Grid Lock-in and AI Development*

Xinming Du† Jingyuan Guo‡

October 30, 2026

Abstract

Countries compete for leadership in AI, but the grid capacity needed to power new data centers may already be committed to other industries. When grid capacity expands slowly, an industry that arrives first can deter the next from investing. We study this grid lock-in using close statewide votes on recreational cannabis legalization in the United States and a new satellite measure of electricity-intensive indoor cultivation. Narrowly passing ballots reduce the active data center stock by 0.70 facilities per county, more than half its counterfactual level, and operating capacity by 16.2 megawatts. Five years after the ballot, the increase in detected cultivation implies 22 megawatts of additional electricity demand per county. The decline in data centers extends to counties in other states on the same grid, where the ballot does not legalize cultivation. In narrowly passing states, higher rates where cultivation expands imply 0.7 billion dollars a year in additional bills for customers other than data centers. In a calibrated siting model, national data center capacity would be 30 percent above its 2024 level without legalization by ballot if displaced facilities are not built elsewhere.

JEL Classifications: Q41, L94, L86, R11, K42

Keywords: Electricity markets, cannabis legalization, data centers, AI infrastructure, transmission capacity

*We thank Mark Colas, Roger Fouquet, David Lagakos, Chen Liu, Haoming Liu, Joris Mueller, Jonathan Norris, Ivan Png, and Wolfram Schlenker for helpful comments. We used AI-assisted tools for editing in accordance with journal guidelines. All analysis, arguments, and conclusions are our own.

†Department of Economics, National University of Singapore. Email: xdu@nus.edu.sg.

‡Department of Economics, National University of Singapore. Email: jingyuan.guo@u.nus.edu.

1 Introduction

AI has become a field of competition between nations. The U.S. government describes leadership in AI as a race with China and lists the rapid construction of data centers and the power plants that serve them among its priorities (The White House, 2025). Just as Britain did with the steam engine of the 1760s, the winner of this competition may set the rules of the world order that follows.

This competition depends on an electricity grid that expands more slowly than the demand placed on it. Data centers consume 4-5 percent of U.S. electricity and could reach 9-17 percent by 2030 (Electric Power Research Institute, 2026). Only 322 miles of transmission at 345 kilovolts and above were completed in 2024, compared with nearly 4,000 miles in 2013, and generation and transmission take years to build (Shreve et al., 2025; Rand et al., 2024).

When a shared input is slow to expand, an industry that arrives first can deter the next from investing. Existing work has studied how labor reallocation and housing constraints shape local economic development (Allcott and Keniston, 2018; Hsieh and Moretti, 2019). This paper studies how competition for grid capacity constrains investment in data centers, which support AI and other computing services. We call this mechanism grid lock-in.

Identifying grid lock-in is difficult, because grid conditions and industrial investment are jointly determined. Places that attract one large electricity user may also differ in ways that affect later entry. A lack of subsequent investment does not establish that grid constraints deterred entry, since developers may never have planned a facility there. Identification therefore requires a shock to local electricity demand and a credible counterfactual for subsequent investment. Recreational cannabis legalization provides such a shock by expanding an electricity-intensive industry. Indoor cultivation draws 1 percent of total U.S. electricity, and per square foot it uses eight times the energy of a typical commercial building (Mills, 2012). Close legalization ballots generate plausibly exogenous variation in this localized electricity load. Indoor cultivation and data centers have little overlap in their workforces, technologies, or product markets. We compare counties in states where a ballot barely passed with counties in states where it barely failed.

We model how cultivation and data centers compete for limited grid capacity. Transmission capacity is fixed in the short run. Retail rates rise with load in the utility's service territory as utilities recover costs through rate cases. Large new customers also face delays from connection studies and required upgrades. Developers consider connection costs and expected electricity bills over a facility's operating life. In this framework, the ballot outcome informs developers about future cultivation load. After passage, cultivation expands, retail rates rise, and data center investment falls. The investment decline is larger where cultiva-

tion can expand further and extends to counties across state lines on the same network. We examine competition for workers, land, and water separately.

We assemble a national county-year panel spanning 2000-2024 that links data centers, electricity markets, cannabis ballots, and indoor cultivation. Administrative data on cultivation are incomplete, inconsistent across state regulatory regimes, and unavailable where cultivation remains illegal. We therefore detect indoor cultivation in satellite imagery from 2017 onward, combining nightlight intensity, surface heat, and rooftop features (Henderson et al., 2012; Donaldson and Storeygard, 2016). We apply the same detection rule across states and over time, without relying on state reports.

We identify the effect with a panel difference-in-discontinuities, which compares the change after narrowly passing ballots with the change after narrowly failing ones (Grembi et al., 2016). A preexisting outcome gap at the threshold is allowed, so long as the gap would not have changed had the ballot failed. Before the ballot, counties in passing and failing states do not differ measurably in grid conditions, including generator interconnection requests and outages, or in local economic conditions.

In counties in states where a recreational ballot barely passed, the active data center stock falls by 0.70 facilities and operating capacity by 16.2 megawatts (MW) per county, relative to counties in states where a ballot barely failed. The decline in the stock is more than half of what a county in a narrowly passing state would otherwise have held over the post-ballot years. Passing and failing counties do not diverge measurably in the eight years before the ballot. The average post-ballot decline remains distinguishable from zero when we allow differences in trends as large as the largest seen before the ballot (Rambachan and Roth, 2023). The decline persists through the eighth post-ballot year. At U.S. construction costs, the lost capacity represents over 150 million dollars of forgone investment per county.

Detected indoor cultivation rises by 3.6 times the sample mean within five post-ballot years. At the energy intensity of indoor cultivation, the five-year increase implies an average electricity demand of 22 MW per county, of the same scale as the 16.2 MW of nameplate data center capacity lost. In legalizing counties, retail electricity prices rise 2.2 percent more after the ballot for each square kilometer of detected cultivation. Total electricity consumption, which we observe only as utility sales allocated across the counties each utility serves, does not move measurably.

Data centers appear to respond not only to the capacity absorbed by cultivation, but also to the expected pressure on the shared grid. When counties are split by pre-ballot cannabis supply, the price increase is larger in the high-supply group, while data centers decline in both groups, and by more in the high-supply group. Data center activity also declines in counties in other states that share the grid, where the ballot does not legalize cultivation.

Dropping failing-state counties served by a utility that also serves a passing state leaves the main estimate essentially unchanged. Competition for workers, land, or water does not appear to drive the decline, and neither does crime.

We estimate a discrete-choice model of county-level siting using 2,121 decisions from 2015-2024. In the model, siting odds fall with distance to the high-voltage transmission network. We calibrate the legalization parameter to the 53 percent decline in active stock estimated by the ballot design. Without legalization by ballot, national data center capacity in the model would be 10.6 gigawatts, or 30 percent, above its 2024 level if displaced facilities are not built elsewhere, and the forgone capacity represents over 100 billion dollars of facility investment. With national capacity fixed, legalization instead shifts 6.0 gigawatts to other states, mainly ones that never held a recreational ballot, such as Virginia, Texas, and Georgia. Higher rates in the counties where cultivation expands add 0.7 billion dollars a year to electricity bills in the narrowly passing states.

Our results make three contributions. First, we add to work on how the entry of one industry affects the local development of others. Existing work emphasizes labor reallocation, agglomeration, and local sectoral adjustment (Greenstone et al., 2010; Moretti, 2010; Suárez Serrato and Zidar, 2016; Glaeser and Gottlieb, 2008; Allcott and Keniston, 2018; Black et al., 2005), as well as knowledge spillovers and product-market competition (Bloom et al., 2013; Moretti, 2021). We document a channel operating through shared electricity infrastructure. When grid capacity adjusts slowly, demand from one electricity-intensive industry can deter later entrants from investing in the region. The closest evidence is on cryptocurrency mining, whose electricity use raises the bills of existing households and small businesses where miners enter (Benetton et al., 2023). This channel connects to work on how transportation infrastructure (Donaldson and Hornbeck, 2016; Donaldson, 2018; Faber, 2014; Asher and Novosad, 2020) and housing supply constraints (Hsieh and Moretti, 2019) shape local economic development, and to the broader literature on lock-in and path dependence (David, 1985; Arthur, 1989; Unruh, 2000).

Second, we contribute to work on investment in data centers and their local effects. Existing work documents their energy and water demands (Masanet et al., 2020; Siddik et al., 2021; Mytton, 2021; Shehabi et al., 2024) and their implications for electricity grids and ratepayers (International Energy Agency, 2024; Electric Power Research Institute, 2024), while recent studies examine their local economic effects and the determinants of their siting (Alvarez et al., 2026; Ancona et al., 2026; Du et al., 2026; Zhang, 2026). Our design uses close cannabis ballots to generate exogenous variation in electricity demand from another industry, which allows us to study whether competition for local grid capacity deters subsequent data center investment. Studies of how data centers affect electricity prices (Zhang, 2026; Feher et

al., 2025) and local economies (Alvarez et al., 2026; Du et al., 2026) must account for local conditions that influence both these outcomes and developers' location choices. Existing industrial demand on the grid is one such condition. Our results also relate to work on digital technology and AI, which has focused primarily on labor markets, productivity, and market structure and less on the physical infrastructure that computing requires (Acemoglu and Restrepo, 2018, 2020; Brynjolfsson et al., 2021; Goldfarb and Tucker, 2019).

Third, we identify an electricity-market spillover of recreational cannabis legalization. Existing work has studied the effects of cannabis laws on traffic safety and substance use (Anderson et al., 2013; Wen et al., 2015), crime (Adda et al., 2014; Gavrilova et al., 2019), and housing markets (Cheng et al., 2018). We show that legalization can also affect investment in an unrelated industry through competition for a shared local input. The decline in data centers also reaches counties on the same grid in states where the ballot did not legalize cultivation. We construct a satellite-based measure of indoor cultivation that does not depend on state reporting regimes (Henderson et al., 2012; Donaldson and Storeygard, 2016; Colas and Reynier, 2026).

2 Background and Institutional Setting

2.1 Data Centers and AI Infrastructure

Data centers are the buildings, power systems, and cooling plants that house the servers behind model training, inference, and cloud services. Hyperscale campuses operated by Amazon, Google, Microsoft, Meta, and Oracle have grown from loads of tens of megawatts a decade ago to several hundred megawatts, and announced AI training campuses exceed 1 gigawatt. Hyperscale facilities use large amounts of electricity and capital but relatively little labor. A 100-megawatt facility operating near nameplate capacity consumes 800 gigawatt-hours per year, making its annual power bill tens of millions of dollars at typical U.S. industrial rates. A campus drawing several hundred megawatts is operated by a permanent staff in the low hundreds, made up of technicians, network engineers, and security personnel, far fewer workers than a comparably sized industrial plant. Construction costs are 9.5-13.3 million dollars per megawatt in U.S. markets (Turner & Townsend, 2025), so the shell, power, and cooling of a gigawatt-scale AI campus cost 9.5-13.3 billion dollars before any computing hardware. For these facilities, electricity prices and the depreciation of long-lived capital matter more for profitability than local labor costs.

U.S. data center capacity has historically concentrated in a few regional clusters. Northern Virginia, centered on Loudoun County, is the largest, with several gigawatts of installed

capacity. These clusters reflect dense fiber-optic interconnection, cheap industrial electricity, and state and local tax incentives for capital-intensive infrastructure. Panel A of Figure 1 shows the geographic distribution of U.S. data center facilities. Electricity dominates their operating cost, and industrial retail rates vary by a factor of two or more across U.S. counties. Energy-intensive manufacturing concentrates in counties with low electricity prices (Kahn and Mansur, 2013). A one-cent change in the price per kilowatt-hour changes the lifetime electricity cost of a 100-megawatt facility by more than 100 million dollars over a 15-year lease.

Data center firms and other industrial users disagree over who should pay for grid upgrades. The number of federal lobbying reports mentioning data centers rose from fewer than 100 a year in 2008-2023 to 581 in the first three quarters of 2026. Of these 581 reports, 410 also included electricity-related terms or issue codes. Before 2024, no year had more than 54 such reports.[1] Panels A and B of Appendix Figure B1 plot the annual counts. In 2026, 91 clients reported lobbying on the Ratepayer Protection Act. Appendix Table C1 groups them by type. The bill would require state utility regulators to consider making large new loads pay the full cost of the generation and transmission upgrades built to serve them. Its first version covered every project with a peak demand of 100 megawatts or more. Before the House Energy and Commerce Committee approved it in July 2026, committee leaders narrowed the bill to data centers. The Data Center Coalition, the main trade group of data center firms, had backed the original bill and opposed the narrower one. The American Iron and Steel Institute and the Industrial Energy Consumers of America endorsed the change (Davidson, 2026). As Panel C of Appendix Figure B1 shows, the Data Center Coalition's reported lobbying and political expenditures rose from 43,000 dollars in 2019 to 1.6 million dollars in 2024. The Edison Electric Institute, the trade association of investor-owned electric utilities, reported 8-11 million dollars a year over the same period. Amazon reported lobbying on federal cannabis legalization bills in 2021-2024,[2] and campaign finance records for the ten recreational cannabis ballots examined in Section 2.3 list no contributions from

[1]Each LD-2 report filed under the Lobbying Disclosure Act covers one registrant, one client, and one quarter. We keep a report if its specific-issue text names data centers, server farms, hyperscale or colocation facilities, or cloud infrastructure, and we count it as electricity-related if the same issue line names electricity, energy, power, utilities, the grid, units of power, or the Federal Energy Regulatory Commission, or carries the energy or utilities issue code. Reports through 2019 are from the Senate Office of Public Records bulk files. Later reports are from the Lobbying Disclosure Act database as of October 22, 2026. The Ratepayer Protection Act is H.R. 9340 in the House and S. 5028 in the Senate, and we count each client once across its 2026 reports.

[2]Amazon.com Services LLC filed 14 LD-2 reports between the second quarter of 2021 and the third quarter of 2024 whose specific-issue text names the MORE Act of 2021, the States Reform Act, or the Cannabis Administration and Opportunity Act. LD-2 reports list the bills lobbied on but not the position taken.

data center, cloud, or AI firms. We study whether electricity demand from another industry deters data center investment before grid capacity can expand.

2.2 Indoor Cultivation and Electricity Demand

Indoor cannabis production relies on high-intensity lighting, cooling, dehumidification, and continuous air circulation to maintain a controlled growing environment. Plants receive 12-18 hours of light a day depending on growth stage, making electricity a major variable cost. Mills (2012) estimates that indoor cultivation uses eight times as much energy per square foot as a typical commercial building. Electricity use by cannabis cultivation and infused-product manufacturing grew from 1 percent of Denver's total in 2013 to 4 percent in 2018 (Denver Department of Public Health and Environment, 2019). Nationally, cannabis cultivation used 35 terawatt-hours of electricity in 2023, comparable to U.S. cryptocurrency mining (Mills, 2025; U.S. Energy Information Administration, 2024). Environmental control and lighting dominate the load, and the energy needed for an artificial indoor climate makes cannabis more emissions-intensive than conventional crop production (Summers et al., 2021).

Outdoor cultivation uses less electricity, but regulation and product-market incentives concentrate licensed production indoors (Mills, 2025). Security, video-surveillance, and seed-to-sale tracking requirements are more easily met in enclosed facilities. Outdoor cultivation is limited to a single growing season. Indoor cultivation allows several harvests a year, tighter control over potency, and protection from weather and pests. California and Oregon retain a meaningful outdoor share, whereas licensed production in Colorado is predominantly indoor. Recreational legalization expands the most power-intensive form of production.

2.3 U.S. Recreational Cannabis Legalization

U.S. cannabis policy has changed in two waves over the past three decades. Medical legalization began with California's Proposition 215 in 1996 and spread to most states by 2025, allowing licensed cultivation and dispensary sales restricted to patients with a physician's recommendation and a state-issued medical card for qualifying conditions. Recreational legalization began in 2012, when Colorado's Amendment 64 and Washington's Initiative 502 passed by statewide ballot, and as of 2025 has spread to 24 states and the District of Columbia. Legalization extends legal purchase and possession beyond registered medical patients to any adult 21 years or older, sharply expanding the legal market. In Colorado, recreational dispensary sales overtook medical sales within 12 months of the first stores opening in January 2014 and far exceeded medical sales within a few years, with similar trajectories in Washington, Oregon, California, and Nevada. Lawful cultivation expands well

beyond the medical-only baseline, and this expansion is the shock we study. Stores typically open within two years of the ballot. Most recreational adoptions have come through citizen ballot initiatives, a smaller number through state legislation. Ballot timing is set through signature collection and certification before election day, and passage is determined by a statewide vote rather than a county-level policy decision.[3] Campaign finance records for ten recreational cannabis ballots list no contribution from a data center, cloud, or AI firm to either side.[4]

Figure 2 shows where and how often states have voted. Panel A shades each state by the Yes share on its most recent recreational ballot and Panel B plots all thirty-two ballots held since 2002. Six states (Alaska, Arizona, California, Nevada, Ohio, and Oregon) rejected legalization at least once before approving it, and South Dakota rejected it in 2022 and 2024 after a court overturned its 2020 approval. Even where voters approve it, recreational cannabis remains a Schedule I controlled substance under the federal Controlled Substances Act. Interstate shipment remains illegal, so each state with recreational legalization must build out within-state cultivation capacity to supply its retail market. State regulations typically require licensed cultivators to register each plant in a track-and-trace system, maintain perimeter security, and meet building-code requirements most easily satisfied by warehouses built for indoor cultivation. Recreational legalization expands indoor cultivation within the state and adds demand to the local grid. We use the ballot date to mark the policy change.

2.4 U.S. Electricity Markets and Slow Supply Adjustment

U.S. electricity is delivered through regional grids and retail markets that differ in structure, and all of them adjust supply slowly. Wholesale generation in most of the country is coordinated through one of seven independent system operators or regional transmission organizations, which run nodal or zonal wholesale markets. Retail service comes from regulated vertically integrated utilities in most of the country. In the restructured states customers may choose a competitive supplier, and many stay with the local distribution company (Borenstein and Bushnell, 2015). Across both structures, adding new generation or transmission takes years. Only 322 miles of transmission at 345 kilovolts and above were completed in

[3]State law sets the window for signature collection, which reaches two years in some states, and signatures are filed with the secretary of state months before the vote. Once certified, the measure is placed on a statewide ballot, in most states the next regular general election.

[4]The ten ballots are California Proposition 19 in 2010, Colorado Amendment 64, Oregon Measure 80, and Washington Initiative 502 in 2012, District of Columbia Initiative 71 and Oregon Measure 91 in 2014, California Proposition 64 and Massachusetts Question 4 in 2016, Oklahoma State Question 820 in 2023, and Florida Amendment 3 in 2024. We screen the contributors that the committees for and against each measure reported in their campaign finance filings.

2024, compared with nearly 4,000 miles in 2013 (Shreve et al., 2025). Interconnection from request to commercial operation took nearly five years for projects built in 2023, compared with three years in 2015 (Rand et al., 2024). Where transmission constraints limit imports, short-run supply expansion cannot absorb a localized step increase in electricity demand, and it shows up in higher local prices.

How a demand shock reaches local prices depends on market structure. Retail rates in regulated states are reset through rate cases that pass capacity, transmission, and fuel costs through to consumers with multiyear lags, so a sustained demand shock raises prices gradually. Each state's public utility commission decides those cases for service inside its own borders, so a utility whose territory crosses a state line files separate rates on each side. In restructured states, wholesale locational marginal prices respond faster to transmission congestion, so binding transmission constraints can create large price differences within a single regional market. Where a regional market makes capacity payments, they are allocated to load-serving entities in proportion to their share of system peak, so higher local peak demand raises the per-kilowatt-hour cost borne by all local customers, including industrial users such as data centers. A localized cultivation shock reaches local retail prices over several years, mainly through energy costs and the rate cases that pass them through.

A balancing authority coordinates resource plans, balances electricity supply and demand within its area, and helps maintain grid frequency in real time (North American Electric Reliability Corporation, 2025). It dispatches generators to follow load, holds operating reserves, and schedules power flows with neighboring areas (U.S. Energy Information Administration, 2016b; Federal Energy Regulatory Commission, 2023). The United States had 66 balancing authorities in 2016 (U.S. Energy Information Administration, 2016a). Most are utilities, and every regional transmission organization is also a balancing authority (U.S. Energy Information Administration, 2016b). The Southwest Power Pool operates one balancing authority for its members in Kansas, Missouri, and other states (Southwest Power Pool, 2026; Federal Energy Regulatory Commission, 2023). It also assesses resource adequacy using its members' load forecasts and plans transmission across the area (Southwest Power Pool, 2024; Federal Energy Regulatory Commission, 2023). Evergy Metro serves customers in Kansas and Missouri from the same generators and sets its capacity needs from a joint peak-demand forecast for both states (Evergy Metro, Inc., 2025). Empire District Electric plans its Missouri, Kansas, Oklahoma, and Arkansas service areas as a single system (The Empire District Electric Company, 2020). A large load anywhere on such a system draws on the same generators, reserves, and transmission.

A large new load is connected on the same slow timetable. Before agreeing to serve a prospective data center, the utility assesses whether the local substation, feeder, and up-

stream transmission can support the requested load. Any network upgrades required by the study enter the rate base. Developers use these connection studies and the load queues utilities publish to assess how much local capacity is already committed. These sources reveal capacity constraints years before their costs are reflected in retail rates. Grid capacity has limited new load in recent episodes. The 2025-2026 base residual capacity auction run by PJM Interconnection, the regional transmission organization serving the Mid-Atlantic, cleared at nine times the previous auction's price. PJM attributes the increase in part to generator retirements and to a jump in projected peak load in its footprint (PJM Interconnection, 2024). The independent market monitor puts the data center load included in that auction's forecast at 4,654 megawatts above existing levels, from new and expanding facilities (Monitoring Analytics, 2025). In eastern Loudoun County, Dominion Energy stopped connecting new data centers in 2022, and connections resumed only in part while the transmission upgrades that relieve the constraint continued into 2026.

