California's cap-and-trade program and emission leakage in the electricity sector: an empirical analysis

Chiara Lo Prete[*] Ashish Tyagi[†] Cody Hohl[‡]

January 3, 2019

Abstract

We conduct the first econometric analysis of leakage in the electricity sector from California's cap-and-trade program. The paper presents three sets of empirical results that support the hypothesis of leakage. First, we measure the policy impact on baseload power plant operations in the Western Interconnection applying a differences-in-differences estimator to a novel dataset at the monthly level from 2009 to 2016. Second, we preprocess the data to improve balance between treated and control plants by matching on hour-of-day specific variables, and explore treatment effect heterogeneity across daytime and nighttime hours using daily measures of plant utilization. Third, we test for leakage from the cap-and-trade program by examining the relationship between emission allowance prices and scheduled power imports into California. Results suggest a policy-induced reduction in natural gas combined cycle generation in California and an increase in coal-fired generation in the Western U.S., corresponding to a leakage rate of about 70%.

1 Introduction

California has been at the forefront of U.S. environmental policies for years. The Global Warming Solutions Act of 2006 (also known as Assembly Bill 32 or AB 32) set the state's target to reduce greenhouse gas (GHG) emissions to 1990 levels by 2020 [1]. In September 2016, California passed Senate Bill 32 (SB 32), which limited emissions to 40% below 1990 levels by 2030 [2]. Further, Executive Order S-3-05 set a GHG emission reduction target of 80% below 1990 levels by 2050 [3]. In order to achieve these ambitious goals, the

[*]John and Willie Leone Family Department of Energy and Mineral Engineering, The Pennsylvania State University, University Park, PA, USA, chiaraloprete@psu.edu.

[†]Frankfurt School of Finance & Management, Frankfurt am Main, Germany, a.tyagi@fs.de.

[‡]John and Willie Leone Family Department of Energy and Mineral Engineering, The Pennsylvania State University, University Park, PA, USA, cbh21@psu.edu.

state relies on a suite of complementary policies, including a multi-sector cap-and-trade program that covers about 80% of the state's emissions from the electricity sector, large industrial facilities, and fuel distribution sector, and is expected to drive roughly 22% of emission reductions by 2020 [4].

A central issue in the implementation of cap-and-trade programs is represented by the choice of the point of regulation. For example, the Regional Greenhouse Gas Initiative (RGGI), an emission trading system for CO_2 emissions from electricity generation in U.S. Northeastern and mid-Atlantic states, adopted a production-based (or source-based) approach where the point of regulation is at the generator level. In contrast, given its reliance on imports to satisfy electricity consumption,[1] California opted for a first deliverer approach, whereby entities that own electricity at the first point of delivery in the state represent the point of regulation: in-state generators must monitor and report their emissions following a source-based paradigm, while electricity importers are responsible for emissions associated with in-state sales.

The introduction of a border adjustment mechanism for the electricity sector was intended to mitigate concerns of leakage, defined as the shift in production and associated emissions from the region where climate regulations apply to surrounding unregulated jurisdictions [6, 7, 8]. However, energy modeling studies have concluded that the possibility of reshuffling contracts may enable substantial leakage under the AB 32 cap-and-trade system [6, 7, 9, 10, 11]. Under resource shuffling, electricity contracts are rearranged so that production from low emission sources serving out-of-state load is directed to California, while production from higher emission sources is assigned to serve out-of-state load [12]. This would result in apparent emission reductions due to changes in the composition of imports to California, although emissions in exporting regions are unchanged or even increase. It is worth noting that, in recent years, the decrease in GHG emissions from the electric power sector in California has been attributed primarily to measured reductions in emissions from imports [13]. This underscores the importance of assessing whether leakage has occurred in practice and considering potential policy modifications to mitigate its impacts.

We contribute to the literature by conducting the first econometric analysis of leakage in the electricity sector from California's cap-and-trade program. While earlier studies were prospective (*ex ante*) and employed numerical models, our study is retrospective (*ex post*) and employs statistical analysis of historical data. The paper presents three sets of empirical results. First, we measure the impact of California's carbon policy on baseload power plant operations in the Western Interconnection applying a differences-in-differences estimator to a novel dataset at the monthly level from 2009 to 2016. After controlling for key determinants of plant capacity factors like fuel costs, electric demand, nuclear and renewable generation, temperature and

[1] California imports about a third of its total electricity consumption from out of state [5].

precipitation levels, the estimated average treatment effects suggest a policy-induced reduction in NGCC generation by 14% in California and increase in coal-fired generation by about 4% in the Western U.S., corresponding to a leakage rate of about 70%. Results are robust to the choice of leaker and control groups, clustering methods and sample definition. In our second set of empirical results, we preprocess the data to improve balance between treated and control groups by matching plants on coarsened hour-of-day variables, and carry out parametric inferences using daily measures of plant utilization. This approach changes the estimand to a local average treatment effect for the plants that were matched, and allows us to explore heterogeneity across daytime and nighttime periods. Results from the matched sub-samples are broadly consistent with those from the full sample, and robust to the inclusion of matching variables and choice of cut points. In our final set of analyses, we test for leakage from the policy by examining the relationship between the AB 32 allowance price and scheduled power imports into the California Independent System Operator (CAISO), which centrally dispatches generation and coordinates the movement of wholesale electricity in much of California and part of Nevada. Specifically, we estimate a model of daily scheduled power flows into CAISO, and test for leakage based on the statistical significance of the AB 32 allowance price as one of the explanatory variables. The analysis of daily scheduled flows across major CAISO interfaces further supports the hypothesis of leakage.

The remainder of the paper is organized as follows. Section 2 reviews the literature on emission leakage and Section 3 provides background on California's cap-and-trade program. Section 4 describes the data, while Section 5 outlines the research design and econometric approach. Section 6 presents the empirical results, and Section 7 offers concluding remarks.

2 Literature review

The potential for emission leakage in the electricity sector under regional climate policies has been analyzed using numerical models. A first strand of the literature employs simulation-based models of the electricity sector. [7, 9, 14, 10] explore leakage in the context of California's *proposed* cap-and-trade program for GHG emissions (i.e., before regulations were finalized). [7] uses partial equilibrium analysis to determine the extent of leakage potential under incomplete, market-based regulation of CO_2 emissions in California's electricity sector. Results indicate that emission leakage is greater when emission rates per unit of production are high, demand is more elastic, and the industry is more competitive. The theoretical framework is used to investigate related welfare effects in the electricity sector under a range of assumptions regarding market competitiveness. [9] formulate a market equilibrium model to compare source-based, load-based and first deliverer approaches for cap-and-trade regulation in California, and examine related economic and emission implications on the electricity market. They find that leakage is substantial (85%)[2] and largely due to reshuffling: emission reductions due to changes in the composition of electricity imports to California are illusory, because emissions in the rest of the Western Interconnection increase under regulation. [14] and [10] examine the impacts of alternate cap-and-trade designs applying to Western U.S. states and California, respectively, and conclude that a first deliverer approach in California is vulnerable to leakage due to laundering and reshuffling of import resources. In contrast to the studies cited above, [11] consider California's *actual* cap-and-trade program. The authors simulate distributions of emission allowance prices assuming that price containment mechanisms may be binding or market participants engage in withholding strategies. With respect to RGGI, [15] develop a model of the Eastern U.S. and Canada to analyze the effects of a CO_2 price in RGGI on emissions, electricity prices, and generator entry and exit decisions, finding that leakage represents a likely market outcome.

Computable general equilibrium (CGE) models have also been used to examine the impacts of regional climate policies [16]. CGE models can account for several potential leakage channels, but are sensitive to assumptions about the parameters. One such study is [17], that develops a multi-state CGE model of the U.S. economy. In the context of RGGI, the authors estimate that, if the program's cap was fully binding, power imports to New York (a RGGI state) from Pennsylvania (a non-RGGI state) could result in emission leakage rates of more than 50%. [18] develop a modified version of the GTAP-E model to assess the economic and carbon emission effects of alternative trade policy measures aimed at reducing carbon leakage. [19] conduct a general equilibrium assessment of leakage from sub-national climate policies, using California's cap-and-trade program as an example. When imported electricity is included in the cap and provisions to prevent reshuffling are enforced, the estimated leakage rate is only 9%.

Empirical analyses of leakage are less common in the literature. [20], [21] and [22] examine leakage in the context of the Kyoto Protocol. In particular, [22] develop a gravity model to calculate the carbon content of bilateral trade flows for forty countries between 1995 and 2007, and find that binding commitments under the Kyoto Protocol led to emission leakage to noncommitted countries. With respect to RGGI, [23] analyze the relation between CO_2 permit prices and transmission power flows on seven high-voltage interties between New York and Pennsylvania between 2008 and 2010. Higher net flows from Pennsylvania to New

[2]This percentage (also referred to as "leakage rate") is given by one minus the ratio of aggregate emission decrease in the regulated and unregulated regions, relative to a baseline in which no emission cap applies, over emission decrease in the regulated region, relative to the no cap scenario and including emissions associated with electricity imports.

York associated with a higher RGGI allowance price would indicate leakage. The authors do not find a significant impact of RGGI permit prices on PA-NY transmission flows, but prices may have been too low in the early years of the program to affect leakage. [24] conduct two complementary analyses of RGGI-induced leakage. First, they use plant-level data to assess operational impacts of the carbon policy. They estimate that RGGI induced a reduction in coal-fired generation in the regulated region and an increase in cleaner NGCC generation in the unregulated region, leading to an aggregate emission reduction across regions. Their second analysis examines changes in electricity transmission flows, finding that power imports to New York from outside the RGGI region increased substantially since the policy was implemented: this supports the generation substitution pattern identified previously.

Finally, a growing body of research in economics assesses the potential for leakage risk across sectors (e.g., [25]), and explores how environmental regulation affects trade flows and the location choice of firms in the long run ("pollution haven" effect) [26, 27, 28, 29, 30, 31].

3 Policy background

California's cap-and-trade program regulates GHG emissions from large industrial facilities, electricity generators and importers, and transportation fuel suppliers. Covered entities emit at least 25,000 metric tons of CO_2e per year and are responsible for about 80% of the state's GHG emissions [32]. The first phase of compliance for the program began on January 1, 2013. The 2013 emission cap was set at approximately 98% of forecast 2012 emissions, with an annual decline of 2% in 2014 and 3% from 2015 through 2020. In July 2017, the scheme was extended through 2030 with bipartisan support: emission reductions in the carbon market are expected to deliver about 40% of the state's total mitigation efforts [32].

The California Air Resources Board (CARB) is responsible for implementing AB 32 and designed the cap-and-trade system. CARB issues annual emission allowances equal to the cap, and each allowance represents a permit to emit one ton of carbon dioxide equivalent. Entities must monitor and annually report their emissions, and return an amount of allowances equivalent to their GHG emissions each year. Capped sources that keep emissions below the allowance amount can sell excess permits on the market, while sources that cannot cover total emissions may take measures to reduce pollution and/or buy allowances on the market.[3] Emission allowances are distributed through a mix of free allocation and quarterly auctions, transitioning over time to greater auctioning of allowances: in 2016, 46% of allowances were auctioned and 50% were

[3] Covered entities may also use carbon offsets (e.g., GHG emission reduction projects undertaken by entities not subject to the carbon policy) to cover up to 8% of their emissions.

given away for free [33].[4] CARB allows banking and borrowing of allowances, and the risk of unexpected price changes and excess volatility is mitigated through the use of a price collar; secondary market allowance prices have generally hovered at or near the auction price floor from market launch to 2016 [34].

Two provisions in the current design of AB 32 cap-and-trade system are intended to mitigate concerns about leakage. First, free emission allowances are allocated to energy-intensive, trade-exposed industries as an incentive to keep production in California. Electric distribution utilities are granted free allowances to ensure that their ratepayers do not experience sudden increases in electricity prices as a result of emission costs, although the share of free permits declines over time [35]. In particular, investor-owned utilities (IOUs) are required to auction their allowances and credit the resulting revenues to their electricity customers [33]. Further, electricity generators operating under long-term contracts are eligible to apply for free allowances for emissions related to contracted power [36].

Second, the cap-and-trade program features a first deliverer approach, whereby both in-state electricity generators and electric utilities that import power from out-of-state are subject to the carbon policy. In-state generators must monitor and report their emissions following a source-based paradigm, while electricity importers must acquire emission allowances (and possibly offsets) equal to measured or estimated emissions of generation resources supplying their power imports. Since energy entering the grid flows over the path of least resistance (rather than directly from an injection point to a withdrawal point), the CO_2 intensity of electricity imported in California from the rest of the Western Interconnection cannot be determined unambiguously.[5] To address the issue, CARB classifies imports as specified or unspecified source power.

Specified sources include generation resources owned by or under long-term contract to California's load serving entities, as well as generation resources owned by non-California entities that are approved and registered by CARB [37]. First deliverers may claim facility-specific emission factors for power imports from out-of-state generation resources that are owned or under long-term contract. Further, CARB has developed the designation of Asset-Controlling Suppliers for out-of-state electric power entities that operate interconnected generating facilities. Once approved and registered by CARB, Asset-Controlling Suppliers are assigned a system emission factor for wholesale electricity procured from their systems and imported into California. For example, specified source power from Bonneville Power Administration (BPA) and Powerex (a subsidiary of BC Hydro) must be reported using CARB-approved emission factors reflecting the hydro-

[4]The remaining 4% were made available at predetermined prices to reduce price volatility.

[5]California is part of the Western Interconnection, a synchronous electric grid that encompasses all or parts of 14 Western states in the U.S., the Canadian provinces of Alberta and British Columbia, and Northern Baja California in Mexico. Since reliability within the area is overseen by the Western Electric Coordinating Council, this synchronous grid is commonly referred to as WECC. Figure 1 presents the U.S. portion of WECC.

dominant resource portfolio of these systems [38]. Specified sources mainly consist of coal, natural gas and nuclear power from the Pacific Southwest, and of hydro and wind power from the Pacific Northwest [5].[6]

In contrast, unspecified source power (representing about 40% of total imports, as of 2016) corresponds to wholesale market purchases from power plants that do not have a contract with a California utility and have not gone through the CARB process to become specified. Since in this case the generation source is unknown, unspecified sources are assigned a default emission factor of 0.428 metric ton CO_2/MWh, which was set by CARB based on the generation technology expected to be at the margin in WECC (i.e., a fairly clean natural gas plant) [10]. According to the California Energy Commission, much of the Pacific Northwest spot market purchases are served by surplus hydro and gas-fired plants, while Southwest spot market purchases generally come from coal and natural gas combined cycles [5]. The presence of a default emission factor creates an incentive for electricity importers to not report the emission content of out-of-state higher-emitting generation resources, in order to attain the lower default emission factor ("laundering"). This has been identified as one of the primary types of resource shuffling [39], defined by CARB as "any plan, scheme, or artifice undertaken by a First Deliverer of Electricity to substitute electricity deliveries from sources with relatively lower emissions for electricity deliveries from sources with relatively higher emissions to reduce its emissions compliance obligation" (Cal. Code Regs., Title 17, Article 5, § 95802(a) [40]). As discussed in Section 1, contract shuffling would lead to apparent emission reductions due to changes in the composition of imports to California, although emissions in the exporting regions are unchanged or even increase. As a result, contract shuffling creates potentially severe leakage risks for the electricity sector in California. In response to these concerns, CARB released a guidance document listing a number of "safe harbor" exceptions to the regulatory ban on resource shuffling (i.e., transactions deemed not to be resource shuffling) (Cal. Code Regs., Title 17, Article 5, § 95852(b)(2) [40]). This approach has been controversial because it is difficult to identify all potential violations *ex ante* [10]. Further, allowance prices hovering near the auction floor have been interpreted as evidence that contract shuffling is taking place, enabling regulated entities to avoid a significant part of their carbon liability and reducing demand for allowances [11, 34].

4 Data

We use a novel panel dataset built from publicly available sources including the U.S. Department of Energy's Energy Information Administration (EIA), the U.S. Environmental Protection Agency (EPA), the Federal

[6] According to the California Energy Commission, the Pacific Northwest includes Alberta, British Columbia, Idaho, Montana, Oregon, South Dakota, Washington and Wyoming. The Pacific Southwest includes Arizona, Baja California, Colorado, Mexico, Nevada, New Mexico, Texas and Utah [5].

Energy Regulatory Commission (FERC) and the California Independent System Operator. The period of our study spans January 2009 through December 2016, including four years before and four years after the treatment date (January 2013, when compliance obligations began).

