Regulation, Allocation, and Leakage in Cap-and-Trade Markets for CO2

James Bushnell and Yihsu Chen*

March, 2009

Abstract

Among the most contentious elements of the design of cap-and-trade systems for emissions trading is the allocation or assignment of the emissions credits themselves. Policy-makers usually try to satisfy a range of goals through the allocation process, including easing the transition costs for high-emissions firms, reducing leakage to unregulated regions, and mitigating the impact of the regulations on product prices such as electricity. In this paper we develop a detailed representation of the US western electricity market to assess the potential impacts of various allocation proposals. Several proposals involve the "updating" of permit allocation, where the allocation is tied to the ongoing output, or input use, of plants. These allocation proposals are designed with the goals of limiting the pass-through of carbon costs to product prices, mitigating leakage, and of mitigating the costs to high-emissions firms. However, allocation updating can greatly inflate permit prices, thereby limiting the benefits of such schemes to high emissions firms. Rather than mitigating the impact on high carbon producers, the net operating profit of such firms can actually be lower under input-based updating than under auctioning. This is due to the fact that product prices (and therefore revenues) are lower under input-based updating, but overall compliance costs are relatively comparable between auctioning and input-based updating. Thus, the anticipated benefits from allocation updating are greatly reduced and further distortions are introduced into the trading system.

Bushnell: University of California Energy Institute, Haas School of Business, and NBER. Email: bushnell@haas.berkeley.edu. *Chen:* Sierra Nevada Research Institute, UC Merced. Email: yihsu.chen@ucmerced.edu.

1 Introduction

As climate change policy in the United States begins to move from concept to implementation, there is increasing focus on the details of that implementation. There is widespread support for focusing much of the regulatory emphasis for greenhouse gas (GHG) reductions on cap-and-trade as the mechanism to achieve these reductions. Although the concept of capping GHG emissions and allowing trading for compliance has broad acceptance, many details about the design of a cap-and-trade system are still hotly debated. Policy-makers must often balance equity considerations with the desire to achieve cost-effective and meaningful reductions in emissions. The differential impacts of GHG regulation on various industries, regions, and consumers make the design of those regulations very contentious.

To date, most of the policy initiative has come from the state or regional level in the United States. The Regional Greenhouse Gas Initiative, a cap-and-trade program which covers the electricity sector of the northeastern U.S., began operating in 2009. California's Assembly Bill 32 (AB 32) requires that all sectors of its economy reduce their aggregate GHG emissions to 1990 levels by 2020. The California initiative is proceeding in parallel with the broader-based Western Climate Initiative (WCI). The WCI will establish a regional cap-and-trade program that will initially encompass large stationary sources (primarily electricity) and then expand to include other sources, including transportation fuels in a second phase.[1]

The fact that GHG policy is being driven at the local, rather than national level, has created concern over the geographic limitations of the regulations. Environmental targets can be undermined if production is able to shift away from the jurisdictional reach of the regulator through either leakage or reshuffling of production sources.[2] These concerns over regional US policies reflect similar, more general concerns with leakage as a challenge even for international climate agreements. Because of this, a significant fraction of the implementation efforts have been devoted to policies and design choices that can mitigate this leakage problem. The choice of the level of the supply chain at which emissions are measured and regulations applied, known as the "point of regulation," can play a role in limiting the circumvention of the regulations, although previous analysis has argued that these benefits can be overstated and that other efficiency problems can be created.[3] Here, we focus on the impact of alternative allocation proposals on market performance.

In this paper, we develop a detailed model of the power sector in the western United

[1] WCI, 2008.

[2] See Bushnell, et al, 2008, Fowlie, 2008, and Chen (2008).

[3] See the recommendations of the Market Advisory Committee to the California Air Resources Board (2007), Burtraw (2008), Chen, et al., (2008).

States, and examine the impacts of alternative cap-and-trade designs on the operations, emissions, and prices in this region. Our research is motivated by several important economic and policy questions relating to cap-and-trade design. First, there is the practical question of just how severe leakage would be under varying levels of WCI participation. Obviously the more states that participate, the less severe the potential for leakage may be, but we hope to quantify those risks by examining the benefits of expansion from California to the rest of the WCI. Second, we take advantage of current policy debates over cap-and-trade design to examine the general relationships between specific design elements and market outcomes such as leakage and overall market efficiency. Last, we provide some quantitative, but necessarily qualified, estimates of the impacts of cap-and-trade on permit and power prices in the western U.S.

We find that leakage of electricity production to unregulated regions is a significant concern, even under the expanded WCI program. One of the key design elements under consideration is the allocation of emissions credits through contingent allocation, or "updating." This is a form of allocation that rewards firms with permits for continued production. The environmental economics literature often focuses on "output-based" updating, which ties allocation of permits to quantities of product produced. Our results show that output-based updating substantially reduces leakage, and produces relatively low electricity prices compared to an exogenous form of allocation, such as auctioning. Much of the allocation seen in practice, however, in fact ties allocations to *inputs* or even emissions, rather than outputs. This form of "input-based" updating has drawn supporters because of the view that it can help ease the transition to carbon regulation by allocating disproportionately more permits to relatively high carbon producers, as well as limit the permit windfall that may be reaped by a low carbon producer under a purely output-based scheme. However, as we demonstrate in the context of the WCI market, input, or equivalently, "fuel-based" updating in fact reverses many of the effects seen under output-based updating. One implication of this result is that, while it is motivated by a desire to provide financial relief to carbon intensive firms, fuel-based updating can in fact leave high polluting firms little better off than if all the permits were auctioned.

2 Design and Modeling of Cap-and-Trade Markets

As is necessarily the case with markets that are created by regulatory fiat, the choices made by regulators in designing those markets go a long way toward influencing their outcomes. This is certainly true in the case of cap-and-trade (C&T) markets for the trading of emissions compliance obligations. Such markets exist because regulators create a demand for emissions credits through the setting of the emissions cap. By establishing tradable emissions credits, regulators also create a valuable new property right. Not

surprisingly, the distribution or allocation of those valuable rights is often a source of great contention.[4]

Traditionally, the allocation of permits has been held to be an issue limited to economic transfers that need not affect the efficiency of the resulting emissions trading market. Certainly it is the equity concerns that dominate the discussions and debates amongst policymakers and the affected industries. The impacts on efficiency can be negligible if the allocation is truly exogenous to the ongoing operations of the industries subject to the emissions cap, as is the case with the US SO2 trading program (Ellerman, et al., 2000). However, in many cases the allocation of emissions permits has either been endogenous, or contingent upon market outcomes.

One proposal that has been of increasing interest is to allocate emissions credits using *output-based updating*. Under output-based updating each firm receives an allocation of emissions credits that is proportional to its total product production. In the electricity context, for example, this means each firm receives an allocation that is proportional to the MWh generated within the regulatory jurisdiction.[5] The effects of output-based updating have been a subject of much research.[6] In general, it is believed that output-based updating would help to mitigate leakage, as firms would be rewarded (in the form of permits) for domestic production. Output-based updating is also widely believed to result in lower product prices than alternative forms of allocation. While the academic literature has focused on the detrimental efficiency effects of such a price impact, it has an appeal to regulators. For example, the design recommendations of both the California Public Utilities Commission and the WCI include the minimization of the impacts of carbon regulations on consumers as a prominent objective of the allocation process. Despite the appeal of the product price effect, these "lower" prices can lead to inefficient over-consumption as the externality cost of the pollution is not reflected in product prices.[7]

Further, there is a concern that output-based updating, if applied symmetrically to all producers (or at least to all fossil-fueled producers), would exacerbate equity concerns. For example, there is a fear that low-carbon producers will experience a "windfall" under

[4]See, for example Kanter and Mouawad, 2008.

[5]We use the term generated somewhat loosely here, as in the California and WCI context, the point of regulation is mixed. Emissions from sources within the cap-and-trade region will be regulated along the lines of traditional source-based cap-and-trade systems. In additions imports into the cap-and-trade region will also be regulated under a system known as the "first-deliverer." In effect the importer of the product will be required to surrender emissions credits equivalent to the emissions required to produce that product, even though the production itself occurs physically outside of the cap-and-trade region. Under first-deliverer, the importers would also be eligible for allocations under an updating scheme. These allocations would be proportional to the MWh imported into the cap-and-trade region.