3 A Framework for Grid Lock-in

3.1 Grid Capacity and Two Industries

Consider a county c on a transmission network g whose capacity K_g is fixed in the short run, since new generation and transmission take years to build and connect, as Section 2.4 documents. The network carries a baseline load D_g^0, the load L_g that licensed cultivation adds after the ballot, and the load of N_g data centers, each drawing m megawatts. The baseline load includes any cultivation licensed before the ballot, and the utilization of the network is

$$u_g = \frac{D_g^0 + L_g + m\,N_g}{K_g} \tag{1}$$

where each load sums over the counties on the network, on both sides of any state line. Two costs rise with utilization. The first is the retail rate. A utility recovers the capacity, transmission, and fuel costs of serving a larger load through rate cases, and each state's commission hears the cases for service inside its own borders. The rate in county c is therefore set for its rate jurisdiction r, the utility's service territory within the state, and rises with the utilization u_r of the capacity that serves that territory,

$$p_r = p_r^0 \left[1 + \phi\,(u_r - u_r^0)\right] \tag{2}$$

where u_r is defined as in equation (1) over the loads of the territory and the pre-ballot capacity whose cost its rates recover. Here, u_r^0 is the utilization that the baseline load

implies, p_r^0 is the rate at that baseline, and $\phi > 0$ is the pass-through of utilization into the rate. The parameter ϕ captures how the costs of capacity and upgrades for new load enter retail rates. Load added on one side of a state line increases u_r and the retail rate there. It affects rates across the line only through shared costs. In the restructured states the energy and capacity charges follow the wholesale price of the zone behind a constraint, which rises with the same load, so equation (2) applies there with the zone in place of r.

The second cost is borne at connection. A new large customer waits for the utility's connection study and for the upgrades the study triggers before service begins. Upgrades that enter the rate base reach the retail rate in equation (2). We denote the connection cost per megawatt by $\kappa(u_c, u_g)$. It includes the operating margin forgone while the customer waits for studies and upgrades, and any upgrade costs charged directly to the customer. This cost rises in the utilization u_c of the substation and feeders serving the county and in the utilization u_g of the network, since a more heavily loaded system requires larger upgrades and longer studies. Through u_g, the connection cost depends on load anywhere on the network the customer joins, whichever state sets its retail rate. A generator seeking to connect waits for its own study of the same network, which takes longer as network utilization rises, so its cost of joining the network also rises with u_c and u_g.

In year t, developer i chooses a county for a facility that draws m megawatts for H hours per year and operates for T years after connection. The developer maximizes

$$\Pi_{ict} = A_c - mH \sum_{j=t+\ell}^{t+\ell+T-1} \delta^{j-t}\, \mathbb{E}_t\Big[p_{r(c)}(j)\Big] - \delta^{\ell} m\, \kappa\Big(\mathbb{E}_t\Big[u_c(t+\ell)\Big], \mathbb{E}_t\Big[u_{g(c)}(t+\ell)\Big]\Big) + \varepsilon_{ict} \quad (3)$$

where A_c collects the fiber, land, and tax attributes of the county, ℓ is the time from the siting decision to connection on a system with spare capacity, δ is the discount factor, $p_{r(c)}(j)$ is the rate per megawatt-hour in year j in the county's rate jurisdiction, and ε_{ict} is an idiosyncratic match term. We write the two cost terms in equation (3) as $m\tilde{p}_{ct}$, where $\tilde{p}_{ct}$ is the expected lifetime cost of power per megawatt. The developer therefore maximizes $A_c - m\tilde{p}_{ct} + \varepsilon_{ict}$. The electricity-bill component depends on expected load in the rate jurisdiction over the facility's operating life. The connection-cost component depends on expected load at local substations and feeders and on the shared network when the facility connects, rather than load at the time of the siting decision.

A recreational ballot in the county's state passes at date B. Licensed cultivation then builds up over several years. The load L_c it adds in county c rises from zero toward a long-run level $\bar{L}_c$. This level is higher where cannabis supply was already established, because existing growers are the first to obtain licenses and expand. The network load L_g in equation (1)

sums these county loads. Before the vote, a developer weighs the chance of passage, which is similar in states where the ballot narrowly passes and states where it narrowly fails. A developer deciding at $t \geq B$ knows the outcome, so the expected utilization and the expected cost of power in equation (3) rise at B, while cultivation is still building toward its long-run level and before retail rates adjust to the additional load.

3.2 Predictions

Three predictions follow. First, the cultivation load and the retail rate rise together after passage. Growers must obtain licenses, build facilities, and connect to the grid before drawing power, so their load arrives with a lag after the vote and keeps building as facilities come online. By equation (2), for a given data center load, the rate in a jurisdiction rises in proportion to the cultivation it carries, and the increase is concentrated in the utility's service territory in the legalizing state, where the rate case is heard. Residential rates are reset in the same rate case, so they rise with the cultivation load as well. The change in total consumption equals the additional cultivation load minus the data center load forgone.

Second, data center investment falls once the ballot passes. Developers consider expected power costs, so entry can fall where retail rates have not risen. As fewer facilities enter, the active stock falls below its counterfactual level. The decline is larger where $\bar{L}_c$ is higher, because local cultivation adds load to the county's substations, its rate jurisdiction, and the shared network. Because the connection cost κ rises with the utilization of the whole network, the fall also reaches counties on the same network across a state line, where another state's commission sets the retail rate. Cultivation load also raises connection costs for new generators on the same network.

Third, the siting decision in equation (3) is a discrete choice over counties in which the cost of power enters through $\tilde{p}_{ct}$. Legalization raises $\tilde{p}_{ct}$ most in the counties of a legalizing state, which lowers the log odds that a facility locates there by an amount proportional to that rise and to the developer's sensitivity to the cost of power. For a county, legalization carries three costs. It raises the bills of the county's captive electricity customers (all users other than data centers) through equation (2), raises the connection cost κ for the facilities that still connect, and reduces local investment as facilities locate elsewhere or are not built. A developer builds only if the highest payoff across counties exceeds the payoff from not building. Legalization can therefore reduce the number of facilities built.

We abstract from competition for workers and land. Cultivation employs farmworkers and occupies warehouse space, while a data center employs software developers, technicians, and engineers on a parcel chosen for its substation and fiber, so the entry of one industry

need not raise the wages or land costs the other pays. Competition for these inputs, and for water, capital, and knowledge, remains an alternative explanation to the grid channel.

4 Data

4.1 Data Center Facilities

Facility-level data on U.S. data centers come from the Aterio data center database, a commercial industry tracker that compiles announcement, construction, activation, and withdrawal dates for each facility. Each record also gives the county location, the nameplate power capacity in megawatts, the provider name, an indicator of provider type, and an indicator of on-site generation equipment. The extract is current as of October 2025 and covers 5,140 facilities across 461 U.S. counties. Aggregated to the county-year level over 2000-2024, the records give the year-end stock at each development stage and the annual flows of entries and exits. An active facility is one that has entered operation by year end and has not been withdrawn by that date. We calculate the corresponding nameplate capacity totals in megawatts, imputing the national median for facilities with missing capacity. The aggregation also records provider type as hyperscaler, colocation, or unknown, taken from the Aterio classification and cross-checked by hand against publicly disclosed hyperscaler portfolios for the largest operators.

4.2 Cannabis Legalization and Ballot Outcomes

We compile state recreational cannabis ballot outcomes from Ballotpedia and cross-check them against secretary-of-state records. The dataset contains thirty-four statewide recreational ballots in twenty states and the District of Columbia over 1972-2024, with the ballot year, the Yes share, the No share, the binary outcome, the ballot origin as either citizen initiative or legislative referral, and the topic. The empirical strategy uses the Yes share minus 50 percent as the running variable, so crossing the simple-majority threshold provides the intent-to-treat assignment. For states with multiple ballots, the cross-sectional design keeps the most recent ballot, while the stacked and panel designs treat each of the thirty-four ballots as a separate event. The 0.10 bandwidth used throughout Section 5 retains twenty-six ballots, twelve failing and fourteen passing. Panel B of Figure 2 marks them. The twelve failing ballots come from eight states, and all but four of them (three in North Dakota and one in Arkansas) are in states that pass a ballot at another election. All twenty-six enter the panel design. Twenty-three are observed in the first post-ballot year the pooled design uses,

and thirteen enter the cross-sectional design, which keeps one ballot per state. Treatment begins in the ballot year, and licensed retail typically opens 12-24 months later. The running variable remains the ballot vote share in every post-ballot year. A county enters once for each ballot it is matched to, so the unit of observation in the stacked and panel designs is the county-ballot-year.

Two majority outcomes did not translate into retail sales. South Dakota's 2020 Amendment A received 54 percent of the vote but was overturned by a state court in 2021 before retail sales began. Florida's 2024 measure also received a majority but fell short of the 60 percent supermajority required by state law. Treatment follows the same simple-majority rule for every ballot, so we retain both majority outcomes in the intent-to-treat assignment. Section 5.4 reports the estimate with South Dakota's overturned initiative counted as legalization. Dropping the Florida ballot leaves the decline in the data center stock at 0.70 facilities per county, within the 0.53-0.84 range of the estimates that drop one ballot at a time.

Maryland and New Jersey legalized through legislative referrals tied to a ballot question, not citizen initiatives, and the running variable still applies because the referral required popular approval at the same statewide threshold. Nine states legalized through state legislation without any ballot, namely Connecticut, Delaware, Illinois, Minnesota, New Mexico, New York, Rhode Island, Vermont, and Virginia. We code them with the year of legislative passage and a missing ballot vote share. These states have no ballot vote share and are excluded from the discontinuity sample. Their counties enter the location-choice model in Section 7 as locations in states that never held a ballot.

4.3 Cannabis Cultivation Activity

We measure indoor cannabis cultivation activity with a satellite detection algorithm that combines three remote-sensing signals at the county-year level. Nightlight intensity from the Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band on the Suomi NPP satellite captures the high-intensity grow lights typical of indoor cultivation. We use the NASA Black Marble monthly composite, the VNP46A3 product, which grids the band's 750-meter observations at 15 arc-seconds, and resample it to the 0.005-degree grid on which the three signals are combined. Surface thermal anomalies come from Landsat 8 thermal infrared band 10 at 100-meter native resolution resampled to 30 meters, with a 16-day revisit cycle. We use the band 10 surface temperature from the U.S. Geological Survey (USGS) Collection 2 Level-2 product. Rooftop indicators come from Sentinel-2 multispectral imagery at 10-meter resolution, distributed by the Copernicus program with a five-day revisit. Brightness and vegetation-index bands separate greenhouse, warehouse, and industrial-cultivation rooflines

from residential and conventional commercial roofs, and the rooftop data begin in 2017.

We combine the three signals using inverse-variance weights estimated on state-licensed cultivation facilities in Colorado, Oregon, and Washington. The composite achieves an area under the receiver operating characteristic curve of 0.87, compared with 0.76-0.82 for the individual signals. The primary continuous outcome is the area, in square kilometers, of pixels where all three signals exceed their detection thresholds in the same county-year, formed by summing flagged pixel area and converting from square meters. We also report the composite activity index. For the mechanism analysis, we classify a county-year as having detected cultivation if at least one pixel passes all three thresholds.

Apart from the licensed-facility validation, the measure does not depend on state administrative reporting, so the same detection rule applies in every county and year, including the states where cultivation is never licensed. The three-signal panel covers 2017-2024, since the rooftop signal begins in 2017. For sample splits that require pre-2017 observations, we use a two-signal nightlight-thermal measure available from 2000 onward and divide counties at its sample mean. This measure uses DMSP-OLS nightlight and MODIS land surface temperature in the earlier years, and monthly VIIRS composites from the NOAA Earth Observation Group with Landsat thermal data in the later years. We omit the rooftop signal throughout, so both groups are measured in the same way. Appendix A documents the three-signal processing pipeline and the licensed-facility validation, and Panel B of Figure 1 maps the geographic distribution of detected cultivation.

4.4 Electricity Prices and Consumption

Retail electricity prices and quantities come from the U.S. Energy Information Administration (EIA) Form 861 utility-by-state annual data. Form 861 reports retail revenues and sales by customer class, covering residential, commercial, and industrial service. For each customer class, we divide revenue by sales to obtain the nominal price per kilowatt-hour. Prices calculated this way are unaffected by changes in the shares of residential, commercial, and industrial sales. Calendar-year fixed effects absorb inflation common to all counties.

We map utilities to counties with the service-territory schedule of Form 861, which lists the counties each utility serves in each state and year, and we match those county names to Census FIPS codes, so the assignment is rebuilt for every year of the panel. Franchised retail service territories are set by state commissions and change rarely, so the mapping does not move with the outcomes. Service territories routinely cross county and state lines. Form 861 reports revenue and sales by utility and state, so we allocate each utility's state total across the counties it serves that year in proportion to their 2000 Census population, a benchmark

fixed before every ballot the electricity panel covers. We then sum allocated revenue and sales across all utilities serving a county, and the ratio of the two defines its price.

Local consumption is the same allocated Form 861 panel, aggregated to county-year megawatt-hours overall and by customer class. Because sales are reported by utility and state, the quantity variation a county inherits is that of the service territories serving it. We also build a county-to-balancing-authority crosswalk from the balancing-authority schedule of Form 861, which defines the wider grid footprint used in Section 6.4 and screens which other-state counties enter a ballot's grid. The full electricity panel covers 2000-2024.

4.5 Labor Markets

Labor-market outcomes used in the alternative-mechanism tests come from the BLS Quarterly Census of Employment and Wages (QCEW), which reports establishment counts, covered employment, and wages by county, industry, and ownership, compiled from establishment unemployment-insurance records. QCEW covers 95 percent of U.S. wage and salary employment. We obtain the QCEW at the six-digit North American Industry Classification System (NAICS) level and aggregate it to the county-year level for total employment and two-digit sectors. We report the broad sectors of the local economy separately, along with the data center, cannabis cultivation, and shared building-trade industries themselves, taken as NAICS 518210 and 541513 for data centers, NAICS 111419 and 111998 for cannabis cultivation, and NAICS 2382 for the building trades. We code county-years with no covered firm and cells withheld for disclosure as zero. The industry-specific estimates therefore reflect changes in whether activity is reported as well as changes in wages or employment. We report these estimates alongside the county aggregate to assess the labor-market explanation. Occupational composition comes from the BLS Occupational Employment and Wage Statistics, using the May 2024 national release by industry.

5 Empirical Strategy and Baseline Effects

Legalization may be correlated with economic conditions, political preferences, and industry composition that also affect data center location choices. A simple comparison of legalizing and other states therefore need not identify the causal effect of legalization. We address this endogeneity with close statewide ballot margins on recreational legalization. Whether a ballot barely passes or barely fails is plausibly unrelated to the local economic conditions that drive data center siting, yet it sharply shifts whether the state legalizes recreational cannabis cultivation. We use the majority threshold to compare changes in data center activity after

narrowly passing and narrowly failing ballots. Let s index states. For a state's ballot we define the running variable as the Yes share minus the 50 percent simple-majority threshold

$$R_s = YesShare_s - 0.50 \tag{4}$$

and passage as $Pass_s = \mathbf{1}\{R_s \geq 0\}$. Throughout this section and the mechanism analysis of Section 6, we hold the bandwidth at $|R_s| \leq 0.10$ around each cutoff and cluster by ballot.

5.1 A Single-Ballot Cross-Section

We first estimate a cross-sectional regression discontinuity (RD) using outcomes measured one year after each state's most recent ballot,

$$Y_{cs} = \alpha + \beta\, Pass_s + \gamma_1 R_s + \gamma_2 (R_s \times Pass_s) + \varepsilon_{cs} \tag{5}$$

where c indexes counties and α is a constant. The running variable enters linearly with a separate slope on each side of the cutoff, and the model has no fixed effects. The coefficient β is then the level contrast in active data center stock between barely-passing and barely-failing counties. States that never held a recreational cannabis ballot have no running variable and drop out of the sample. The single-ballot column of Appendix Table C2 and the full outcome set in Appendix Table C3 report these estimates. Counties in barely-passing states host 1.36 fewer active data centers than counties in barely-failing states, install 20 fewer megawatts of operating capacity, have fewer new entries, and host fewer facilities with on-site generation. Restricting each state to its single most recent ballot leaves thirteen ballots in the estimation sample, and the next design pools every ballot.

5.2 Pooling Across Ballots

To increase the number of ballots near the cutoff, in the spirit of the multiple-cutoff pooling of Cattaneo et al. (2020b), we next stack every ballot event, indexed by a, and estimate the pooled cross-section,

$$Y_{csa} = \alpha + \beta\, Pass_{sa} + \gamma_1 R_{sa} + \gamma_2 (R_{sa} \times Pass_{sa}) + \varepsilon_{csa} \tag{6}$$

still without fixed effects. Each ballot keeps its own re-centered running variable. A ballot contributes the counties of a single state, on one side of the threshold, so the pooled coefficient compares counties in states whose ballots landed just above the threshold with counties in states whose ballots landed just below it, across ballot events. Pooling brings

twenty-three ballots into the cross-section and makes every estimate more precise.[5] Appendix Table C4 reports the pooled estimates. Counties in barely-passing states host 1.46 fewer active data centers, install 19 fewer megawatts of capacity, have fewer new entries, and host fewer generation-equipped facilities.

The pooled estimate identifies a causal effect only if there is no selection into treatment or control. Barely-passing and barely-failing states must be comparable in data center activity apart from the ballot outcome itself, which the close margin makes plausible. We examine this by estimating equation (6) separately at each event time, where a pre-ballot gap between passing and failing states reveals a preexisting difference, not an effect of passage. Figure 3 shows that passing states host 0.8-1.0 fewer facilities across the pre-ballot years. The gap widens after passage, reaching 1.5 fewer facilities by the first post-ballot year and 4.5 fewer by the eighth. Because the cross-sectional design blends this preexisting difference with the effect of passage, it overstates the causal effect. Because the pre-ballot gap moves little from year to year, we use a panel difference-in-discontinuities specification to remove its fixed component (Grembi et al., 2016).

5.3 The Panel Difference-in-Discontinuities

Our baseline specification absorbs the fixed pass-fail gap by combining the discontinuity with a difference-in-differences (DiD) panel

$$\begin{aligned} Y_{csta} = \beta\,(Pass_{sa} \times Post_{sta}) + \gamma_1\, Post_{sta} + \gamma_2\,(R_{sa} \times Post_{sta}) \\ + \gamma_3\,(R_{sa} \times Pass_{sa} \times Post_{sta}) + \alpha_{ca} + \lambda_t + \varepsilon_{csta} \end{aligned} \tag{7}$$

where α_{ca} are county-by-ballot fixed effects, λ_t are fixed effects for calendar year t, and $Post_{sta} = \mathbf{1}\{t \geq B_{sa}\}$ indicates the years at or after ballot year B_{sa}. The county-by-ballot fixed effects absorb every time-invariant county characteristic specific to a ballot's comparison group, so β identifies the post-ballot change in the pass-fail contrast, not its level. We restrict the sample to eight years before and after each ballot (Cengiz et al., 2019; Wing et al., 2024).[6]

Identification now rests on a parallel-trends assumption, conditional on the running-variable polynomial. The design allows a preexisting discontinuity at the threshold, provided

[5] Because this cross-section is measured one year after each ballot and the panel ends in 2024, the three ballots decided in November 2024, in Florida, North Dakota, and South Dakota, have no 2025 observation to enter it.

[6] Each stacked ballot contributes counties from a single state, so $Pass_{sa} \times Post_{sta}$ is constant within a ballot-year and would be absorbed by year-by-ballot fixed effects. We therefore use calendar-year fixed effects alongside the county-by-ballot fixed effects. These absorb shocks common to every ballot in a year but keep the variation across ballots that identifies β, since in any calendar year some ballots have already passed, others have failed, and others have not yet been held.

the discontinuity would not have changed had the ballot failed. Section 5.2 documented a gap of that kind. It also requires the stable unit treatment value assumption (SUTVA) that a county's outcomes do not depend on ballot outcomes in other states. We examine that assumption by dropping every control county served by a utility that also serves a passing state, which leaves the estimate essentially unchanged, as reported in Section 6.4. To test this parallel-trends assumption and date the onset of the effect, the dynamic form of the same panel replaces the single post-ballot term with a full set of event-time interactions, with event time defined as $\tau_{sta} = t - B_{sa}$

$$\begin{aligned} Y_{csta} = & \sum_{k \neq -1} \beta_k \left(Pass_{sa} \times \mathbf{1}\{\tau_{sta} = k\} \right) + \gamma_1 \, Post_{sta} + \gamma_2 \left(R_{sa} \times Post_{sta} \right) \\ & + \gamma_3 \left(R_{sa} \times Pass_{sa} \times Post_{sta} \right) + \alpha_{ca} + \lambda_t + \varepsilon_{csta} \end{aligned} \tag{8}$$

with the year before the ballot as the omitted reference, so the coefficients β_k trace the effect of passage at each lead and lag. The lead coefficients test for differences in trends before the ballot, while the lag coefficients trace the post-ballot response.

Figure 4 plots these coefficients. The pre-ballot coefficients are small and jointly indistinguishable from zero, so the test does not reject parallel trends. After the ballot, the active data center stock stays 0.53-0.69 facilities lower at every horizon through the eighth post-ballot year. Appendix Table C2 reports the static specification, where the active data center stock falls by 0.702 facilities per county. The same decline appears in every other measure of data center presence in Appendix Table C5. Operating capacity falls by 16.2 MW, entrant capacity by 4.4 MW, and generation-equipped facilities by 0.445, and new entries fall by 0.089 per year. At the construction cost of 9.5-13.3 million dollars per megawatt reported in Section 2.1, the lost 16.2 MW is over 150 million dollars of investment that a county in a narrowly passing state does not receive. This forgone investment is not an unambiguous welfare loss, since data centers bring few local jobs and add to the local electricity load, but the county still forgoes real economic activity and the property-tax base a data center provides.