4.1 EIA data

U.S. electric generating facilities with more than one MW of capacity are required to complete an annual survey to report plant characteristics. Form EIA-860 collects information on the status of existing plants in the U.S., while EIA-923 gathers information on plant operations. Relying on these surveys, we assemble a dataset for power plants within the U.S. portion of four NERC regions (WECC, MRO, SPP, TRE) from 2009 to 2016 (Figure 1). A plant consists of at least one, but typically several, generating units, which may be added to or retired from service over its lifetime. Although energy output, operating capacity and fuel input are available at the unit level, we aggregate this information for units of the same technology within a plant because our analysis relies on capacity factors and heat rates. For natural gas combined cycle plants, in particular, energy output is reported separately for the steam and combustion parts of the plants, but both are needed to calculate capacity factors and heat rates accurately. The advantage of EIA data is that its coverage is comprehensive, including not only large thermal plants, but also nuclear, hydro and renewable facilities. Plant-level characteristics reported at the annual level include primary fuel type, operating capacity, month and year when each unit was in service, type and number of emission abatement controls, EIA regulatory status,[7] NERC region and subregion, balancing authority and planning area.[8] In addition, the EIA provides monthly plant operating statistics like energy output (measured by megawatt-hours or MWh of net electricity generation),[9] consumption and heat content by fuel type, and cost of fuel delivered to the plant. We rely on EIA Form 860 for primary fuel type and operating capacity [42], and EIA Form 923 for other plant characteristics [43]. Plants with operating capacity below 25 MW are excluded for

[7]For the purpose of EIA's data collection efforts, regulated entities include investor-owned electric utilities that are subject to rate regulation, municipal utilities, federal and state power authorities, and rural electric cooperatives. Facilities that qualify as cogenerators, small power producers under the Public Utility Regulatory Power Act (PURPA) and other nonutility generators (including independent power producers) are non-regulated plants.

[8]Each power plant falls under the operational control of a balancing authority, which is responsible for dispatching generation units and maintaining consumption-interchange-generation balance within a region of the electric grid [41]. Balancing authority areas and electric utilities with a planning area annual peak demand greater than 200 MW must file FERC Form 714. Electric utilities charged with carrying out resource planning and demand forecasts for a planning area ("planning authorities") are required to report actual hourly demand in their planning area in Part 3 of FERC Form 714. The definition of balancing authority and planning authority is similar, and the footprint of most planning authorities in WECC coincides with that of balancing authorities. WECC balancing authorities in the U.S. are presented in Figure 2. Summary statistics by balancing authority and fuel type are provided in Tables A1 and A2 in the Appendix.

[9]Net generation excludes power consumption for plant operations.

consistency with CEMS data (Section 4.2).[10]

Plant fuel costs are not publicly available for non-regulated plants and plants with nameplate capacity below 50 MW. In these instances, we use state average costs of fossil fuels for electricity generation provided by the EIA [45]. If state average coal costs are also not available, we impute these costs assuming the same growth rate of Rocky Mountain Colorado Rail prices (Section 4.4). Fuel costs are used to construct monthly ratios for assessing power plant competitiveness. For coal plants, the coal-to-gas cost ratio divides plant-specific variable cost of generation by state average variable cost of natural gas for power generation. Similarly, for natural gas plants the gas-to-coal ratio divides plant-specific variable cost of generation by state average variable cost of coal for power generation. After January 2013, variable costs for California include emission allowance prices (Section 4.4).

4.2 CEMS data

To complement monthly data from the EIA, we assemble a database of hourly gross electricity generation, heat input and CO_2 emissions for NGCC and coal-fired plants from the EPA's Continuous Emissions Monitoring System (CEMS) [44]. CEMS represents the only publicly available information on high frequency operating data for thermal power plants in the U.S., and has been widely used in empirical studies [46, 47, 48, 49, 50, 51, 52]. We match units in CEMS to EIA generators using a 2015 crosswalk provided by the EPA (personal communication), and aggregate unit level information from CEMS at the plant level by EIA site code and technology type. This step allows us to assign operating capacity to each power plant for which EPA data is available. Following [52], we apply a 5% reduction to gross generation from CEMS to obtain an implied measure of net generation that can be compared to net generation from the EIA. Finally, as noted above only thermal plants with capacity above 25 MW are required to report to CEMS; cogeneration, industrial and commercial facilities are also generally not in CEMS. These exceptions do not result in a substantial loss of coverage for our analysis: net generation of NGCC (coal-fired) plants from CEMS represents about 84% (79%) of EIA generation in WECC over the period of our study.

4.3 CAISO data

We collect hourly data on available transmission capacity and scheduled net power flows on twelve major transmission interfaces connecting the California ISO to the rest of WECC from the ISO's Open Access

[10] 25 MW corresponds to the minimum size of generators subject to requirements for monitoring and reporting emissions under EPA's Continuous Emissions Monitoring System [44]. Plants with capacity below 25 MW generally use renewable energy sources and represent less than 5% of generating capacity in our sample.

Same-time Information System (OASIS) [53]. Grid interfaces are identified based on [54] and the analysis of CAISO annual reports detailing the frequency of import congestion on each intertie [55]. We also obtain hourly aggregate generation by technology, including wind and solar production, from CAISO [56]. It should be noted that transmission capacity and scheduled net energy from imports/exports are only available from April 2009 to October 2015, while hourly aggregate generation by technology is only available from April 2010.

4.4 FERC, NOAA and price data

We complement detailed information on the operations and status of electric power plants with data from other sources. Electricity consumption (or load) comes from the Federal Energy Regulatory Commission (FERC). FERC Form 714 provides hourly load information by planning area [57]. We aggregate load to the monthly and daily level, and assign it to power plants based on their planning area. Monthly population-weighted heating and cooling degree days, as well as measures of water scarcity (like the Standardized Precipitation Index or SPI [58]) by state climate division are from the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information [59]. Finally, we obtain daily natural gas prices at four locations in WECC (Sumas, PG&E Citygate, SoCal Border and El Paso San Juan [60]) and weekly Rocky Mountain Colorado Rail coal prices (with a heat rate of 11,700 Btu/lb and a sulfur content of 0.8 lb/MMBtu) from SNL Energy,[11] and daily carbon futures prices for year vintage allowances expiring in December of the same year, in $/ton, from California Carbon Dashboard [61].

5 Research Design

5.1 Empirical Framework

Based on the potential outcome framework that is commonly used in the treatment evaluation literature [62], each observation has two potential outcomes $Y_{it}(d_i)$ depending on treatment status. Let $d_i = 1$ if observation i is treated, and $d_i = 0$ if i is not treated. Potential outcomes for observation i at time t are denoted by $Y_{it}(1)$ and $Y_{it}(0)$ for treatment and non treatment status, respectively. We are interested in estimating the average treatment effect on the treated (ATT), defined as the average difference between treated and untreated (or

[11] There exists no public regional price for coal. S&P Global Market Intelligence's physical market survey details spot prices for coal traded for physical delivery in forward quarters. Weekly prices are "assessments" based on direct supplier-consumer transaction data collected from utility buyers and sellers through a weekly survey.

control) outcome, conditional on treatment:

$$\alpha = E[Y_{it'}(1) - Y_{it'}(0)|d_i = 1] \tag{1}$$

where α measures the average treatment effect, E is the expectation operator, and t' represents any period after treatment. Outcomes after treatment can be used to identify $E[Y_{it'}(1)|d_i = 1]$. In contrast, $[Y_{it'}(0)|d_i = 1]$ represent counterfactual outcomes to be estimated based on the outcomes for control observations.

The objective of our study is to investigate the leakage effects of the AB 32 cap-and-trade program in the electricity sector. The primary leakage mechanism consists in supplanting power generation in the regulated region (California) by increased generation in the unregulated regions ("leakers"). Thus, α_C is the treatment effect in California and α_L is the treatment effect in potential leaker region L outside California. The choice of potential leakers and controls is a key point of our empirical framework. In principle, all plants in the Western Interconnection may be leakers because contract shuffling could create knock-on effects. In practice, however, some balancing authorities in WECC have transmission capabilities allowing plants in their footprint direct access to California load. These plants are more likely to adjust their generation in response to policy changes in California. Therefore, in the baseline specification we identify potential leakers based on the CARB's Greenhouse Gas Emission Inventory, which reports annual estimated CO_2 emissions from power plants supplying specified source power to California [63]. Specifically, we designate as leakers WECC balancing authorities in the U.S. that dispatch power plants supplying specified source power to California. We emphasize two points here. First, we do not intend to suggest that our approach identifies the only leakers unequivocally. Rather, the approach identifies generation resources outside California that are deemed likely to provide exports to California. Other methods for identifying potential leakers in the Western Interconnection are possible and worth exploring for further empirical analyses; we consider alternate leakers in one of the robustness checks, and return to this point in Section 7. Second, Western Canada and the northern portion of Baja California in Mexico may also be leakers because they are part of WECC. In particular, British Columbia is a net exporter of power to the Western U.S., and a large share of its power export sales are directed to California [64]. However, since we do not observe Canadian and Mexican monthly generation, we restrict the scope of our study to intranational leakage.

After identifying potential leakers, we group balancing authorities into regions of contiguous connected electrical components, following the classification in [65]. As discussed in Section 5.1.1, we focus on two baseload technology types (natural gas combined cycle and coal-fired plants) that are most likely affected by the carbon policy. Region definition differs slightly by technology, depending on the location of plants supplying specified source power. For NGCC plants, the Northwest region includes plants in BPAT, PACE and PACW, and the Southwest region includes plants within the CAISO footprint but located in Arizona and Nevada, as well as plants in AZPS, HGMA, NEVP, SRP and WALC (Figure 3). For coal-fired plants, the Northwest region includes plants in BPAT, LDWP in Utah, NWMT, PACE and PGE; the Eastern region includes plants in PSCO and WACM; the Southwest region includes plants in AZPS, NEVP, PNM, SRP, TEPC and WALC (Figure 4). In our baseline specification, the set of treated plants consists of plants of a given technology type that are either in California or one of the leaker regions. For NGCC plants, the set of controls consists of WECC plants that are not in California or one of the leaker regions; for coal-fired plants, the number of controls in WECC outside of the leaker regions is limited, and thus we extend the set of controls to include plants in MRO, SPP and TRE. Alternate definitions of the set of treated and control groups are considered in the robustness checks.

5.1.1 Differences-in-differences regressions

Our first set of empirical results obtains estimates of the treatment effects of interest with the following differences-in-differences (DID) model specification:

$$Y_{it} = \alpha_C TREAT_{it}^C + \sum_L \alpha_L TREAT_{it}^L + \mathbf{X}_{it}^{'}\underline{\beta} + \gamma_i + \gamma_y + \gamma_{sm} + \epsilon_{it} \tag{2}$$

where i indexes a plant-technology, t indicates month, L denotes a leaker region, and s, y, m stand for state, year, and month-of-year respectively. Our dependent variable Y_{it} is the capacity factor of plant-technology i in month t, defined as the ratio of net generation over operating capacity multiplied by total number of hours. We focus on two baseload technology types that are most likely affected by the carbon policy (natural gas combined cycle (NGCC) plants and coal-fired plants), and run separate regressions by technology type.[12] $TREAT_{it}^C$ is a dummy equal to 1 if plant i is in California and t is January 2013 or later; $TREAT_{it}^L$ is similarly defined for plants in leaker region L. Assuming treated and control facilities would have followed parallel trajectories in the absence of the carbon policy (as discussed in Section 5.2), the treatment effects of interest, α_C and α_L, measure the average effect of the cap-and-trade program on capacity factors of power plants in California and each leaker region, respectively, conditional on observable covariates. It is worth

[12] Natural gas steam turbines represent a small fraction of generating capacity in the WECC region (Figure A1 in the Appendix). Other technology types like natural gas combustion turbines and oil turbines that are used as peaker plants during high load periods are unlikely to have responded to California's carbon policy, given the modest level of permit prices in our sample period.

noting that California has limited coal generation capacity. Hence, leakage would result in lower natural gas generation in state and higher coal and/or natural gas generation out of state. In terms of the model in (2), in the presence of leakage we would expect negative and statistically significant treatment effects for California plants, and positive and statistically significant treatment effects for plants in the leaker regions.

$\mathbf{X}'_{it}$ represents a broad set of determinants of capacity factors. First, in the baseline specification we include the natural log of electricity consumption in the plant's planning area, and the log of nuclear and renewable generation (including hydro) in the plant's state. This functional form implies low responsiveness of capacity factors when demand or non-thermal power generation is high, in line with [51, 6, 24]; however, estimation results do not critically hinge on this assumption, and are robust to a linear specification. Second, we include variables that may affect plant productivity, like temperature (measured by heating and cooling degree days in the plant's climate division) and precipitation (measured by the Standardized Precipitation Index in the plant's climate division). Third, we consider a measure of plant competitiveness, the fuel cost ratio described in Section 4.1, including both linear and quadratic terms to account for potential nonlinear responses to input prices [50]. It is worth noting that, besides the cap-and-trade program, AB 32 relies on a suite of complementary policies to achieve emission reductions, such as renewable electricity policies and energy efficiency standards; similar programs are also implemented in other WECC states. These policies clearly affect capacity factors of baseload generation technologies (e.g., through the merit order effect of renewables dispatched before thermal units). However, their impacts are accounted for through some of our covariates (load and nonhydro renewable energy production), while the estimated treatment effects measure changes that are specifically induced by California's cap-and-trade program. Finally, we include individual and time fixed effects in the regressions. Plant specific effects, γ_i, may be associated with time invariant differences in plant characteristics, like ownership (private utilities or political subdivision) and vintage. Year fixed effects, γ_y, capture differential changes in average utilization that are common to all plants in a given year, while state by month-of-year fixed effects γ_{sm} allow us to account for seasonality within the vast WECC footprint and control for differential changes that are common to all plants within a state in a given month. The error term ϵ_{it} is assumed independent of the covariates and treatment indicators.

5.1.2 Matching and differences-in-differences

The regression approach described above has some potential drawbacks [66]. First, temporal aggregation at the monthly level may bias results [67]; thus, using higher frequency measures of generation would be advantageous. Another concern is that plants with similar monthly average capacity factors may be operated very differently. As a result, counterfactual outcomes may be estimated incorrectly. To illustrate, consider two periods; a plant with zero net generation in period 1 and operated at 80% capacity factor in period 2 would have the same average capacity factor of a plant operated at 40% capacity factors in both periods. Yet, the two plants would hold different positions in the dispatch order of their respective balancing authority, and thus not represent a suitable pair of treated-control observations. More accurate estimates of the counterfactual outcomes may be constructed based on control plants that had similar utilization and efficiency levels to the treated units before policy implementation. The treatment effect of interest could then be obtained estimating a DID model in which the impact of other policies potentially affecting capacity factors (e.g., renewable electricity policies) are subsumed in the covariates. In order to mitigate potential bias, we take two steps. First, we use hourly generation data to construct daily measures of plant utilization. Increased generation from renewable energy sources affects the operations of fossil fuel units during the day. Thus, we explore treatment effect heterogeneity across daytime and nighttime. We average capacity factors across a twelve hour period (7am to 7pm) to form a daily "daytime" capacity factor for plants reporting to the EPA's Continuous Emissions Monitoring System, and average over the remaining hours to obtain a daily "nighttime" capacity factor. Second, we preprocess the data to improve balance between treated and control groups by matching on high frequency pre treatment variables. The basic idea of matching is to find untreated units that are similar to the treated ones in terms of variables that influence the outcome of interest (i.e., so called "matching variables"), except for treatment status. Counterfactual outcomes for treated plant i are then inferred using a weighted average of the outcomes of units that are comparable to i, but receive a different treatment. Control units whose observable characteristics are closer to those of plant i are weighed more heavily in the construction of the counterfactual estimate. While earlier empirical work in energy and environmental economics employed parametric and semi-parametric matching methods,[13] we explore the use of coarsened exact matching (CEM) to improve balance between treated and control observations [70, 71, 72]. CEM is a nonparametric method that bounds the maximum imbalance between treated and control groups with respect to the full joint distribution of the covariates *ex ante*. Unlike model dependent methods, CEM does not extrapolate counterfactual outcomes when there is limited overlap in the distributions of covariates across treatment and control groups, because matched data are restricted to areas of common empirical support. Recent applications of this matching method are presented in [73, 74, 75].

[13] In the context of Southern California's RECLAIM program, [66] employ a semi-parametric DID matching estimator of the ATT that compares differences between post and pre treatment NO_x emissions across treated and control plants, and use a regression-based adjustment to mitigate bias introduced by poor match quality [68, 69]. As a robustness check, the authors implement a propensity score matching estimator, which relies on a parametric regression model to estimate the propensity score. One disadvantage of this approach is that a misspecified matching model may produce greater imbalance in variables that are omitted from the matching procedure.

The objective of our matching procedure is to achieve statistically indistinguishable distributions between treated and control plants across a set of exogenous covariates that are highly correlated with the outcomes of interest (i.e., daily daytime and nighttime capacity factors). Hour-of-day specific capacity factors are clearly correlated with time-of-day capacity factors; further, more efficient plants tend to be used more heavily [76]. Therefore, we choose 2009 and 2010 as pre treatment period,[14] and use hour-of-day specific capacity factors and heat rates (averaged over this two-year period) as matching variables. Next, we coarsen the average hourly variables into discrete bins that identify strata corresponding to different levels of plant utilization and efficiency.[15] The following step is to perform exact matching on these discrete bins and discard observations from bins that do not contain both treated and control observations. It should be noted that the matched control sample varies by treated region and time of day, and CEM produces weights for each matched unit in each stratum [71].[16] Finally, we measure the impact of the cap-and-trade policy on plant operations by estimating the following differences-in-differences models with weighted least squares:

$$Y_{it} = \alpha_C TREAT_{it}^C + \mathbf{X}_{it}^{'}\underline{\beta} + \gamma_i + \gamma_y + \gamma_{sm} + \epsilon_{it} \tag{3}$$

$$Y_{it} = \alpha_L TREAT_{it}^L + \mathbf{X}_{it}^{'}\underline{\beta} + \gamma_i + \gamma_y + \gamma_{sm} + \epsilon_{it} \tag{4}$$

where t indicates daytime or nighttime, and Y_{it} is the capacity factor of plant-technology i reporting to CEMS in period t. Daily electric load by planning area is obtained from FERC Form 714; other covariates in $\mathbf{X}_{it}^{'}$ are invariant at the monthly level (as in Section 5.1.1), since we do not observe higher frequency data for the entire WECC region.

