where $j \in i$ is used to indicate the set of components $q_{n-1,j}^h$ associated with $q_{n,i}^h$, and where $\sigma_{n,i}^h$ denotes the elasticity of substitution between commodities $j \in i$. Note that we write the nested utility function in calibrated share form (Rutherford, 2002); θ and $\bar{q}$ denote the value share and consumption in the benchmark equilibrium, respectively.

The decomposition algorithm is implemented by replacing in each region the household side with an artificial agent whose utility function exhibits the identical structure as household utility in Equation (B22):

$$Q_{n,i} = \left[\sum_{j \in i} \Theta_{n-1,j} \left(\frac{Q_{n-1,j}}{\bar{Q}_{n-1,j}} \right)^{\tilde{\rho}_{n,i}} \right]^{\frac{1}{\tilde{\rho}_{n,i}}} \quad , \quad \tilde{\rho}_{n,i} = \frac{\tilde{\sigma}_{n,i} - 1}{\tilde{\sigma}_{n,i}} \tag{B23}$$

where $\Theta_{n,j}$ and $Q_{n,j}$ denote the respective counterparts for the artificial agent to individual households as defined in Equation (B22). A key insight from Rutherford and Tarr (2008) is that the choice of $\tilde{\sigma}_{n,i}$ is entirely innocuous as this parameter bears no economic significance for the behavior of "real" households in the underlying economic model (it can, however, affect the convergence speed of the iterative solution procedure).

Given benchmark data on observable household demand $\bar{q}_i^h$ and prices $\bar{p}_i$, we initialize the artificial agent general equilibrium model such that commodity demands are consistent with the aggregate of benchmark household demands. This is achieved by calibrating consumption ($\bar{Q}$) and value share (Θ) parameters as:

$$\bar{Q}_{n,j} = \sum_{h=1}^{H} \bar{q}_{n,j}^h \, , \tag{B24}$$

$$\Theta_{n,j} = \frac{\bar{p}_{n,j} \bar{Q}_{n,j}}{\sum_{j' \in i} \bar{p}_{n,j'} \bar{Q}_{n,j'}} \, . \tag{B25}$$

Solving for a CO_2 policy shock involves first solving the artificial agent model to obtain a candidate vector of general equilibrium prices $\mathbf{p}^k$. k denotes an iteration index. The second step solves a partial equilibrium relaxation of the underlying economy by evaluating household demand functions $q_{n,i}^{h,k}(\mathbf{p}^k, y^k)$, where household income y^k is updated sequentially at prices in iteration k. The key step in each iteration involves "re-calibrating" preferences of the artificial agent based on partial equilibrium households' quantity choices:

$$\bar{Q}_{n,i}^{k+1} = \sum_{h=1}^{H} q_{n,i}^{h,k}(\mathbf{p}^k, y^k) \, , \tag{B26}$$

$$\Theta_{n,j}^{k+1} = \frac{\bar{p}_{n,j}^{k} \sum_{h=1}^{H} q_{n,j}^{h,k}(\mathbf{p}^k, y^k)}{\sum_{j' \in i} \bar{p}_{n,j'}^{k} \sum_{h=1}^{H} q_{n,j'}^{h,k}(\mathbf{p}^k, y^k)}\,. \tag{B27}$$

Note that this iterative procedure never alters preferences of the "real" households; it simply "re-benchmarks" successively the utility function of the artificial household to be consistent with the aggregated choices of individual households in each iteration.