The decline of 0.70 facilities and 16.2 MW per county is large relative to the stock these counties would otherwise have held. Data centers are a young industry, and 91 percent of county-years in the sample have no active facility. Small absolute changes in facility counts can therefore imply large percentage changes. The sample mean of 0.446 facilities includes pre- and post-ballot years, passing and failing states, and counties with no facilities. We instead compare the decline with two benchmarks. The first is what counties in narrowly passing states would have held without legalization. Assuming a common effect across these counties, we add the estimated declines to their observed post-ballot means of 0.63 facilities

and 9.4 MW, which gives 1.33 facilities and 25.6 MW. The second is what counties in narrowly failing states hold eight years after their ballot, 1.47 facilities and 29.4 MW. The decline in the stock is 53 percent of the first benchmark and 48 percent of the second, and the decline in operating capacity is 63 and 55 percent.

Annual entry falls by 0.089 facilities per county, or 74 percent of the counterfactual flow of 0.12. The proportional decline is larger for entry than for the stock.

Data centers need few workers and serve their users over the internet, so developers have wide latitude in choosing a location (Pan Fang and Greenstein, 2025; Powell et al., 2025). The capacity of local power plants is the main predictor of where data centers locate (Ancona et al., 2026). In Virginia, land accounted for 6 percent of data center capital investment in fiscal years 2021-2023, compared with 68 percent for equipment (Joint Legislative Audit and Review Commission, 2024). Where power has run short, development has moved to other places (International Energy Agency, 2025). After Singapore stopped approving large new data centers in 2019, capacity in neighboring Johor grew from almost nothing in 2020 to 397 MW in 2024 (Vichyanond, 2026).

New generation can take five or more years to build, and major transmission projects often take 5-7 years (Joint Legislative Audit and Review Commission, 2024). Developers file connection requests in several places to find where they can build soonest (Powell et al., 2025). If cultivation demand remains high while grid capacity is constrained, developers may build elsewhere or not at all rather than wait for upgrades.

5.4 Robustness

The close-election design requires that no state can place its vote share precisely on either side of the 50 percent threshold, so that barely-passing and barely-failing states are comparable at the margin (Lee, 2008). Campaigns and interest groups are unlikely to control the final margin precisely in a statewide ballot. Caughey and Sekhon (2011) find that bare winners and losers of close U.S. House races differ on predetermined characteristics, though such imbalance is the exception across the more than forty thousand close elections examined by Eggers et al. (2015). Our design does not require balance in levels. Section 5.2 finds a pre-ballot gap in the data center stock between passing and failing states, and the panel of Section 5.3 differences it out. We also test for sorting around the threshold using the density of ballot vote shares. The local-polynomial density test of Cattaneo et al. (2020a), a refinement of McCrary (2008), finds no discontinuity at the threshold on the twenty-six ballots within the bandwidth and on all thirty-four. Appendix Figure B2 plots the estimated density with its confidence bands.

Counties in passing and failing states do not differ measurably in grid or economic conditions before the vote. Figure 5 compares fifteen county characteristics in the year before the ballot, each scaled by its standard deviation. The four grid measures are generator interconnection requests entering the queue, requests withdrawn from it, outage minutes per customer, and outages per customer. None of the four differs measurably between passing and failing counties. The other eleven cover output, income, wages, home values, unemployment, poverty, and the employment shares of construction, manufacturing, and information. All fifteen differences are indistinguishable from zero.

A cannabis ballot may also signal voter willingness to adopt other regulation that independently affects data center siting. We test for such policy bundling by re-estimating the panel difference-in-discontinuities of Section 5.3 with five other state policies as outcomes, namely a binding mandatory Renewable Portfolio Standard (DSIRE), Regional Greenhouse Gas Initiative membership (RGGI), a right-to-work statute (NCSL), full retail electricity choice (EIA), and the top statutory corporate income tax rate (ITEP).[7] As Appendix Table C6 shows, only RGGI membership moves after passage, by 6 percentage points. The right-to-work and renewable-portfolio coefficients are within two hundredths of zero. Retail-choice status does not change in any bandwidth state during the window, so that coefficient is a mechanical zero, and the corporate rate moves by less than a percentage point. The estimates provide little support for a bundled-policy explanation.

A pre-trend test that fails to reject is not by itself evidence that trends are parallel (Rambachan and Roth, 2023). The pre-ballot coefficients in Figure 4 are individually small but drift upward at the farthest leads, so we ask whether a lingering difference in trends could account for the post-ballot decline. Even allowing a difference in trends as large as the largest one seen across the seven pre-ballot coefficients, the robust confidence set for the average post-ballot effect still excludes zero, as Figure 6 shows. The drift also reflects which ballots reach the farthest leads. The panel begins in 2000 and ends in 2024, so each point of Figure 4 rests on a different set of ballots. Figure 7 restricts the sample to the twelve ballots observed in every year from five years before to five years after the vote. In this balanced sample, the pre-ballot coefficients are jointly indistinguishable from zero, and the largest is 0.060 facilities. The effect on the same sample is 0.709 fewer facilities, compared with 0.702 in the full panel.

The control group is itself partly treated because three control states legalize within their own comparison windows. California legalizes in 2016 after its 2010 ballot failed, Oregon in

[7] We select these as the state policies that could accompany a cannabis ballot while also shifting data center location through data centers' main input costs, namely electricity, labor, and taxes. The renewable-portfolio standard, RGGI carbon pricing, and retail-choice deregulation bear on electricity prices, right-to-work status on labor cost, and the corporate income tax rate on the tax burden.

2014 after its 2012 ballot, and Arizona in 2020 after its 2016 ballot. Dropping each control state's observations from the year it legalizes onward, in the spirit of the not-yet-treated comparisons of Callaway and Sant'Anna (2021), leaves the decline at 0.46 fewer facilities, and at 0.56 when South Dakota's overturned 2020 initiative is also counted as legalization. Appendix Table C7 reports the first. At a three-year window the restricted and baseline samples give nearly identical estimates, 0.192 and 0.196, and they diverge only at the longest horizons, where the excluded states supply three of the six control ballots.

We next vary the event window, the bandwidth, and the fixed effects. The level effect on the active stock grows monotonically with the event window, from 0.196 fewer facilities per county at a three-year window to 0.496 at six years and 0.702 at eight, and remains precisely estimated at every window, as Appendix Table C8 shows. A wider window takes in more post-ballot years of a persistent effect, and it reaches the longer horizons that only the earlier ballots supply. The estimate is likewise stable across bandwidths, staying at 0.64-0.84 fewer facilities over the 0.07-0.15 range and tightly estimated throughout, and it grows to 1.19 fewer facilities at the mean-squared-error optimal bandwidth of 0.04. Appendix Table C9 reports the full bandwidth path.[8] It is also stable across fixed-effect structures, spanning 0.68-0.72 fewer facilities across county, county-by-ballot, and county-by-ballot-plus-year fixed effects once the ballot-specific post indicator is included, as Appendix Table C10 shows. The estimate is similar across ballots. Appendix Table C11 estimates a separate effect for each ballot-year cohort against the same pooled control group, and all eight are negative and significant, at 0.62-0.97 fewer facilities per county. Dropping one of the twelve ballot years at a time holds the pooled estimate at 0.56-0.96 fewer facilities, and dropping one ballot at a time at 0.53-0.84, significant in all twenty-six. Because the design fits a straight line on each side of the cutoff, a curved relationship between vote share and data center activity could imitate a jump. Adding a quadratic term on each side moves the estimate to 0.99 fewer facilities, compared with 0.702 at baseline, as Appendix Table C12 shows. We go no higher than a quadratic, since high-order polynomials are unreliable in this design (Gelman and Imbens, 2019).

A county's own supply of electricity could shape the prospects of both industries, so we control for pre-ballot supply. Utility-scale generating capacity sited in the county in the year before its own ballot, total and renewable, and the capacity it already had waiting in the interconnection queue, are each interacted with the post-ballot indicator and then with the full set of calendar years, which lets counties differing in supply follow different paths. The decline stays at 0.536-0.616 fewer facilities across the four specifications, compared with

[8]The optimal bandwidth of 0.04 leaves only four or five ballots on each side of the cutoff, too few for reliable inference, so we hold the baseline at 0.10, which keeps twelve failing and fourteen passing ballots.

0.702 without them. Appendix Table C13 reports them. Capacity is measured before the ballot because generation responds to the same demand shock, and interconnection requests are themselves an outcome in Section 6.3. Conditioning on either contemporaneously would net out part of the mechanism the paper is measuring. The design compares twelve failing and fourteen passing ballots. With so few clusters, tests based on conventional cluster-robust standard errors can over-reject (Cameron and Miller, 2015). We therefore also report wild cluster bootstrap inference (Cameron et al., 2008), including the null-imposed variant (MacKinnon and Webb, 2018). Appendix Table C14 collects the estimate under analytical and bootstrap inference, clustered by ballot and by state, and the active-stock estimate remains distinguishable from zero under each procedure.

6 Mechanism

6.1 The Cannabis Cultivation First Stage

We first examine how legalization changes the electricity demand that cultivation places on the local grid. We apply the dynamic panel design of equation (8) to satellite-detected indoor cultivation area, the three-signal measure of Section 4.3.[9] Detected cultivation is zero in most county-years. The event-study coefficients measure changes in mean detected area. Figure 8 shows the first stage. Detected cultivation rises from the ballot year onward, with the estimated effect rising from 0.096 square kilometers at passage to 0.176 square kilometers five years after the ballot. With a sample mean of 0.0485 square kilometers, the fifth-year increase is 3.6 times the average detected area. The nightlight, thermal, and rooftop criteria are designed to identify indoor cultivation. Indoor grow operations use eight times the energy per square foot of a typical commercial building, as Section 2.2 documents.

6.2 The Electricity-Market Channel

We relate retail electricity prices to detected cultivation within county, with county-by-ballot and calendar-year fixed effects. We then estimate the panel design of equation (7) with a cultivation-intensity interaction. The interaction lets the post-ballot price effect vary with the square kilometers of detected cannabis in a legalized county. In the specifications without the interaction, ballot passage identifies the average effect, and the interaction relates the

[9] A post-ballot rise in detected area could also reflect a detector that improves over time, but calendar-year fixed effects absorb any change in detector performance common to all counties, and the estimate compares passing to failing states within the same year.

price response to realized cultivation intensity.[10]

Within county, retail electricity prices rise with detected cultivation. In Table 1, the coefficient on the inverse hyperbolic sine of detected cultivation area is 0.051 for the log average retail price in Column (1), 0.046 for the log commercial price in Column (2), and 0.070 for the log industrial price in Column (3). Commercial and industrial service are the two customer classes from which a data center buys electricity, and prices rise with cultivation in both. Table 2 shows a similar pattern in the ballot design. The average price rises 2.2 percent more after the ballot for each additional square kilometer of detected cultivation, the commercial price 2.3 percent more, and the industrial price 2.6 percent more. Legalized counties with more cultivation thus face larger post-ballot price increases. A uniform statewide price increase alone would not explain the larger increases in counties with more cultivation.

Panel B shows no detectable change in aggregate electricity consumption. Commercial consumption is the exception, rising 3.7 percent per square kilometer of detected cannabis. Licensed grow operations buy power as commercial customers. Figure 9 likewise shows no detectable change in total electricity consumption. The pre-ballot coefficients are flat, and the post-ballot estimates become slightly positive but remain statistically indistinguishable from zero. Commercial buildings use 12.6 kilowatt-hours of electricity per square foot a year on average (U.S. Energy Information Administration, 2022), so at eight times this intensity, as reported for indoor cultivation in Section 2.2, the fifth-year cultivation increase of 0.176 square kilometers, or 1.9 million square feet, implies an average electricity demand of 22 MW, of the same scale as the 16.2 MW of nameplate data center capacity a passing county loses.

Appendix Tables C15, C16, C17, and C18 report the corresponding estimates for the full panel, for cultivation area in levels, and for the eight-year window. The detection measure begins in 2017, so the within-county and cultivation-intensity estimates use 2017-2024 county-years, and the interaction is defined only on that sample. On the full panel the ballot-price estimate is nearly unchanged. The within-county estimates remain positive using cultivation area in levels and at the eight-year window, and at that window total consumption shows no detectable change in either the static or the dynamic design, as Appendix Figure B3 shows.

[10]Detected cultivation is itself a post-ballot outcome, so the interaction is not an exogenously assigned dose. All columns of Table 2 use the 2017-2024 detection sample, and Appendix Table C15 also reports the estimates on the full sample.

6.3 Anticipatory Responses

Data center developers consider electricity costs over a decade or more when choosing sites, so they may respond to the scarcity they expect after legalization as well as to current prices. A response to current prices alone would be concentrated where power has already become more expensive. A response to expected scarcity could also appear in counties whose current prices do not change. We estimate price and data center responses separately for counties with higher and lower pre-ballot cannabis supply. Pre-ballot cultivation is a proxy for the scope for expansion after legalization, so electricity costs should rise more where supply was already established.

The decline in data centers is largest where cannabis supply was already established, and so is the price increase. Figure 10 splits on satellite-detected cultivation before the ballot, with counties in each panel classified by the cultivation measure that covers that panel's estimation window. Both outcomes in Figure 11 are grouped on a single measure, pre-ballot arrests for cannabis sale and manufacture, so each county falls in the same group for both.[11] The pattern is similar under both measures of pre-ballot supply. Counties with little cultivation before the ballot also lose data centers in Figure 10, which points to a response to the load expected on the shared grid and not only to cultivation already in the county. Splitting on the year before the ballot, the average over available pre-ballot years, or the per-capita arrest rate gives the same pattern.

Interconnection-queue applications by generators provide an additional forward-looking measure of demand for capacity on the shared network, because developers apply years before their projects begin operating. These applications indicate developers' interest, while transmission capacity and connection costs influence whether proposed projects proceed. Johnston et al. (2023) separate the interconnection-cost channel from the wholesale-price channel and find little effect of a new generator beginning operation on wholesale prices at its location in PJM. Limited transmission capacity raises the network-upgrade costs facing new requests, which lifts withdrawal rates and may deter developers from filing at all (Rand et al., 2025). Solar and gas generator applications fall with cultivation on the county's utility grid within the state, both in the county and in the other counties the grid serves. A generator pays for the network upgrades its own connection triggers, so capacity claimed anywhere on the grid raises its cost of connecting, and cultivation elsewhere on the grid should deter applications in the county as well. Appendix Table C19 reports county-year counts of solar and gas generator interconnection-queue applications and dated withdrawals from the

[11] The arrest data extend back to the 1970s and provide pre-ballot measures for all twenty-six ballots, including the successful 2012 ballots in Colorado and Washington. The three-signal satellite measure begins in 2017.

LBNL Queued Up records, in the panel difference-in-discontinuities of Section 5. Cultivation is measured only from 2017 onward, so these estimates use the cultivation-dose subsample. Applications fall by 0.042 a year for each additional square kilometer of cultivation on the grid. When own-county and rest-of-grid cultivation enter separately, applications fall by 0.199 per square kilometer in the county and by 0.029 per square kilometer elsewhere on the grid, each conditional on the other. We find no detectable change in withdrawals for either own-county or rest-of-grid cultivation. Fewer applications also reduce the number of projects that could later withdraw.

6.4 Grid Transmission and Spillovers

We next examine the spatial extent of the price response across counties on the same electricity grid. These spillovers also bear on SUTVA, because a shared grid can transmit the effects of legalization to control counties in other states. For each ballot we take every county served by a utility that also serves the balloting state, keeping an out-of-state county only when its balancing authority also operates in that state, and partition this grid footprint into counties that legalize and cultivate, counties that legalize but do not cultivate, and counties in other states on the same grid where the ballot does not legalize cultivation, whose composition is summarized in Table 3. The price response is concentrated within the utility's own service territory. In Table 4, a county's retail price rises 0.83 percent more after the ballot for each additional square kilometer of cultivation in the county. The corresponding estimate for cultivation in other counties within the same utility's in-state service territory is 0.78 percent per square kilometer. A single tariff covers the whole service territory. Across the state line the estimate falls to 0.37 percent. Table 5 reports the own, within-state, and cross-state estimates on the utility and balancing-authority grids. The within-state and cross-state coefficients are both smaller under the broader balancing-authority definition, which includes many distant counties that may not face the same local grid constraints. We therefore use the utility-based estimates as our main results. Each state's public utility commission decides rate cases for service inside its own borders, so the retail rate a customer pays is set on that customer's own side of the line, even where the two sides share a grid.

Panel A of Figure 12 maps the main balancing authority of each county. The largest balancing authorities are the Midcontinent Independent System Operator, PJM, and the Southwest Power Pool. Each spans many states, so the grid footprint of a ballot can reach across state lines. A county in another state enters a ballot's grid only if its balancing authority also operates in the balloting state. A shared utility alone is not enough, because a utility's service territories may lie on separately operated systems. Appendix Tables C20

and C21 report both sets of results under two alternative rules, one requiring only a shared utility and the other a shared state border. The data center estimates and the within-state price estimate are similar under all three definitions. Form 861 reports revenue and sales by utility and state, so a county's price is constructed from the utility-state prices of the utilities serving it, and Appendix Table C22 reports how closely county prices track utility-state prices. Appendix Table C23 estimates the same relationship at the utility-state level using reported revenue and sales, without allocating them to counties. The estimate is 0.66 percent per square kilometer, compared with 0.78 percent in the county-level specification. The corresponding estimate for residential rates is 0.62 percent, which a shift in the mix of customer classes could not produce, because cultivators buy commercial and industrial service.

Data center activity also declines in other-state counties on the same grid. Table 6 reports a decline of 0.36 facilities in legalizing counties that cultivate, 0.41 in legalizing counties that do not, and 0.45 in other-state counties on the same grid. The within-state group sees both prices rise and the data center stock fall. In these states, cultivation elsewhere on a utility's grid also lowers solar and gas interconnection applications, as Section 6.3 shows. The other-state group loses data centers on a similar scale.[12] Panel B of Figure 12 shows the grid footprint of Missouri's 2022 ballot near Joplin and Kansas City. Empire District Electric serves counties around Joplin on both sides of the Missouri-Kansas line, and Evergy Metro serves the Kansas City area on both sides. The grid footprint includes Cherokee County near Joplin and Johnson, Leavenworth, and Miami counties near Kansas City. It excludes Wyandotte County, whose main provider is a municipal utility.

Grid capacity, unlike retail rates, is shared across the state line. A county outside the balloting state but within its grid footprint shares its main utility's generators, reserves, and transmission with the balloting state. Generators in the Southwest Power Pool and the Midcontinent Independent System Operator respond mainly to load in their own market rather than to load in the neighboring one, and within the Southwest Power Pool they respond to load anywhere in the market (Hausman, 2025). When the Southwest Power Pool ran short of power in February 2021, it spread the load cuts across the transmission operators of its whole balancing authority area in proportion to their prior use, including the utilities that serve both sides of the Kansas-Missouri line (Southwest Power Pool, 2021a,b). Evergy Metro includes new large loads in the joint peak-demand forecast used to plan capacity for both states (Evergy Metro, Inc., 2025). Cultivation in one state can leave less capacity for

[12] The specification includes the same county-by-ballot and calendar-year fixed effects as the main panel. The coefficient therefore measures the change in these counties' data center stock after a ballot in another state, rather than a preexisting level difference.

new data centers elsewhere on the same system. Interference across the shared grid could move the estimate in either direction. A data center displaced from a legalizing state may relocate to a control county, while a decline that reaches control counties on the same grid would work the other way. Appendix Table C24 drops every control county served by a utility that also serves a state passing a ballot within the event window, and the data center estimate moves from -0.702 to -0.735. The restricted sample retains all twelve failing ballots at every horizon, so the change in the estimate reflects the exclusion of counties rather than ballots.

6.5 Labor-Market Channel

Labor competition is a natural alternative explanation. An expanding cultivation sector could raise local wages, and data centers facing a higher wage bill could forgo investment there. We examine it using industry scale, occupational overlap, wages, and employment. Both industries account for very small shares of county employment and draw on different occupations. Appendix Table C25 reports the industry scale. Indoor cannabis cultivation employs 2.6 workers per county-year, or 0.003 percent of county employment, while data centers employ 110.9 workers, or 0.042 percent. These employment shares are much smaller than those of manufacturing at 11.3 percent, retail at 11.4 percent, accommodation and food services at 7.0 percent, and construction at 4.6 percent. The two industries also draw on separate occupational pools. Appendix Table C26 shows data center employment concentrated in computer and business work, with software developers alone at 16.1 percent, and information-systems managers, support specialists, and systems analysts also high on the list. On the cultivation side, employment in Support Activities for Crop Production is concentrated in manual farm work, with farmworkers at 55.4 percent followed by hand packagers and agricultural equipment operators. The overlap among the most common occupations is limited to office clerks and managers, suggesting little competition for operating staff. Construction is the one phase where the two industries might overlap. Electrical and HVAC contractors install lighting and climate-control systems in cultivation facilities and power-distribution and cooling systems in data centers. The overlap is limited to local installation work, because the equipment itself is not made locally.

All four wage estimates in Table 7 are negative. The county-total weekly wage falls by 6.2 dollars from a pre-period mean of 745 dollars, under one percent of the mean, the data center wage by 195.6 dollars, the cannabis cultivation wage by 70.7 dollars, and the HVAC

and electrical contractor wage by 76.9 dollars, none distinguishable from zero.[13] Wages in the broad sectors of the local economy show no increase distinguishable from zero either. Appendix Table C27 reports the weekly-wage effect in construction, manufacturing, retail, information, accommodation and food, leisure, and professional services, and no sector shows the positive response labor competition requires. The employment effects in the county total, data centers, cannabis cultivation, and those contractors are indistinguishable from zero. Figure 13 traces total county employment in the dynamic design and shows no post-ballot change distinguishable from zero. The data center wage declines after the ballot. Figure 14 shows flat pre-ballot coefficients followed by a post-ballot decline of 220 dollars a week that later attenuates. Labor competition would push this wage up, as data centers pay more to keep their workers, whereas a decline is consistent with fewer data center entries. Among county-years with an operating cultivator, the indoor cultivation wage rises by 619 dollars a week after legalization, in Appendix Table C28, reflecting higher labor demand in the sector that legalization expands. We find little evidence that labor competition explains the decline in data center activity.