[6]see Jensen and Rasmussen (2000), Fischer (2003), and Fischer and Fox (2007)

[7]See Burtraw, Palmer, and Kahn (2005) for a discussion of the various impacts of updating.

output-based allocation, while high-carbon producers will suffer most of the cost impacts of GHG regulations. This is because output-based allocation favors cleaner producers. They are rewarded for production, and penalized for emissions. Traditionally, allocation has been used as a tool to "soften the blow" of increased environmental compliance through allocations based upon historic emissions patterns, also known as grandfathered allocations. Under grandfathered allocation, larger polluters receive a larger share of the allocations, while also paying more for compliance due to their higher emissions levels. In this way the total costs to high-emissions producers are mitigated, while the marginal cost of compliance remains the same. The California Public Utilities Commission has recommended (CPUC, 2008) an alternative we will refer to as "fuel-based" updating in order to address this equity concern. Under fuel-based updating, the allocation of emisssions credits per MWh of generation would be higher for high-carbon (*e.g.,* coal-based) producers than it would be for low-carbon (*e.g.,* gas-based) producers.

Fuel-based updating is part of a general class of allocation approaches also known as input-based, or emissions-based updating. It is in fact more common than output-based allocation in practice. This is in part due to the equity concerns described above, and also due to the fact that it is not always easy to either measure or compare the "output" of some plants, particularly in C&T programs that span multiple industries. Some states participating in the US NOx budget program utilize a form of input-based updating that is based upon the heat-input of power plants. Most notably, allocations the ETS market for CO_2 in the European Union have contained, at least implicitly, several aspects of updating. (Grubb and Neuhoff, 2006, Ahman, et al., 2006).

The analysis of updating proposals, both input and output based, has focused on the efficiency implications of these approaches. In addition to inefficient over-consumption, updating can result in a productive inefficiency by distorting relative production decisions, as well as distort long-term investment signals.[8] Several papers have examined the interaction of allocation policy with leakage and efficiency for specific industries, including electricity (Neuhoff, Martinez, and Sato, 2006) and cement (Demailly and Quirion, 2006).

However, these papers tend not to emphasize the aspects of updating that motivate their application in practice. These are the impacts of updating on permit and product prices, as well as the equity effects for firms. Using a theoretical model, Bohringer and Lange (2005) consider a "closed" trading system where the cap is fixed and there is no opportunity for trading with other emissions markets. When allocations are exactly proportional (but not equivalent) to emissions, input-based updating recreates the "first-best" product prices and emissions of auctioning. The permit prices, however, rise in a closed system. This is because emissions-based allocation reduces the perceived marginal

[8]See Jensen and Rasmussen (2000), Ahman and Holmstrom (2006), and Sterner and Muller (2008)

rate of emissions for firms. In order to achieve the reductions required by the cap, permit prices must rise to offset this perceived discounting of emissions rates. Although higher-emissions firms receive larger allocations, they pay disproportionately more for the emissions that are not covered by allocation, due to the higher equilibrium permit prices. In an "open" trading system, the effect of updating tends to push the abatement to regions or industries that are not receiving the implicit production subsidies in the form of updated permit allocations.

The western market we examine here has characteristics of both closed and open systems. The allocation rules are aimed at market shares, and therefore the industry level cap would not change with allocation results. In the initial years of the WCI, when the allocated shares will be the largest, emissions will be dominated by the electric sector. As mentioned before, however, leakage is also a concern. There will therefore be opportunities for trading product, if not emissions permits with neighboring regions. The goal of this paper is to try to sort through these factors and establish the relative impacts of them on market outcomes.

2.1 Analysis of Cap-and-Trade Design

For the most part, empirical projections of cap-and-trade markets for GHG have tended to take the "long-view" approach of simulating outcomes 10 to 20 years in the future. The studies commissioned by the regulatory agencies responsible for implementing California's AB 32, for example, emphasize the year 2020 (see discussion of the E3 model in CPUC, 2008). As such this work tends to be focused on the important assumptions one must make about the trajectory of future electricity demand and trends in the investment of new generation technologies. Chen et al. (2008) examine the economic and emissions implications of the three C&T proposals considered under AB32: source-, load- based and first-seller. When cross-boarder electricity sales are subject to an emissions cap, they show that these three proposals produce the same market outcomes, and all of them are prone to emissions leakage. None of these studies consider how outcomes might change under different permits allocation schemes.

Because of our focus on the specific design of the cap-and-trade mechanism, and its impact on the operation of electricity markets, we instead take the "near-view" approach. We base our analysis upon actual market data drawn from the year 2007, and look at the counter-factual question of how those markets would have functioned under a cap-and-trade regime. In this sense the work follows in the spirit of Fowlie (2008), who also studies the potential for leakage from a California-only market, and also that of Bushnell, Mansur, and Saravia (2008) who deploy similar techniques to examine competition and vertical contracting issues. In a fashion similar to Schuliken, et al., (2008),

we formulate the joint equilibrium outcomes of the emissions and electricity market as a linear-complementarity problem.

Our study differs from previous work in several important ways. While Fowlie and the E3 study employ an empirical model of portions of the western electricity market, we model the emissions credit prices as endogenous to the cap-and-trade market. This is central to our work given our focus on the endogenous impact of allocation policies on permit prices. Second, in addition to California's CO2 policies, we examine the broader western market proposed under the WCI. Last, we explicitly consider how allocation policies can affect firm behavior in the western US. Previous work examining the impacts of allocation have either taken a general equilibrium approach (Bohringer and Lange (2005), Sterner and Muller (2008), Fischer and Fox (2008)), or more complex formulations applied to stylized market data (Shuliken, et al., 2008, Neuhoff, et al., 2006).

We also take a different approach than most similar analyses to developing our market elements used for the simulation. One difficulty with simulating electricity markets in a high level of detail is that, while data on most fossil-fuel based generation units is quite extensive and reliable, there are far less data on the activities of hydro-electric plants, renewable generation, and the substantial amount of power generated from combined heat & power or "cogeneration" plants. When building a counter-factual recreation of an electricity market, these data gaps make assumptions about the missing production necessary.

We take the approach of restricting our construction of a counter-factual market outcome to the portion of resources for which we do have detailed data. In effect we are assuming that, under our counter-factual, the operations of non-modeled generation plants would not have changed. This is equivalent to assuming that compliance with the carbon reduction goals of a cap-and-trade program will be achieved through the reallocation of production within the set of modeled plants. We feel that this is a reasonable assumption for two reasons. First the vast majority of the carbon emissions from this sector come from these modeled resources. Indeed, data availability is tied to emissions levels since the data are reported through environmental compliance to existing regulations. Second, the total production from "clean" sources is unlikely to change in the short-run. The production of low carbon electricity is driven by natural resource availability (rain, wind, sun) or, in the case of combined heat and power (CHP), to non-electrcity production decisions. The economics of production are such that these sources are already producing all the power they can, even without additional carbon regulation. To a first-order, short-run emissions reductions will have to come either from shifting production from among conventional sources, a reduction in end-use electricity demand, or through substitution with unregulated imports, *i.e.*, leakage or reshuffling.[9]

[9]It is important to recognize that our modeling approach not only assumes that existing zero-carbon

3 Model

In this section, we describe first our equilibrium model and then discuss how we apply data from various sources to arrive at our calculations.

Although this simulation approach is capable of representing imperfect competition in the product market (*i.e.*, electricity) we assume here that firms act in a manner consistent with perfect competition with regards to both the electricity and emissions permit markets. We still model these markets as a series of equilibrium conditions for each of the individual firms represented, as the incentive effects on individual firms from policies such as updating are still relevant here.

The key variables and parameters of the model are grouped according to four important indices; the firm, location, technology, and time period of production. The total production of firm i from generation technology j, and location l at time t is represented by $q^i_{j,l,t}$. Firm and technology level emissions rates are denoted $e^i_{j,l}(q^i_{j,l,t})$. We assume that emissions rates can be increasing in quantity, but are unchanging over time. Production costs $C^i_{j,l}(q^i_{j,l,t})$ vary by firm, technology, and location, and as described below are assumed to be quadratic in output $q^i_{j,l,t}$.

For each firm $i \in \{1, ..., N\}$, locations $l \in \{1..L\}$ and time period $t \in \{1, ..., T\}$, a perfectly competitive, or cost-minimizing firm i maximizes profits:

$$\pi_{i,l,t}(q^i_t) = \sum_l [p_{l,t} \cdot q^i_{j,l,t} - \sum_j C^i_{l,j}(q^i_{j,l,t})] - \lambda \sum_{l \in REG} q^i_{j,l,t} \cdot e^i_{j,l}(q^i_{j,l,t}), \tag{1}$$

where $p_{l,t}$ and λ are the wholesale prices of electricity and CO2 permits, respectively. Permit prices are assumed to be uniform across the regulated (capped) region. Wholesale electricity is assumed to be a homogenous commodity for purposes of setting wholesale prices, although prices are assumed to vary by location subject to transmission constraints as described below. However, electricity production falls into two categories, that within the region covered by the emissions cap and that outside the reach of the regulation. The set REG represents those plants located inside the cap and trade region.

sources will not change *how much* they produce but also *when* they produce it. An interesting question is whether a redistribution of hydro-electric power across time could lower carbon emissions by enabling a better management of fossil generation sources. Such an analysis would require a co-optimization of hydro and thermal electric production and is beyond the scope of this paper.