5.1.3 Scheduled power flow regressions

In our final set of empirical results, we switch from analyzing plant operations to examining the relationship between the AB 32 allowance price and scheduled power imports into CAISO. Building on [23], we estimate a model of daily scheduled power flows into CAISO, and test for leakage based on the statistical significance of the AB 32 allowance price as one of the explanatory variables. As detailed in Section 4, we identify the major

[14] We exclude 2011, a wet hydrological year in which NGCC plants ran at much lower capacity factors than usual [77], and 2012, the year before compliance obligations began. We also remove from the matching dataset outliers (i.e., plants for which generation from CEMS is greater than generating capacity from EIA) and plants that were not operating over the entire period of our study.

[15] A detailed discussion of matching hour sets and binning strategies is presented in Section 6.2.

[16] Matched units receive a weight of 1 if they belong to the treatment group, and $\frac{m_C}{m_T}\frac{m_T^s}{m_C^s}$ if they belong to the control group, where m_C is the total number of control units, m_T is the total number of treated units, and m_C^s and m_T^s are their counterparts in stratum s. Weights normalize the variance in distribution of attribute bins across treatment and control observations. Unmatched units receive a weight of 0.

electricity grid interfaces into CAISO based on [54] and the analysis of annual CAISO reports detailing the frequency of import congestion on each intertie [55]. Scheduled flows on major interties are aggregated into two regions: the Northwest region including lines from the BPAT, PACE and PACW balancing authorities, and the Southwest region with interties from the AZPS, SRP, TEPC and WALC balancing authorities.[17] To account for the highly interconnected nature of power flows in WECC and mitigate potential endogeneity bias in our estimates, we model net scheduled energy flows from the Northwest and Southwest regions as a system of equations, and estimate it using maximum likelihood:

$$
\begin{aligned}
Z_{NW_t} &= \mathbf{X}_t^{'}\underline{\beta} + \delta_{NW}\mathrm{CO}_{2_t} + \gamma_d + \gamma_q + \epsilon_t \\
Z_{SW_t} &= \mathbf{X}_t^{'}\underline{\theta} + \delta_{SW}\mathrm{CO}_{2_t} + \eta_d + \eta_q + \zeta_t
\end{aligned} \tag{5}
$$

where t indicates day, d denotes day-of-week, q indexes a quarter, Z_{NW_t} refers to scheduled flows from the Northwest region, and Z_{SW_t} refers to scheduled flows from the Southwest region.

A key determinant of net scheduled flows is given by the electric demand in CAISO and exporting balancing authorities. Higher load in CAISO increases net imports, while higher load in Northwest or Southwest balancing authorities is expected to reduce export availability from each region to California. Further, net scheduled flows in each export region may be reduced by scheduled flows into CAISO from the other region; as a result, we include net flows from the competing export region as a covariate in each equation. We also control for daily nuclear, wind and solar generation in CAISO. These technology types are unlikely to have responded to the carbon policy because they operate at near maximum capacity most of the time (nuclear), or their output depends on resource availability (wind and solar). Higher renewable and nuclear output in CAISO would reduce the need for imports. On the other hand, higher production from hydro and renewable energy sources in the Northwest and Southwest regions is expected to increase electricity imported from outside of California, displacing in-state natural gas-fired generation. Since daily aggregate production by technology is not publicly available for WECC balancing authorities other than CAISO, we include monthly generation from EIA-923; further, we only consider the most significant non-fossil energy source for each region, i.e., hydro in the Northwest and solar in the Southwest. Fuel prices in California and other WECC regions are also likely to affect electricity imports into CAISO. Since fuel prices at the plant level are only available from EIA-923 at the monthly level, we use wholesale natural gas prices at four locations in WECC (Sumas, PG&E Citygate, SoCal Border and El Paso San Juan) [60] to control for daily price dynamics in

[17] Northwest interties include Cascade, Pacific AC Intertie, Nevada-Oregon Border, COTPISO, Summit, IPP DC Adlanto, Mona IPP DC, Mead, and El Dorado; Southwest interties include Palo Verde and IID-SCE [54].

the Northwest, CAISO and Southwest regions. Higher PG&E Citygate prices in Northern California and SoCal Border prices in Southern California are likely to make power imports more economically viable. In contrast, higher natural gas prices at Sumas in the Northwest and El Paso San Juan in the Southwest are expected to favor in-state electricity generation relative to power imports.[18] As an alternate measure of fuel prices, we use ratios of PG&E Citygate/Sumas prices and SoCal Border/El Paso San Juan, expecting that higher fuel price ratios increase imports in CAISO. To account for seasonal effects, we include day-of-week and quarter dummies in each regression.

Finally, the daily price of AB 32 CO_2 allowances is equal to zero until the beginning of compliance obligations on January 1, 2013. A positive and statistically significant coefficient associated with the CO_2 price in the Northwest equation, δ_{NW}, would suggest empirical evidence of emission leakage from the Northwest region of WECC. Similarly, a positive and statistically significant δ_{SW} would support the hypothesis of leakage from the Southwest region.

5.2 Identifying Assumptions

Our estimation strategy relies on several identifying assumptions. First, treatment is exogenous: participation in the cap-and-trade program does not depend on the outcomes (i.e., plant capacity factors). Second, in order to interpret α_C and α_L as estimates of the effect of California's cap-and-trade program on plant-level capacity factors, an important identifying assumption is that treated and control plants would have followed parallel trajectories, absent the AB 32 cap-and-trade program [78]. We assess the parallel trends assumption by testing the equivalence of time trends between treatment and control groups before program implementation [79].[19] Specifically, we test the significance of the interaction term between the time trend and the treatment group: estimated parameters associated with group specific time trends that are not statistically different from zero indicate that pre treatment trends are similar for treated and control groups, and are consistent with a causal interpretation of the results in (2). We estimate the following equation:

$$Y_{it} = \alpha_{Ct} D_t TREAT_i^C + \sum_L \alpha_{Lt} D_t TREAT_i^L + \mathbf{X}'_{it}\underline{\beta} + \gamma_i + \gamma_y + \gamma_{sm} + \epsilon_{it} \tag{6}$$

[18] We only observe weekly Rocky Mountain Colorado Rail coal prices for the Western Interconnection (Section 4.4). Absent regional variation, we do not control for coal prices in the scheduled power flow regressions.

[19] In [79], court decisions altering common law occur at different times providing multiple experiments (i.e., multiple treatment periods); in our setting, the treatment period is instead unique and corresponds to January 1, 2013, when the California cap-and-trade program began its compliance obligation.

where D_t is a quarterly dummy equal to 1 after January 2013 and 0 otherwise, $TREAT_i^C = 1$ if plant i is in California, $TREAT_i^L = 1$ if plant i is in one of the leaker regions, and α_{Ct} and α_{Lt} are the estimated coefficients associated with group specific time trends. Other variables are defined as above. If the parallel trends assumption is satisfied, α_{Ct} and α_{Lt} are not statistically different from zero before the implementation of the cap-and-trade program. Not all the α_{Ct} can be identified as the $TREAT_i^C$ dummies are perfectly collinear in the presence of state effects. Hence, similarly to [24] we omit the first year for all groups in our tests.

Matching relies on selection on observables (ignorability assumption) and common support (overlap assumption). Based on the ignorability assumption, once we control for matching variables, treatment is randomly assigned, experimental conditions are re-established, and biases in the DID estimator are removed because capacity factors at treated and control plants would have followed parallel paths over the study period. We run balancing tests to compare matching variable means in the treated and control groups before and after matching. Finally, the overlap assumption requires that the support of the distribution of covariates in the treated group overlaps the support of the distribution of these covariates in the control group. Coarsened exact matching automatically restricts the matched data to areas of common support, as discussed in Section 5.1.2: this helps avoid making inferences based on extrapolation, which are known to be highly model dependent.

6 Results

6.1 Differences-in-differences regressions

Tables 1 and 2 present our differences-in-differences regression results from the analysis of capacity factors. We run the baseline specification with monthly data in equation (2) for NGCC and coal plants separately, and present estimation results in Column (1) of the two tables. Our dependent variable is the capacity factor of plant-technology i in month t. The unit of observation in the analysis is a plant-month, and standard errors are clustered at the plant level. As noted in Section 5.1, leaker regions and the set of controls differ by technology type. For NGCC plants, treatment effects are estimated for California and two leaker regions (Northwest and Southwest) including plants that supply specified source power to California. In Table 1, the "California" line presents the treatment effect estimate for α_C, while the "Northwest" and "Southwest" lines present the estimated effects for α_L in equation (2). The set of controls consists of all NGCC power plants in WECC, but outside of California or any of the leaker regions. Treated and control regions are presented in Figure 3, and summary statistics by region are given in Table A3 in the Appendix. Figure 5(a) reports the results of the parallel trend tests. Specifically, we plot the estimated α_{Ct} and α_{Lt} from equation (6) with 95% confidence intervals: with one exception, pre treatment period coefficients are not statistically different from zero for California and the leaker regions, indicating that treated and control regions have similar pre treatment trends.

For coal-fired plants, we do not include a treatment effect for California. Generating capacity from coal is limited, and capacity factors have declined steadily in the three quarters before the introduction of the cap-and-trade system, resulting in a rejection of the parallel path assumption before treatment. However, coal plants supplying specified source power to California are located throughout WECC. After grouping them into three regions (Northwest, Southwest and Eastern) according to [65], the number of control plants in WECC is limited (Table A4). As a result, in our baseline coal specification we extend the set of controls to include coal plants in the MRO, SPP and TRE regions besides WECC facilities that are outside California or any of the leaker regions. Treated and control regions are presented in Figure 4, and summary statistics by region are given in Table A4. Figure 5(b) reports the results of the parallel trend tests, indicating that treated and control regions exhibit similar pre treatment trends.

Estimated treatment effects in the baseline specification suggest that NGCC generators in California had a statistically significant policy-induced reduction in capacity factors of about 14%. In contrast, California's cap-and-trade led to a 4% increase in coal capacity factors in the Northwest and Eastern leaker regions (statistically significant at the 5% and 10% significance level, respectively). We find a small and insignificant response for NGCC plants in the Northwest and Southwest regions, and coal plants in the Southwest regions. Overall, this result suggests empirical evidence of leakage: the policy induced a reduction in NGCC generation in California and an increase in coal-fired generation in WECC balancing authorities that dispatch plants supplying specified source power to California.

We conduct a back of the envelope calculation to examine whether our estimated coefficients result in realistic magnitudes of leakage. First, we find the estimated generation leakage by multiplying the treatment effect (when statistically significant) by the average annual generation capacity in each region and number of hours in a month. That is, we multiply the estimated treatment effect of 14.1% by the average NGCC generation capacity in California between 2009 and 2016 (19,369 MW), and the estimated treatment effects of 4.25% and 3.93% by the average coal generation capacity in the Northwest and Eastern leaker regions (13,259 MW and 5,723 MW, respectively). Due to the cap-and-trade policy, NGCC generation in California decreased by approximately 2 million MWh per month, while coal generation in the leaker regions increased by about 0.6 million MWh per month.[20] In turn, this implies that emissions increased by about 8.5 million tons per year in the Northwest and Eastern regions of WECC, and decreased by about 12 million tons per year in California, corresponding to an aggregate decrease of about 3.6 million tons per year and an estimated policy-induced leakage of about 70%.[21] This estimate is within the range of predictions obtained from simulation-based studies considering California's cap-and-trade program, although direct comparisons between *ex ante* and *ex post* analyses are difficult [24]. On the lower end of the spectrum, [19] find leakage of 9% when provisions to prevent resource shuffling are enforced. [10] simulate the effects of a first deliverer approach with default emission factor of 0.428 tons CO_2/MWh applied to imports, assuming a 15% and a 25% reduction in California utility power sector emissions from 2007 levels. Compliance with the lower cap yields no change in emissions, relative to the no cap scenario. Assuming a 25% reduction, the first deliverer approach would lead instead to an emission decrease of 4.5 million tons in California and an increase of 1.6 million tons in the rest of WECC relative to a no cap scenario, i.e. an implied leakage of about 35%. Closer to our estimate, under a first-deliverer approach [9] predict a leakage rate of about 85% (corresponding to an emission decrease of 0.7 million tons in California and 0.1 million tons in the Western U.S.), while [14] find that carbon regulation in California would result in leakage of about 89% (corresponding to an emission decrease of 5.5 million tons in California and 0.6 million tons in the Western U.S.). Finally, a recent econometric study of leakage in the context of the Regional Greenhouse Gas Initiative [24] estimates a leakage rate of approximately 50% for the policy (corresponding to an emission decrease of 8.8 million tons per year in RGGI and an aggregate decrease of 4.3 million tons per year).

Robustness checks. We conduct several robustness checks to assess the sensitivity of our estimates. First, results are robust to alternate clustering (e.g., at the balancing authority or state level, to account for the likely dependence in capacity factor innovations across plants in the same power control area/state, and arbitrary time series correlation within the same power control area/state), and the inclusion of different fixed effects (plant, year, and month-of-year; and plant, state-by-year, and month-of-year).

Second, we adjust standard errors to allow for correlation along two dimensions (plant and time). Standard errors clustered at the plant level assume that correlation of the residuals within the cluster may be nonzero, but residuals across clusters are uncorrelated. This may lead to incorrect standard errors and t statistics, if residual correlation exists both within a plant across time and across plants at a moment in time

[20]Between 2009-2012 and 2013-2016, monthly power generation from solar and wind in California increased, on average, by about 1 million MWh and 0.6 million MWh per month, respectively.

[21]Our calculations are based on an average heat rate of 12,046 Btu/kWh and CO_2 emission rate of 207.87 lb/MMBtu for coal plants in the Northwest region, heat rate of 12,163 Btu/kWh and CO_2 emission rate of 207.82 lb/MMBtu for coal plants in the Eastern region, and heat rate of 8,645 Btu/kWh and CO_2 emission rate of 118.88 lb/MMBtu for NGCC plants in California. Average heat rates and CO_2 emission rates over the period of our study (2009-2016) are from Tables A3 and A4.

[80, 81]. In Column (2) of Tables 1 and 2, we allow for arbitrary correlation of the error terms at the plant and month level, and find that estimation results are robust to double clustering.[22]

In Column (3) of Tables 1 and 2, we restrict our sample to facilities that were operating over the entire period of our study to account for entry and attrition at the plant level.[23] For NGCC plants, 10 facilities out of 121 operating in WECC in 2009 were no longer in service in 2016 (about 2% of 2009 operating capacity), and 14 plants out of 125 operating in 2016 were not active in 2009 (about 10% of 2016 operating capacity). For coal-fired plants, 15 plants out of 55 operating in WECC in 2009 were no longer in service in 2016 (about 6% of 2009 operating capacity), and 3 plants out of 43 operating in 2016 were not active in 2009 (about 2% of 2016 operating capacity). Estimated treatment effects are in line with those from the full sample, and suggest that capacity factors of NGCC generators in California decreased by about 13%, while capacity factors of coal generators in the Northwest and Eastern leaker regions increased by about 6% and 5%, respectively. The estimated coal treatment effect in the Northwest region is statistically significant at the 1% level (rather than 5%, as in the baseline regressions).

In Column (4) of Table 1, we estimate equation (2) for NGCC plants with a broader set of controls including plants in SPP and TRE, in addition to the rest of WECC.[24] Figure A2 in the Appendix shows that the parallel trend assumption holds for treated and control groups before the introduction of the cap-and-trade program. Estimated treatment effects confirm a statistically significant reduction in capacity factors for NGCC plants in California, but indicate a smaller effect (about 11%) than in the baseline specification.

We also test the sensitivity of our estimates to alternate definitions of the leaker regions. Concerns have been raised about resource shuffling in the context of the California ISO's Energy Imbalance Market (EIM), a real-time power market to meet short-term supply imbalances in the Western United States [83, 13], and CAISO has been experimenting with approaches to mitigate leakage [84, 85]. We designate as leakers WECC balancing authorities in the U.S. that are active or pending (as of 2018) participants of the Western EIM. These entities include AZPS, IPCO, NEVP, PACE, PACW, PGE, PSEI, SCL and SRP, and are divided into two leaker regions, as shown in Figure A3 in the Appendix. We control for the same variables of the baseline specification, and cluster standard errors at the plant level. Results for the parallel trend tests are given in Figure A4. Estimated treatment effects in column (5) of Table 1 and column (4) of Table 2 confirm a policy-induced reduction of about 12% in capacity factors for NGCC power plants in California. The estimated increase in coal capacity factors for Northwest plants is lower than in the baseline regressions

[22] Two-way clustering relies on asymptotics in the smaller number of clusters (i.e., the dimension with fewer clusters), and is thus likely to produce unbiased estimates when each dimension has many clusters [82].