Cannabis laws can affect local crime (Adda et al., 2014; Gavrilova et al., 2019), which could in turn affect data center siting. Appendix Table C29 shows no detectable change after the ballot in arrests for violent offenses, property offenses, non-drug offenses, or drugs other than cannabis. County employment and the number of establishments also do not decline, as Table 7 and Column (5) of Appendix Table C29 show.

6.6 Competition for Water

Many data centers use evaporative cooling (Mytton, 2021; Siddik et al., 2021), and U.S. data centers consumed 66 billion liters on site in 2023 (Shehabi et al., 2024). Training and running AI models add to this demand (Li et al., 2025). Data center growth also raises public-supply water withdrawals in U.S. counties (Alvarez et al., 2026). Cannabis cultivation requires substantial water (Carah et al., 2015; Wartenberg et al., 2021), so the two industries may compete for local supply.

We test whether water deliveries and water-utility revenue rise after the ballot, using the panel difference-in-discontinuities of equation (7). Total public-supply deliveries and deliveries to commercial, industrial, institutional, and irrigation customers come from USGS estimates for 2000-2020. Water-utility revenue of local governments comes from the Census

[13] In the data center, cannabis, and contractor columns, county-years with no operating firm and cells withheld for disclosure are coded as zero. These estimates therefore reflect changes in reported industry activity as well as changes in wages. Appendix Tables C27 and C28 report wages among county-years with a disclosed wage.

of Governments and the Annual Survey of State and Local Government Finances. Appendix Table C30 shows no detectable change after the ballot in total deliveries, deliveries to these customers, or water-utility revenue. The event studies in Appendix Figure B4 show no detectable change in any of the three outcomes, before or after the ballot.

6.7 Other Channels

We also consider land, capital, and knowledge spillovers. Data centers locate on industrial and commercial parcels chosen for proximity to high-voltage substations and fiber backbone, where land is a small share of a capital outlay dominated by the shell, power, and cooling. Licensed indoor cultivation occupies warehouse space under the security and building-code requirements that Sections 2.2 and 2.3 describe. Large data center developers, including the national and global hyperscalers that Section 2.1 describes, draw on national capital markets, so cannabis cultivation in a county does not compete for the funds that build their facilities. Hyperscaler capital spending reaches hundreds of billions of dollars a year. Cannabis cultivation and data centers share little common technology or knowledge base through which one could raise the other's productivity, and the two do not compete in the same product markets (Bloom et al., 2013; Moretti, 2021).

7 Data Center Capacity and Electricity Bills

We use a conditional logit model of county choice to calculate data center capacity without legalization by ballot. The reallocation counterfactual holds national capacity fixed, while the development counterfactual allows a developer not to build and assumes that facilities displaced by legalization are not built elsewhere. We separately calculate the increase in captive customers' electricity bills using the cultivation-intensity coefficient from Section 6.2.

7.1 Data Center Siting and Counterfactual Capacity

7.1.1 Model and Estimation

The model uses facility-level data from the Aterio database described in Section 4.1. The estimation sample contains 2,121 facilities with siting decisions dated 2015-2024, each choosing one county from the set of contiguous U.S. counties in the year of that decision.[14] We normalize equation (3) by the scale σ of the match term. County population, distance to

[14] Where the announcement date is missing we date the decision by the activation date, and where both are missing by the construction start date.

the high-voltage transmission network, and state constants summarize A_c, and legalization shifts the expected cost of power. The utility for developer i choosing county c in year t is

$$U_{ict} = \theta \, Legal_{ct} + \eta \log pop_{ct} + \gamma \log(1 + dist_c) + \alpha_{s(c)} + \varepsilon_{ict} \tag{9}$$

Here, $Legal_{ct}$ equals one if a recreational ballot passed in the county's state in or before year t and legalization followed. We represent legalization as a common shift θ in siting utility, consistent with the increase in the expected cost of power in Section 3. The parameter η is the elasticity of siting odds with respect to county population, pop_{ct}. The variable $dist_c$ is the distance in kilometers from the county's 2010 center of population to the nearest transmission line of 345 kilovolts or more, and γ is the coefficient on $\log(1 + dist_c)$. The state constant $\alpha_{s(c)}$, population, and distance summarize A_c and pre-legalization power costs after normalization by σ. These constants absorb average differences across states in tax incentives, the structure of electricity markets, and other state conditions. Within a state, population and distance, which proxy for access to fiber networks and to the high-voltage grid, govern how the model divides sitings among counties. The normalized match term ε_{ict} is an independent type-I extreme value shock. The choice probability takes the conditional logit form

$$\Pr(c \mid i) = \frac{\exp\big(V_{ct}\big)}{\sum_{c' \in \mathcal{C}} \exp\big(V_{c't}\big)} \tag{10}$$

where V_{ct} is the deterministic part of utility and $\mathcal{C}$ is the choice set of counties. We estimate the model on a sampled choice set of fifty counties per siting, the chosen county plus forty-nine non-chosen counties drawn uniformly without replacement from the counties the model covers in that year. This sampling keeps the conditional logit estimates consistent (McFadden, 1978), and we resample states for inference. The counterfactuals compute choice probabilities over the full set of covered counties. The conditional logit imposes independence of irrelevant alternatives. A price wedge in one county lowers its siting probability, and under this restriction the choice probabilities of other counties rise in proportion to their initial shares. This restriction determines where displaced capacity goes when national capacity is fixed. Section 7.1.2 therefore uses the reallocation counterfactual to describe which states gain and which lose capacity, not the relocation of individual facilities.

We take θ from the ballot design. Counties in legalizing states differ from other counties in their attractiveness to data centers. Section 5.2 documents a pre-ballot gap between passing and failing states. The panel difference-in-discontinuities in Section 5.3 removes this gap near the threshold. Because both counterfactuals start from the operating stock, we use the post-ballot stock ratio as the calibration target. This is the observed stock in narrowly

passing states divided by the stock estimated under ballot failure, using passing and failing counties from the same post-ballot years. A county in a narrowly passing state holds 0.63 active facilities over the post-ballot years, compared with 1.33 under ballot failure, so we set $\theta = \log(0.63/1.33) = -0.75$. In the development counterfactual, this shift reproduces the proportional stock decline of the ballot design, and the reallocation counterfactual uses the same shift. Using the entry ratio of 0.26 would imply a more negative value of θ. We then estimate the population elasticity, the distance coefficient, and the state constants on the 2,121 siting choices with θ held at this value.

Table 8 reports the estimates. In Column (2), which the counterfactuals use, population enters with an elasticity of 0.97, so a county with twice the population has nearly twice the siting odds of an otherwise identical county in the same state. Siting odds also fall with distance to the nearest high-voltage line. Without the distance term, the population elasticity in Column (1) is 1.05. The state constants capture the regional clusters of Section 2.1. The interval for θ maps the interval of the ballot estimate through the same log ratio, and the intervals for the population elasticity and the distance coefficient come from resampling states. With θ fixed, observed siting choices determine how capacity is allocated across counties.

7.1.2 Fixed National Capacity

We construct two counterfactuals that remove the legalization shift from the utility of siting in counties of states where a recreational ballot passed and legalization took effect, which leaves out South Dakota's overturned 2020 initiative and Florida's 2024 majority. We apply the near-threshold calibration of θ to all of these states and hold siting utility elsewhere at its estimated value, using the estimates in Column (2) of Table 8. Active facilities in the counties the model covers have 35.9 gigawatts of nameplate capacity in 2024. We allocate this total across counties in proportion to the model's 2024 choice probabilities, with the same average facility size in every county. Counterfactual utility in each legalizing county is $V_{ct}^{\text{cf}} = V_{ct} - \theta\, Legal_{ct}$. Let π_c be the 2024 choice probability of county c with the shift in place, M this 35.9 gigawatt total, and $Legal_c$ the county's 2024 legalization indicator. The choice data determine where facilities are sited, not how many are built, so the two counterfactuals differ in what becomes of the facilities that legalization displaces (Schmidheiny and Brülhart, 2011). Both counterfactuals are static and start from the model's allocation of the 2024 operating stock. Table 9 reports the results.

The reallocation counterfactual holds national capacity fixed. Developers build to meet national demand for cloud and AI services, and a cannabis ballot in one state does not change that demand. Because data centers need few workers and serve their users over the internet,

developers can meet the same demand from many counties (Pan Fang and Greenstein, 2025; Powell et al., 2025). The largest developers finance their facilities in national capital markets (Section 6.7). A facility that a legalizing county loses can then be built in another county with spare grid capacity, as development moved to neighboring Johor after Singapore stopped approving large new data centers in 2019 (Vichyanond, 2026).

County shares are renormalized after the shift is removed, so the capacity that legalizing counties gain comes from other counties in proportion to π_c. Capacity without legalization by ballot is

$$M_c^{\mathrm{R}} = M \frac{\pi_c \exp(-\theta\, Legal_c)}{\sum_{c'} \pi_{c'} \exp(-\theta\, Legal_{c'})} \tag{11}$$

In the reallocation counterfactual, the legalizing states hold 6.0 gigawatts more and the other states 6.0 gigawatts less, so legalization moves capacity away from the legalizing states toward states that did not legalize by ballot. Panel A of Appendix Figure B5 maps the direction of the state-level shift. Virginia, Texas, and Georgia have the largest predicted gains from legalization by ballot, and Nevada, Ohio, and Oregon the largest losses.[15] These gains reflect the proportional reallocation rule of equation (11).

7.1.3 Flexible National Capacity

Reallocation requires spare grid capacity in the counties that would receive a displaced facility, and that capacity has become scarce across the country. Generation projects built in 2023 took nearly five years from interconnection request to commercial operation, compared with three years in 2015, and major transmission projects often take 5-7 years (Rand et al., 2024; Joint Legislative Audit and Review Commission, 2024). In eastern Loudoun County, at the center of the largest U.S. cluster, Dominion Energy stopped connecting new data centers in 2022 (Section 2.4). Developers file connection requests in several places and build where they can connect soonest (Powell et al., 2025). When the networks that would receive a displaced facility are themselves full, the facility can go unbuilt, and legalization then lowers national capacity.

The development counterfactual relaxes the fixed national total and assumes that displaced capacity is not built elsewhere. Removing the shift raises the capacity of each legalizing county by the factor $\exp(-\theta)$ and leaves other counties unchanged, so national capacity rises by the amount the legalizing states gain. Capacity without legalization by ballot is

$$M_c^{\mathrm{D}} = M\, \pi_c \exp(-\theta\, Legal_c) \tag{12}$$

[15] The three gaining states did not legalize by ballot, so the shift the counterfactuals remove does not apply to them.

This is the Poisson limit of the siting model with the outside option of not building, in which expected sitings in a county are proportional to $\exp(V_{ct})$ and do not depend on the utility of other counties, and the conditional logit estimates of Table 8 apply (Guimarães et al., 2003; Schmidheiny and Brülhart, 2011). In the development counterfactual, the legalizing states hold 10.6 gigawatts more capacity without legalization by ballot, and national capacity is 30 percent above its 2024 level. At the construction costs of Section 2.1, the 10.6 gigawatts represent over 100 billion dollars of investment in shell, power, and cooling. Panel B of Appendix Figure B5 maps the forgone capacity by state. Across the 95 percent interval of the ballot estimate, the forgone capacity is 5.0-14.0 gigawatts, and the lower end exceeds the 2024 stock of every state but Virginia. At each end of the interval, we recompute the shares π_c with that value of θ and scale them to M, with the population elasticity, the distance coefficient, and the state constants at their central estimates. Appendix Table C31 reports state trade in data center equipment, using the classification of Waugh (2026). After a ballot passes, the state's imports in the combined compute-hardware category decline relative to its other imports, and the decline holds when the measure is restricted to products under Harmonized System (HS) heading 8471, automatic data processing machines and units.

7.2 Electricity Bills and Forgone Investment

In legalizing states, we apply the cultivation-intensity component of the post-legalization price increase to the 2024 electricity consumption of captive customers, all users other than data centers. The annual increase in bills is

$$\Delta B_c = \left(p_c - p_c^{\mathrm{b}}\right) \cdot q_c^{\mathrm{captive}} \tag{13}$$

where q_c^{captive} is 2024 county electricity consumption in kilowatt-hours, net of the estimated annual electricity use of active data centers operating at nameplate capacity. In equation (13), p_c is the county's 2024 retail rate per kilowatt-hour and $p_c^{\mathrm{b}} = p_c \exp(-0.0219\, Area_c)$ is the benchmark rate, where $Area_c$ is the county's 2024 area of detected indoor cultivation in square kilometers and 0.0219 is the cultivation-intensity coefficient in Column (2) of Table 2, the change in the log retail rate per square kilometer. The benchmark removes the cultivation-interaction term and keeps the rest of the 2024 rate, including the price component common to all counties in a legalizing state. The calculation measures the change in electricity expenditure at fixed consumption. Applied to the legalizing counties whose ballots fall inside the estimation bandwidth, these price differences imply an annual bill increase of 0.68 billion dollars across 581 counties. Most of the increase is in Middlesex, Worcester, and Hampden counties in Massachusetts and in Mesa County, Colorado, where detected culti-

vation is largest. A uniform 1 percent increase in retail rates across the 581 counties raises their captive customers' bills by 1.13 billion dollars a year. Panel C of Appendix Figure B5 gives the increase by state. The 16.2 MW decline per county in the close-ballot sample corresponds to more than 150 million dollars of facility investment in shell, power, and cooling, before any computing hardware. This is a one-time measure of the local investment forgone when facilities locate elsewhere or are not built.

8 Conclusion

After narrowly passing recreational cannabis ballots, indoor cultivation expands and data center activity declines. The active data center stock falls by 0.702 facilities per county and operating capacity by 16.2 MW. The decline exceeds half of the stock these counties would otherwise have held over the post-ballot years. Five years after the ballot, the added cultivation implies an average electricity demand of 22 MW per county, of the same scale as the nameplate data center capacity lost. In legalizing counties, retail prices rise 2.2 percent more after the ballot for each square kilometer of detected cultivation and 2.6 percent more for industrial users, while total electricity consumption does not move measurably. We find little evidence for labor, land, capital, water, or crime channels.

The decline in data centers is larger where cannabis supply was already established, and it also appears in counties in other states that share the grid, where the ballot does not legalize cultivation. Solar and gas generators file fewer interconnection applications where cultivation expands on their utility's grid. These patterns are consistent with siting decisions that respond to expected grid scarcity over a facility's lifetime. The siting model is calibrated to the ballot estimates. Without legalization by ballot, national data center capacity would be 10.6 gigawatts, or 30 percent, above its 2024 level if displaced facilities are not built elsewhere. When national capacity is fixed, legalization instead shifts 6.0 gigawatts to other states, mainly Virginia, Texas, and Georgia. Higher rates where cultivation expands cost captive customers in the narrowly passing states 0.7 billion dollars a year.

Utilities and commissions decide in which order large loads connect and who pays for the capacity they claim. When grid capacity adjusts slowly, the effect on later investment depends on the connection charges faced by new customers and the timing of transmission expansion. The same grid lock-in may matter for electricity-intensive industries, including cryptocurrency mining, semiconductor fabrication, and battery and electrolytic hydrogen production, wherever new load arrives faster than transmission is built.

References

Acemoglu, Daron and Pascual Restrepo, "The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 2018, *108* (6), 1488–1542.

_____ **and** _____, "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

Adda, Jérôme, Brendon McConnell, and Imran Rasul, "Crime and the Depenalization of Cannabis Possession: Evidence from a Policing Experiment," *Journal of Political Economy*, 2014, *122* (5), 1130–1202.

Allcott, Hunt and Daniel Keniston, "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America," *Review of Economic Studies*, 2018, *85* (2), 695–731.

Alvarez, Fernando E., David Argente, Joyce Chow, and Diana Van Patten, "Data Centers and Local Economies in the Age of AI: A Shift-Share Approach," Working Paper 35194, National Bureau of Economic Research May 2026. Revised September 2026.

Ancona, Camilla, Ofek Lauber Bonomo, Anton Rozhkov, and Maurizio Porfiri, "Urban Infrastructure and Fossil Fuel Industrial Legacy Drive US Data Center Siting," *Nature Cities*, 2026, *3*, 756–767.

Anderson, D. Mark, Benjamin Hansen, and Daniel I. Rees, "Medical Marijuana Laws, Traffic Fatalities, and Alcohol Consumption," *Journal of Law and Economics*, 2013, *56* (2), 333–369.

Arthur, W. Brian, "Competing Technologies, Increasing Returns, and Lock-In by Historical Events," *Economic Journal*, 1989, *99* (394), 116–131.

Asher, Sam and Paul Novosad, "Rural Roads and Local Economic Development," *American Economic Review*, 2020, *110* (3), 797–823.

Benetton, Matteo, Giovanni Compiani, and Adair Morse, "When Cryptomining Comes to Town: High Electricity-Use Spillovers to the Local Economy," Working Paper 31312, National Bureau of Economic Research 2023.

Black, Dan, Terra McKinnish, and Seth Sanders, "The Economic Impact of the Coal Boom and Bust," *Economic Journal*, 2005, *115* (503), 449–476.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Borenstein, Severin and James Bushnell, "The US Electricity Industry After 20 Years of Restructuring," *Annual Review of Economics*, 2015, *7*, 437–463.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson, "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies," *American Economic Journal: Macroeconomics*, 2021, *13* (1), 333–372.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

_____ **and Douglas L. Miller**, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 2015, *50* (2), 317–372.

Carah, Jennifer K., Jeanette K. Howard, Sally E. Thompson, Anne G. Short Gianotti, Scott D. Bauer, Stephanie M. Carlson, David N. Dralle, Mourad W. Gabriel, Lisa L. Hulette, Brian J. Johnson, Curtis A. Knight, Sarah J. Kupferberg, Stefanie L. Martin, Rosamond L. Naylor, and Mary E. Power, "High Time for Conservation: Adding the Environment to the Debate on Marijuana Liberalization," *BioScience*, 2015, *65* (8), 822–829.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020a, *115* (531), 1449–1455.

_____**, Rocío Titiunik, and Gonzalo Vazquez-Bare**, "Analysis of Regression-Discontinuity Designs with Multiple Cutoffs or Multiple Scores," *Stata Journal*, 2020b, *20* (4), 866–891.

Caughey, Devin and Jasjeet S. Sekhon, "Elections and the Regression Discontinuity Design: Lessons from Close U.S. House Races, 1942–2008," *Political Analysis*, 2011, *19* (4), 385–408.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Cheng, Cheng, Walter J. Mayer, and Yanling Mayer, "The Effect of Legalizing Retail Marijuana on Housing Values: Evidence from Colorado," *Economic Inquiry*, 2018, *56* (3), 1585–1601.

Colas, Mark and Emmett Reynier, "Optimal Subsidies for Residential Solar," *Journal of Political Economy Microeconomics*, 2026, *4* (1), 213–264.

David, Paul A., "Clio and the Economics of QWERTY," *American Economic Review*, 1985, *75* (2), 332–337.

Davidson, Amelia, "Data Center Lobby Grumbles about Reworked Ratepayer Protection Act," E&E News, July 22, 2026.

Denver Department of Public Health and Environment, "Cannabis Environmental Best Management Practices Guide: Energy," Technical Report, City and County of Denver, Denver, CO October 2019.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 2018, *108* (4-5), 899–934.

_____ **and Richard Hornbeck**, "Railroads and American Economic Growth: A 'Market Access' Approach," *Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

_____ **and Adam Storeygard**, "The View from Above: Applications of Satellite Data in Economics," *Journal of Economic Perspectives*, 2016, *30* (4), 171–198.

Du, Rui, Xiaojin Sun, and Siqi Zheng, "Bits and Mortar: The Microspatial Economics of Data Centers," Research Paper 26/05, MIT Center for Real Estate 2026.

Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder, "On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races," *American Journal of Political Science*, 2015, *59* (1), 259–274.

Electric Power Research Institute, "Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption," White Paper 3002028905, Electric Power Research Institute, Palo Alto, CA 2024.

_____, "Powering Intelligence 2026: Updated Scenarios of U.S. Data Center Electricity Use and Power Strategies," White Paper 3002034696, Electric Power Research Institute, Palo Alto, CA February 2026.

The Empire District Electric Company, "2020 Integrated Resource Plan Annual Update Report," Missouri Public Service Commission File No. EO-2020-0284 March 2020.

Evergy Metro, Inc., "Evergy Metro 2025 Annual Update Integrated Resource Plan, Public Version," Filed with the Missouri Public Service Commission March 2025.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Federal Energy Regulatory Commission, "Energy Primer: A Handbook for Energy Market Basics," Staff Report, Federal Energy Regulatory Commission, Office of Energy Policy and Innovation, Washington, DC December 2023.

Feher, Adam, Emilia Garcia-Appendini, and Roxana Mihet, "Is AI Trained on Public Money? Evidence from U.S. Data Centers," CEPR Discussion Paper 20758, Centre for Economic Policy Research 2025.

Gavrilova, Evelina, Takuma Kamada, and Floris Zoutman, "Is Legal Pot Crippling Mexican Drug Trafficking Organisations? The Effect of Medical Marijuana Laws on US Crime," *Economic Journal*, 2019, *129* (617), 375–407.

Gelman, Andrew and Guido W. Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Glaeser, Edward L. and Joshua D. Gottlieb, "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 2008, *39* (1), 155–239.