3.1 Cap-and-Trade Design

The profit function described in the previous section assumes a standard source-based cap-and-trade market, where the compliance obligation rests explictly on the producer (in this case the electricity generator). As the focus of C&T design turned to allocation, however, much of the regulatory emphasis was devoted to mitigating consumer prices, smoothing the cost impacts to firms (at least somewhat), and mitigating the "windfall" profits that might be earned by low-carbon producers (CPUC, 2008). These goals were to be addressed primarily through allocation policies. In particular, two specific alternative implementations of output-based updating are considered here.

3.1.1 Output-based Updating

As discussed above, one mechanism that can depress product market prices and at least partially combat leakage is output-based updating. In this context, the allocation of emissions credits would be tied to the electricity production of firms. Each MWh of production would earn a fraction of an emissions credit.

In practice, the above model would be part of a larger multi-period cycle of emissions compliance and allocation. Because of the data-intensive nature of our approach, it is difficult to explicitly model multiple years. We instead represent the allocation decision as part of a "closed-loop," to a single cap-and-trade compliance cycle. This is sufficient to capture the key qualitative impacts of updating on the incentives of firms. In effect, the allocations are given out at the end of the cycle, just before credits are required to be surrendered. These allocations are then linked to the actual output of producers during the compliance cycle that is about to conclude. We therefore suppress the effect of interest rates or other dynamic considerations.

Following this assumption, we can rewrite the profit maximization problem for each firm to include the prospect of output-based allocation of emissions credits. Let $\delta_t(q^i_{l,t}) = \delta \cdot q^i_{l,t}$ be the allocation of emissions credits earned for use in the compliance cycle $1..T$ from producing $q^i_{l,t}$ units of electricity in regulated region l during period $t \in 1..T$. Note that we assume that the overall cap does not change, only that the distribution of (zero-cost) emissions credits across firms varies with the relative output of firms and their facilities. In other words $\delta = \tilde{\delta}\frac{CAP}{\sum_{1..T} Q_t}$. where Q_t is the aggregate production (market demand) in period $t, \tilde{\delta}$ is the overall fraction of carbon credits that are allocated through updating, and $t \in 1..T$ is the cycle of the compliance period. Thus, the program of output-based updating would not take the form of a "tradable performance standard."[10]

[10] See Fischer, 2003.

Under a performance standard, the subsidy for output is not limited by an overall cap. Even if the performance standard were a regulatory mandate, rather than an allocation of emissions credits, there is an implicit subsidy of production. Compliance with a mandate, when specified as an intensity per unit of output, can be advanced both through limiting the undesirable input *and* expanding total output.[11]

The profit for firm i will now include consideration of the additional permits earned from additional production.

$$\pi_{i,l,t}(q_t^i) = \sum_l [p_{l,t} \cdot q_{j,l,t}^i - \sum_j C_{l,j}^i(q_{j,l,t}^i)] - \lambda \sum_{l \in REG} q_{j,l,t}^i \cdot [e_{j,l}^i(q_{j,l,t}^i) - \delta]. \tag{2}$$

This profit equation highlights how the updating weakens the marginal cost impact of the cap-and-trade requirement *for a given permit price, lambda*. For facilities with an emissions rate higher than the allocation rate, the cap-and-trade still effectively taxes output, although at a lower rate. For facilities with an emissions rate that is lower than the allocation rate, $e_{j,l}^i < \delta$, there is now a production subsidy.

3.1.2 Fuel-based Updating

The other approach to updating under consideration would distinguish between the inputs of various production sources. Motivated by a desire to limit the cost impacts of cap-and-trade on utilities heavily reliant on coal-based sources of power, this proposal would allocate emissions credits to generation from differing fuel sources in a ratio roughly aligned with the average GHG emission rate from each fuel source. In the notation of our model, this approach would provide δ_j emissions credits to each MWh of generation from a source of technology type j. In other words, each technology could in theory be subject to a separate allocation ratio. The resulting equilibrium condition for production for firm i would be

$$\pi_{i,l,t}(q_t^i) = \sum_l [p_{l,t} \cdot q_{j,l,t}^i - \sum_j C_{l,j}^i(q_{j,l,t}^i)] - \lambda \sum_{l \in REG} q_{j,l,t}^i \cdot [e_{j,l}^i(q_{j,l,t}^i) - \delta_j]. \tag{3}$$

As with the output-based allocation, the allocation component δ_j weakens the impact of permit prices on the perceived marginal cost of production. The strength of this effect is now asymmetric, and its net impact will depend upon the specific value of δ_j. The general intent of the fuel-based updating is to weaken the impact on higher emission technologies

[11] See Fullerton and Heutel, 2007. One current proposal that exhibits this characteristic is the "low carbon fuel standard" for transportation fuels (see Holland, Hughes, and Knittel, 2008).

and therefore soften the blow of implementing the cap. A somewhat extreme version of this allocation would arise if emissions rates within each technology class were constant and equivalent across all firms and locations (*i.e.,* $e^i_{j,l} = e_j \quad \forall i, l$) and the allocation factors were applied proportionately to emission rates $\delta_j = \tilde{\delta} \cdot e_j$. Then equation (3) can be rewritten as

$$\pi_{i,l,t}(q^i_t) = \sum_l [p_{l,t} \cdot q^i_{j,l,t} - \sum_j C^i_{l,j}(q^i_{j,l,t})] - \lambda(1-\tilde{\delta}) \sum_{l \in REG} q^i_{j,l,t} \cdot e_j. \tag{4}$$

Note that (4) is essentially equivalent to equation (1), again assuming that emissions rates are constant over technologies and firms, except for the fact that the permit price has now been scaled by $1-\tilde{\delta}$ for all firms. The Bohringer and Lange (2005) result implies that in a closed cap-and-trading system this results in the same outcomes that would be produced by a grandfathered allocation of permits, except for the fact that permit prices are increased by $1/(1-\tilde{\delta})$. In this paper we examine the impact of this kind of updating in a much more complex production environment, with leakage and capacity constraints. In addition, the actual updating policies proposed for the WCI do not reach the level of perfectly matching emissions rates, although some would come close. As the updating policy moves toward better correlation with emissions rates, we would expect these effects to become more pronounced. An empirical analysis such as this one is necessary to determine exactly how pronounced these impacts would be.

3.2 Transmission Network Management

We assume that the transmission network is managed efficiently in a manner that produces results equivalent to those reached through centralized locational marginal pricing (LMP). For our purposes this means that the transmission network is utilized to efficiently arbitrage price differences across locations, subjet to the limitations of the transmission network. Such arbitrage could be achieved through either bilateral transactions or a more centralized operation of the network. For now we simply assume that this arbitrage condition is achieved.

Mathematically, we adopt an approach utilized by Metzler,et al. (2003), to represent the arbitrage conditions as another set of constraints of the market equilibrium. Under the assumptions of a DC load-flow model, the transmission 'flow' induced by a marginal injection of power at location l can be represented by a power transfer distribution factor $PTDF_{lk}$, which maps injections at locations, l, to flows over individual transmission paths k. Within this framework, the arbitrage condition will implicitly inject and consume power, $y_{l,t}$ to maximize available and feasible arbitrage profits as defined by

$$\sum_{l \neq h} (p_{h,t} - p_{l,t})\, y_{l,t}.$$

In the above arbitrage equation, the location h is the arbitrarily assigned "hub" location from which all relative transmission flows are defined. Thus an injection of power, $y_l \geq 0$, at location l is assumed to be withdrawn at h. This arbitrage condition is subject to the flow limits on the transmission network, particularly the line capacities, T_k.

$$-\overline{T}_k \leq PTDF_{l,k} \cdot y_{l,t} \leq \overline{T}_k$$

This combination of arbitrage pressure and physical transmission constraints are resolved in the solution to the following langrangian.

$$\max_{y_{l,t}} \sum_{l \neq h} \left[(p_{h,t} - p_{l,t})\, y_{l,t} - \left(PTDF_{l,k} \cdot y_{l,t} - \overline{T}_k \right) \tau_k \right].$$

Where τ_k is the shadow value of capacity on transmission path k.