[23] Entry and attrition at the unit level are limited, as plant capacity did not vary significantly over the period of our study.

[24] We exclude MRO from the set of controls because capacity factors are much lower than in the rest of WECC.

(about 3%), and statistically significant at the 10% level.

6.2 Matching and differences-in-differences

Our second set of empirical results explores the use of coarsened exact matching methods to prune observations that have no close matches on pre treatment variables in both treated and control groups. We then run a differences-in-differences model on matched plants only to measure policy impact using high frequency measures of generation and load for a subset of power plants reporting to the EPA's Continuous Emissions Monitoring System.

Our baseline specification matches treated and control plants based on four pre treatment, hour-of-day specific capacity factors in each period. We consider three sets of matching hours: in hour set 1, matching is based on hours 7,10,13,16 (Daytime) and 19,22,1,4 (Nighttime); in hour set 2, matching is based on hours 8,11,14,17 (Daytime) and 20,23,2,5 (Nighttime); in hour set 3, matching is based on hours 9,12,15,18 (Daytime) and 21,0,3,6 (Nighttime). We coarsen each matching variable according to manually defined cut points that identify strata corresponding to different levels of plant utilization. In general, fewer strata yield more matches, but result in more diverse observations within the same stratum. To balance this trade-off, we define cut points based on the empirical distribution of capacity factors by technology type (0.3, 0.5, and 0.7 for NGCC plants, 0.6 and 0.8 for coal-fired plants).[25] Balancing tests confirm that matching achieves statistically indistinguishable distributions between treated and control plants. Tables 3 and 4 present the t statistics of tests of identical means in the treated and control groups for hour set 2. Before matching, there exist significant differences between the covariates, particularly in California and the Northwest and Eastern regions in WECC; after matching, the null of identical means in both group is no longer rejected for all variables.[26] Further, it is worth noting that in each period matching helps to reduce bias not only for specific hours used as matching variables, but across all hours, bringing us closer to a quasi-experimental dataset.[27]

Tables 5 and 6 present the treatment effects from the differences-in-differences regressions estimated using weighted least squares.[28] Results from the matched sub-samples are broadly consistent with the ones

[25] Coal-fired plants (particularly in the Southwest region of WECC) tend to be more heavily utilized than NGCC plants, motivating our choice of higher capacity factors as cut points.

[26] In the Southwest region, raw data does not show statistical difference across treated and controls, thus matching does not yield substantial benefits.

[27] To illustrate, consider Table 3. Columns 2 to 7 present t statistics of a balancing test of identical means when matching is based on hours 8,11,14,17, while columns 8 to 13 present balancing tests when matching is done on hours 20,23,2,5. In Columns 2-7, balance improves not only for matching covariates (i.e., hours 8,11,14,17), but for all hours, including those omitted from the matching procedure.

[28] Estimated coefficients for other covariates in the DID regressions have the expected sign and statistical significance, and are available from the authors upon request. Further, in Table 5 the presence of one additional outlier for hour set 1 leads to

from the full sample. For California, the effect is smaller (in absolute value) compared to the OLS result, suggesting a policy-induced reduction of NGCC capacity factors in California by about 7% only during daytime hours. Leakage from coal plants in the U.S. Northwest region of WECC is confirmed across all periods, while nighttime leakage for the Eastern and Southwest regions are more sensitive to the choice of matching hours and statistically significant at the 10% level. A decrease in NGCC generation in California during daytime hours and an increase in coal generation in the leaker regions over the entire day may be due to heavy utilization of Western U.S. coal plants, which tend to ramp more slowly than NGCC plants [86]. It should be noted that our matching procedure is based on quite small subsamples of plants reporting to CEMS, which represent 81% of NGCC generation in California and 96%, 100%, and 62% of coal-fired generation in the Northwest, Eastern and Southwest regions, respectively.[29] Further, the estimated treatment effects are only averaged over the subset of treated units for which good matches exist among available controls (i.e., constitute local ATTs [71]), and do not account for correlation across daytime/nighttime hours or treated groups. Thus, we believe that the causal effects defined on the full sample are better suited for obtaining an estimate of policy-induced leakage.

Robustness checks. Empirical findings are robust to the inclusion of matching variables and choice of cut points. We present results for hour set 2. In our first robustness check (Table A5 in the Appendix), we include hourly heat rates as matching variables and set cut points at the percentiles of the distribution of matching variables. For NGCC, we use the 40th, 60th and 80th percentiles, while for coal we use the 20th, 40th, 60th and 80th percentiles. These cut points define four strata for NGCC capacity factors and heat rates, and five strata for coal matching variables, respectively. We choose a different number of strata for each technology type to balance the trade-off of matching described above, since five strata for NGCC yield few matches. In our second robustness check (Table A6 in the Appendix), we match based on hour-of-day specific capacity factors in each period, as in the baseline specification, but coarsen the matching variables using the statistical-based binning algorithm that returns the lowest value for the L_1 statistic, a measure of imbalance with respect to the full joint distribution [72].[30] Results are in line with those from the baseline specification, but do not suggest leakage from the Southwest region in WECC.

fewer control plants (72), relative to hour sets 2 and 3 (73).

[29]These percentages represent region-specific average shares of CEMS generation over EIA generation in 2009-2016.

[30]The Scott and Freedman-Diaconis automatic binning algorithms perform well in our sample [87, 88] and result in more strata, relative to coarsening by fixed cut points or percentiles. Note that there exists a trade-off between bin width and value of the L_1 statistic. Coarsening by user choice results in values of L_1 close to zero and more overlap between the distribution of covariates in the treated and control groups; automated coarsening yields narrower bin width that better approximates each distribution, but results in higher values of L_1 and less overlap between the two distrbutions.

6.3 Scheduled power flow regressions

Our last set of empirical results examines changes in daily net power flows across major CAISO interfaces after the introduction of California's cap-and-trade program [23]. We estimate a model of scheduled power flows into CAISO, including the AB 32 emission allowance price as one of the explanatory variables. As discussed in Section 5.1.3, the main drivers of scheduled flows are electric demand in CAISO and exporting balancing authorities, hydroelectric, nuclear and renewable generation in CAISO, fuel prices in California and exporting regions, and electricity imports from the competing region. After controlling for these variables, a positive and statistically significant coefficient associated with the allowance price (δ_{NW} or δ_{SW}) would support empirical evidence of leakage from the Northwest or Southwest region of WECC into CAISO.

Model (1) in Table 7 represents the system of equations with fuel prices as covariates, while model (2) includes fuel price ratios. All covariates have the expected sign. Further, in both models the CO_2 allowance price is highly significant as an explanatory variable for Northwest flows, but does not have a statistically significant effect on power flows from the Southwest. Therefore, results suggest that net scheduled flows into California increased from the Northwest region of WECC in response to the carbon policy, further supporting the hypothesis of leakage.

7 Discussion and conclusions

California has pledged to reduce its greenhouse gas emissions to 1990 levels by 2020, 40% below 1990 levels by 2030, and 80% below 1990 levels by 2050. These ambitious goals are being accomplished through a suite of complementary policies, including a multi-sector cap-and-trade program that covers about 80% of the state's emissions and applies to in-state electricity generation and imports. To mitigate leakage in the electricity sector, California opted for a source-based regulation applied to in-state sources, with first deliverer measures for imports into California. However, the possibility of reshuffling contracts may enable substantial leakage under the AB 32 cap-and-trade system. Under resource shuffling, electricity contracts are rearranged so that production from low emission sources serving out-of-state load is directed to California, while production from higher emission sources is assigned to serve out-of-state load. This would result in apparent emission reductions due to changes in the composition of imports to California, although emissions in exporting regions are unchanged or even increase.

Simulation-based studies have concluded that resource shuffling represents a significant potential conduit for emission leakage in the electricity sector under California's cap-and-trade program [9, 14, 10, 11]. Our paper brings empirical evidence to bear on this issue, using a novel dataset from 2009 to 2016. We present three sets of results that support the hypothesis of leakage. First, we analyze monthly operations of baseload power plants in WECC applying a differences-in-differences estimator. Regression results point to a policy-induced reduction in NGCC generation by about 14% in California and an increase in coal-fired generation by about 4% in regions of the Western Interconnection that supply specified source power to California. In turn, these estimated treatment effects imply a policy-reduced leakage of about 70%, which is within the range of *ex ante* predictions and in line with recent econometric estimates for the Regional Greenhouse Gas Initiative. Results are robust to the choice of leaker and control groups, clustering methods and sample definition. In particular, direction and magnitude of the estimated treatment effects is confirmed when we designate as leakers WECC balancing authorities in the U.S. that are active or pending participants of the California ISO's Energy Imbalance Market, a real-time power market in the Western U.S. Our second set of empirical estimates combines differences-in-differences with matching methods to ensure common support in the covariates across treated and control groups. We preprocess the data by matching units on coarsened hourly variables and carry out parametric inferences using daily measures of plant utilization. This approach changes the estimand to a local average treatment effect for the plants that were matched. Importantly, results from the matched sub-samples are broadly consistent with those from the full sample, and robust to the inclusion of matching variables and choice of cut points. In our final set of analyses, we test for leakage from the policy by examining the relationship between the AB 32 allowance price and scheduled power imports into CAISO. Specifically, we estimate a model of daily scheduled power flows into CAISO, and test for leakage based on the statistical significance of the AB 32 allowance price as one of the explanatory variables. The CO_2 allowance price is highly significant as an explanatory variable for Northwest flows, but does not have a statistically significant effect on power flows from the Southwest. This suggests that net scheduled flows from the Northwest region of the Western Interconnection into California have increased in response to the carbon policy, in line with the analysis of power plant operations with differences-in-differences and matching methods.

While the consistency of results across statistical approaches supports the hypothesis of leakage, our study is subject to limitations. For example, one caveat is that we do not observe power contracts between California utilities and out-of-state power plants. Absent this information, we are unable to control, for example, for the divestiture from long-term contracts with coal facilities [89]. Yet, if coal-fired production was redirected to out-of-state electricity consumers, resource shuffling (and leakage) would have happened. Another caveat relates to our approach for identifying potential leakers. Due to the specific features of electric power systems, identifying out-of-state generation resources that are deemed to provide exports to California represents a challenge. For example, until recently the California ISO was testing a "two-pass solution" of the Energy Imbalance Market algorithm to identify out-of-state resources dispatched to California in response to the carbon price [12]. This two-stage solution has been subject to criticism because it introduces discriminatory constraints applying in the second stage economic dispatch, and may lead market participants to distort their offers relative to the true generation costs [84]. As noted in Section 5.1, alternate approaches for identifying leakers in the Western Interconnection are possible and worth exploring for further empirical analyses. Ongoing work considers this important issue.

References

[1] Official California Legislative Information, "Assembly Bill No. 32 - Chapter 488," 2006. `http://www.leginfo.ca.gov/pub/05-06/bill/asm/ab_0001-0050/ab_32_bill_20060927_chaptered.pdf`.

[2] Official California Legislative Information, "Senate Bill No. 32 - Chapter 249," 2016. `https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201520160SB32`.

[3] California Climate Change Executive Orders, "Executive Order S-3-05," 2005. `https://www.gov.ca.gov/news.php?id=1861`.

[4] California Air Resources Board, "Semi-Annual Report to the Joint Legislative Budget Committee on Assembly Bill 32," 2013. `https://www.arb.ca.gov/cc/jlbcreports/july2013jlbcreport.pdf`.

[5] California Energy Commission, "Total System Electric Generation (2016 data)," 2017. `http://www.energy.ca.gov/almanac/electricity_data/total_system_power.html`.

[6] J. B. Bushnell, C. Peterman, and C. Wolfram, "Local Solutions to Global Problems: Climate Change Policies and Regulatory Jurisdiction," *Review of Environmental Economics and Policy*, vol. 2, no. 2, pp. 175–193, 2008.

[7] M. Fowlie, "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage," *American Economic Journal: Economic Policy*, vol. 1, no. 2, pp. 72–112, 2009.

[8] R. N. Stavins, J. Borck, and T. Schatzki, "Options for Addressing Leakage in California's Climate Policy," 2010. `http://www.analysisgroup.com/uploadedfiles/content/insights/publishing/options_addressing_leakage_california_climate_policy_feb_2010.pdf`.

[9] Y. Chen, A. L. Liu, and B. F. Hobbs, "Economic and Emissions Implications of Load-Based, Source-Based, and First-Seller Emissions Trading Programs under California AB 32," *Operations Research*, vol. 59, no. 3, pp. 696–712, 2011.

[10] J. B. Bushnell, Y. Chen, and M. Zaragoza-Watkins, "Downstream Regulation of CO_2 Emissions in California's Electricity Sector," *Energy Policy*, vol. 64, pp. 313–323, 2014.

[11] S. Borenstein, J. B. Bushnell, F. Wolak, and M. Zaragoza-Watkins, "Report of the Market Simulation Group on Competitive Supply/Demand Balance in the California Allowance Market and the Potential for Market Manipulation," 2014. `https://www.arb.ca.gov/cc/capandtrade/simulationgroup/msg_final_v25.pdf`.

[12] D. Burtraw, A. Carlson, D. Cullenward, Q. Foster, and M. Fowlie, "2018 Annual Report of the Independent Emissions Market Advisory Committee." `https://calepa.ca.gov/wp-content/uploads/sites/6/2018/10/Final_2018_IEMAC_Annual_Report_10-22-2018.pdf/`, 2018.

[13] California Air Resources Board, "California Greenhouse Gas Emissions for 2000 to 2016 - Trends of Emissions and Other Indicators," 2018. `https://www.arb.ca.gov/cc/inventory/pubs/reports/2000_2016/ghg_inventory_trends_00-16.pdf`.

[14] J. B. Bushnell and Y. Chen, "Allocation and Leakage in Regional Cap-and-Trade Markets for CO_2," *Resource and Energy Economics*, vol. 34, pp. 647–668, 2012.

[15] D. L. Shawhan, J. T. Taber, D. Shi, R. D. Zimmerman, J. Yan, C. M. Marquet, Y. Qi, B. Mao, R. E. Schuler, W. D. Schulze, and D. Tylavsky, "Does a Detailed Model of the Electricity Grid Matter? Estimating the Impacts of the Regional Greenhouse Gas Initiative," *Resource and Energy Economics*, vol. 36, no. 1, pp. 191–207, 2014.

[16] J. C. Carbone and N. Rivers, "The Impacts of Unilateral Climate Policy on Competitiveness: Evidence from Computable General Equilibrium Models," *Review of Environmental Economics and Policy*, vol. 11, no. 1, pp. 24–42, 2017.

[17] I. Sue Wing and M. Kolodziej, "The Regional Greenhouse Gas Initiative: Emission Leakage and the Effectiveness of Interstate Border Adjustments," 2009. `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1448748`.

[18] A. Antimiani, V. Costantini, C. Martini, L. Salvatici, and M. Tommasino, "Assessing Alternative Solutions to Carbon Leakage," *Energy Economics*, vol. 36, pp. 299–311, 2013.

[19] J. Caron, S. Rausch, and N. Winchester, "Leakage from Sub-National Climate Policy: the Case of California's Cap-and-Trade Program," *The Energy Journal*, vol. 36, no. 2, pp. 167–190, 2015.

[20] R. Aichele and G. Felbermayr, "Estimating the Effects of Kyoto on Bilateral Trade Flows Using Matching Econometrics," *The World Economy*, pp. 303–330, 2013.

[21] R. Aichele and G. Felbermayr, "The Effect of the Kyoto Protocol on Carbon Emissions," *Journal of Policy Analysis and Management*, vol. 32, no. 4, pp. 731–757, 2013.

[22] R. Aichele and G. Felbermayr, "Kyoto and Carbon Leakage: an Empirical Analysis of the Carbon Content of Bilateral Trade," *Review of Economics and Statistics*, vol. 97, no. 1, pp. 104–115, 2015.

[23] A. Kindle, D. Shawhan, and M. Swider, "An Empirical Test for Inter-State Carbon-Dioxide Emissions Leakage Resulting from the Regional Greenhouse Gas Initiative," 2011. `https://www.researchgate.net/publication/322477345_An_Empirical_Test_for_Inter-State_Carbon-Dioxide_Emissions_Leakage_Resulting_from_the_Regional_Greenhouse_Gas_Initiative`.

[24] H. Fell and P. Maniloff, "Leakage in Regional Environmental Policy: the Case of the Regional Greenhouse Gas Initiative," *Journal of Environmental Economics and Management*, vol. 87, pp. 1–23, 2018.

[25] M. Fowlie and M. Reguant, "Challenges in the Measurement of Leakage Risk," *AEA Papers and Proceedings*, vol. 108, pp. 124–129, 2018.

[26] A. Levinson and M. S. Taylor, "Unmasking the Pollution Haven Effect," *International Economic Review*, vol. 49, no. 1, pp. 223–254, 2008.