Goldfarb, Avi and Catherine Tucker, "Digital Economics," *Journal of Economic Literature*, 2019, *57* (1), 3–43.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, 2016, *8* (3), 1–30.

Guimarães, Paulo, Octávio Figueiredo, and Douglas Woodward, "A Tractable Approach to the Firm Location Decision Problem," *Review of Economics and Statistics*, 2003, *85* (1), 201–204.

Hausman, Catherine, "Power Flows: Transmission Lines, Allocative Efficiency, and Corporate Profits," *American Economic Review*, 2025, *115* (8), 2574–2615.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, *102* (2), 994–1028.

Hsieh, Chang-Tai and Enrico Moretti, "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 1–39.

International Energy Agency, "Electricity 2024: Analysis and Forecast to 2026," Technical Report, International Energy Agency, Paris 2024.

_____, "Energy and AI," World Energy Outlook Special Report, International Energy Agency, Paris April 2025.

Johnston, Sarah, Yifei Liu, and Chenyu Yang, "An Empirical Analysis of the Interconnection Queue," Working Paper 31946, National Bureau of Economic Research December 2023. Revised August 2026.

Joint Legislative Audit and Review Commission, "Data Centers in Virginia 2024," Report 598 to the Governor and the General Assembly of Virginia, Joint Legislative Audit and Review Commission, Richmond, VA December 2024.

Kahn, Matthew E. and Erin T. Mansur, "Do Local Energy Prices and Regulation Affect the Geographic Concentration of Employment?," *Journal of Public Economics*, 2013, *101*, 105–114.

Lee, David S., "Randomized Experiments from Non-Random Selection in U.S. House Elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

Li, Pengfei, Jianyi Yang, Mohammad A. Islam, and Shaolei Ren, "Making AI Less 'Thirsty'," *Communications of the ACM*, 2025, *68* (7), 54–61.

MacKinnon, James G. and Matthew D. Webb, "The Wild Bootstrap for Few (Treated) Clusters," *Econometrics Journal*, 2018, *21* (2), 114–135.

Masanet, Eric, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey, "Recalibrating Global Data Center Energy-Use Estimates," *Science*, 2020, *367* (6481), 984–986.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

McFadden, Daniel, "Modelling the Choice of Residential Location," in Anders Karlqvist, Lars Lundqvist, Folke Snickars, and Jörgen W. Weibull, eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland, 1978, pp. 75–96.

Mills, Evan, "The Carbon Footprint of Indoor Cannabis Production," *Energy Policy*, 2012, *46*, 58–67.

_____, "Energy-Intensive Indoor Cultivation Drives the Cannabis Industry's Expanding Carbon Footprint," *One Earth*, 2025, *8* (3), 101179.

Monitoring Analytics, "Analysis of the 2025/2026 RPM Base Residual Auction, Part G," Technical Report, Monitoring Analytics, Independent Market Monitor for PJM June 2025.

Moretti, Enrico, "Local Multipliers," *American Economic Review*, 2010, *100* (2), 373–377.

_____, "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, 2021, *111* (10), 3328–3375.

Mytton, David, "Data Centre Water Consumption," *npj Clean Water*, 2021, *4* (1), 11.

North American Electric Reliability Corporation, "Glossary of Terms Used in NERC Reliability Standards," North American Electric Reliability Corporation, Atlanta, GA 2025. Updated July 10, 2025.

Pan Fang, Tommy and Shane Greenstein, "Where the Cloud Rests: The Economic Geography of Data Centers," *Strategy Science*, 2025, *10* (4), 404–420.

PJM Interconnection, "2025/2026 Base Residual Auction Report," Technical Report, PJM Interconnection, Audubon, PA 2024.

Powell, Devyn, Shahnawaz Siddiqui, Jessica Kerby, Alan Cooke, Sara Murphy, and Thomas Benson, "A Framework for Streamlining Large Load Interconnection: An Assessment of Current Practices and Emerging Strategies for Fair and Efficient Load Interconnection Processes," Report PNNL-38441, Pacific Northwest National Laboratory, Richland, WA September 2025.

Rambachan, Ashesh and Jonathan Roth, "A More Credible Approach to Parallel Trends," *Review of Economic Studies*, 2023, *90* (5), 2555–2591.

Rand, Joseph, Nick Manderlink, Will Gorman, Ryan Wiser, Joachim Seel, Julie Mulvaney Kemp, Seongeun Jeong, and Fritz Kahrl, "Queued Up: 2024 Edition, Characteristics of Power Plants Seeking Transmission Interconnection As of the End of 2023," Technical Report, Lawrence Berkeley National Laboratory, Berkeley, CA April 2024.

_____, _____, **Steven Zhang, Chris Talley, Will Gorman, Ryan Wiser, Joachim Seel, Julie Mulvaney Kemp, Seongeun Jeong, and Fritz Kahrl**, "Queued Up: 2025 Edition, Characteristics of Power Plants Seeking Transmission Interconnection As of the End of 2024," Technical Report, Lawrence Berkeley National Laboratory, Berkeley, CA December 2025.

Schmidheiny, Kurt and Marius Brülhart, "On the Equivalence of Location Choice Models: Conditional Logit, Nested Logit and Poisson," *Journal of Urban Economics*, 2011, *69* (2), 214–222.

Shehabi, Arman, Sarah J. Smith, Alex Hubbard, Alex Newkirk, Nuoa Lei, Md Abu Bakar Siddik, Billie Holecek, Jonathan Koomey, Eric Masanet, and Dale Sartor, "2024 United States Data Center Energy Usage Report," Technical Report LBNL-2001637, Lawrence Berkeley National Laboratory, Berkeley, CA December 2024.

Shreve, Nathan, Zach Zimmerman, Gretchen Kershaw, and Rob Gramlich, "Fewer New Miles: Strategic Industries Held Back by Slow Pace of Transmission," Technical Report, Grid Strategies July 2025.

Siddik, Md Abu Bakar, Arman Shehabi, and Landon Marston, "The Environmental Footprint of Data Centers in the United States," *Environmental Research Letters*, 2021, *16* (6), 064017.

Southwest Power Pool, "A Comprehensive Review of Southwest Power Pool's Response to the February 2021 Winter Storm: Analysis and Recommendations," Technical Report, Southwest Power Pool, Little Rock, AR July 2021a.

_____, "Update of 2021 Winter Weather Event Comprehensive Review," Presentation to the Oklahoma Corporation Commission, August 17, 2021b.

_____, "Planning Reserve Margin Recommendation, with Attachment B: 2023 SPP Loss of Load Expectation Report," Technical Report, Southwest Power Pool, Little Rock, AR July 2024.

_____, "Markets and Operations," Southwest Power Pool web page, accessed October 17, 2026.

Suárez Serrato, Juan Carlos and Owen Zidar, "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms," *American Economic Review*, 2016, *106* (9), 2582–2624.

Summers, Hailey M., Evan Sproul, and Jason C. Quinn, "The Greenhouse Gas Emissions of Indoor Cannabis Production in the United States," *Nature Sustainability*, 2021, *4* (7), 644–650.

Turner & Townsend, "Data Centre Construction Cost Index 2025-2026," Technical Report, Turner & Townsend, Leeds, United Kingdom November 2025.

Unruh, Gregory C., "Understanding Carbon Lock-in," *Energy Policy*, 2000, *28* (12), 817–830.

U.S. Energy Information Administration, "Hourly Information on U.S. Electricity Supply, Demand, and Flows Is Now Available," Today in Energy, July 25, 2016a.

_____, "U.S. Electric System Is Made Up of Interconnections and Balancing Authorities," Today in Energy, July 20, 2016b.

_____, "2018 Commercial Buildings Energy Consumption Survey, Table C14: Electricity Consumption and Expenditure Intensities, 2018," Technical Report, U.S. Energy Information Administration, Washington, DC December 2022.

_____, "Tracking Electricity Consumption from U.S. Cryptocurrency Mining Operations," Today in Energy, February 1, 2024.

Vichyanond, Jade, "Data Center Boom in Malaysia: Recent Developments and Challenges," Selected Issue, AMRO Annual Consultation Report: Malaysia 2026, ASEAN+3 Macroeconomic Research Office, Singapore May 2026.

Wartenberg, Ariani C., Patricia A. Holden, Hekia Bodwitch, Phoebe Parker-Shames, Thomas Novotny, Thomas C. Harmon, Stephen C. Hart, Marc Beutel, Michelle Gilmore, Eunha Hoh, and Van Butsic, "Cannabis and the Environment: What Science Tells Us and What We Still Need to Know," *Environmental Science & Technology Letters*, 2021, *8* (2), 98–107.

Waugh, Michael E., "Trade in AI-Related Products," Working Paper 35053, National Bureau of Economic Research April 2026.

Wen, Hefei, Jason M. Hockenberry, and Janet R. Cummings, "The Effect of Medical Marijuana Laws on Adolescent and Adult Use of Marijuana, Alcohol, and Other Substances," *Journal of Health Economics*, 2015, *42*, 64–80.

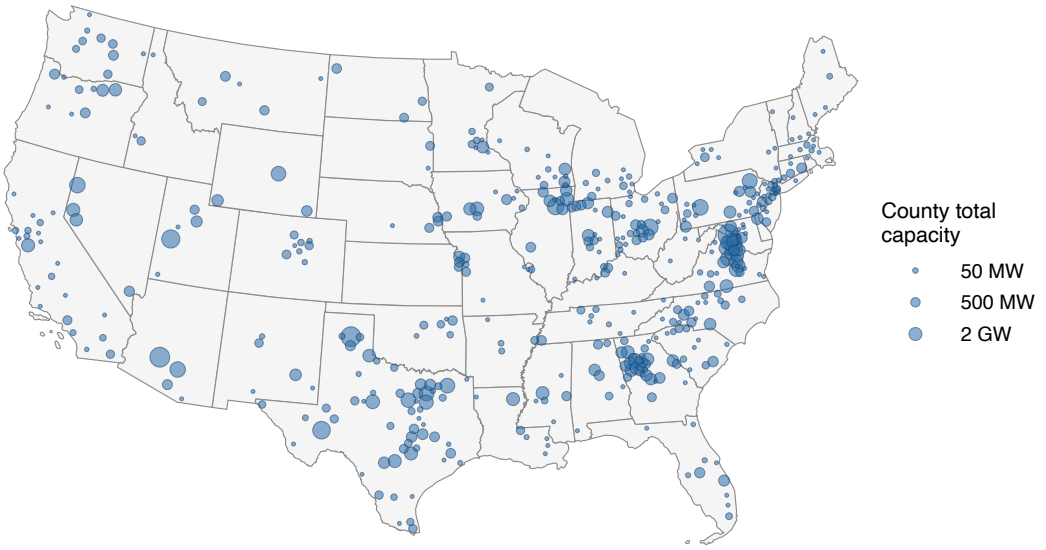
The White House, "Winning the Race: America's AI Action Plan," Technical Report, Executive Office of the President, Washington, DC July 2025.

Wing, Coady, Seth M. Freedman, and Alex Hollingsworth, "Stacked Difference-in-Differences," Working Paper 32054, National Bureau of Economic Research 2024.

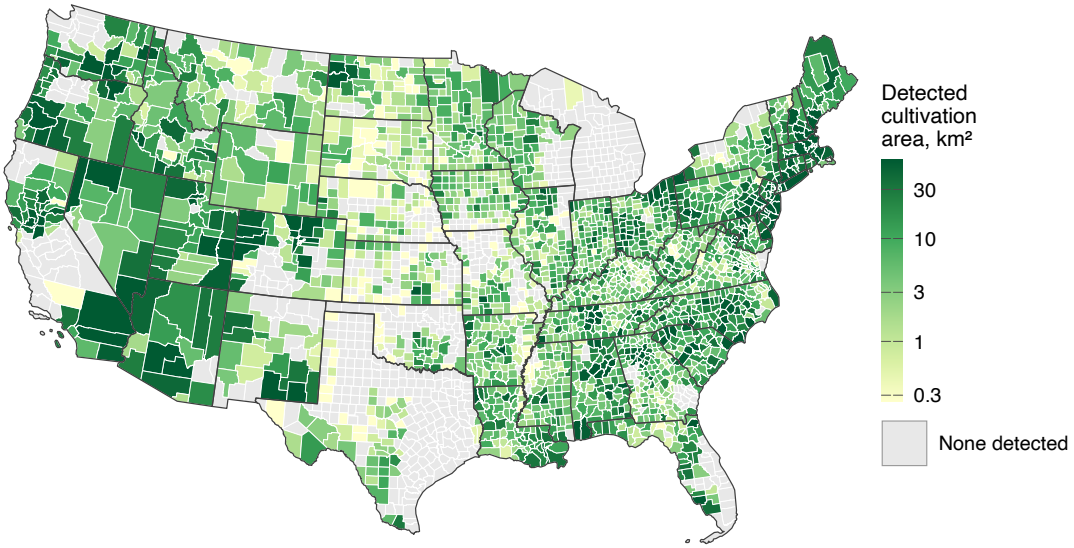
Zhang, Calvin, "AI Data Centers and Electricity Prices: Much Ado About Something?," Working Paper 6930258, SSRN 2026.

Figure 1. U.S. data center locations and detected cannabis cultivation

Panel A. Data center locations



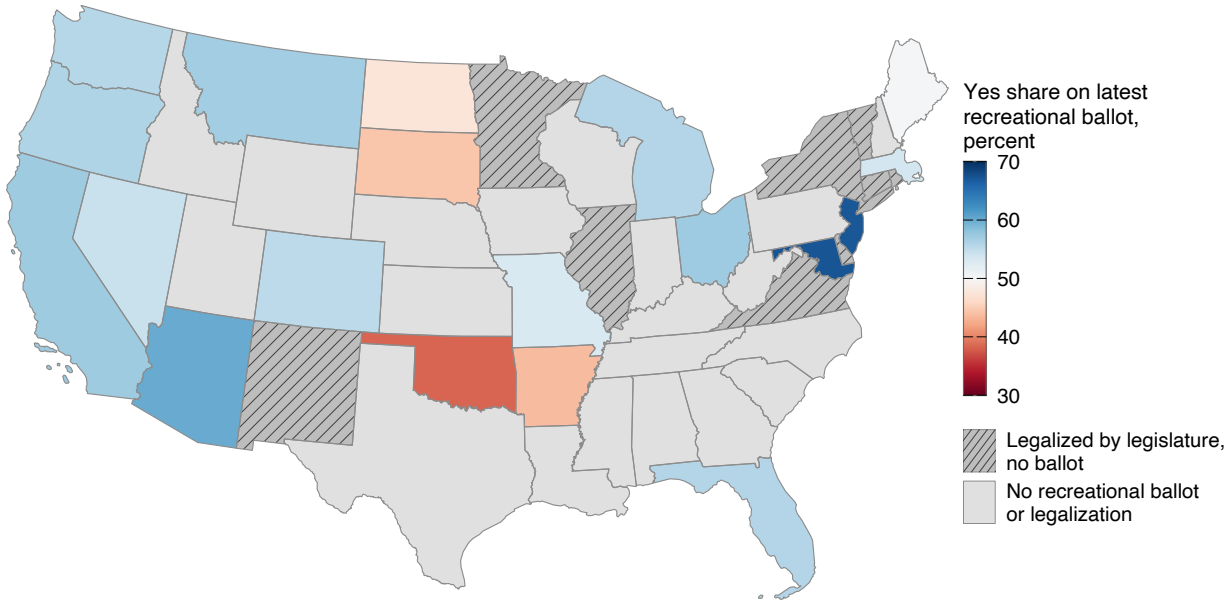
Panel B. Detected cannabis cultivation, 2024



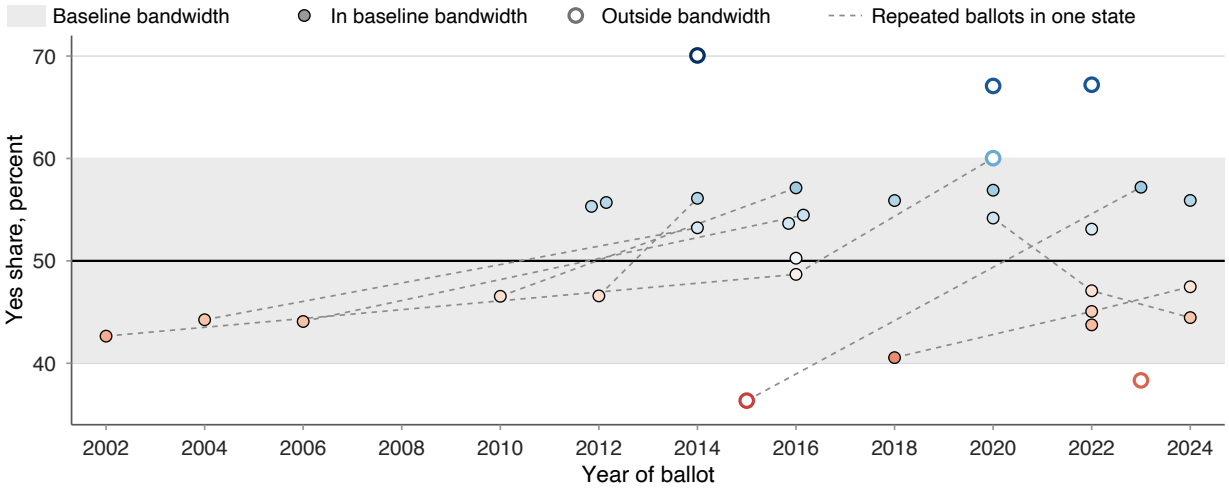
Notes: Both panels show the contiguous United States, and Alaska and Hawaii are omitted. In Panel A, each marker is placed at the centroid of a county with at least one data center facility, and its area is proportional to the square root of the total nameplate power capacity of the facilities in that county in megawatts. Counties whose facilities total less than 10 megawatts are drawn at the 10-megawatt size. The sample includes 5,140 facilities in 461 counties at every stage recorded in the database. Facilities without a reported capacity enter the sample but add no capacity to their county. Data for Panel A come from the Aterio facility-level data center database. In Panel B, counties are shaded by the detected high-confidence cultivation area in 2024 in square kilometers on a logarithmic color scale. The scale spans the 5th-95th percentiles of the area among counties with detected cultivation, and counties outside this range take the color at the nearer end of the scale. Counties with no detected cultivation are shown in light gray. The area comes from the three-signal detection algorithm described in Appendix A. We compute this area from VIIRS nightlight, Landsat thermal, and Sentinel-2 imagery.

Figure 2. Recreational cannabis ballots

Panel A. State outcomes

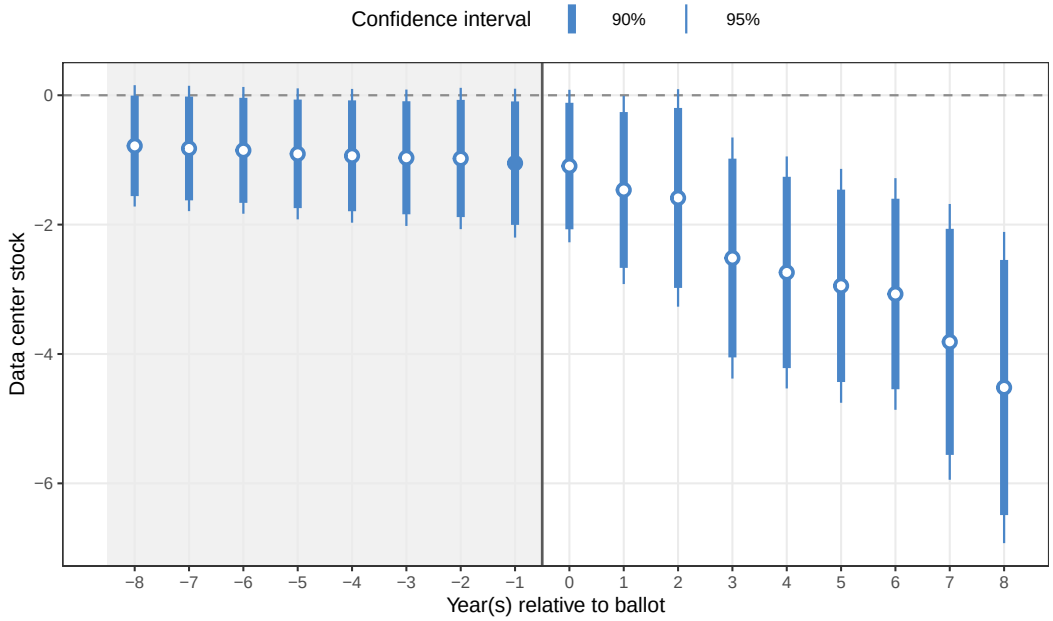


Panel B. Every ballot in the design



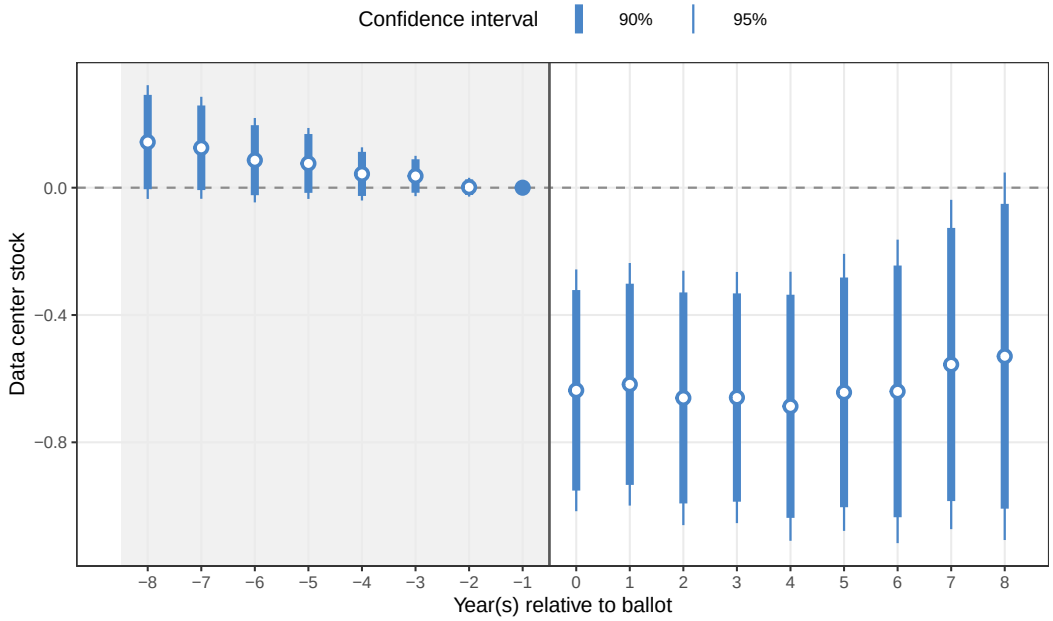
Notes: In Panel A, states are shaded by the Yes share on the most recent recreational cannabis ballot. The diverging color scale spans 30-70 percent and is centered at the 50 percent simple-majority threshold, so blue shades indicate a Yes majority and red shades a No majority. States that legalized recreational use through legislative action rather than ballot are shaded gray with diagonal hatching. States with no recreational ballot and no recreational legalization are shaded light gray. Panel A shows the continental United States, and Alaska and Hawaii are omitted from the map. Panel B plots every ballot from 2002 to 2024 by election year and Yes share, with markers colored on the scale of Panel A. The horizontal line marks the 50 percent threshold, and the shaded band marks the baseline bandwidth of ten percentage points on either side of it. Filled markers are the 26 ballots within ten percentage points of the cutoff that enter the baseline estimates, 14 passing and 12 failing in eight states; hollow markers lie outside this bandwidth. Dashed lines join ballots in the same state. Where two ballots in the same year have Yes shares within one percentage point of each other, their markers are offset horizontally. Two earlier ballots, California in 1972 and Oregon in 1986, lie outside the bandwidth and are not shown. Ballot data come from the Ballotpedia recreational cannabis ballot history.