3.3 Equilibrium Conditions

Each firm has a limited capacity of each technology type in each location, which we denote by $\overline{q^i_{j,l}}$. Given the above framework, we can represent the resulting equilibrium as the set of quantities that simultaneously satisfy the following first order conditions. We represent as a complementarity condition, where the symbol $\perp$ signifies complementarity. For each firm i and period t:

$$q^i_{j,l,t} \geq 0 \;\; \perp \;\; p_t - C^{i\prime}_{l,j}(q^i_{j,l,t}) - \lambda e^i_{j,l,t} - \gamma^i_{j,l,t} \leq 0 \quad \forall i, j, t, \; l \in REG. \tag{5}$$

and

$$q^i_{j,l,t} \geq 0 \;\; \perp \;\; p_t - C^{i\prime}_{l,j}(q^i_{j,l,t}) - \gamma^i_{j,l,t} \leq 0 \quad \forall \; i, j, t, l \notin REG. \tag{6}$$

Here $\gamma^i_{j,l,t}$ is the shadow value of the capacity constraint on technology $q^i_{j,l,t}$.

$$\gamma^i_{j,l,t} \geq 0 \perp \overline{q^i_{j,l}} - q^i_{j,t} \geq 0 \quad \forall \; i, j, t, l. \tag{7}$$

Each firm, taking prices as exogenous, sets its production so that marginal costs equal the price at the location of the production. This marginal cost component includes the costs of emissions credits in locations subject to the emissions cap as well as the shadow price of the limited capacity of that technology.

Output-based Updating

As described above, output-based updating would allocate δ permits per MWh to each firm. Differentiating the profit function (2) yields the following.

$$q^i_{j,l,t} \geq 0 \;\; \perp \;\; p_t - C^{i\prime}_{l,j}(q^i_{j,l,t}) - \lambda(e^i_{j,l,t} - \delta) - \gamma^i_{j,l,t} \leq 0 \quad \forall\, i,j,t,l \in REG. \tag{8}$$

If the updating is instead fuel or technology specific, then the above condition is modified so that the allocation quantity, now δ_j, can be unique to a technology type j.

$$q^i_{j,l,t} \geq 0 \;\; \perp \;\; p_t - C^{i\prime}_{l,j}(q^i_{j,t}) - \lambda(e^i_{j,l,t} - \delta_j) - \gamma^i_{j,l,t} \leq 0 \quad \forall\, i,j,t,l \in REG. \tag{9}$$

Environmental Constraint

Along with equilibrium conditions (5) and (6), the equilibrium for a combined electricity and emissions market will include the following condition defining the permit price for the overall compliance period.

$$\lambda \geq 0 \perp \sum_{i,l \in REG,t} e^i_{j,l,t}(q^i_{j,l,t}) - e^{MAX} \leq 0. \tag{10}$$

Where, again, the symbol $\perp$ signifies complementarity between the constraint on available emissions permits and the permit price, which is the shadow price of that constraint. If there are excess emissions permits, the price is zero; otherwise λ is positive.

Network Constraints

Prices at individual locations will be determined by the production decisions of firms and the flows over the transmission network. The arbitrage minimization assumption described above produces the following condition.

$$p_{h,t} - p_{l,t} - \sum_k PTDF_{l,k} \cdot \tau_{l,t} = 0.$$

This reflects the general condition from an efficiently utilized network, that the prices between locations differ only by the additional costs of congestion of a shipment between those locations, as measured by the flows over lines times their shadow prices.

$$\tau_k \geq 0 \perp PTDF_{l,k} \cdot y_{l,t} - \overline{T}_k \leq \quad \forall k.$$

When the inverse demand, marginal cost, and emissions functions are linear, as they are described below, the equilibrium conditions for each of the possible cap-and-trade regimes, along with the respective conditions for network operations and emissions market balance, combine to form a linear complementarity problem (Cottle, Pang, and Stone, 1992) with variables $q^i_{j,l,t}, y_{l,t}$ and dual values $\tau_{l,t}, \lambda$, and $\gamma^i_{j,l,t} \geq 0$. The solution to this complementarity problem constitutes a perfectly competitive equilibrium to this market, subject to the respective definitions of the cap region and allocation policy. Using the data sources and functional forms described in the following section, we calculate these equilibrium outcomes using the PATH solver algorithm (Dirske and Ferris, 1995) implemented through the AMPL math programming language.

4 Data Sources and Assumptions

Our primary data source is the BASECASE dataset from Platts, which is in turn derived primarily from the continuous emissisions monitoring system (CEMS) utilized by the U.S. Environmental Protection Agency (EPA) to monitor the emissions of large stationary sources. Almost all large fossil-fired electricity generation sources are included in this dataset. However, hydro-electric, renewable, and some small fossil generation sources are missing. The CEMS reports hourly data on several aspects of production and emissions. Hourly data on nuclear generation plants are included with fossil generation data in the BASECASE dataset. Here we utilize the hourly generation output and carbon emissions for available facilities.

These hourly output data are aggregated by firm and region to develop the "demand" in the simulation model. As described above, this is in fact a residual demand; the demand that is left after applying the output from non-CEMS plants. Plant cost, capacity, and availability characteristics and regional fuel prices are then taken from the Platts POWERDAT dataset. These data are in turn derived from mandatory industry reporting to the Energy Information Administration (EIA) and the National Electric Reliability Council (NERC).

These data are then combined to create a demand profile and supply functions for periods in the simulation. Although hourly data are available, for computational reasons we aggregate these data into representative time periods. There are 20 such periods for each of the four seasons, yielding 80 explicitly modeled time periods. As California policy was the original focus of this work, the aggregation of hourly data was based upon a sorting of the California residual demand. California aggregate production was sorted into 20 bins based upon equal MW spreads between the minimum and maximum production levels observed in the 2007 sample year. The number of season-hour observations in

each bin is therefore unbalanced, there are relatively few observations in the highest and lowest production levels, and more closer to the median levels. The demand levels used in the simulation are then based upon the mean production levels observed in each bin. In order to calculate aggregate emissions, the resulting outputs for each *simulated* demand level was multiplied by the number of *actual* market hours used to produce the input for that simulated demand level.

In the following sub-sections, we describe further the assumptions and functional forms utilized in the simulation.

4.1 Market Demand

End-use consumption in each location is represented by the demand function $Q_{l,t} = \alpha_t - \beta_l p_{l,t}$, yielding an inverse demand curve defined as

$$p_{lt} = \frac{\alpha_t - \sum_{i,j} q_{i,j,t} - y_i}{\beta_l}$$

where y_i is the aggregate net transmission flow into location l. The intercept of the demand function is based upon the actual production levels calculated as described above. The mean hourly demand is summarized by GHG regulatory region in Table 1.[12] In each representative hour, demand is assumed to be at the levels reflected in Table 1 when market prices are equal to the levels observed in the actual market hours from which the demand numbers are taken. In other words, we model a linear demand curve that passes through the observed price-quantity pairs for each hour. As electricity is an extremely inelastic product, we utilize an extremely low value, 8 MW/$ for the slope of this demand curve. In most hours and regions, this yields a price elasticity of less than -.05.[13]

4.2 Fossil-Fired Generation Costs and Emissions

We explicitly model the major fossil-fired thermal units in each electric system. Because of the legacy of cost-of-service regulation, relatively reliable data on the production costs of thermal generation units are available. The cost of fuel comprises the major component of the marginal cost of thermal generation. The marginal cost of a modeled generation

[12] As described below, supply and demand regions can be characterized as belonging in one of 5 electrical zones or one of the four zones distinguished by climate regulation listed in Table 1.

[13] When the market is modeled as perfectly competitive, as it is here, the results are relatively insensitive to the elasticity assumption, as price is set at the marginal cost of system production and the range of prices is relatively modest.

Table 1: Demand by Region and Season

Season	AZNM	CA	NWPP	non-WCI
Winter	10925	11641	10781	16407
Spring	12130	11369	8394	15604
Summer	14705	16314	12823	18766
Fall	10943	13504	12878	16622

unit is estimated to be the sum of its direct fuel, environmental, and variable operation and maintenance (VO&M) costs. Fuel costs can be calculated by multiplying the price of fuel, which varies by region, by a unit's 'heat rate,' a measure of its fuel-efficiency.

The capacity of a generating unit is reduced to reflect the probability of a forced outage of each unit. The available capacity of generation unit i, is taken to be $(1 - fof_i) * cap_i$, where cap_i is the summer rated capacity of the unit and fof_i is the forced outage factor reflecting the probability of the unit being completely down at any given time.[14] Unit forced outage factors are taken from the generator availability data system (GADS) data that is collected by the North American Reliability Councils. These data aggregate generator outage performance by technology, age, and region.