[27] M. E. Kahn and E. T. Mansur, "Do Local Energy Prices and Regulation Affect the Geographic Concentration of Employment?," *Journal of Public Economics*, vol. 101, pp. 105–114, 2013.

[28] J. E. Aldy and W. A. Pizer, "The Competitiveness Impacts of Climate Change Mitigation Policies," *Journal of the Association of Environmental Resource Economists*, vol. 2, no. 4, pp. 565–595, 2015.

[29] M. Panhans, L. Lavric, and N. Hanley, "The Effects of Electricity Costs on Firm Re-location Decisions: Insights for the Pollution Haven Hypothesis," *Environmental and Resource Economics*, vol. 68, no. 4, pp. 893–914, 2017.

[30] A. Saussay and M. Sato, "The Impacts of Energy Prices on Industrial Foreign Investment Location: Evidence from Global Firm Level Data," 2018. https://personal.lse.ac.uk/satom/publication/currentss/.

[31] M. Fowlie, M. Reguant, and S. Ryan, "Market-based Environmental Regulation and Industry Dynamics," *Journal of Political Economy*, vol. 124, no. 1, pp. 249–302, 2016.

[32] California Air Resources Board, "California's 2017 Climate Change Scoping Plan," 2017. https://www.arb.ca.gov/cc/scopingplan/scoping_plan_2017.pdf.

[33] Legislative Analyst's Office, "Cap-and-Trade," 2017. http://www.lao.ca.gov/reports/2017/3553/cap-and-trade-021317.pdf.

[34] D. Cullenward and A. Coghlan, "Structural Oversupply and Credibility in California's Carbon Market," *The Electricity Journal*, vol. 29, pp. 7–14, 2016.

[35] California Air Resources Board, "Annual Allocation to Electrical Distribution Utilities (EDU) under the Cap-and-Trade Regulation," 2015. https://www.arb.ca.gov/cc/capandtrade/allowanceallocation/edu-ng-allowancedistribution/electricity-allocation.pdf.

[36] California Energy Commission, "Allowance Allocation," 2017. https://www.arb.ca.gov/cc/capandtrade/allowanceallocation/allowanceallocation.htm#other.

[37] California Air Resources Board, "Mandatory GHG Reporting - Electric Power Entities," 2018. https://ww2.arb.ca.gov/mrr-epe.

[38] California Air Resources Board, "Mandatory GHG Reporting - Asset Controlling Supplier," 2018. https://ww2.arb.ca.gov/mrr-acs.

[39] T. Alcorn, "The Constitutionality of California's Cap-and-Trade Program and Recommendations for Design of Future State Programs," *Michigan Journal of Environmental and Administrative Law*, vol. 3, no. 1, pp. 87–177, 2013.

[40] California Air Resources Board, "California Code of Regulations," 2017. https://www.arb.ca.gov/cc/capandtrade/capandtrade/unofficial_ct_100217.pdf.

[41] National Electric Reliability Council, "Glossary of Terms Used in NERC Reliability Standards," 2017. http://www.nerc.com/files/glossary_of_terms.pdf.

[42] U.S. Energy Information Administration, “Form EIA-860.” https://www.eia.gov/electricity/data/eia860/, 2017.

[43] U.S. Energy Information Administration, “Form EIA-923.” https://www.eia.gov/electricity/data/eia923/, 2017.

[44] U.S. Environmental Protection Agency, “Air Market Program Data,” 2018. https://ampd.epa.gov/ampd/.

[45] U.S. Energy Information Administration, “Average Cost of Fossil Fuels for Electricity Generation,” 2017. https://www.eia.gov/electricity/data.cfm.

[46] S. L. Puller, “Pricing and Firm Conduct in California’s Deregulated Electricity Market,” *Review of Economics and Statistics*, vol. 89, no. 1, pp. 75–87, 2007.

[47] E. T. Mansur, “Do Oligopolists Pollute Less? Evidence from a Restructured Electricity Market,” *Journal of Industrial Economics*, vol. 55, no. 4, pp. 661–689, 2007.

[48] P. L. Joskow and E. Kahn, “A Quantitative Analysis of Pricing Behavior in California’s Wholesale Electricity Market During Summer 2000,” *The Energy Journal*, vol. 23, no. 4, pp. 1–35, 2002.

[49] J. S. Graff Zivin, M. J. Kotchen, and E. T. Mansur, “Spatial and Temporal Heterogeneity of Marginal Emissions: Implications for Electric Cars and Other Electricity-Shifting Policies,” *Journal of Economic Behavior and Organization*, vol. 107 (Part A), pp. 248–268, 2014.

[50] J. A. Cullen and E. T. Mansur, “Inferring Carbon Abatement Costs in Electricity Markets: a Revealed Preference Approach using the Shale Revolution,” *American Economic Journal: Economic Policy*, vol. 9, no. 3, pp. 106–133, 2017.

[51] L. Davis and C. Hausman, “Market Impacts of a Nuclear Power Plant Closure,” *American Economic Journal: Applied Economics*, vol. 8, no. 2, pp. 92–122, 2016.

[52] M. J. Kotchen and E. T. Mansur, “How Stringent Are the U.S. EPA’s Proposed Carbon Pollution Standards for New Power Plants?,” *Review of Environmental Economics and Policy*, vol. 8, no. 2, pp. 290–306, 2014.

[53] California Independent System Operator, “Open Access Same-time Information System (OASIS),” 2018. http://oasis.caiso.com/mrioasis/logon.do.

[54] U.S. Department of Energy, “Transmission Constraints and Congestion in the Western and Eastern Interconnections,” 2014. https://www.energy.gov/oe/downloads/transmission-constraints-and-congestion-western-and-eastern-interconnections-2009-2012.

[55] California Independent System Operator, “Annual and Quarterly Numbers and Performance Reports,” 2018. http://www.caiso.com/market/Pages/MarketMonitoring/AnnualQuarterlynumbersPerfomanceReports/Default.aspx.

[56] California Independent System Operator, “Daily Renewables Watch,” 2018. http://www.caiso.com/market/Pages/ReportsBulletins/RenewablesReporting.aspx.

[57] Federal Energy Regulatory Commission, “Form No. 714 - Annual Electric Balancing Authority Area and Planning Area Report,” 2018. https://www.ferc.gov/docs-filing/forms/form-714/data.asp.

[58] World Meteorological Organization, “Standardized Precipitation Index User Guide,” 2012. http://www.droughtmanagement.info/literature/WMO_standardized_precipitation_index_user_guide_en_2012.pdf.

[59] National Oceanic and Atmospheric Administration, “National Centers for Environmental Information,” 2018. https://www.ncdc.noaa.gov/.

[60] U.S. Energy Information Administration, “Electricity Monthly Update - Regional Wholesale Markets: July 2018,” 2018. https://www.eia.gov/electricity/monthly/update/wholesale_markets.php#tabs_wh_price-3.

[61] Climate Policy Initiative, “California Carbon Dashboard,” 2018. http://calcarbondash.org/.

[62] D. B. Rubin, “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies,” *Journal of Educational Psychology*, vol. 66, no. 5, pp. 688–701, 1974.

[63] California Air Resources Board, “Documentation of California’s 2000-2015 GHG Inventory — Index,” 2018. https://www.arb.ca.gov/cc/inventory/doc/doc_index.php.

[64] National Energy Board of Canada, “Commodity Statistics,” 2018. https://apps.neb-one.gc.ca/CommodityStatistics/Statistics.aspx?language=english.

[65] WECC Staff, “WECC Data Preparation Manual for Steady-State and Dynamic Base Case Data,” 2014. https://www.wecc.biz/Reliability/WECC_Data_Preparation_Manual.docx.

[66] M. Fowlie, S. P. Holland, and E. T. Mansur, "What Do Emissions Markets Deliver and To Whom? Evidence from Southern California's NO_x Trading Program," *American Economic Review*, vol. 102, no. 2, pp. 965–993, 2012.

[67] J. Geweke, "Temporal Aggregation in the Multiple Regression Model," *Econometrica*, vol. 46, no. 3, pp. 643–661, 1978.

[68] J. J. Heckman, H. Ichimura, and P. E. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, vol. 64, no. 4, pp. 605–654, 1997.

[69] J. J. Heckman, H. Ichimura, J. Smith, and P. E. Todd, "Characterizing Selection Bias Using Experimental Data," *Econometrica*, vol. 66, no. 5, pp. 1017–1098, 1998.

[70] M. Blackwell, S. Iacus, G. King, and G. Porro, "CEM: Coarsened Exact Matching in Stata," *The Stata Journal*, vol. 9, no. 4, pp. 524–546, 2009.

[71] S. M. Iacus, G. King, and G. Porro, "Multivariate Matching Methods that are Monotonic Imbalance Bounding," *Journal of the American Statistical Association*, vol. 106, no. 493, pp. 345–361, 2011.

[72] S. M. Iacus, G. King, and G. Porro, "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis*, vol. 20, pp. 1–24, 2012.

[73] T. Simcoe and M. W. Toffel, "Government Green Procurement Spillovers: Evidence from Municipal Building Policies in California," *Journal of Environmental Economics and Management*, vol. 68, pp. 411–434, 2014.

[74] D. Guignet, R. Jenkins, M. Ranson, and P. J. Walsh, "Contamination and Incomplete Information: Bounding Implicit Prices using High-Profile Leaks," *Journal of Environmental Economics and Management*, vol. 88, pp. 259–282, 2018.

[75] C. Ek and J. Miliute-Plepiene, "Behavioral Spillovers from Food-Waste Collection in Swedish Municipalities," *Journal of Environmental Economics and Management*, vol. 89, pp. 168–186, 2018.

[76] J. Linn, E. Mastrangelo and D. Burtraw, "Regulating Greenhouse Gases from Coal Power Plants under the Clean Air Act," 2013. `http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-13-05.pdf`.

[77] M. Nyberg, "Thermal Efficiency of Natural Gas-Fired Generation in California: 2017 Update," 2018. https://www.energy.ca.gov/2017publications/CEC-200-2017-003/CEC-200-2017-003.pdf.

[78] J. D. Angrist and J.-S. Pischke, *Mostly Harmless Econometrics*. Princeton University Press, 2009.

[79] D. Autor, "Outsourcing at Will: the Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing," *Journal of Labor Economics*, vol. 21, no. 1, pp. 1–42, 2003.

[80] M. Bertrand, E. Duflo, and S. Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, pp. 249–275, 2003.

[81] M. A. Petersen, "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches," *Review of Financial Studies*, vol. 22, no. 1, pp. 435–480, 2009.

[82] A. C. Cameron and D. L. Miller, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, vol. 50, no. 2, pp. 317–372, 2015.

[83] Western Energy Imbalance Market, "Market Information," 2018. https://www.westerneim.com/Pages/About/default.aspx.

[84] W. W. Hogan, "An Efficient Western Energy Imbalance Market with Conflicting Carbon Policies," *The Electricity Journal*, vol. 30, pp. 8–15, 2017.

[85] California Independent System Operator, "EIM Greenhouse Gas Enhancements: 3rd Revised Draft Final Proposal," 2018. https://www.caiso.com/informed/Pages/StakeholderProcesses/RegionalIntegrationEIMGreenhouseGasCompliance.aspx.

[86] I. Herrero, P. Rodilla, and C. Batlle, "Enhancing Intraday Price Signals in U.S. ISO Markets for a Better Integration of Variable Energy Resources," *The Energy Journal*, vol. 39, no. 3, pp. 141–165, 2018.

[87] S. M. Iacus, G. King, and G. Porro, "Matching for Causal Inference Without Balance Checking," 2008. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1152391.

[88] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley and Sons, 1992.

[89] D. Cullenward, "Leakage in California's Carbon Market," *The Electricity Journal*, vol. 27, pp. 36–48, 2014.

Figures and Tables

Figure 1: NERC regions in the United States

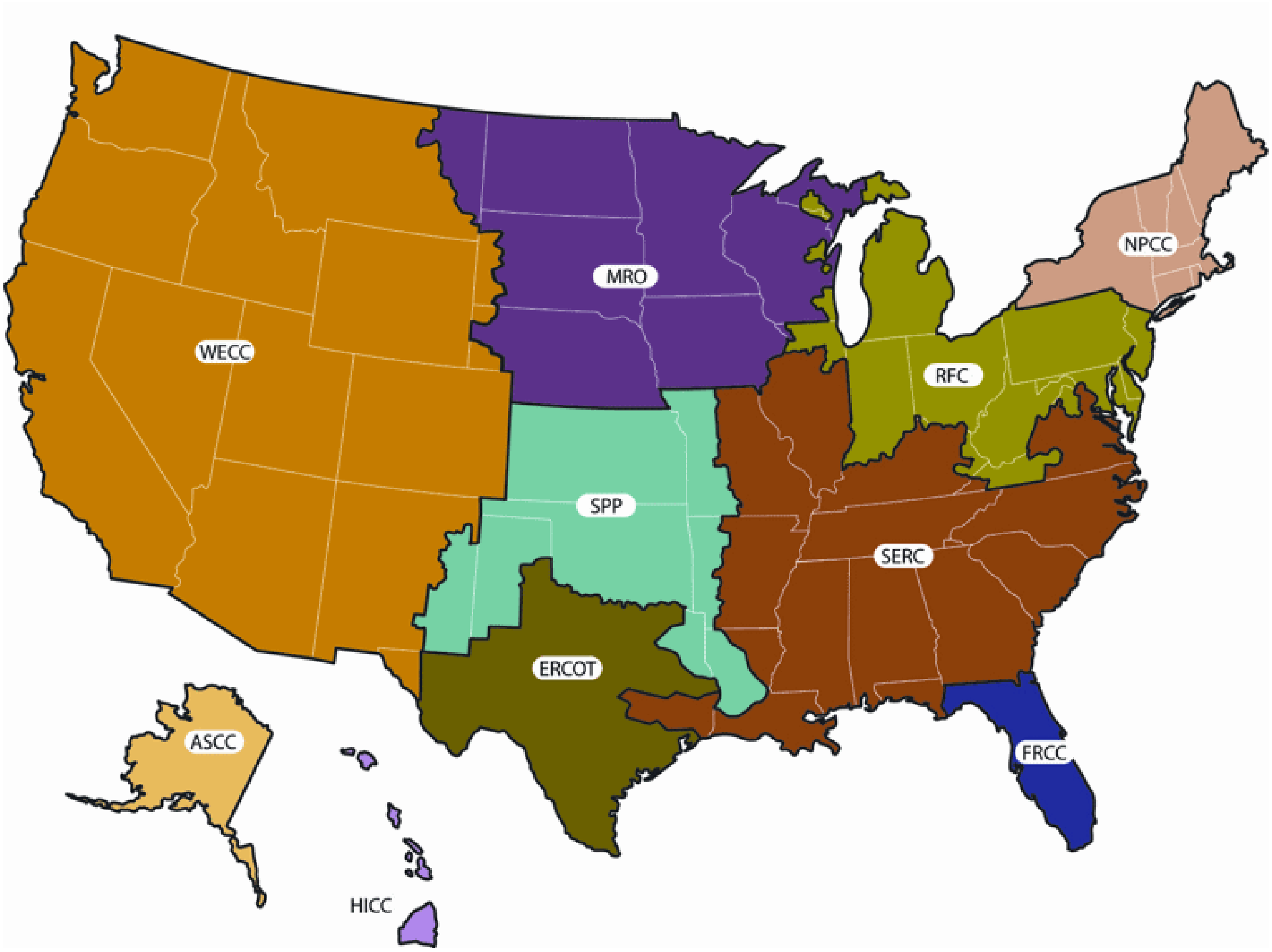
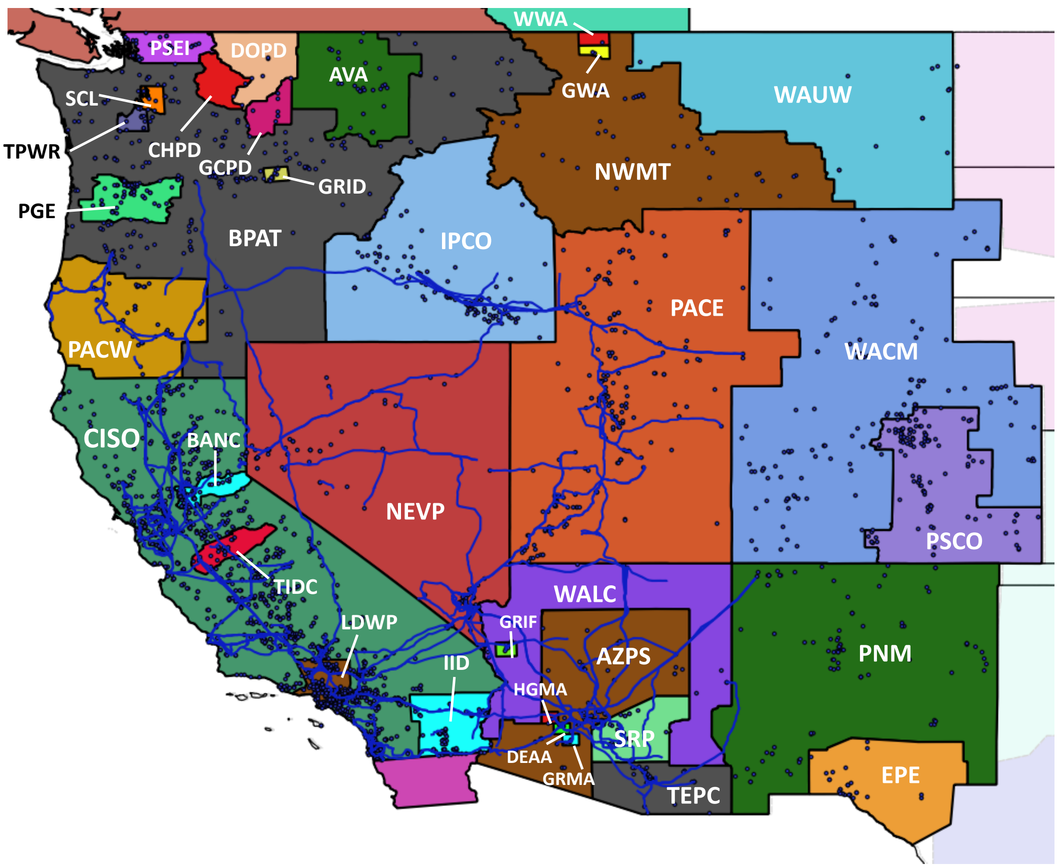
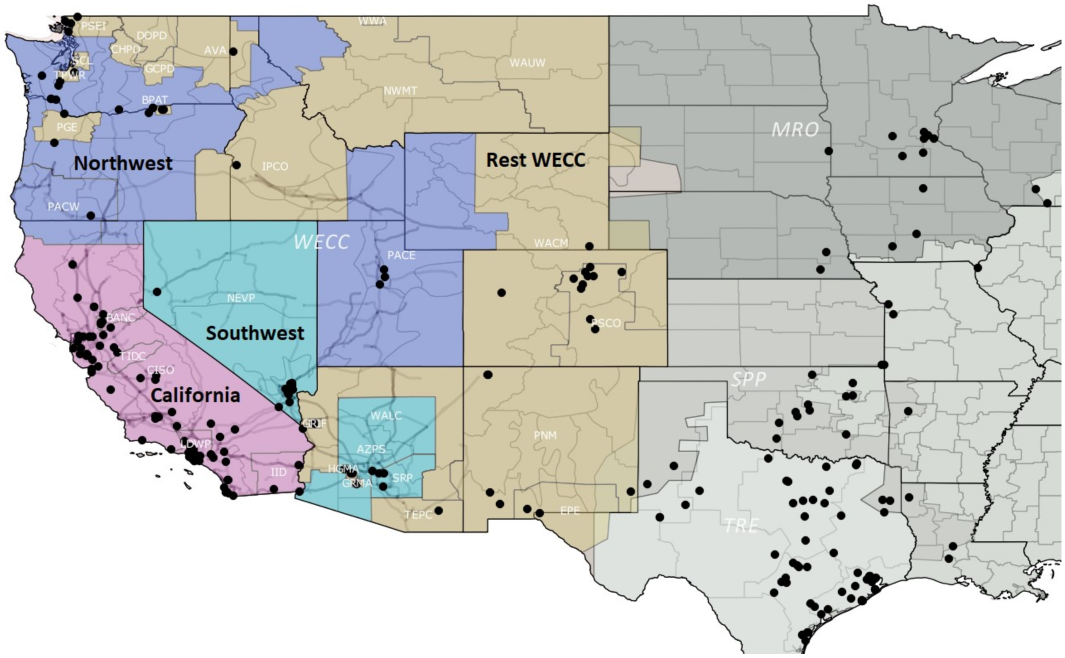