Figure 3. Cross-sectional data center stock gap by event time



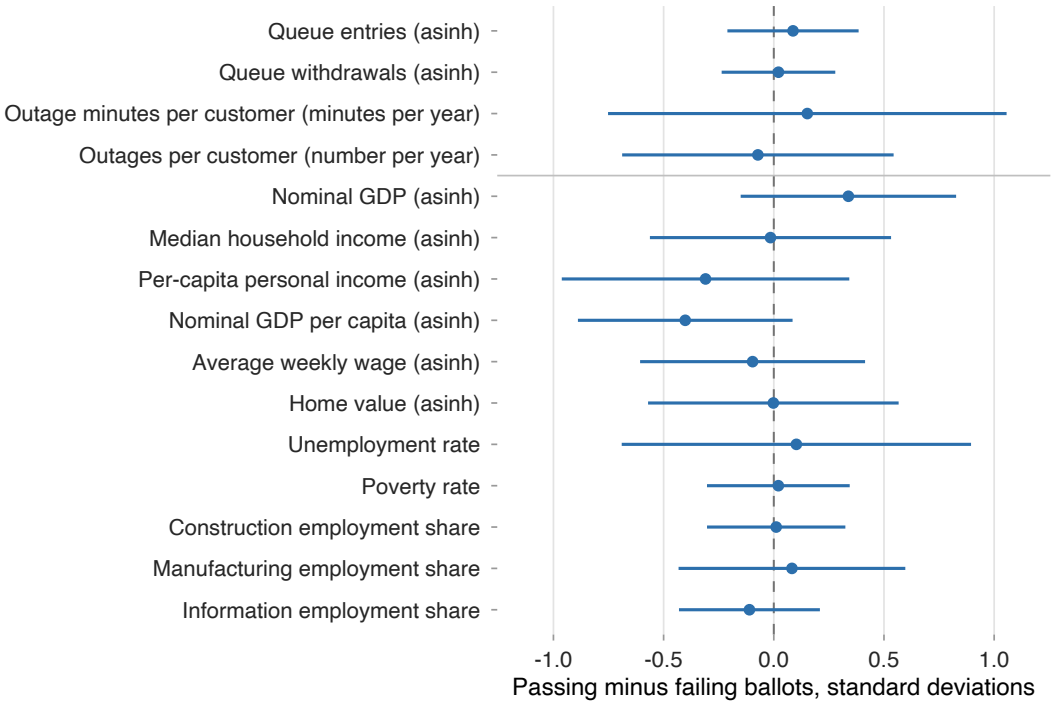
Notes: The figure plots the pooled cross-sectional RD coefficient on *Pass*, estimated separately at each event time k. Bars show 90 and 95 percent confidence intervals, based on standard errors clustered at the ballot level. The estimation sample is restricted to $|R_{sa}| \leq 0.10$. The outcome has a mean of 0.446 active facilities per county-year.

Figure 4. Data center stock around cannabis legalization



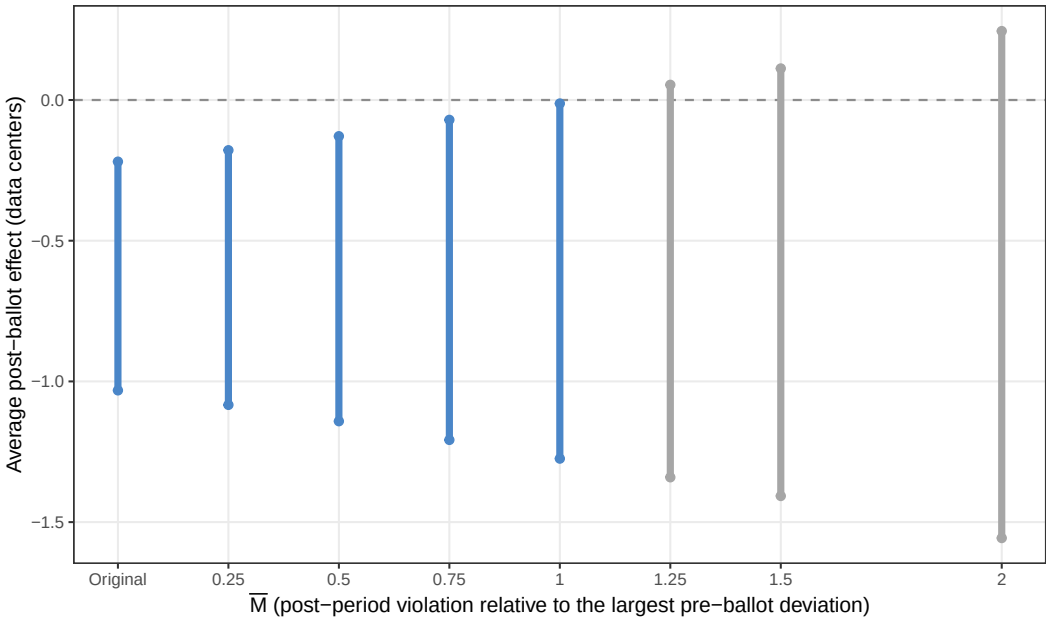
Notes: The figure reports a panel difference-in-discontinuities event study in which the single post-ballot term is replaced by interactions of *Pass* with event time, and the year before the ballot is the omitted reference period. All specifications include county-by-ballot and calendar-year fixed effects. Bars show 90 and 95 percent confidence intervals, based on standard errors clustered at the ballot level. The estimation sample is restricted to $|R_{sa}| \leq 0.10$. The outcome has a mean of 0.446 active facilities per county-year.

Figure 5. County characteristics before the ballot



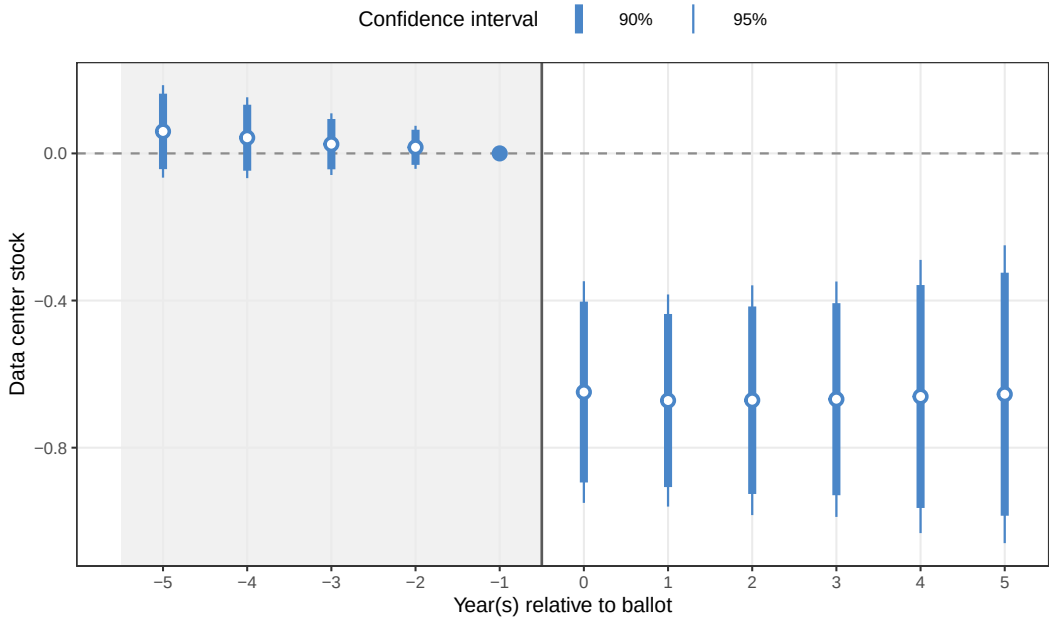
Notes: Each row reports the coefficient from a regression of a county characteristic on the pass indicator alone, with the characteristic standardized using its sample mean and standard deviation. The unit is the county-ballot cell in the year before each ballot, for the 26 ballots with $|R_{sa}| \leq 0.10$, 14 passing and 12 failing. Queue entries and withdrawals are the inverse hyperbolic sine of the number of generator interconnection requests entering the queue and the number withdrawn from it, from the LBNL Queued Up data, which do not cover Alaska, so these rows use 24 ballots. Outage minutes and outages per customer are the system average interruption duration and frequency indices from EIA Form 861, counting interruptions longer than five minutes and excluding major event days. Each utility's statewide customers and customer-minutes are allocated across the counties it serves in proportion to their 2000 Census population, as for retail prices. Outage data begin in 2013, so these rows use the 19 ballots from 2014 on. Output and per-capita personal income come from the Bureau of Economic Analysis, median household income and poverty from the Small Area Income and Poverty Estimates, unemployment from the Local Area Unemployment Statistics, home values from Zillow, and wages and employment shares from the Quarterly Census of Employment and Wages. Dollar-valued variables are expressed in current dollars and transformed using the inverse hyperbolic sine. Outage indices, rates, and employment shares enter in levels before standardization. Horizontal lines are 95 percent confidence intervals based on standard errors clustered at the ballot level. The gray line across the full width of the panel separates the four grid measures at the top from the eleven economic measures.

Figure 6. Sensitivity of the post-ballot effect to violations of parallel trends



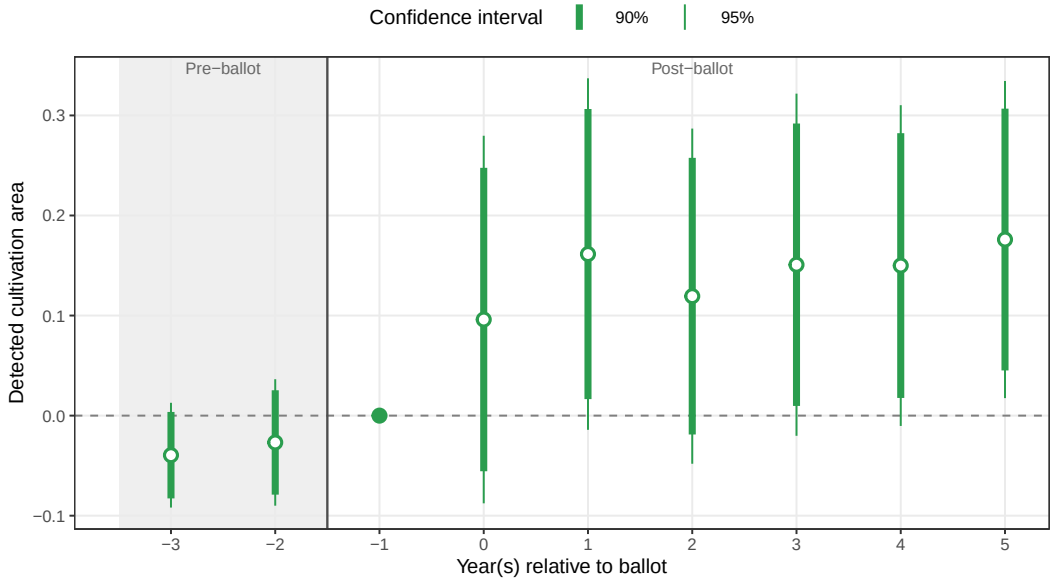
Notes: The figure reports 95 percent confidence sets for the average post-ballot effect on the active data center stock under the relative-magnitudes restrictions of Rambachan and Roth (2023). The original interval imposes exact parallel trends after the ballot, $\bar{M} = 0$. The remaining intervals allow post-ballot violations up to $\bar{M}$ times the largest violation observed across the seven pre-ballot coefficients, for values of $\bar{M}$ up to 2. The average is taken over the nine post-ballot event-study coefficients. Intervals that exclude zero are shown in blue and those that admit zero in gray. The event study is the panel difference-in-discontinuities specification of Figure 4, with standard errors clustered at the ballot level and the sample restricted to $|R_{sa}| \leq 0.10$.

Figure 7. Data center stock around cannabis legalization, balanced five-year window



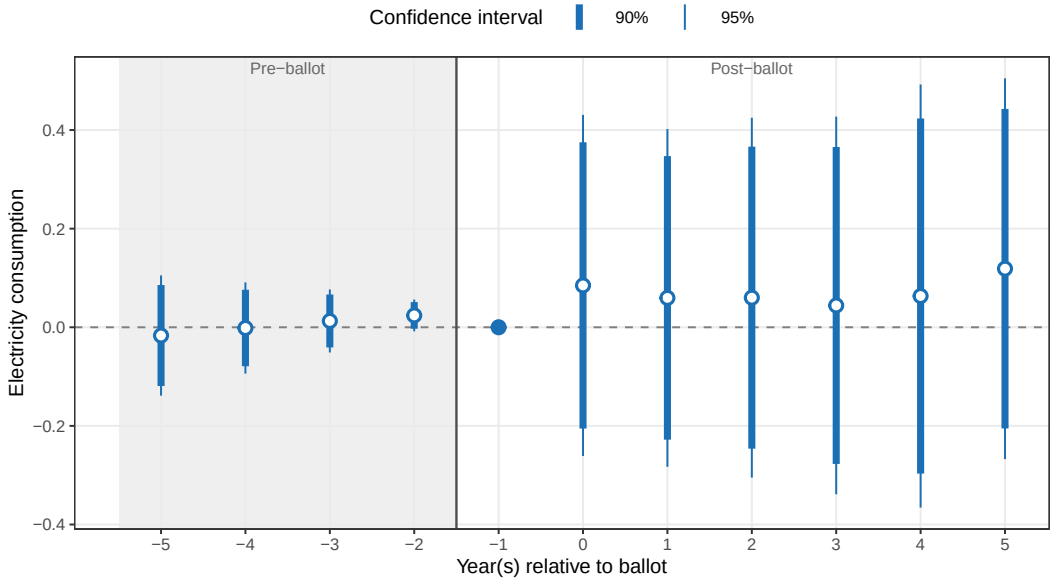
Notes: The sample is the 12 ballots whose counties are observed in every year from five before to five after their own ballot, so the same ballots contribute to every event-time estimate. Nine of these ballots passed and three failed, and the sample covers 409 counties and 5,324 county-ballot-years. Arizona's 2016 ballot and Oregon's 2012 ballot are excluded from the control group because Arizona legalized in 2020 and Oregon in 2014, inside their own windows. The legalization dates are those of Section 5.4, taken from every passing ballot at any margin. Each point is the passing-minus-failing gap in the active data center stock at that year relative to the ballot, with the year before the ballot as the reference. Bars show 90 and 95 percent confidence intervals from the clustered-t distribution, based on standard errors clustered at the ballot level. The specification includes county-by-ballot and calendar-year fixed effects, and the estimation sample is restricted to $|R_{sa}| \leq 0.10$.

Figure 8. Detected cannabis cultivation around legalization



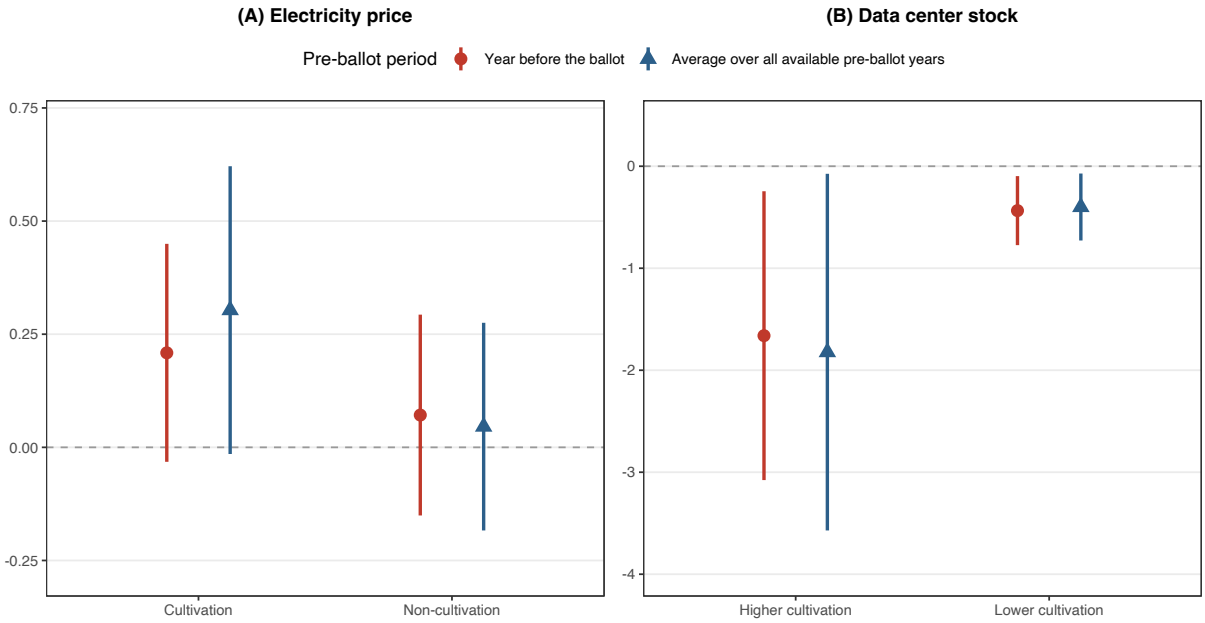
Notes: The figure reports a panel difference-in-discontinuities event study for cultivation area detected by all three signals in square kilometers, estimated over a plus or minus five-year window, and the year before the ballot is the omitted reference period. Bars show 90 and 95 percent confidence intervals, based on standard errors clustered at the ballot level. The estimation sample is restricted to $|R_{sa}| \leq 0.10$ and contains 6,927 county-ballot-year observations. The display is trimmed to three years before the ballot; the regression uses the full plus or minus five-year panel. The outcome has a mean of 0.0485 square kilometers per county-year.

Figure 9. Total electricity consumption around legalization



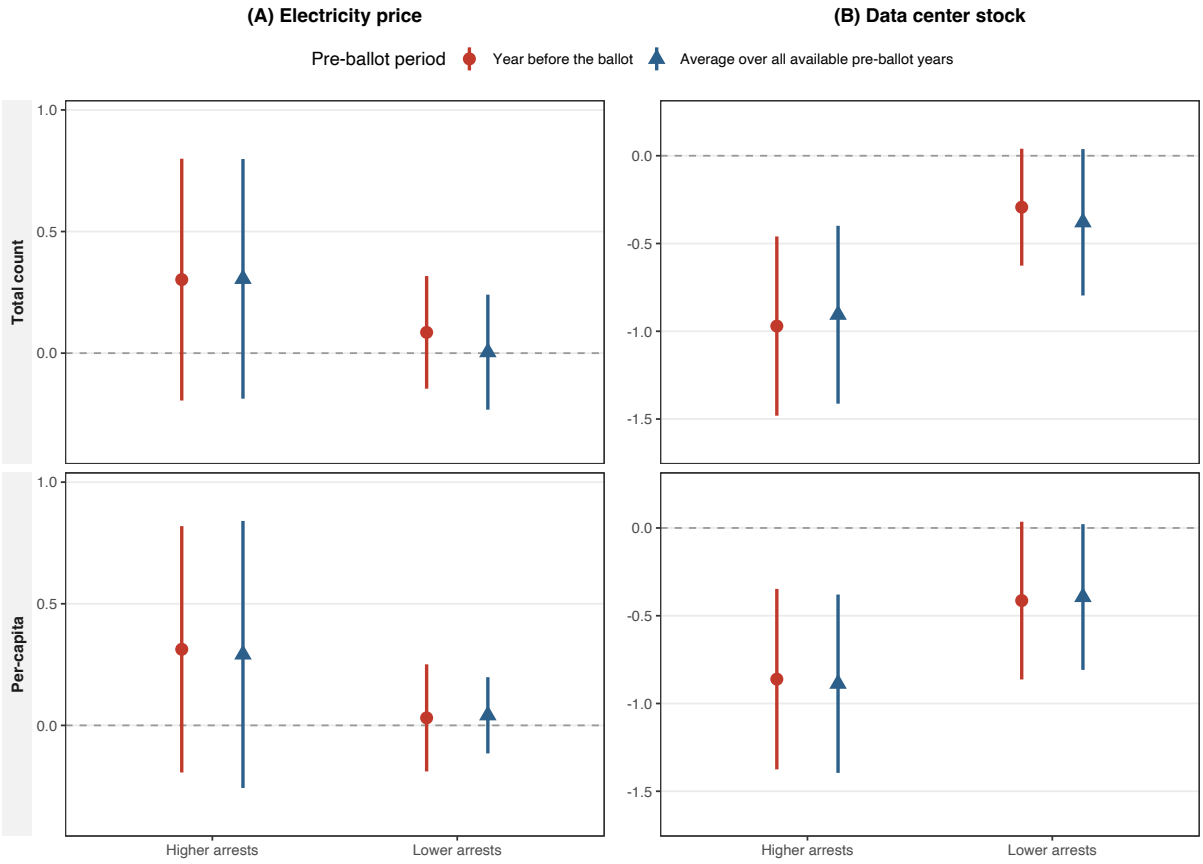
Notes: The figure reports a panel difference-in-discontinuities event study for population-weighted total electricity consumption in inverse hyperbolic sine of megawatt-hours, estimated over a plus or minus five-year window, and the year before the ballot is the omitted reference period. Bars show 90 and 95 percent confidence intervals, based on standard errors clustered at the ballot level. The estimation sample is restricted to $|R_{sa}| \leq 0.10$ and contains 11,563 county-ballot-year observations. The plus or minus eight-year version appears in the appendix as Figure B3. The outcome has a mean of 1.43 million megawatt-hours per county-year.