Generation marginal costs are derived from the costs of fuel and variable operating and maintenance costs for each unit in our sample. Platts provides a unit average "heat-rate" or fuel-efficiency value for each of these units. These heat-rates are multiplied by a regional average fuel cost for each fuel and region, also taken from Platts. Costs for each technology type are then aggregated by firm and region, and then represented with a single quadratic function for each of five technology types, further separated by firm and region. Marginal cost of technology j at location l for firm i is therefore an affine function.

$$C^{i\prime}_{l,j}(q^i_{j,l,t}) = k^i_{l,j} + c^i_{l,j} q^i_{l,j,t}$$

These cost functions are derived by aggregating the generation of each firm by region and technology type. The five technology categories are coal, gas combined cycle (CCGT), conventional (steam) gas, gas combustion turbine (CT), and oil.

There are ten firms consisting of the nine largest fossil generation producers and a "fringe" firm derived from the aggregation of the generation from all remaining firms.

[14] This approach to modeling unit availability is similar to Wolfram (1999) and Bushnell, Mansur and Saravia (2008).

Table 2: Generation by Ownership and Fuel Type

Firm	Coal	CCGT	Gas St	Gas CT	Oil
BRKA	6104	629	235	319	0
CPN	0	4802	0	915	0
DYN	0	2120	2875	0	0
EIX	720	1373	0	237	0
LADWP	2117	1303	1929	282	0
PW	1741	1569	430	486	0
SALTRP	1802	1537	407	0	0
SEMPRA	0	2366	0	46	0
XCEL	2593	690	107	0	0
Others	14153	16338	12049	4840	629

The generation capacity of each of these firms is summarized by technology type in Table 2.

Emissions Rates

Emissions rates are based upon the fuel-efficiency (heat-rate) of a plant and the carbon intensity of the fuel burned by that plant. They are modeled as affine functions, with rates differentiated by firm, location, and technology. This yields a functional form of

$$e^i_{l,j}(q^i_{l,j,t}) = E^i_{l,j} + \epsilon^i_{l,j} q^i_{l,j}.$$

4.3 Transmission Network

Our regional markets are highly aggregated geographically. The region we model is the electricity market contained within the U.S. portion of the Western Electricity Coordinating Council (WECC). The WECC is the organization responsible for coordinating the planning investment, and general operating procedures of electricity networks in most states west of the Mississipi. The multiple sub-networks, or control areas, contained within this region are aggregated into the four "sub-regions." Between (and within) these regions are over 50 major transmission interfaces, or paths. Due to both computational and data considerations, we have aggregated this network into a simplified, 5 region network consisting primarily of the 4 major subregions.[15] Figure 1 illustrates the

[15]The final "node" in the network consists of the Intermountain power plant in Utah. This plant is connected to southern California by a high-capacity DC line, and is often considered to electrically be

areas covered by these regions. The states in white, plus California, constitute the US participants in the WCI.

Given aggregated level of the network, we model the relative impedance of each set of major pathways as roughly inverse to their voltage levels. The network connecting AZNM and the NWPP to CA is higher voltage (500 KV) than the predominantly 345 KV network connecting the other regions. For our purposes, we assume that these lower voltage paths yield roughly twice the impedance of the direct paths to CA.

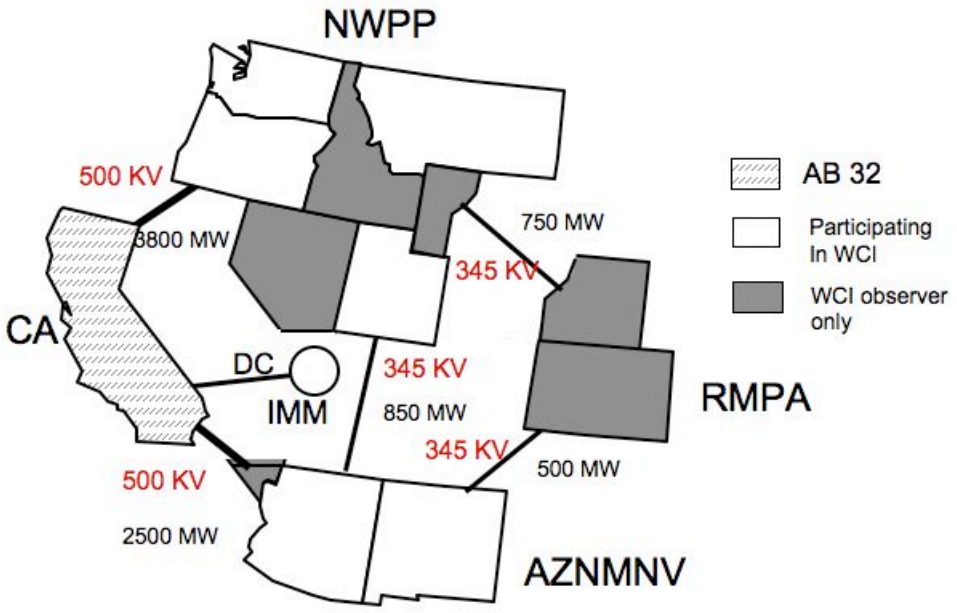

Figure 1: Western Regional Network and Cap-and-Trade Regions

There are sub-regions with both the NWPP and AZNM areas that would also not be subject to the currently organized WCI agreement. These include the states of Nevada and Idaho, as well as power plants located on tribal lands in the desert southwest. In

part of California. Because under some regulatory scenarios, it would not in fact be part of California for GHG purposes, it is represented as a separate location that connects directly to California.

each case these regions were considered to electrically be part of the region in which they were located, but for purposes of GHG regulation were treated as separate regions.

The flow capacity over the regional interfaces would ideally be based upon the amount of "slack" capacity remaining over these interfaces under actual market conditions. Unfortunately hourly data for these values are not currently available. Rather aggregate studies of the levels of congestion over the key paths in the western system are available. These typically characterize the usage of transmission paths by frequency of congestion or by data on the distribution of flows relative to capacity. There are not currently data on the flows over a specific path for a specific hour however.

For now we characterize the capacity over the key paths according to a single value based upon a conservative value of a 30% availability level. This is the amount of slack capacity that is left over those paths less than 30% of the hours in 2007, presumably during the higher usage periods in which those capacities are likely to be relevant.[16] In future versions, we hope to better match transmission capacity conditions to specific market conditions.

5 Results

Following the assumptions described above, we simulate the electricity production for the western electricity market under a variety of assumptions about the scope and design of cap-and-trade for CO2. For the geographic scope of the regulation, we first simulate operations under no-cap at all to establish a reference level for the other simulation results. Then we examine CO2 caps applied to California-only, to all (US) states participating in the WCI, and finally to all states (and tribal areas) in the western market. For each of these cap-and-trade scenarios, we assume that the cap is set at 85% of the CO2 emissions from the "no-cap" scenario. For all of the results in this section, we assume that permits are allocated exogenously and therefore do not effect the output decisions of firms. As described above, the simulation encompasses 8760 hours of actual market data that were aggregated into 80 representative hours, 20 for each season. These representative hourly results were then multiplied by the number of actual hours in each of the "bins" from which these hours were based upon. The results reported below are therefore annual totals, based upon 8760 hours of production.

Table 3 summarizes the aggregate annual CO2 emissions for each of the key regulatory regions. Results are reported for each of the simulated scenarios, as well as the actual

[16]The WECC periodically reports on the congestion level of its transmission network, and currently the only data available are the distribution of overall slack capacity on the key interfaces, rather than an hourly decomposition of the slack capacity.

(2007) aggregate emissions, as measured by CEMS, for each of these regions. First note that simulated emissions under the "no cap" scenario are about 6% lower than measured actual emissions. This difference is most pronounced in the California region. As illustrated in Figure 2, these differences are driven by the relative production of combined cycle (CCGT) to less efficient (CT and ST) gas plants. Production from less efficient plants is lower, and from CCGT plants higher, in our simulation than in actuality. This is most likely due to two factors. First, by aggregating actual hourly observations into representative market hours, we in effect truncate the peak demand levels of the system into a single level representing the average of a set of high demand hours. The operation of these less efficient plants is usually concentrated in these very high demand hours. Second, our simulation ignores inter-temporal operating constraints on plants, and CCGT plants are in fact less nimble than our simulation implicitly assumes them to be.