Figure 2: WECC balancing authorities in the United States, 2016



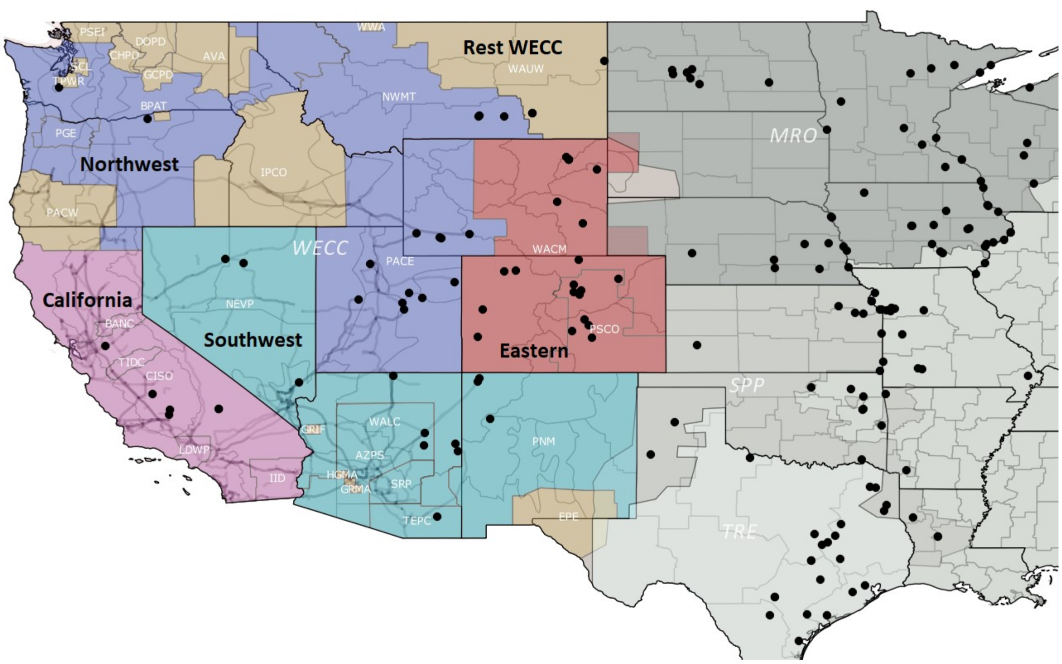
Notes: black dots represent NGCC and coal-fired power plants, blue lines indicate transmission lines.

Figure 3: Treated and control regions in WECC, NGCC plants



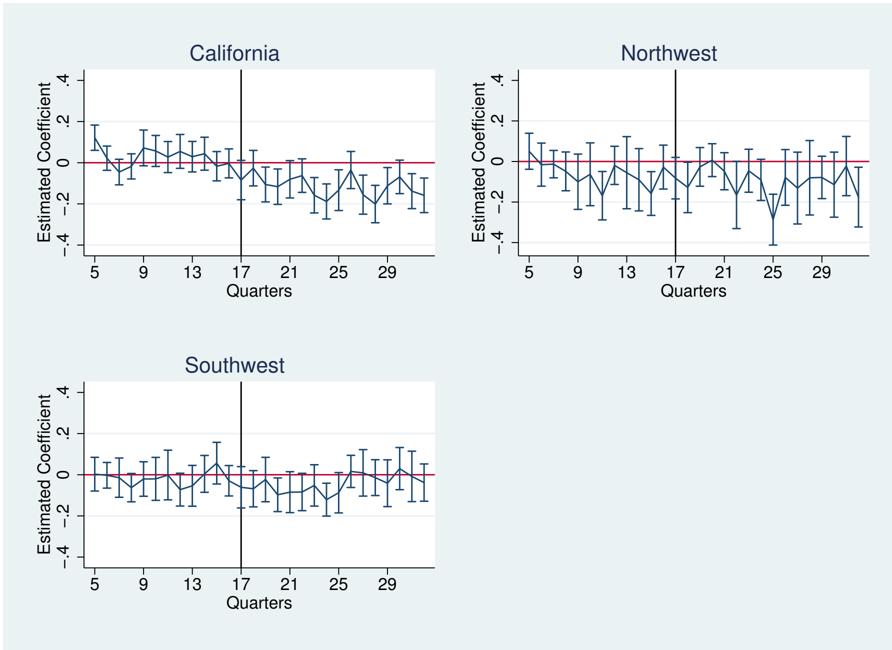
Note: black dots represent power plants.

Figure 4: Treated and control regions in WECC, coal-fired plants

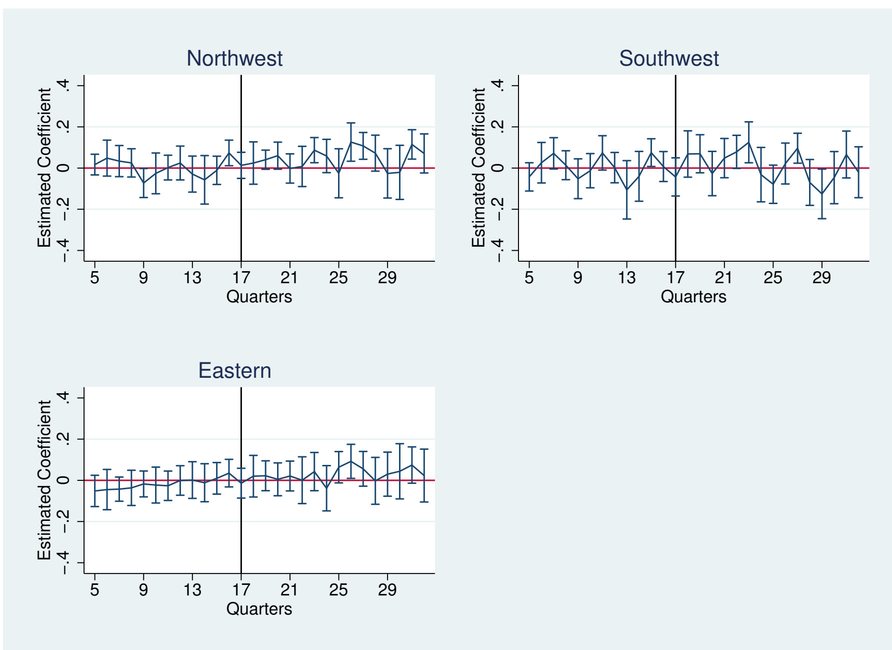


Note: black dots represent power plants.

Figure 5: Estimated quarterly treatment effects - baseline



(a) NGCC plants



(b) Coal-fired plants

Notes: vertical lines represent 95% confidence intervals. Controls include NGCC plants in the rest of WECC, and coal-fired plants in the rest of WECC, MRO, SPP and TRE.

Table 1: Differences-in-differences: Estimated effects of California's cap-and-trade program on NGCC plant capacity factors in WECC

	(1)	(2)	(3)	(4)	(5)
California	−0.1414***	−0.1414***	−0.1338***	−0.1036***	−0.1222***
Northwest	−0.0483	−0.0483	−0.0432	−0.0266	−0.0315
Southwest	−0.0321	−0.0321	−0.0280	−0.0138	0.0111
Electric Demand	0.0757	0.0757	0.0815	0.1170***	0.0605
Nuclear and Renewable Generation	−0.2202***	−0.2202***	−0.2209***	−0.1656***	−0.2225***
CDDs	0.0002***	0.0002***	0.0002***	0.0003***	0.0003***
HDDs	0.0001*	0.0001	0.0001*	0.00001	0.0001*
SPI	−0.0074***	−0.0074**	−0.0073***	−0.0043***	−0.0077***
NG-to-Coal	−0.0602***	−0.0602***	−0.0586***	−0.0660***	−0.0605***
NG-to-Coal^2	0.0022***	0.0022***	0.0021***	0.0026***	0.0022***
Intercept	1.4686***	1.4686***	1.4280***	0.3470	1.5938***
Plant FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
State x Month-of-Year FE	Yes	Yes	Yes	Yes	Yes
N	11,938	11,938	10,858	19,368	11,938
R^2	0.6705	0.6705	0.6683	0.6636	0.6703
Controls	Rest of WECC	Rest of WECC	Rest of WECC	Rest of WECC, SPP, TRE	Rest of WECC
S.E. Clustering	Plant	Plant & Month	Plant	Plant	Plant

Notes: In specifications (1)-(4), we designate as leakers WECC balancing authorities in the U.S. that dispatch power plants supplying specified source power to California, and divide them into two regions, based on [65]. Northwest includes plants in BPAT, PACE and PACW, and Southwest includes non-California plants in AZPS, AZ CAISO, NV CAISO, HGMA, NEVP, SRP and WALC. In specification (5), we designate as leakers WECC balancing authorities that are active or pending participants of the California ISO's Energy Imbalance Market: these include AZPS, IPCO, NEVP, PACE, PACW, PGE, PSEI, SCL and SRP. LDWP and BANC are also pending participants of the Western EIM, but are not considered leakers as their footprint is entirely within California. *, **, and *** indicate statistical significance at 10%, 5% and 1% level, respectively. The unit of observation for these regressions is plant-month.

Table 2: Differences-in-differences: Estimated effects of California's cap-and-trade program on coal-fired plant capacity factors in WECC

	(1)	(2)	(3)	(4)
Northwest	0.0425**	0.0425**	0.0559***	0.0276*
Eastern	0.0393*	0.0393*	0.0478*	-
Southwest	0.0088	0.0088	0.0160	−0.0298
Electric Demand	0.0318	0.0318	0.0285	0.0367
Nuclear and Renewable Generation	−0.0315**	−0.0315**	−0.0316**	−0.0299**
CDDs	0.0002***	0.0002***	0.0002***	0.0002***
HDDs	0.0001***	0.0001***	0.0001***	0.0001***
SPI	−0.0066***	−0.0066***	−0.0073***	−0.0064***
Coal-to-NG	−0.2843***	−0.2843***	−0.2808***	−0.2847***
Coal-to-NG^2	0.0878***	0.0878***	0.0889***	0.0896***
Intercept	0.8888***	0.8888***	0.8773***	0.8752***
Plant FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
State x Month-of-Year FE	Yes	Yes	Yes	Yes
N	14,298	14,298	12,615	14,298
R^2	0.6029	0.6029	0.5769	0.6022
Controls	Rest of WECC, MRO, SPP, TRE	Rest of WECC, MRO, SPP, TRE	Rest of WECC, MRO, SPP, TRE	Rest of WECC, MRO, SPP, TRE
S.E. Clustering	Plant	Plant & Month	Plant	Plant

Notes: In specifications (1)-(3), we designate as leakers WECC balancing authorities in the U.S. that dispatch power plants supplying specified source power to California, and divide them into three regions, based on [65]. Northwest includes plants in BPAT, UT LDWP, NWMT, PACE and PGE. Eastern includes plants in PSCO and WACM. Southwest includes plants in AZPS, NEVP, PNM, SRP, TEPC, and WALC. In specification (4), we designate as leakers WECC balancing authorities that are active or pending participants of the California ISO's Energy Imbalance Market: these include AZPS, IPCO, NEVP, PACE, PACW, PGE, PSEI, SCL and SRP. LDWP and BANC are also pending participants of the Western EIM, but are not considered leakers as their footprint is entirely within California. *, **, and *** indicate statistical significance at 10%, 5% and 1% level, respectively. The unit of observation for these regressions is plant-month.

Table 3: Balancing tests, NGCC plants

	Daytime						Nighttime					
	California		Northwest		Southwest		California		Northwest		Southwest	
	Before matching	After matching	Before matching	After matching	Before matching	After matching	Before matching	After matching	Before matching	After matching	Before matching	After matching
Hour 0	2.753***	−0.473	2.088**	−0.736	0.788	1.724	2.753***	0.077	2.088**	−0.154	0.788	0.195
Hour 1	2.691***	−0.571	2.240**	−0.634	0.710	1.682	2.691***	0.168	2.240**	−0.034	0.710	0.194
Hour 2	2.644***	−0.628	2.293**	−0.605	0.629	1.613	2.644***	0.189	2.293**	−0.001	0.629	0.111
Hour 3	2.643***	−0.627	2.299**	−0.604	0.536	1.520	2.643***	0.164	2.299**	−0.009	0.536	−0.011
Hour 4	2.752***	−0.543	2.216**	−0.690	0.351	1.340	2.752***	0.126	2.216**	−0.155	0.351	−0.289
Hour 5	2.580**	−0.711	2.134**	−0.858	0.059	1.054	2.580**	−0.077	2.134**	−0.436	0.059	−0.550
Hour 6	2.486**	−0.727	2.213**	−0.889	−0.055	0.770	2.486**	−0.146	2.213**	−0.609	−0.055	−0.831
Hour 7	2.575**	−0.560	2.543**	−0.644	−0.134	0.524	2.575**	−0.062	2.543**	−0.365	−0.134	−0.971
Hour 8	2.489**	−0.545	2.682***	−0.522	−0.278	0.351	2.489**	−0.107	2.682***	−0.165	−0.278	−1.209
Hour 9	2.249**	−0.599	2.562**	−0.539	−0.406	0.208	2.249**	−0.282	2.562**	−0.185	−0.406	−1.439
Hour 10	2.044**	−0.666	2.417**	−0.593	−0.401	0.177	2.044**	−0.418	2.417**	−0.267	−0.401	−1.470
Hour 11	1.884*	−0.767	2.333**	−0.644	−0.319	0.152	1.884*	−0.548	2.333**	−0.332	−0.319	−1.377
Hour 12	1.799*	−0.820	2.256**	−0.692	−0.247	0.144	1.799*	−0.615	2.256**	−0.389	−0.247	−1.294
Hour 13	1.786*	−0.821	2.241**	−0.694	−0.171	0.166	1.786*	−0.606	2.241**	−0.409	−0.171	−1.195
Hour 14	1.764*	−0.833	2.266**	−0.670	−0.089	0.252	1.764*	−0.620	2.266**	−0.395	−0.089	−1.107
Hour 15	1.742*	−0.850	2.289**	−0.649	−0.040	0.308	1.742*	−0.637	2.289**	−0.373	−0.040	−1.053
Hour 16	1.666*	−0.920	2.268**	−0.666	−0.015	0.342	1.666*	−0.722	2.268**	−0.378	−0.015	−1.008
Hour 17	1.687*	−0.899	2.232**	−0.691	0.034	0.398	1.687*	−0.717	2.232**	−0.396	0.034	−0.936
Hour 18	1.726*	−0.856	2.169**	−0.728	0.082	0.441	1.726*	−0.761	2.169**	−0.483	0.082	−0.936
Hour 19	1.837*	−0.770	2.131**	−0.764	0.164	0.501	1.837*	−0.669	2.131**	−0.530	0.164	−0.865
Hour 20	1.943*	−0.735	2.205**	−0.727	0.292	0.591	1.943*	−0.735	2.205**	−0.727	0.292	0.591
Hour 21	2.289**	−0.515	2.271**	−0.702	0.619	0.962	2.289**	−0.332	2.271**	−0.336	0.619	−0.335
Hour 22	2.739***	−0.231	2.303**	−0.663	0.842	1.430	2.739***	−0.097	2.303**	−0.115	0.842	0.041
Hour 23	2.929***	−0.219	2.040**	−0.805	0.897	1.724	2.929***	0.071	2.040**	−0.221	0.897	0.234

Notes: the table provides t statistics of a two-sided t test of mean comparisons between treated and control groups. *, **, and *** indicate statistical significance at 10%, 5% and 1% level, respectively.