Figure 10. Price and data center responses by pre-ballot cultivation exposure



Notes: The figure reports the difference-in-discontinuities Pass $\times$ Post effect on the average retail electricity price (Panel A, log) and the active data center stock (Panel B, levels), with counties split by pre-ballot satellite-detected cultivation status. Panel A groups counties by whether the three-signal measure detects any cultivation and uses an estimation window of three years before and after the ballot. Panel B groups counties by whether the two-signal cultivation measure exceeds the sample mean and uses an eight-year window on either side of the ballot. The two panels use different vertical scales. The two series within each panel use the cultivation measure in the year before the ballot $(t-1)$ and the average over all available pre-ballot years, shown as red circles and blue triangles respectively. The bandwidth is $|R_{sa}| \leq 0.10$, and the year before the ballot is the omitted reference period. All specifications include county-by-ballot and calendar-year fixed effects. The 95 percent confidence intervals use standard errors clustered at the ballot level.

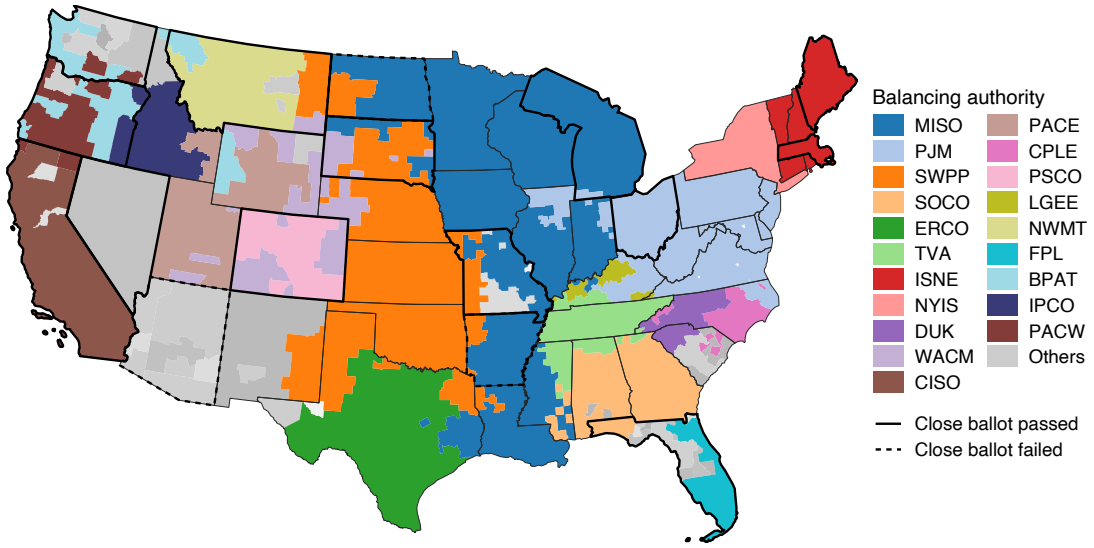
Figure 11. Price and data center responses by pre-ballot cannabis-supply enforcement



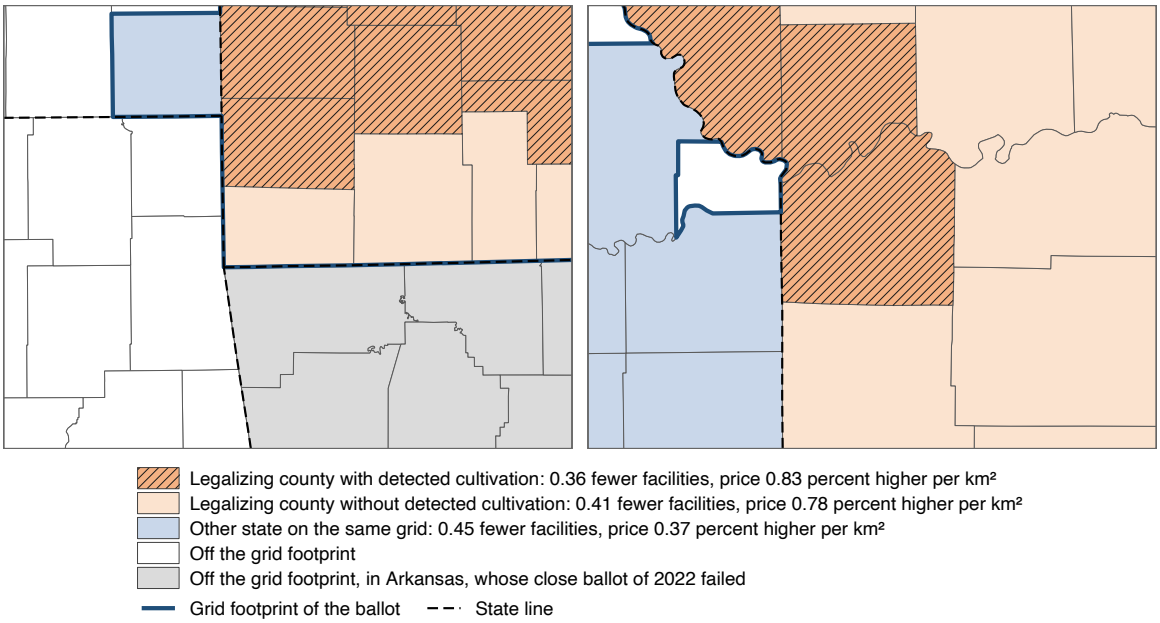
Notes: The figure reports the difference-in-discontinuities Pass × Post effect on the average retail electricity price (Panel A, log) and the active data center stock (Panel B, levels), with counties split by pre-ballot cannabis-supply enforcement, namely marijuana sale and manufacturing arrests from the Uniform Crime Reporting Arrests by Age, Sex, and Race series (Kaplan, openICPSR #102263). Unlike the three-signal satellite measure in Figure 10, which begins in 2017, the arrest data extend back to the 1970s and provide pre-ballot measures for all 26 ballots, including the 2012 ballots in Colorado and Washington. The top row uses the total count of pre-ballot sale and manufacturing arrests as the grouping measure, and the bottom row uses the per-capita rate. Counties are split at the sample-wide median of each grouping measure. Each column uses a common vertical scale across its two rows. Red circles classify counties using the year before the ballot $(t-1)$. Blue triangles use the average over all available pre-ballot years. The estimation window is plus or minus eight years, the bandwidth is $|R_{sa}| \leq 0.10$, and the year before the ballot is the omitted reference period. All specifications include county-by-ballot and calendar-year fixed effects. The 95 percent confidence intervals use standard errors clustered at the ballot level.

Figure 12. Balancing authorities and the grid footprint of a ballot

Panel A. Balancing authorities

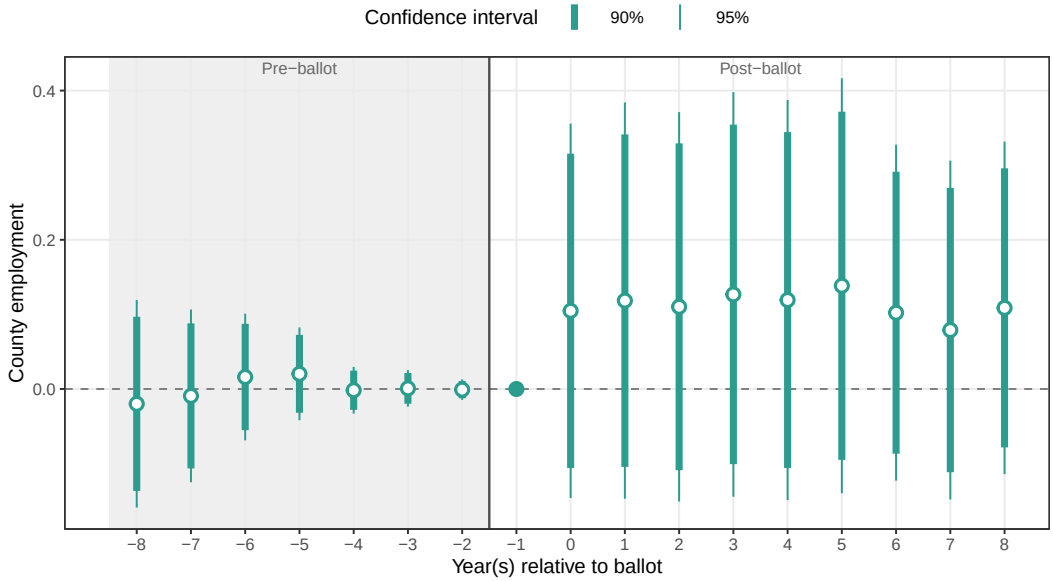


Panel B. The grid footprint of one ballot at two state borders



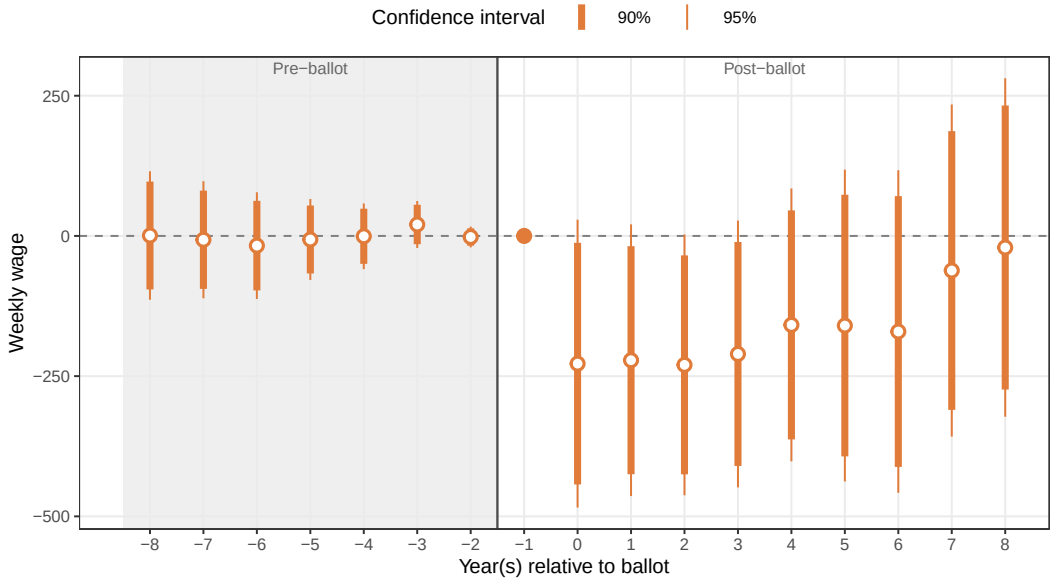
Notes: Panel A shades each county by the balancing authority that serves the largest share of its customers in EIA Form 861 over 2012-2024, with colors for the twenty largest authorities by county count and gray for the rest, and the legend gives the EIA code of each colored authority. Solid outlines mark states whose close recreational ballot passed, and dashed outlines mark states whose close ballots all failed. A close ballot falls within ten percentage points of the majority threshold. Panel B shows the grid footprint of Missouri's 2022 ballot near Joplin and Kansas City, in two windows drawn at different scales. The footprint defined in Section 6.4 is outlined in dark blue, and dashed black lines mark state borders. Within legalizing states, hatched counties have detected cultivation and plain orange counties do not. Blue counties lie in another state on the same grid, and white counties are outside the grid footprint. Light gray counties are in Arkansas, whose close ballot of 2022 failed. The legend reports the estimates for the active data center stock from Column (2) of Table 6 and for electricity prices from Column (2) of Table 4.

Figure 13. Total county employment around legalization



Notes: The figure reports a panel difference-in-discontinuities event study for total county employment in inverse hyperbolic sine, estimated over a plus or minus eight-year window, and the year before the ballot is the omitted reference period. Bars show 90 and 95 percent confidence intervals, based on standard errors clustered at the ballot level. The estimation sample is restricted to $|R_{sa}| \leq 0.10$ and contains 16,870 county-ballot-year observations. Total county employment has a pre-period mean of 57,136 and a median of 6,701 workers per county-year in the estimation sample. Missing county-years are set to zero before the inverse-hyperbolic-sine transform. See Table 7.

Figure 14. Data center sector wages around legalization



Notes: The figure reports a panel difference-in-discontinuities event study for data center sector wages in dollars per week, with non-operating county-years set to zero, estimated over a plus or minus eight-year window, and the year before the ballot is the omitted reference period. Bars show 90 and 95 percent confidence intervals, based on standard errors clustered at the ballot level. The estimation sample is restricted to $|R_{sa}| \leq 0.10$ and contains 16,870 county-ballot-year observations. The outcome has a mean of 1,817 dollars per week among operating facilities.

Table 1. Electricity prices and detected cultivation within county

	Average (1)	Commercial (2)	Industrial (3)
$\hat{\gamma}$: Inverse hyperbolic sine of detected cultivation area (km^2)	0.0508***	0.0457***	0.0700***
	(0.0071)	(0.0134)	(0.0074)
Fixed effects			
County × ballot	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes
Fit statistics			
Observations	6,868	6,868	6,841
R^2	0.84735	0.85561	0.89380
Within R^2	0.00585	0.00443	0.00514

Notes: The table reports within-county estimates of the relationship between detected cultivation and electricity prices, at the county-ballot-year level. The outcome in every column is the natural log of the population-weighted retail electricity price in dollars per kWh for a customer class, namely the all-customer average in Column (1), commercial in Column (2), and industrial in Column (3). The regressor is the inverse hyperbolic sine of detected cultivation area in km^2 from the three-signal satellite measure. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the sample is restricted to 2017-2024. Standard errors are clustered at the ballot level, with 26 or fewer clusters, and are reported in parentheses.

Table 2. Cannabis cultivation, electricity prices, and consumption

	Total		Commercial		Industrial	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Electricity price						
$\hat{\beta}_1$: Pass×Post	0.1199	0.1197	0.1823	0.1823	0.1449	0.1444
	(0.1064)	(0.1051)	(0.1341)	(0.1331)	(0.1195)	(0.1179)
$\hat{\beta}_2$: km^2×Pass×Post		0.0219***		0.0231***		0.0264*
		(0.0072)		(0.0066)		(0.0136)
Observations	6,868	6,868	6,868	6,868	6,841	6,841
Panel B. Electricity consumption						
$\hat{\beta}_1$: Pass×Post	0.0485	0.0492	-0.0579	-0.0574	0.0108	0.0114
	(0.3660)	(0.3650)	(0.3940)	(0.3930)	(0.2085)	(0.2078)
$\hat{\beta}_2$: km^2×Pass×Post		0.0295		0.0369**		0.0266
		(0.0170)		(0.0163)		(0.0220)
Observations	6,868	6,868	6,868	6,868	6,868	6,868
km^2 interaction	No	Yes	No	Yes	No	Yes

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on electricity prices and consumption, at the county-ballot-year level. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus five years around the ballot. The sample is restricted to 2017-2024, the period over which the three-signal satellite measure is available. Standard errors are clustered at the ballot level, with 26 or fewer clusters, and are reported in parentheses. The "Total" columns report the average retail price in Panel A and total consumption in Panel B. Price outcomes are in logs and consumption outcomes are in the inverse hyperbolic sine of MWh. Odd-numbered columns omit the interaction between detected cultivation area in km^2 and Pass×Post. Even-numbered columns include this interaction and report its coefficient, $\hat{\beta}_2$.

Table 3. Sample composition of the utility and balancing-authority grid footprints

Diagnostic	Utility grid	Balancing-authority grid
Footprint coverage		
Footprint counties	1,173	2,429
Distinct grid units	138	35
Ballots	26	26
Panel rows (county-ballot-year)	38,556	192,933
Spatial composition		
Share in ballot state	54%	11%
Share out of ballot state	46%	89%
Cultivation dose		
Mean own cultivation (km^2)	0.024	0.019
Mean rest-of-grid cultivation (km^2)	0.859	11.413
Treated-group counts (post-ballot)		
In-state, cultivating (county-ballots)	112	112
In-state, non-cultivating (county-ballots)	609	609
Other state, same grid (county-ballots)	386	4,177

Notes: The table reports the composition of the spillover sample on the utility and balancing-authority grids. Each grid footprint is constructed for the 26 ballots within the bandwidth $|R_{sa}| \leq 0.10$ as the union of counties served by a grid unit, a utility or a balancing authority, that also serves the balloting state. The treated-group counts are post-ballot county-ballots, split into in-state cultivating counties, in-state non-cultivating counties, and counties in other states on the same grid. The two in-state groups are the same on both grids by construction. Rest-of-grid cultivation is summed over the other counties in the grid-unit footprint.

Table 4. Spillover of cannabis cultivation onto the county electricity price

Grid	Utility		Balancing authority	
Spillover	Pooled	Split	Pooled	Split
	(1)	(2)	(3)	(4)
Variables				
Pass×Post	0.0364	0.0454	-0.0137	-0.0141
	(0.0908)	(0.0890)	(0.0231)	(0.0241)
Own-county km^2×Pass×Post	0.0096***	0.0083**	0.0040	0.0040
	(0.0034)	(0.0037)	(0.0029)	(0.0030)
Pooled spillover km^2×Pass×Post	0.0062***		0.0005**	
	(0.0020)		(0.0002)	
Within-state km^2×Pass×Post		0.0078**		0.0004
		(0.0030)		(0.0003)
Cross-state km^2×Pass×Post		0.0037*		0.0005**
		(0.0020)		(0.0002)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	14,630	14,630	88,561	88,561

Notes: The table reports panel difference-in-discontinuities estimates of the spillover of cannabis cultivation onto the county electricity price, at the county-ballot-year level. The outcome is the log population-weighted average retail price. Own-county cultivation and rest-of-grid cultivation, both in km^2, each enter interacted with Pass×Post. The pooled columns enter rest-of-grid cultivation as one term, and the split columns separate it into a within-state and a cross-state component. An out-of-state county counts as a grid neighbor only if its balancing authority also operates in the balloting state. In-state counties are unaffected by this rule. The rule is redundant on the balancing-authority footprint, where the condition holds by construction. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. The sample is restricted to 2017-2024, the period over which the three-signal satellite measure is available. Standard errors are clustered at the ballot level and are reported in parentheses.

Table 5. Own, within-state, and cross-state spillovers onto the retail electricity price

Grid	Dose	Estimate (%)	SE (%)	95% CI (%)	p-value	MDE (%)
Utility	Own county	0.830**	0.373	[0.098, 1.562]	0.037	1.046
Utility	Within-state	0.778**	0.295	[0.200, 1.356]	0.015	0.826
Utility	Cross-state	0.374*	0.200	[-0.019, 0.766]	0.076	0.561
BA	Own county	0.403	0.301	[-0.187, 0.992]	0.194	0.842
BA	Within-state	0.041	0.028	[-0.013, 0.096]	0.148	0.077
BA	Cross-state	0.049**	0.019	[0.012, 0.085]	0.016	0.052

Notes: The table reports the continuous dose specification of Table 4 on the utility and balancing-authority (BA) footprints, at the county-ballot-year level. Own cultivation, cultivation elsewhere on the grid inside the state, and cultivation on the grid across the state line each enter interacted with Pass×Post, alongside the running-variable polynomial. The outcome is in logs, so the coefficient times one hundred is the percent change per square kilometer of cultivation. The minimum detectable effect (MDE) is $(z_{0.975} + z_{0.80}) \times \text{SE} \approx 2.80 \times \text{SE}$, the smallest absolute effect detectable with 80 percent power in a two-sided test. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. The sample is restricted to 2017-2024, the period over which the three-signal satellite measure is available. Standard errors are clustered at the ballot level.