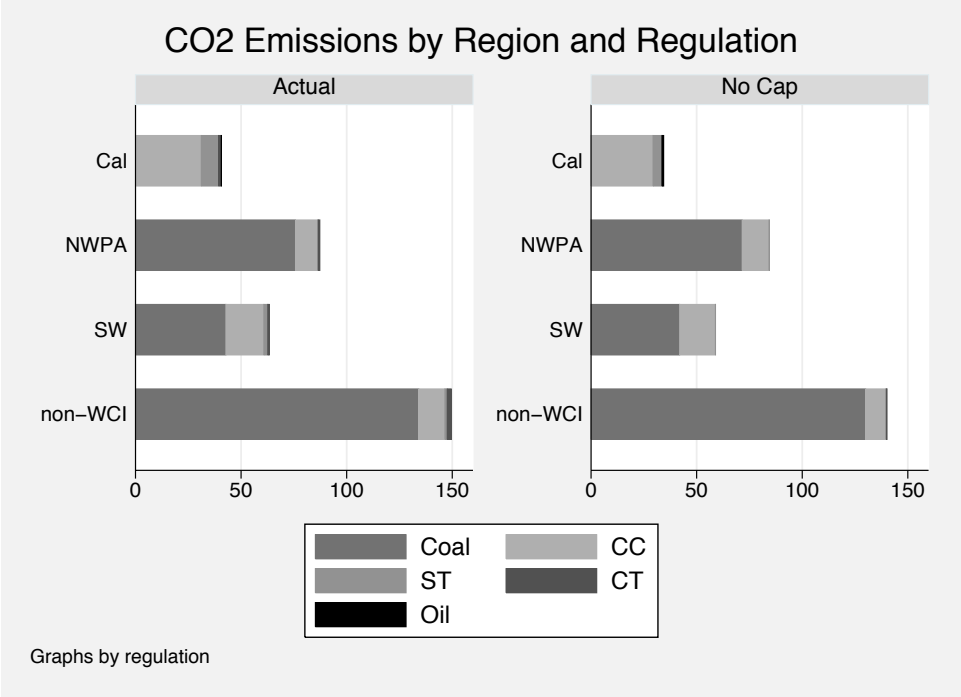

Figure 2: Actual vs. Baseline Emissions

Setting aside these differences for the moment, we turn our focus to the impacts of cap-and-trade regulations relative to our simulated no-cap case. As would be expected, a cap applied only to California, as originally envisioned under AB 32, would result in significant leakage. Although California emissions decline by 6 mmTons as required by the cap, aggregate west-wide emissions decline by less than one mmTon. Emissions prices

Table 3: Scope of Regulation: Emissions by Region (mmTons)

Regulation	CA	NWPP	SW	non-WCI	Total	Carbon Price
Actual (CEMS)	40.71	87.3	63.37	149.52	340.9	NA
No cap	34.4	84.43	58.95	140.17	317.95	NA
Cal only	28.36	85.52	61.55	141.96	317.39	11
WCI cap	30.89	72.01	48.17	148.82	299.89	39.15
WECC cap	35.43	73.74	50.59	131.45	291.21	42.9

are correspondingly low, at only 11 $/ton, due to the fact that compliance through leakage is a relatively inexpensive option. When the cap is applied to the currently configured WCI, leakage is greatly reduced, but still roughly 1/3 of the 30 mmTon reduction in WCI state emissions is picked up in increased non-WCI emissions. When the cap is applied to the entire market, permit prices rise to about $ 43 /ton. This can be interpreted as the value required to reach a true reduction of 30 mmTons over the entire region without any leakage. One implication of this comprehensive carbon cap is that California emissions *increase*. This is because the generation capacity inside California's borders is relatively clean, and a west-wide reduction in overall emissions is most easily accomplished by reducing output from coal generation in other states, and replacing it with gas output from California.

Table 4 summarizes the net injections into the west-wide network originating from each region. Recall that the "demand" modeled here is based upon actual production, rather than end-use demand, so it is the change in these figures that is relevant, rather than their absolute levels. Consistent with the emissions results, one can see the sizable swing in imports into California (around 1500 MW/h) under a California-only cap, as well as the large increase in net injections from the non-WCI region (about 2400 MW/h) under the WCI cap. Note again that California becomes a sizable net-exporter of power under the comprehensive west-wide cap.

The impact of these regulations on wholesale electricity prices in the various regions is summarized in table 5. Note that the regional breakdown in these columns is slightly different than in the previous tables. These are electricity market areas, rather than CO2 regulatory areas. The NWPP includes both capped and uncapped states, while the RMPA has no WCI states in its region. Prices rise substantially under the more comprehensive CO2 caps, even in regions not covered by the cap. This is due to the increased exports from these regions.

Table 4: Change from Actual Exports by Region (Average MWh)

Regulation	CA	NWPP	SW	non-WCI
No cap	122	551	70	-743
Cal only	-1405	862	785	-241
WCI cap	-468	-419	-793	1680
WECC cap	700	77	-20	-757

Table 5: Scope of Regulation: Electricity Prices by Region (Average $/MWh)

Regulation	Cal	NWPP	AZNMNV	RMPA
No cap	57.22	57.91	58.37	62.24
Cal only	59.88	60.81	61.43	64.26
WCI cap	74.78	74.93	75.03	73.74
WECC cap	78.54	79.55	80.22	84.88

5.1 Impacts of Allocation Policies

We now turn to the question of how the various policies for allocation of permits impacts prices and operations. Table 6 summarizes the emissions by regulatory region for the various permutations of a policy applied to the WCI, including the baseline "no-cap" case. In all cases, except the no-cap case, an identical emissions cap of roughly 150 mmTons, or 85% of the uncapped level, is applied to the WCI region. The row labeled 'WCI cap" applies to any allocation policy, such as auctioning or grandfathering, where allocations are exogenous to ongoing market outcomes. There were also two versions of allocations through updating that we considered. The row "WCI updating" refers to output-based updating. Under this policy, we assumed that 80%, or 120 mmTons, of the permits are allocated under the updating policies, with the remainder either allocated in some exogenous fashion or auctioned off. Similarly, in the "Fuel-based" updating scenario, we also assume that 80% of the permits were allocated, and the remainder auctioned. Under the Fuel-based updating scenario, we follow the CPUC's (CPUC, 2008) proposed allocation ratios. This proposal would allocate twice as much to coal generation as it would to gas generation. These ratios apply only to the fraction of total permits allocated, so that the net allocation received by a coal plant was equivalent

to 0.75 tons/MWh, while the allocation to gas plants would be 0.375 tons/MWh.[17] We established these allocation levels so that the total number of permits assigned under both the fuel-based and output-based allocation proposals was the same. This is truly "fuel-based" updating, with the distinction between updating being based upon fuel, rather than technology or explicit emissions rates. We therefore would not expect as extreme an impact from this allocation as that implied by equation (4). However, differentiation by fuel does capture a significant portion of the emissions rate differences between plants, so some significant differences from output-based updating would be expected.

As seen from Tables 6 and 7 the impacts of the allocation policies are indeed significant. Carbon emissions in uncapped "non-WCI" regions increase by roughly 9 mmTons (or 1/3 of the required reduction), under a WCI cap with no updating. When output-based updating is applied to firms within the WCI, this leakage of emissions is reduced to roughly 2 mmTons (or less than 1/10 of the required reductions). Also note that emissions within California rise substantially with the application of output-based updating. As can be seen from Figure 3, this is due to a large decrease in coal production. This is because output-based allocation favors gas generation relative to coal-generation, and California has no coal-based utility scale generation. The output-based updating therefore had a non-trivial impact on mitigation of leakage from the WCI region. When the updating approach is changed instead to be fuel-based, however, this mitigation of leakage is largely offset. Total emissions are only 2 mmTons lower than when no updating at all is applied.

The most striking impact of the updating policies is on the prices of the emissions permits. Permit prices rise from about $ 40/ton without updating to about $52/ton with output-based updating. As predicted, the fuel-based updating approach has an enormous impact on permit prices, raising them to just under $90/ton. Recall that this model reflects only the electricity sector, and therefore the distortions from these price impacts are contained within this industry and are largely offset by the updating policies that caused them. When one considers that this market will eventually include most major sources of CO2 emissions within the west, and be linked with other regions through trades with other CO2 markets, as well as offset programs, the potential distortions caused by such an inflationary impact on permit prices become a significant concern.

The results summarized in Table 7 tell a similar story, this time in terms of energy exports rather than emissions. The application of a CO2 cap on the WCI states results in a net increase of 2400 MW per hour in net exports from the non-WCI regions, which swing from net importers to net-exporters of power. When output based updating is applied, the WCI region again becomes a net importer. As with emissions, the application of

[17] There are very few oil plants in the sample, and they received allocations in equal ratios as those of gas plants.