Table 4: Balancing tests, Coal-fired plants

	Daytime						Nighttime					
	Northwest		Eastern		Southwest		Northwest		Eastern		Southwest	
	Before matching	After matching	Before matching	After matching	Before matching	After matching	Before matching	After matching	Before matching	After matching	Before matching	After matching
Hour 0	3.801***	2.260**	2.702***	1.364	1.277	1.141	3.801***	0.788	2.702***	0.530	1.277	0.170
Hour 1	3.942***	2.406**	2.906***	1.585	1.368	1.263	3.942***	1.037	2.906***	0.654	1.368	0.261
Hour 2	4.002***	2.471**	2.962***	1.653	1.368	1.264	4.002***	1.165	2.962***	0.683	1.368	0.249
Hour 3	3.933***	2.388**	2.953***	1.651	1.309	1.195	3.933***	1.031	2.953***	0.668	1.309	0.177
Hour 4	3.617***	2.022**	2.787***	1.506	1.197	1.067	3.617***	0.545	2.787***	0.565	1.197	0.062
Hour 5	3.215***	1.561	2.464**	1.210	1.028	0.853	3.215***	0.089	2.464**	0.414	1.028	−0.081
Hour 6	2.716***	0.871	2.006**	0.772	0.658	0.296	2.716***	−0.390	2.006**	0.221	0.658	−0.544
Hour 7	2.445**	0.463	1.763*	0.553	0.566	0.134	2.445**	−0.544	1.763*	0.158	0.566	−0.671
Hour 8	2.289**	0.229	1.600	0.397	0.601	0.162	2.289**	−0.635	1.600	0.091	0.601	−0.625
Hour 9	2.188**	0.093	1.499	0.306	0.670	0.257	2.188**	−0.683	1.499	0.054	0.670	−0.523
Hour 10	2.154**	0.046	1.442	0.251	0.779	0.445	2.154**	−0.685	1.442	0.025	0.779	−0.334
Hour 11	2.138**	0.018	1.401	0.200	0.849	0.565	2.138**	−0.722	1.401	−0.014	0.849	−0.230
Hour 12	2.119**	−0.023	1.384	0.166	0.873	0.601	2.119**	−0.779	1.384	−0.028	0.873	−0.197
Hour 13	2.142**	0.017	1.404	0.186	0.937	0.713	2.142**	−0.777	1.404	−0.022	0.937	−0.090
Hour 14	2.206**	0.131	1.449	0.234	1.006	0.836	2.206**	−0.716	1.449	−0.005	1.006	0.015
Hour 15	2.270**	0.236	1.495	0.282	1.049	0.906	2.270**	−0.643	1.495	0.011	1.049	0.078
Hour 16	2.276**	0.245	1.520	0.309	1.099	0.992	2.276**	−0.634	1.520	0.023	1.099	0.164
Hour 17	2.211**	0.145	1.491	0.293	1.091	0.989	2.211**	−0.660	1.491	0.019	1.091	0.168
Hour 18	2.144**	0.049	1.423	0.229	1.064	0.951	2.144**	−0.721	1.423	−0.020	1.064	0.126
Hour 19	2.141**	0.042	1.440	0.248	1.024	0.871	2.141**	−0.735	1.440	0.004	1.024	0.049
Hour 20	2.332**	0.360	1.583	0.385	1.093	0.983	2.332**	−0.544	1.583	0.074	1.093	0.142
Hour 21	2.674**	0.913	1.803	0.569	1.077	0.920	2.674**	−0.257	1.803	0.132	1.077	0.025
Hour 22	3.121***	1.496	2.123**	0.820	1.164	1.021	3.121***	0.131	2.123**	0.244	1.164	0.086
Hour 23	3.526***	1.963	2.395**	1.037	1.210	1.053	3.526***	0.519	2.395**	0.351	1.210	0.103

Notes: the table provides t statistics of a two-sided t test of mean comparisons between treated and control groups. *, **, and *** indicate statistical significance at 10%, 5% and 1% level, respectively.

Table 5: Matching and Differences-in-differences: Estimated effects of California's cap-and-trade program on NGCC plant capacity factors in WECC

	Daytime			Nighttime		
	California	Northwest	Southwest	California	Northwest	Southwest
Hour set 1						
Before matching						
Control plants	72	72	72	72	72	72
Treated plants	37	11	19	37	11	19
After matching						
Control plants	68	57	69	58	51	67
Treated plants	31	9	18	29	9	17
Estimated treatment effect after matching						
	−0.0627**	−0.0265	−0.0786	−0.0135	−0.0161	0.0218
N	190,632	129,214	168,884	166,544	117,164	162,938
R^2	0.4278	0.3565	0.4276	0.4669	0.3952	0.3993
Hour set 2						
Before matching						
Control plants	73	73	73	73	73	73
Treated plants	37	11	19	37	11	19
After matching						
Control plants	65	59	70	66	57	60
Treated plants	37	11	18	28	9	16
Estimated treatment effect after matching						
	−0.0791***	−0.042	−0.0747	−0.0006	0.0215	0.0301
N	196,491	136,428	170,752	180,466	128,451	146,810
R^2	0.5454	0.5585	0.4246	0.4384	0.5635	0.3579
Hour set 3						
Before matching						
Control plants	73	73	73	73	73	73
Treated plants	37	11	19	37	11	19
After matching						
Control plants	62	62	69	70	57	65
Treated plants	36	11	17	30	9	17
Estimated treatment effect after matching						
	−0.0776**	−0.0387	−0.0403	−0.0247	−0.0236	−0.0003
N	188,554	142,455	166,740	192,562	128,917	158,757
R^2	0.5505	0.5664	0.4465	0.5395	0.5542	0.4640

Notes: matching is on hourly capacity factors. Hour set 1 matches on hours 7,10,13,16 (Day) and 19,22,1,4 (Night). Hour set 2 matches on hours 8,11,14,17 (Day) and 20,23,2,5 (Night). Hour set 3 matches on hours 9,12,15,18 (Day) and 21,0,3,6 (Night). We coarsen each matching variable according to cutpoints 0.3, 0.5 and 0.7, which identify four strata corresponding to different levels of plant utilization. DID regressions include plant FE, year FE and state by month-year FE. Standard errors are clustered at the plant level.

Table 6: Matching and Differences-in-differences: Estimated effects of California's cap-and-trade program on coal-fired plant capacity factors in WECC

	Daytime			Nighttime		
	Northwest	Eastern	Southwest	Northwest	Eastern	Southwest
Hour set 1						
Before matching						
Control plants	88	88	88	88	88	88
Treated plants	13	14	9	13	14	9
After matching						
Control plants	63	76	65	35	62	55
Treated plants	13	12	9	12	13	9
Estimated treatment effect after matching						
	0.0383**	0.0281	0.0256	0.0611*	0.0498*	0.0494*
N	141,225	157,597	137,699	85,500	133,511	117,618
R^2	0.2919	0.3011	0.2682	0.3569	0.4226	0.3275
Hour set 2						
Before matching						
Control plants	88	88	88	88	88	88
Treated plants	13	14	9	13	14	9
After matching						
Control plants	66	82	68	42	69	48
Treated plants	13	13	9	12	14	9
Estimated treatment effect after matching						
	0.0376*	0.0312	0.0259	0.0814***	0.0489*	0.0493*
N	146,874	171,179	143,348	99,869	148,961	104,014
R^2	0.2913	0.3088	0.2731	0.3611	0.4105	0.3368
Hour set 3						
Before matching						
Control plants	88	88	88	88	88	88
Treated plants	13	14	9	13	14	9
After matching						
Control plants	68	83	68	43	79	72
Treated plants	12	13	8	13	14	9
Estimated treatment effect after matching						
	0.0415**	0.0221	0.038	0.0719***	0.0335	0.0446
N	149,289	173,447	141,980	103,544	167,337	143,309
R^2	0.2916	0.3078	0.2611	0.3495	0.4066	0.3503

Notes: matching is on hourly plant capacity factors. Hour set 1 matches on hours 7,10,13,16 (Day) and 19,22,1,4 (Night). Hour set 2 matches on hours 8,11,14,17 (Day) and 20,23,2,5 (Night). Hour set 3 matches on hours 9,12,15,18 (Day) and 21,0,3,6 (Night). We coarsen matching variables according to cutpoints 0.6 and 0.8, which identify three strata corresponding to different levels of plant utilization. DID regressions include plant FE, year FE and state by month-year FE. Standard errors are clustered at the plant level.

Table 7: Scheduled power flow regressions

	(1)		(2)	
	Northwest Flows	Southwest Flows	Northwest Flows	Southwest Flows
Electric Demand CAISO	0.1114***	0.0619***	0.1122***	0.0644***
Electric Demand Northwest	−0.3048***		−0.3357***	
Electric Demand Southwest		−0.0303***		−0.0286**
Northwest Flows		−0.1738***		−0.1932***
Southwest Flows	−0.1987		−0.1522	
Nuclear Generation CAISO	−0.0479	−0.0474*	−0.0392	−0.0461*
Wind Generation CAISO	−0.0754***	−0.0861***	−0.0687***	−0.0884***
Solar Generation CAISO	−0.0022	0.0435	0.0028	0.0287
Hydro Generation Northwest	0.0071***		0.0073***	
Solar Generation Southwest		0.1978***		0.2052***
CO_2 Price	3.3390***	−0.3633	3.4188***	−0.3122
PG&E Citygate NG Price	2.4457			
Sumas NG Price	−4.6487***			
SoCal Border NG Price		11.8540***		
El Paso San Juan NG Price		−10.1424***		
NG-to-NG North			0.6332	
NG-to-NG South				14.0579
Intercept	91.1052***	44.8026***	85.1667***	39.0537***
Day-of-week dummy	Yes	No	Yes	No
Quarter dummy	Yes	Yes	Yes	Yes
N	2,017		2,017	
R^2	0.8570		0.8521	
Log-likelihood	-82,522.57		-77,238.14	
AIC	165,195.1		154,622.3	
BIC	165,615.8		155,031.8	

Appendix

Figure A1: Generation mix by technology in NERC regions and WECC sub-regions, 2016

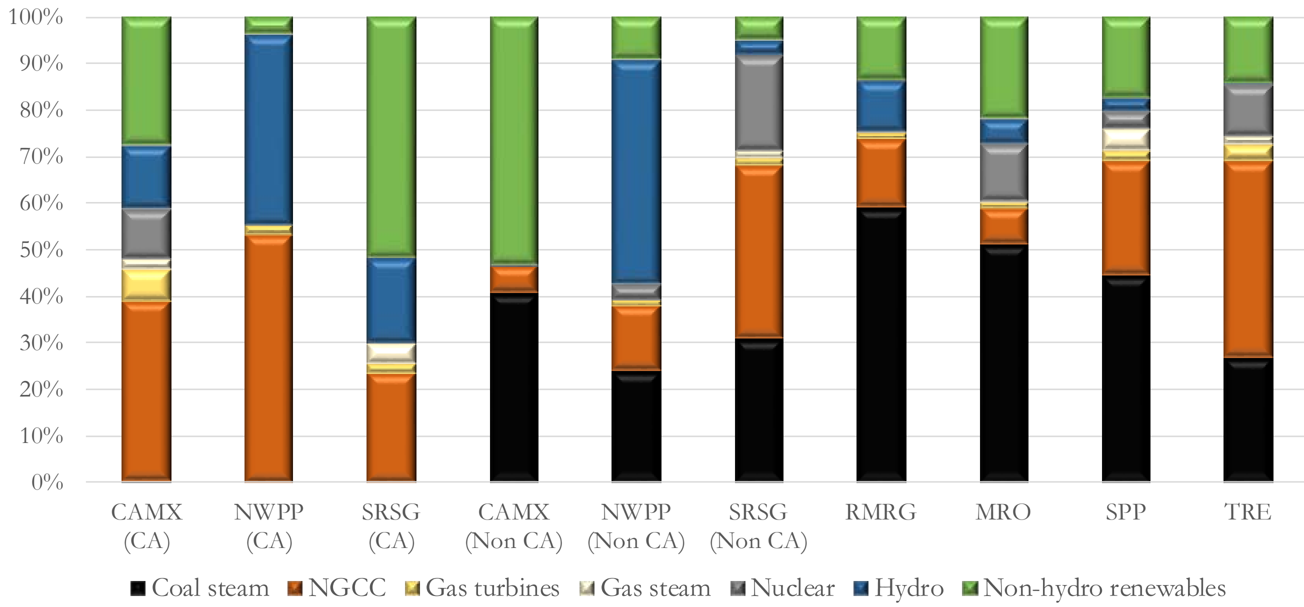
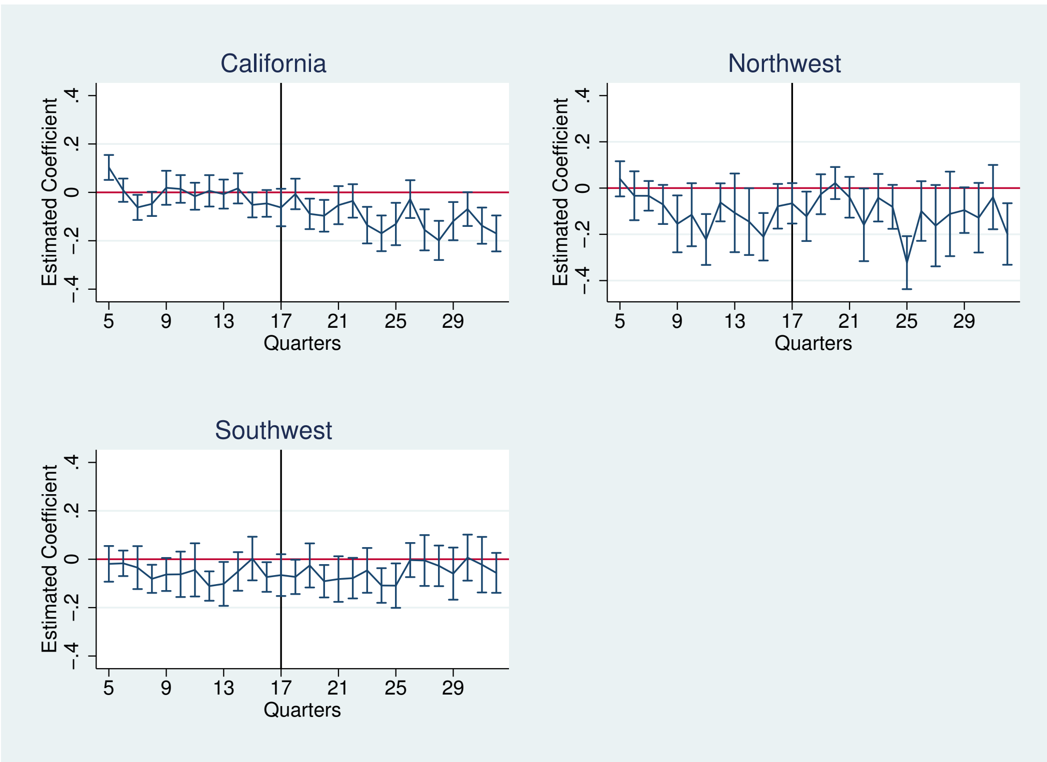
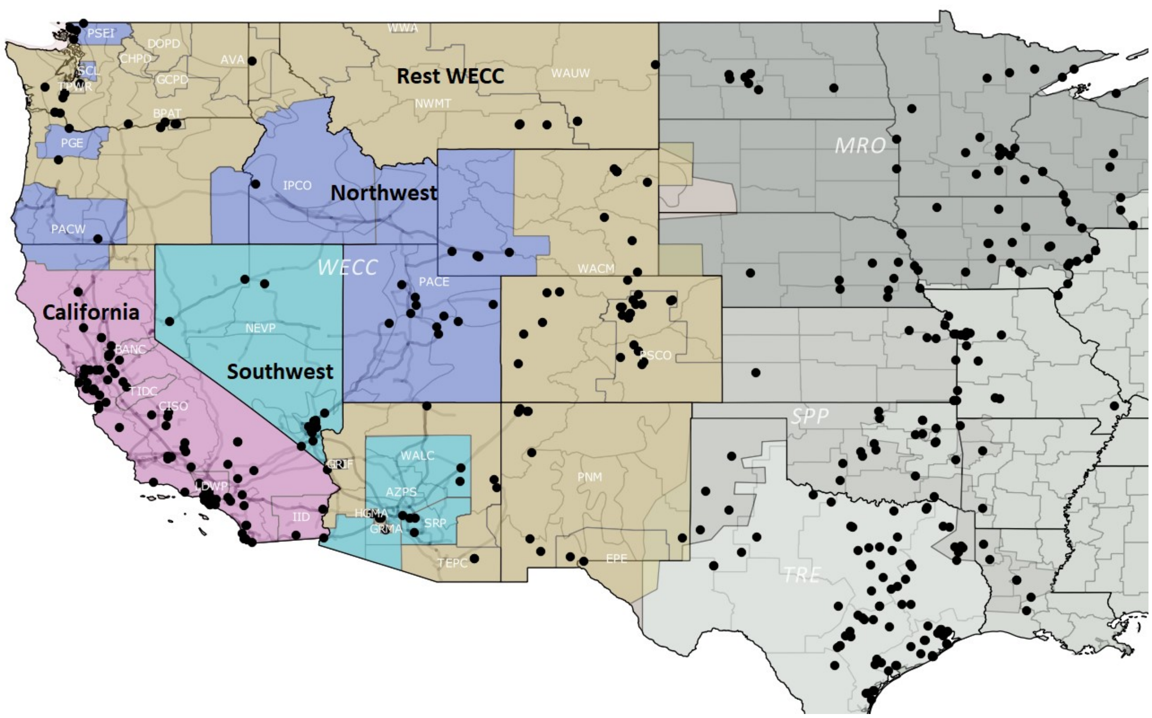