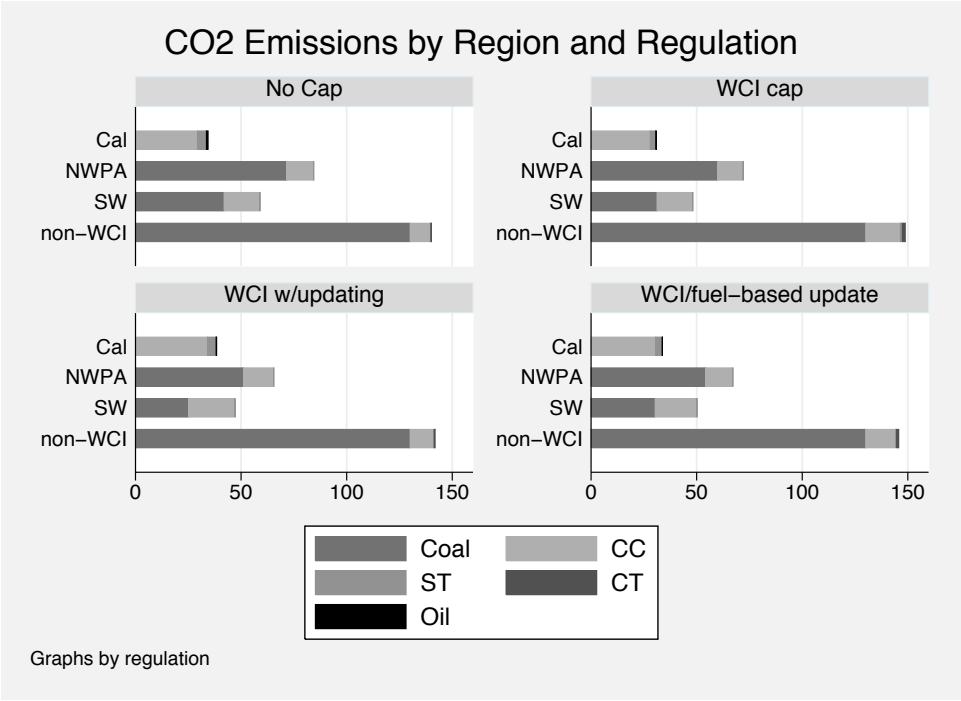

Figure 3: Impact of Updating Policies

Table 6: Effect of Updating: Emissions by Region (mmTons)

Regulation	CA	NWPP	SW	non-WCI	Total	Carbon Price
No cap	34.4	84.4	59.0	140.2	318.0	NA
WCI cap	30.9	72.0	48.2	148.8	299.9	39.15
WCI updating	38.4	65.5	47.2	141.9	293.0	51.55
Fuel-based	33.8	67.1	50.2	145.7	296.8	89.85

Table 7: Change from Actual Exports by Region (Average MWh)

Regulation	CA	NWPP	SW	non-WCI
No cap	122	551	70	-743
WCI cap	-468	-419	-793	1680
WCI updating	1293	-891	-155	-247
Fuel-based	220	-851	-151	781

fuel-based updating reverses the effects of output-based updating, raising net-exports from the non-WCI regions by an average of about 1000 MWh.

Table 8 summarizes the price impacts of the cap, and of the updating policies. The imposition of the cap (again requiring a 15% reduction from the status quo) raises wholesale average prices from around $ 58/MWh to around $ 75 /MWh. The almost $ 20 /MWh increase is consistent with the facts that CO2 costs are about $ 40 /ton in this scenario, and that gas plants, which emit roughly 1/2 ton/MWh are almost always the marginal, price-setting technology. When output-based updating is applied, most of this impact on the market-clearing price is eliminated, as prices "fall" from 75 to under $ 61 /MWh. Yet again the fuel-based updating policy reverses the impacts of the output-based updating. Prices under fuel-based updating average just around $ 68 /MWh.

Profit Impacts of Allocation Policies

We now examine how the allocation policies impact the emissions costs and operating profits of firms. Recall that the the updating schemes are largely motivated by a desire to offset the cost impacts to high emitting firms and limit any perceived windfalls to low emissions firms. Table 9 summarizes the *net* costs of emissions regulations on firms. The net emissions costs is defined here as the costs of emissions permits required by the firm under the cap-and-trade regulation less the value of the emissions permits allocated under

Table 8: Effect of Updating: Electricity Prices by Region (Average $/MWh)

Regulation	Cal	NWPP	AZNMNV	RMPA
No cap	57.22	57.91	58.37	62.24
WCI cap	74.78	74.93	75.03	73.74
WCI updating	60.36	61.14	61.67	64.68
Fuel-based	67.41	68.05	68.48	68.81

the various allocation approaches. As before the fuel-based and output-based contingent allocation schemes assume that 80% of total permits (about 120 mmTons) are allocated to producers. The last column in this table considers an exogenous grandfathered allocation of the same quantity, based upon the emissions under the "no-cap" scenario, which here serves as the proxy for historic emissions.

Note that when a firm receives more in allocation than it must surrender due to its actual emissions, the net costs can be negative. This is in fact the case for largely gas-based producers, such as Calpine (CPN) and Dynegy (DYN) under the output-based allocation approach. In contrast, coal-heavy producers such as PacifCorp (owned by BRKA) and Arizona Public Service (owned by PW) have significant emissions costs under any scenario. Despite the skewing of permit allocation in favor of coal producers under the fuel-based allocation approach, net emissions costs are actually higher for these firms under this approach. The reason is that the higher equilibrium permit prices more than offset the increased allocation quantities these firms receive under the fuel-based approach. These firms are clearly better off under grandfathering, which also skews allocations their way without impacting permit prices.

The picture becomes more complex when one considers the net effects of the allocation scheme on product (electricity) prices as well as emissions costs. Table 10 summarizes the operating profits of the firms under the assumption that each firm were selling all its output at market-clearing prices, rather than at a a regulated cost-based rate. It is important to recognize that several of the firms in this table are in fact either regulated or government-owned.[18] Therefore these results are more a qualitative representation of the general net revenue and cost effects than a literal assessment of each firms bottom line impact. The profits are therefore defined as the total revenues (assuming market-based sales) less the net emissions costs from table 9 as well as the total production costs (fuel and operating expenses).

[18]The results also reflect only revenue of sales from thermal generation sources. Firms with substantial nuclear and hydro generation would benefit disproportionately more from a higher permit price.

Table 9: Net Emissions Costs by Firm (Millions $)

Firm	No Allocation	Fuel-based	Output-based	Grand-fathering
BRKA	892.3	582.2	550.5	321.9
CPN	373.9	57.8	-93.6	137.5
DYN	250.7	-0.2	-78.9	104.8
EIX	153.5	8.3	-43.8	49.2
LADWP	639.7	341.1	325.9	244.3
PW	281.8	154.7	92.0	108.2
SALTRP	339.7	184.5	129.1	124.3
SEMPRA	193.9	-0.7	-68.1	64.2
XCEL	0.0	0.0	0.0	0.0
Others	2789.2	1444.1	892.0	1364.3

The results in table 10 highlight the complex interaction between the allocation policy, permit prices, and electricity prices. Recall that, based on allocation and emissions costs alone, gas-intensive firms appeared to benefit from output-based allocation. However, output-based allocation also greatly limited the pass-through of carbon costs to electricity prices. This results in reduced revenues for all firms. While gas intensive firms such as CPN still prefer contingent allocation to auctioning, they actually do better under fuel-based allocation than output-based. This is despite the fact that fuel-based allocation was intended to limit their perceived windfall benefits from allocation. However, since electricity prices are higher under fuel-based, the increased revenue from this scenario more than offsets the reduced allocation in permits relative to output-based updating for these firms.

For high carbon firms, such as the coal-heavy BRKA and PW, the contingent allocation approaches look even worse. The combination of higher emissions costs due to the inflated permit prices and lower electricity revenues reduce profits under these allocation schemes to below those seen with *no* allocation at all. While fuel-based is in fact preferred to output-based allocation by such firms, neither is particularly appealing. As before, grandfathering is the clear winner from the perspective of such firms.

Table 10: Net Profits Assuming Sales at Market-based Prices (Millions $)

Firm	No Allocation	Fuel-based	Output-based	Grand-fathering	No Cap
BRKA	1681.7	1613.8	1318.7	2329.4	1804.5
CPN	352.9	481.8	450.0	625.0	280.0
DYN	210.4	329.2	301.2	380.4	173.9
EIX	386.2	405.0	364.5	492.9	282.7
LADWP	421.8	542.0	386.9	867.5	649.7
PW	508.5	499.1	407.0	711.4	481.6
SALTRP	505.6	518.9	421.9	742.2	522.0
SEMPRA	233.1	296.3	257.9	370.4	148.6
XCEL	1115.9	954.2	851.2	1115.9	800.2
Others	4047.7	3953.7	3293.8	6599.1	4528.2

6 Conclusions

While the establishment of cap-and-trade regulation, as opposed to command-and-control regulations, is largely motivated by a desire to provide incentives for the efficient mitigation of pollution, many other policy goals are often at play. These goals include mitigating the cost impacts of climate regulation on both consumers and on the firms to which the regulation will be applied. As climate policy advances in the United States, these ancillary goals are playing a prominent role in the design of emissions markets. The allocation of emissions permits is seen as a critical tool for achieving these policy goals.

We have studied these issues in the context of the proposed California and Western Climate Initiative cap and trade programs, by focusing on the electricity market that spans these regions. In this context, the mitigation of regulatory circumvention through leakage is an additional concern. Indeed, we find that even with the expansion of the western cap to 7 states, leakage could still be significant. Here the proposals for the contingent allocation of permits, either linked to the output or the fuel input of a plant, can have a significant impact. Output-based allocation largely achieves the stated goals of policy-makers by effectively mitigating leakage and also electricity prices. However, when the allocation is linked to the fuel of the resource, rather than strictly to its output, most market outcomes closely resemble those seen under an exogenous allocation scheme such as auctioning or grandfathering. Permit prices, however, rise considerably to levels more than double that seen under an exogenous allocation. Although the primary goals of input-based allocation are to insulate high-carbon firms from cost shocks and prevent

"windfalls" to low-carbon producers, these goals are largely unachieved even when 80% of the permits are allocated.

While we believe these results have important practical implications for the design of the western electricity market, we need to note many caveats that limit the interpretation of these results as a forecast of WCI cap-and-trade market results. First, we limit our analysis to the electricity industry, which will dominate the WCI market for its first phase, but will then be combined with several other sectors, including transportation fuels. Second, we model only traditional "source-based" market implementations, where the WCI is pursuing a hybrid design that will combine the source-based regulation of plants located within the WCI with attempts to account for the carbon content of imports into that region.

It is important to examine the specific implications of this "first-deliverer" approach, which will almost certainly limit leakage below the levels we find here. However, we believe that the source based model studied here provides an accurate picture of the implications of allocation policies under either a first-deliverer or source-based design. There are also some factors that may limit the leakage-mitigating impact of the first-deliverer approach. Although regulators can attempt to apply carbon levels to imports, these attempts can be bypassed by a reshuffling of transactions. It is suspected that, under such a system, relatively clean production will be identified as the 'source' of imports into the WCI. Second, and more importantly, the WCI is often a net exporter of power. The first-deliverer design cannot deal with leakage when it is driven by end-use demand located outside the regulators jurisdiction.

When one considers the implications of an integration of the electricity sector with other sectors, the aspect of fuel-based updating that is most problematic is the greatly increased permit price. The concern is that the upward price pressure from the sector receiving updates will lead the mitigation to be concentrated in other sectors that do not. In those sectors, marginal emissions costs will in fact be much higher than in the sectors receiving updated allocations. For example, one would expect the utilization of unconventional "offsets," such credits for retrofitting inefficient facilities, to greatly increase as the result of the inflationary pressure on permit prices caused by updating.

More generally, as discussions concerning national cap-and-trade regime for CO2 advance, these results, consistent with previous work, highlight the potential distortions that updating can introduce into a cap-and-trade market. Just as important from the point of view of policy-makers, careful attention must be paid to the equilibrium effects of any allocation proposal. The "benefits" from more complex allocation schemes may be far less than policy-makers expect, while the negative impacts remain a serious concern.

References

[1] M. Ahman, D. Burtraw, J. Kruger, and L. Zetterberg. A ten-year rule to guide the allocation of eu emission allowances. *Energy Policy*, 35(3):17181730, 2006.

[2] M. Ahman and K. Holmgren. New entrant allocation in the nordic energy sectors: Incentives and options in the eu ets. *Climate Policy*, 6:423–440, 2006.

[3] C. Bohringer and A. Lange. Economic implications of alternative allocation schemes for emission allowances. *The Scandinavian Journal of Economics*, 107(3):563–581, 2005.

[4] C. Bohringer and A. Lange. On the design of optimal grandfathering schemes for emissions allowances. *European Economic Review*, 49:2041–2055, 2005.

[5] D. Burtraw. Regulating co2 in electricity markets: Sources or consumers? *Climate Policy*, 8:588–606, 2008.

[6] D. Burtraw, K. Palmer, and D. Kahn. Allocation of co2 emissions allowances in the regional greenhouse gas cap-and-trade program. Discussion Papers dp-05-25, Resources For the Future, June 2005.

[7] J. Bushnell, E.T. Mansur, and C. Saravia. Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets. *American Economic Review*, 98:237–66, 2008.

[8] J. Bushnell, C. Peterman, and C. Wolfram. Local Solutions to Global Problems: Climate Change Policies and Regulatory Jurisdiction. *Review of Environmental Economics and Policy*, 2008.

[9] California Public Utility Commission. Rulemaking 06-04-09: Final opinion on greenhouse gas regulatory strategies, October 2008. http://docs.cpuc.ca.gov/.

[10] Y. Chen. Regional solution to a global problem: A short-run case analysis of co2 leakage and nox and so2 emissions spillovers. Discussion paper, UC Merced, 2008.

[11] Y. Chen, P. Liu, and B.F. Hobbs. Economic and emissions implications of load-based, source-based, and first-seller emissions programs under california ab32. Discussion paper, The Johns Hopkins University, 2008.

[12] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Academic Press, Boston, 1992.

[13] Damien Demailly and Philippe Quirion. Co2 abatement, competitiveness and leakage in the european cement industry under the eu ets: grandfathering versus output-based allocation. *Climate Policy*, 6(1):93–113, 2006.

[14] S.P. Dirkse and M.C. Ferris. The path solver: a nommonotone stabilization scheme for mixed complementarity problems. *Optimization Methods and Software*, 5(2):123–156, 1995.

[15] A.D. Ellerman, P.L. Joskow, R. Schmalensee, J.P. Montero, and E. Bailey. *Markets for clean air*. Cambridge University Press New York, NY, 2000.

[16] Carolyn Fischer. Output-based allocation of environmental policy revenues and imperfect competition. Discussion Paper 02-60, Resources For The Future, January 2003.

[17] Carolyn Fischer and Alan K. Fox. Output-based allocation of emissions permits for mitigating tax and trade interactions. *Land Economics*, 83(4):575–599, 2007.

[18] M. Fowlie. Incomplete environmental regulation, imperfect competition, and emissions leakage. Working Paper 14421, National Bureau of Economic Research, October 2008.

[19] D. Fullerton and G. Heutel. The general equilibrium incidence of environmental mandates. Working Paper 13645, National Bureau of Economic Research, November 2007.

[20] M. Grubb and K. Neuhoff. Allocation and competitiveness in the eu emissions trading scheme: Policy overview. *Climate Policy*, 6(1):7–30, 2006.

[21] S. Holland, J. Hughes, and C. Knittel. Greenhouse gas reductions under low carbon fuel standards? *The American Economic Journal: Economic Policy*, 1(1):106–146, 2008.

[22] Western Climate Initiative. Design recommendations for the wci regional cap-and-trade program. Technical report, September 2008.

[23] J. Jensen and T.N. Rasmussen. Allocation of co2 emissions permits: A general equilibrium analysis of policy instruments. *Journal of Environmental Economics and Management*, 40:111–136, 2000.

[24] James Kanter and Jad Mouawad. Money and lobbyists hurt european efforts to curb gases. *New York Times*, December 10 2008.

[25] Market Advisory Committee. Market Advisory Committee to the California Air Resources Board: Recommendation for designing a greenhouse gas cap-and-trade system for California, June 2007.

[26] C. Metzler, B. Hobbs, and J.-S. Pang. Nash-Cournot equilibria in power markets on a linearized DC network with arbitrage: Formulations and properties. *Networks & Spatial Economics*, 3(2):123 – 150, 2003.

[27] Karsten Neuhoff, Kim Keats Martinez, and Misato Sato. Allocation, incentives and distortions: the impact of eu ets emissions allowance allocations to the electricity sector. *Climate Policy*, 6(1):73–91, 2006.

[28] J. Shuliken, B. F. Hobbs, and J. S. Pang. Long-run equilibrium modeling of emissions allocation systems in electric power markets. *Operations Research*, forthcoming., 2009.

[29] Thomas Sterner and Adrian Muller. Output and abatement effects of allocation readjustment in permit trade. *Climatic Change*, 86(1-2):33–49, January 2008.