Figure A2: Estimated quarterly NGCC treatment effects - broader control group



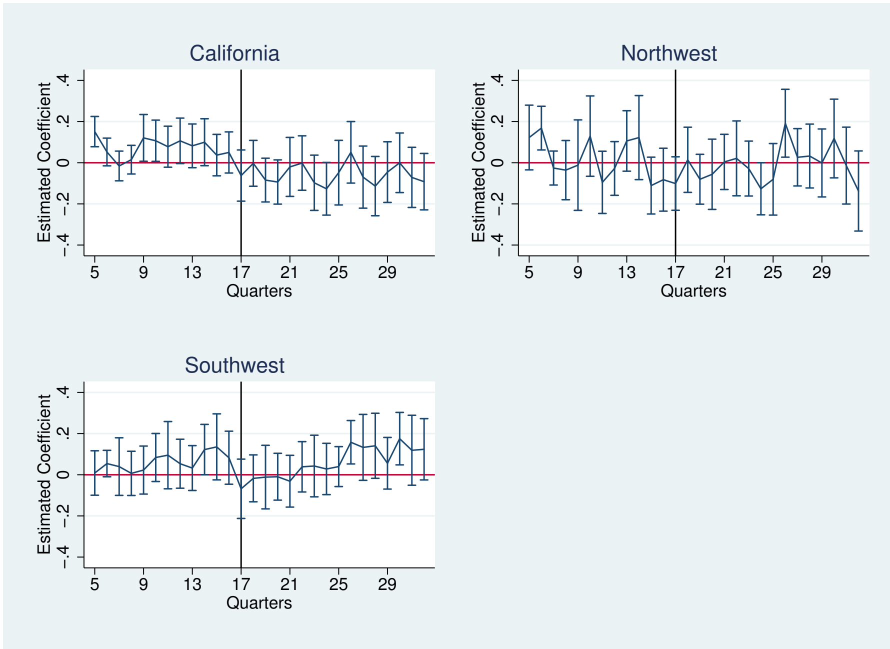
Notes: vertical lines represent 95% CIs. Controls include plants in the rest of WECC, SPP and TRE.

Figure A3: Treated and control regions in WECC, NGCC and coal-fired plants - EIM leakers

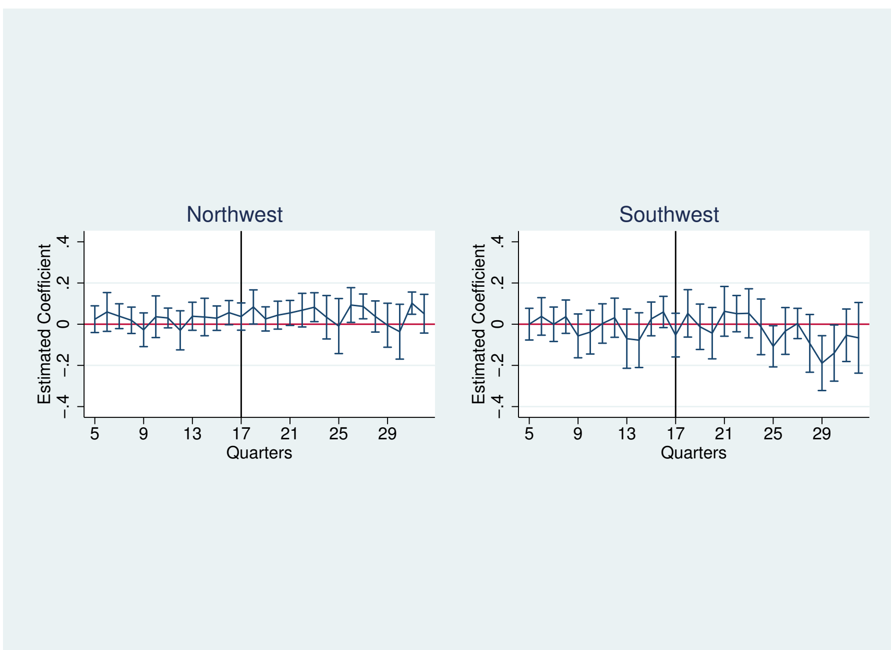


Note: black dots represent power plants.

Figure A4: Estimated quarterly treatment effects - EIM leakers



(a) NGCC plants



(b) Coal-fired plants

Notes: vertical lines represent 95% confidence intervals. Controls include NGCC plants in the rest of WECC, and coal steam plants in the rest of WECC, SPP and TRE.

Table A1: Summary statistics by major WECC balancing authority (BA), NGCC plants

BA	Pre ETS					Post ETS				
	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)
BANC	7	171	0.44	8,124	120.13	7	178	0.55	7,747	118.61
CISO	56	288	0.58	9,009	118.85	57	316	0.49	8,625	118.79
IID	1	154	0.38	7,203	118.90	1	259	0.43	8,112	118.86
LDWP	5	353	0.32	7,497	118.86	6	356	0.31	8,210	117.64
TIDC	1	269	0.60	7,881	118.86	1	269	0.60	7,829	118.86
AVA	2	290	0.49	7,264	118.85	2	294	0.61	6,908	118.86
BPAT	8	441	0.38	6,315	118.86	7	457	0.49	6,737	118.84
IPCO	1	331	0.25	6,713	118.86	1	312	0.57	6,973	118.86
PACE	3	421	0.44	7,660	118.86	3	581	0.42	7,671	118.86
PACW	2	414	0.66	8,816	118.86	1	486	0.54	7,743	118.86
PGE	3	370	0.46	9,292	118.83	4	386	0.46	8,106	118.86
PSEI	5	207	0.37	8,537	118.91	5	215	0.40	9,092	118.99
PSCO	9	280	0.23	8,692	118.86	10	302	0.25	7,704	118.87
WACM	1	497	0.35	8,041	118.86	2	341	0.34	7,779	118.86
AZPS	5	1,074	0.27	7,602	118.77	5	1,074	0.33	7,736	118.79
EPE	1	445	0.36	10,003	118.85	1	522	0.38	9,319	118.85
HGMA	1	1,128	0.17	5,503	118.86	1	1,106	0.20	5,701	118.86
NEVP	12	474	0.52	8,047	118.88	12	484	0.56	7,823	118.89
PNM	3	306	0.18	9,528	118.86	3	307	0.23	8,415	118.86
SRP	3	966	0.41	7,787	118.86	3	964	0.37	7,686	118.87
WALC	4	218	0.49	9,043	118.80	4	240	0.42	7,029	118.89

Notes: pre ETS refers to January 2009-December 2012, post ETS to January 2013-December 2016. Emission rates are only available for a subset of plants from CEMS.

Table A2: Summary statistics by major WECC balancing authority (BA), Coal-fired plants

BA	Pre ETS					Post ETS				
	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)
CISO	6	52	0.78	17,598	-	4	57	0.48	19,133	-
LDWP	1	1,800	0.78	9,728	205.20	1	1,800	0.70	9,497	205.20
BPAT	1	1,358	0.52	8,666	205.62	1	1,340	0.50	9,218	209.76
NWMT	5	493	0.74	12,483	208.75	5	525	0.71	12,776	208.58
PACE	12	630	0.81	11,997	207.73	12	621	0.78	13,343	207.93
PGE	1	585	0.65	8,763	209.57	1	585	0.55	8,259	209.11
PSCO	10	382	0.66	13,825	206.52	8	502	0.70	12,081	206.59
WACM	9	273	0.81	11,679	208.61	8	279	0.76	11,011	208.97
AZPS	3	1,083	0.67	14,573	205.21	2	1,321	0.63	10,315	205.20
NEVP	3	439	0.54	11,165	209.30	3	389	0.48	11,390	209.49
PNM	2	950	0.74	10,687	209.74	2	966	0.66	10,919	209.69
SRP	2	1,510	0.84	10,327	209.76	2	1,506	0.77	10,471	209.76
TEPC	1	1,609	0.73	10,325	209.66	1	1,621	0.72	10,304	209.56
WALC	1	350	0.65	10,944	196.55	1	350	0.72	10,890	203.62

Notes: pre ETS refers to January 2009-December 2012, post ETS to January 2013-December 2016. Emission rates are only available for a subset of plants from CEMS.

Table A3: Summary statistics by region, NGCC plants

Sub-region	Pre ETS					Post ETS				
	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)
California	68	278	0.54	8,779	119.09	70	304	0.49	8,511	118.66
		(293)	(0.31)	(2,819)	(1.96)		(290)	(0.32)	(3,115)	(1.45)
Northwest	13	433	0.43	6,922	118.86	11	494	0.48	7,084	118.85
		(178)	(0.31)	(3,320)	(0.05)		(258)	(0.28)	(2,078)	(0.07)
Southwest	27	625	0.44	7,959	118.83	27	633	0.44	7,600	118.85
		(501)	(0.30)	(2,814)	(0.73)		(496)	(0.30)	(2,429)	(1.18)
Rest of WECC	25	296	0.31	8,731	118.86	28	307	0.36	8,077	118.88
		(190)	(0.29)	(4,809)	(0.85)		(186)	(0.29)	(4,001)	(1.59)
MRO, SPP, TRE	98	555	0.38	8,555	117.73	100	559	0.39	8,575	117.69
		(350)	(0.26)	(3,334)	(11.75)		(367)	(0.26)	(3,286)	(11.45)

Notes: standard deviations listed below means. Pre ETS refers to January 2009-December 2012, and post ETS to January 2013-December 2016. Leaker regions are defined as in Section 4. Emission rates are only available for a subset of plants from CEMS.

Table A4: Summary statistics by region, Coal-fired plants

Sub-region	Pre ETS					Post ETS				
	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)	Number of plants	Operating capacity (MW)	Capacity factor (%)	Heat rate (Btu/kWh)	CO_2 emission rate (lb/MMBtu)
California	6	52	0.78	17,598	-	4	56	0.48	19,133	206.44
		(26)	(0.26)	(11,998)	(-)		(25)	(0.31)	(18,623)	(11.76)
Northwest	20	691	0.76	11,638	207.74	20	695	0.73	12,454	208.00
		(695)	(0.23)	(5,090)	(4.54)		(700)	(0.24)	(7,517)	(5.68)
Eastern	19	332	0.73	12,830	207.70	16	379	0.73	11,495	207.93
		(389)	(0.20)	(5,591)	(6.88)		(418)	(0.22)	(4,259)	(6.70)
Southwest	12	957	0.68	11,700	207.55	11	992	0.64	10,780	208.35
		(732)	(0.20)	(3,743)	(8.09)		(689)	(0.21)	(1,484)	(3.44)
Rest of WECC	2	1,414	0.81	10,287	209.76	2	1,057	0.79	10,362	209.76
		(550)	(0.14)	(255)	(0.05)		(656)	(0.17)	(317)	(0.06)
MRO, SPP, TRE	119	591	0.60	12,447	209.76	115	638	0.56	12,367	207.23
		(578)	(0.26)	(5,332)	(12.93)		(583)	(0.26)	(5,725)	(25.19)

Notes: standard deviations listed below means. Pre ETS refers to January 2009-December 2012, and post ETS to January 2013-December 2016. Leaker regions are defined as in Section 4. Emission rates are only available for a subset of plants from CEMS.

Table A5: Matching and Differences-in-differences: Estimated effects with percentile binning

	Daytime			Nighttime		
NGCC						
	California	Northwest	Southwest	California	Northwest	Southwest
Before matching						
Control plants	73	73	73	73	73	73
Treated plants	37	11	19	37	11	19
After matching						
Control plants	66	49	58	57	27	51
Treated plants	32	7	15	27	7	18
Estimated treatment effect after matching						
	−0.0727**	−0.0154	−0.0276	−0.0107	−0.0054	0.0011
N	188,737	108,498	140,886	160,699	66,415	132,658
R^2	0.4779	0.4669	0.4983	0.5585	0.4529	0.5247
Coal steam						
	Northwest	Eastern	Southwest	Northwest	Eastern	Southwest
Before matching						
Control plants	88	88	88	88	88	88
Treated plants	13	14	9	13	14	9
After matching						
Control plants	44	48	32	13	16	21
Treated plants	10	12	6	6	9	8
Estimated treatment effect after matching						
	0.0456**	0.0209	−0.0386	0.0851**	0.1055***	−0.0005
N	95,396	107,151	68,265	33,932	44,955	51,610
R^2	0.3684	0.3262	0.3001	0.3434	0.4022	0.3875

Notes: matching is on hourly plant capacity factors and heat rates for hours 8,11,14,17 (Day) and 20,23,2,5 (Night). We coarsen matching variables according to cutpoints defined by the 40th, 60th and 80th percentiles for NGCC, and 20th, 40th, 60th and 80th percentiles for coal steam. DiD regressions include plant FE, year FE and state by month-year FE. Standard errors are clustered at the plant level.

Table A6: Matching and Differences-in-differences: Estimated effects with algorithm binning

	Daytime			Nighttime		
NGCC	California	Northwest	Southwest	California	Northwest	Southwest
Before matching						
Control plants	73	73	73	73	73	73
Treated plants	37	11	19	37	11	19
L1	0.43	0.52	0.47	0.52	0.75	0.40
After matching						
Control plants	57	39	52	46	18	46
Treated plants	28	8	15	27	8	17
L1	0.10	0.03	0.31	0.17	0.27	0.25
Algorithm	Scott	Scott	Scott	Scott	Scott	Scott
Estimated treatment effect after matching						
	-0.0754^{*}	-0.0533	-0.0042	0.0013	-0.0122	0.0189
N	163,513	93,656	128,978	139,130	51,803	120,913
R^2	0.4350	0.3721	0.4458	0.4779	0.3935	0.4648
Coal steam	Northwest	Eastern	Southwest	Northwest	Eastern	Southwest
Before matching						
Control plants	88	88	88	88	88	88
Treated plants	13	14	9	13	14	9
L1	0.50	0.59	0.70	0.72	0.74	0.77
After matching						
Control plants	52	44	26	25	23	20
Treated plants	12	13	7	13	12	8
L1	0.35	0.04	0.03	0.40	0.08	0.23
Algorithm	FD	FD	FD	Scott	FD	FD
Estimated treatment effect after matching						
	0.0379^{*}	0.0229	0.0362	0.0633^{**}	0.0508^{*}	0.0502
N	119,770	106,751	57,783	71,559	68,202	49,466
R^2	0.3226	0.3025	0.2772	0.3539	0.3413	0.3641

Notes: matching is on hourly plant capacity factors for hours 8,11,14,17 (Day) and 20,23,2,5 (Night). We coarsen matching variables according to cutpoints defined by the Scott rule and Freedman-Diaconis (FD) rule. DiD regressions include plant FE, year FE and state by month-year FE. Standard errors are clustered at the plant level.