Table 6. Spillover of cannabis cultivation onto the county data center stock

Grid	Utility		Balancing authority	
Spillover	Pooled	Split	Pooled	Split
	(1)	(2)	(3)	(4)
Variables				
In-state, cultivating	-0.3671*	-0.3584*	-0.1361	-0.1318
	(0.1848)	(0.1796)	(0.0973)	(0.0929)
Pooled spillover	-0.4291**		-0.1999**	
	(0.1610)		(0.0772)	
In-state, non-cultivating		-0.4085**		-0.1724
		(0.1569)		(0.1012)
Other state, same grid		-0.4497**		-0.2009**
		(0.1863)		(0.0766)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	25,308	25,308	134,358	134,358

Notes: The table reports panel difference-in-discontinuities estimates of the spillover of cannabis cultivation onto the county data center stock, at the county-ballot-year level. The outcome is the active data center count in levels. The three county groups are in-state cultivating counties, in-state non-cultivating counties, and counties in other states on the same grid. The pooled spillover combines the in-state non-cultivating and other-state groups. An out-of-state county counts as a grid neighbor only if its balancing authority also operates in the balloting state. In-state counties are unaffected by this rule. The rule is redundant on the balancing-authority footprint, where the condition holds by construction. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table 7. Labor-market wages and employment

	Weekly wage, levels ($/wk)				Employment, inverse hyperbolic sine			
	(1) County total	(2) Data center	(3) Cannabis cultivation	(4) HVAC /elec.	(5) County total	(6) Data center	(7) Cannabis cultivation	(8) HVAC /elec.
Pass × Post	-6.2	-195.6	-70.7	-76.9	0.112	-0.054	0.301	-0.027
	(31.2)	(137.9)	(76.1)	(65.6)	(0.135)	(0.156)	(0.412)	(0.116)
Pre-period mean	745.0	240.7	52.4	683.5	57,136	100.9	19.7	778.3
County × ballot FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,870	16,870	16,870	16,870	16,870	16,870	16,870	16,870
R^2	0.932	0.745	0.483	0.849	0.966	0.902	0.689	0.958
Within R^2	0.006	0.006	0.004	0.005	0.002	0.002	0.003	0.002

Notes: Each column is a separate panel difference-in-discontinuities regression of the named outcome on the Pass × Post interaction at the county-ballot-year level, with running-variable controls and the fixed effects listed in the table. Wages in Columns (1)-(4) are measured in dollars per week. Employment in Columns (5)-(8) is transformed using the inverse hyperbolic sine, asinh(workers), which is defined at zero. Missing QCEW cells are filled with zero before transformation, so a county-year with no firm or a disclosure-suppressed cell contributes $0/week or 0 workers, and the full 16,870-observation panel difference-in-discontinuities sample enters every column. Columns (1) and (5) are the NAICS 10 county total, where wage is the employment-weighted county average and employment is total headcount. Columns (2) and (6) are broad data centers in NAICS 518210 and 541513, Columns (3) and (7) cannabis cultivation in NAICS 111419 and 111998, and Columns (4) and (8) HVAC and electrical contractors in NAICS 2382, the contractors shared by indoor-grow and data center construction. Because missing cells are coded as zero, the estimates for these sparser sectors reflect changes in whether industry activity is reported for a county-year, as well as changes in reported wages or employment. Pre-period means are in raw units, dollars per week for wage and workers for employment. All specifications include county-by-ballot and calendar-year fixed effects, and the estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 26 or fewer clusters, and are reported in parentheses.

Table 8. Data center location-choice model

	(1)	(2)
Legalization shift θ, set from Section 5.3	-0.75	-0.75
	[-1.04, -0.34]	[-1.04, -0.34]
log county population	1.05	0.97
	[0.84, 1.20]	[0.77, 1.11]
log(1 + distance to a line of 345 kV or more, km)		-0.56
		[-0.73, -0.38]
State constants	Yes	Yes
Data center sitings (choosers)	2,121	2,121
Alternatives per chooser	50	50
Log likelihood	-4,561.3	-4,405.8

Notes: The table reports conditional logit estimates of data center location choice. The choosers are the 2,121 facilities whose siting decisions fall in 2015-2024, each choosing from a sampled choice set of fifty counties, the chosen county plus forty-nine non-chosen counties drawn uniformly without replacement. Utility is linear in an indicator for legalization by ballot in the county's state, log county population, and state constants. States with fewer than five sitings share one constant. In Column (2), utility also includes the log of one plus the distance in kilometers from the county's 2010 center of population to the nearest transmission line of 345 kilovolts or more. We set the legalization shift to the log ratio of observed post-ballot active stock in narrowly passing states to its counterfactual stock from Section 5.3. Its 95 percent interval applies the same transformation to the ballot-clustered analytical interval in Appendix Table C14. The intervals for the population elasticity and the distance coefficient use 200 bootstrap samples formed by resampling states. Ninety-five percent intervals are in brackets. Transmission lines are from the September 2024 release of the Homeland Infrastructure Foundation-Level Data. The counterfactuals of Table 9 use Column (2).

Table 9. Data center capacity without legalization by ballot, 2024

	Reallocation	Development
National capacity, 2024	35.9	35.9
National capacity without legalization by ballot	35.9	46.5
Change, percent	0.0	29.6
Change in legalizing states	6.0	10.6
Change in other states	-6.0	0.0

Notes: The table reports 2024 active data center capacity in gigawatts under two counterfactuals that remove the legalization shift θ of Table 8 from the counties of states where a recreational ballot passed and legalization took effect. The reallocation counterfactual assumes that displaced capacity is fully redistributed across counties, so national capacity stays fixed and county shares are renormalized after the shift is removed, as in equation (11). The development counterfactual assumes that displaced capacity is not built elsewhere, so removing the shift raises capacity in legalizing counties by the factor $\exp(-\theta)$ and leaves other counties unchanged, as in equation (12). Changes are capacity without legalization by ballot minus 2024 capacity. Capacity is the model's allocation of the 35.9 gigawatts at facilities active in 2024 across the counties the siting model covers. Siting utility uses the estimates in Column (2) of Table 8. Values are the authors' calculations from the Aterio data center database and EIA Form 861.

Online Appendix

A Cannabis Cultivation Detection

This section documents the satellite-based detection algorithm used to construct the county-year measure of indoor cannabis cultivation activity described in Section 4.3. The pipeline combines three remote-sensing signals into a composite county-year activity index. Each signal is processed independently, validated against ground-truth licensed-facility locations where available, and then merged into an inverse-variance-weighted composite. The first signal is monthly cloud-free composite radiance from the VIIRS Day/Night Band on the Suomi NPP satellite, distributed through the NASA Black Marble (VNP46) product. The Day/Night Band has a native ground sample distance of 750 meters at nadir and an effective dynamic range that resolves anthropogenic light sources two orders of magnitude below the saturation level of legacy DMSP-OLS. The monthly VNP46A3 product grids this radiance at 15 arc-seconds, which we resample to 0.005 degrees. We extract the monthly cloud-free, lunar-corrected, atmosphere-corrected at-sensor radiance from VNP46A3, mask water and persistently bright industrial pixels using a baseline mean computed over the 2012 calendar year, and aggregate to the county-year as the sum of pixel-level radiance anomalies above the baseline. Anomaly pixels with values above a county-specific 95th percentile threshold are flagged as candidate cultivation pixels.

The second signal is surface temperature from Landsat 8 thermal infrared band 10, taken from the USGS Collection 2 Level-2 product at 30-meter resampled resolution. Landsat 8 has a 16-day revisit cycle at the equator and a 10:30 a.m. local overpass time, which captures the heat plume from large-scale lighting and cooling operations during the late-morning warm-up period. We compute a deseasonalized surface-temperature anomaly relative to a 2013-2014 baseline at the pixel level, mask cloud and cloud-shadow pixels using the QA_PIXEL band, and aggregate to the county-year as the sum of pixel-level anomalies above a fixed three-Kelvin threshold. Pixels exceeding the threshold for three or more cloud-free scenes in a year are flagged as candidate cultivation pixels.

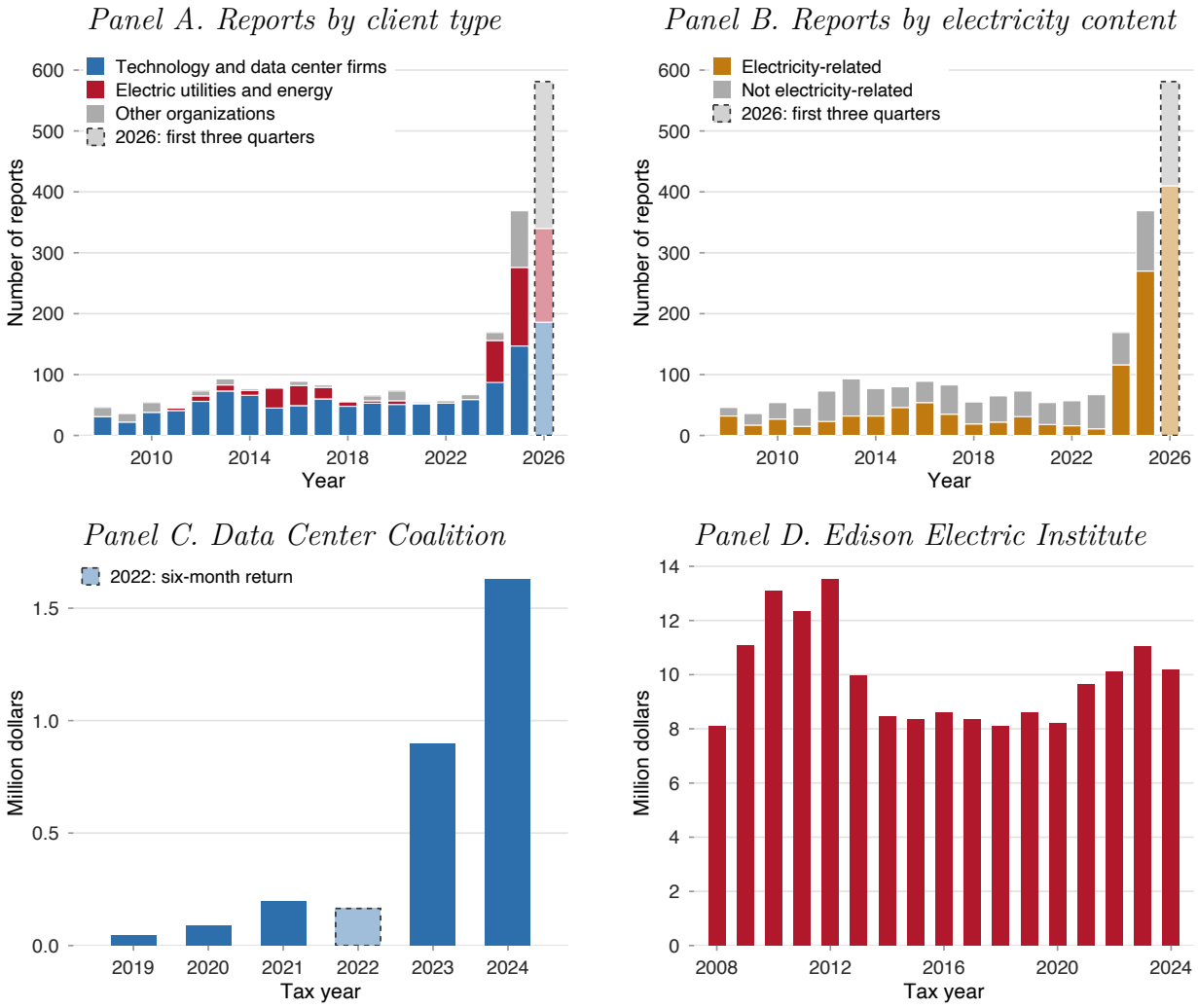
The third signal is optical rooftop detection from Sentinel-2 surface reflectance at 10-meter resolution, distributed through the Copernicus program with a five-day revisit and available since 2017. Greenhouse and warehouse roofs are bright and spectrally flat, so we combine a brightness index across the visible and near-infrared bands with a vegetation index that separates engineered roofing from vegetated and bare ground. Pixels with brightness above a county-specific threshold and a low vegetation index are flagged as possible rooftops of cultivation facilities. Their total area in each county-year provides the third signal. The

three-signal measure begins in 2017, when the Sentinel-2 data become available.

The three candidate signals are merged into a composite county-year activity index using inverse-variance weights estimated on state-licensed cultivation facilities. State seed-to-sale licensing records from Colorado, Oregon, and Washington provide ground-truth cultivation locations through 2024, comprising 9,418 licensed cultivation sites at the time of the most recent extract. We geocode each licensed site to its county and year of license issuance and use this georeferenced sample to evaluate detection at the county-year level. The three-signal composite achieves an area under the receiver operating characteristic curve of 0.87, compared with 0.76 for nightlight alone, 0.78 for thermal anomalies alone, and 0.82 for rooftop classification alone. The analysis uses the county-year measures defined in Section 4.3, namely the area of pixels on which all three signals exceed their detection thresholds, the composite activity index, an indicator for any such pixel in the county-year, and an indicator for two-signal detected area above the sample mean. For sample splits that require pre-2017 observations, Section 4.3 uses a two-signal measure that omits rooftops and extends back to 2000. In the earlier years, it combines DMSP-OLS version 4 stable lights at 30 arc-seconds with MODIS monthly land surface temperature from MOD11C3 at 0.05 degrees. In the later years, it combines the NOAA Earth Observation Group's monthly VIIRS composites at 15 arc-seconds with Landsat 8 and 9 thermal band 10 at 30 meters. A pixel is classified as lit when its DMSP digital number exceeds 10 or its VIIRS radiance exceeds 2 nanowatts per square centimeter per steradian, sensor-specific thresholds for commercial lighting. A pixel is classified as hot when its surface temperature exceeds the local mean by 3 degrees Celsius, the same threshold for both thermal sensors. Panel B of Figure 1 maps the geographic distribution of the high-confidence measure.

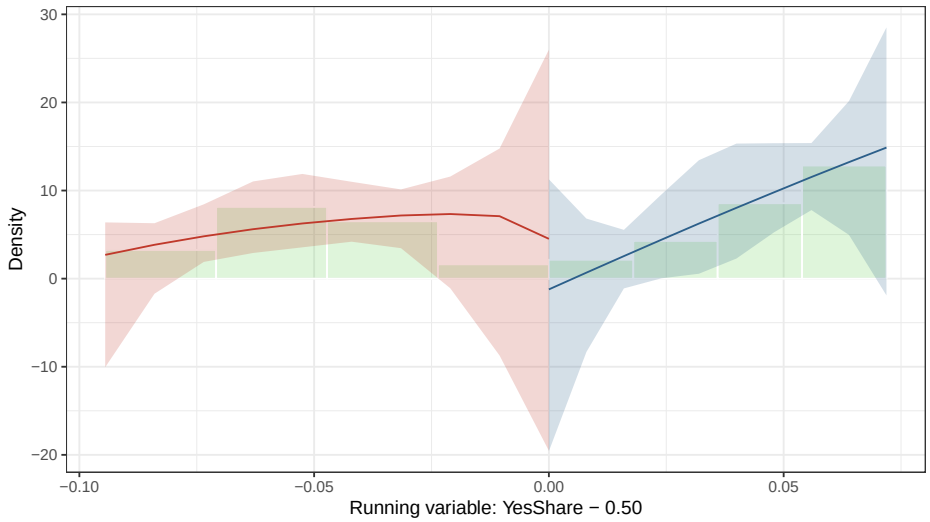
B Additional Figures

Figure B1. Data center lobbying reports and trade group expenditures



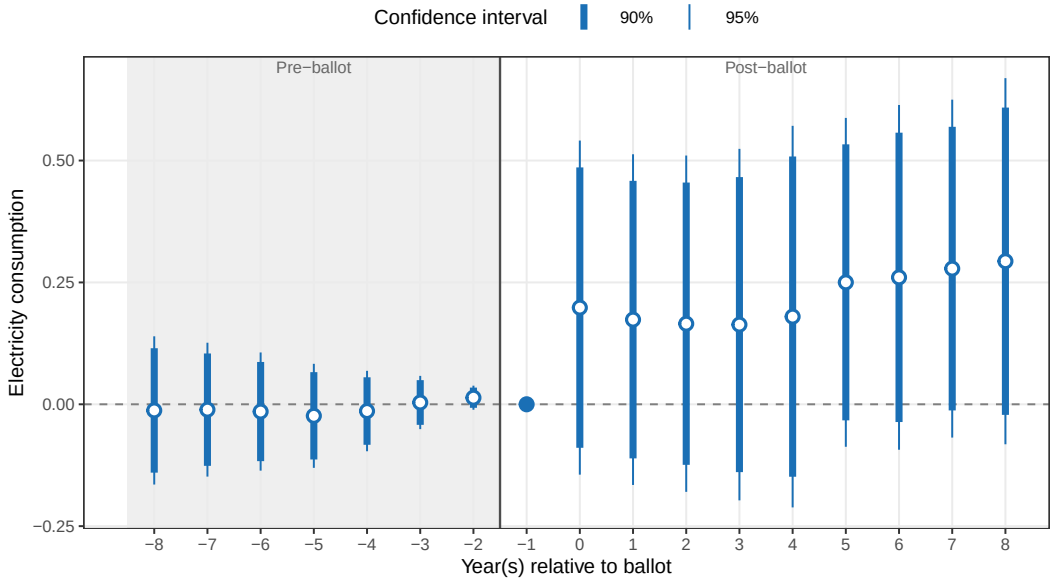
Notes: Panels A and B count quarterly LD-2 reports under the Lobbying Disclosure Act whose specific-issue text names data centers, server farms, hyperscale or colocation facilities, or cloud infrastructure. A report covers one registrant, one client, and one quarter. Panel A groups reports by client type using keyword rules on client names. Panel B separates reports whose data center issue line also names electricity, energy, power, utilities, the grid, units of power, or the Federal Energy Regulatory Commission, or carries the energy or utilities issue code, from all other reports. The lighter 2026 bars with a dashed outline cover the first three quarters. Reports through 2019 are from the Senate Office of Public Records bulk files, and later reports are from the Lobbying Disclosure Act database as of October 22, 2026. Panels C and D plot the amount on line 2a of Schedule C, Part III-B, of the organization's Form 990 return for each tax year, in millions of dollars. Line 2a reports the section 162(e) lobbying and political expenditures of a tax-exempt trade association. Panel C is the Data Center Coalition and Panel D is the Edison Electric Institute, and the two panels use different vertical scales. The lighter 2022 bar of the Data Center Coalition with a dashed outline covers six months, July-December, after the organization moved to a calendar tax year. Its 2019 amount is from an amended return. The amounts in Panels C and D are from IRS e-file returns, supplemented by images of filed returns for three Edison Electric Institute tax years.

Figure B2. Density of the running variable at the cutoff



Notes: The figure plots the local-polynomial estimate of the density of the running variable R_{sa} on each side of the 50 percent cutoff, with 95 percent confidence bands, following Cattaneo et al. (2020a). The estimation bandwidth is $|R_{sa}| \leq 0.10$.

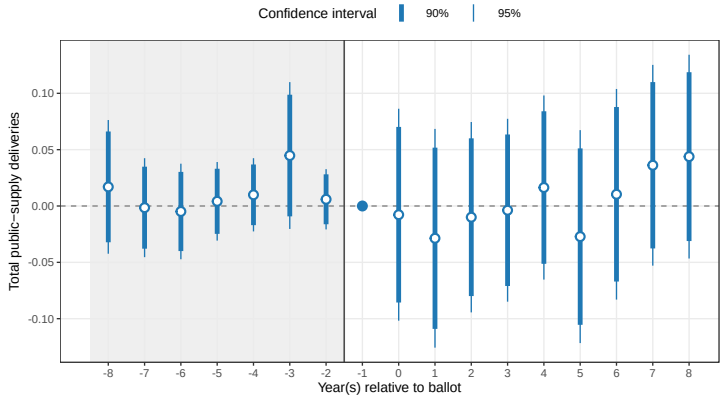
Figure B3. Total electricity consumption around legalization at the eight-year window



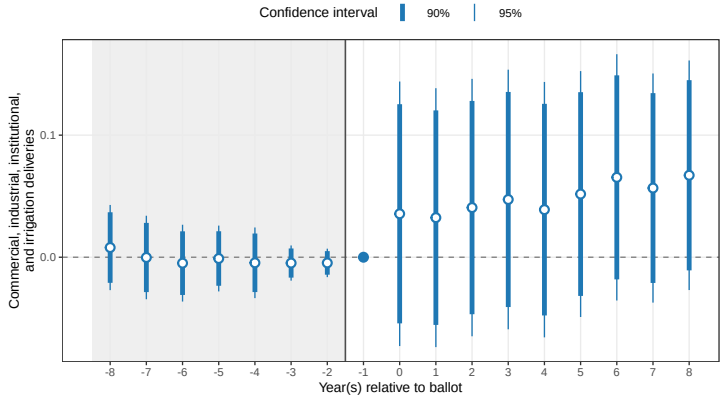
Notes: The figure repeats Figure 9 at a plus or minus eight-year window. The year before the ballot is the omitted reference period, bars show 90 and 95 percent confidence intervals based on standard errors clustered at the ballot level, and the estimation sample is restricted to $|R_{sa}| \leq 0.10$ and contains 16,633 county-ballot-year observations.

Figure B4. Water deliveries and water-utility revenue around legalization

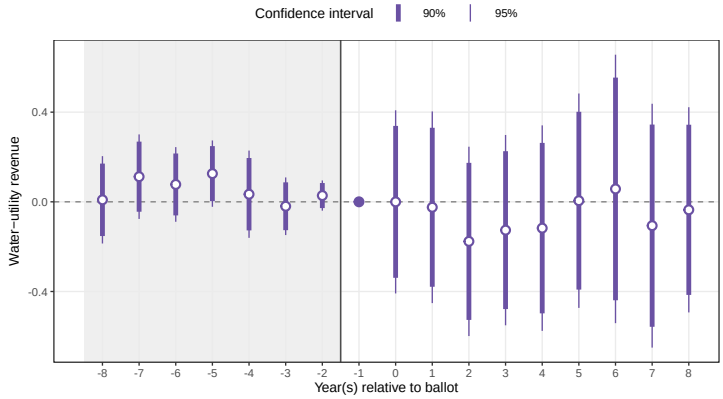
Panel A. Total public-supply deliveries



Panel B. Commercial, industrial, institutional, and irrigation deliveries

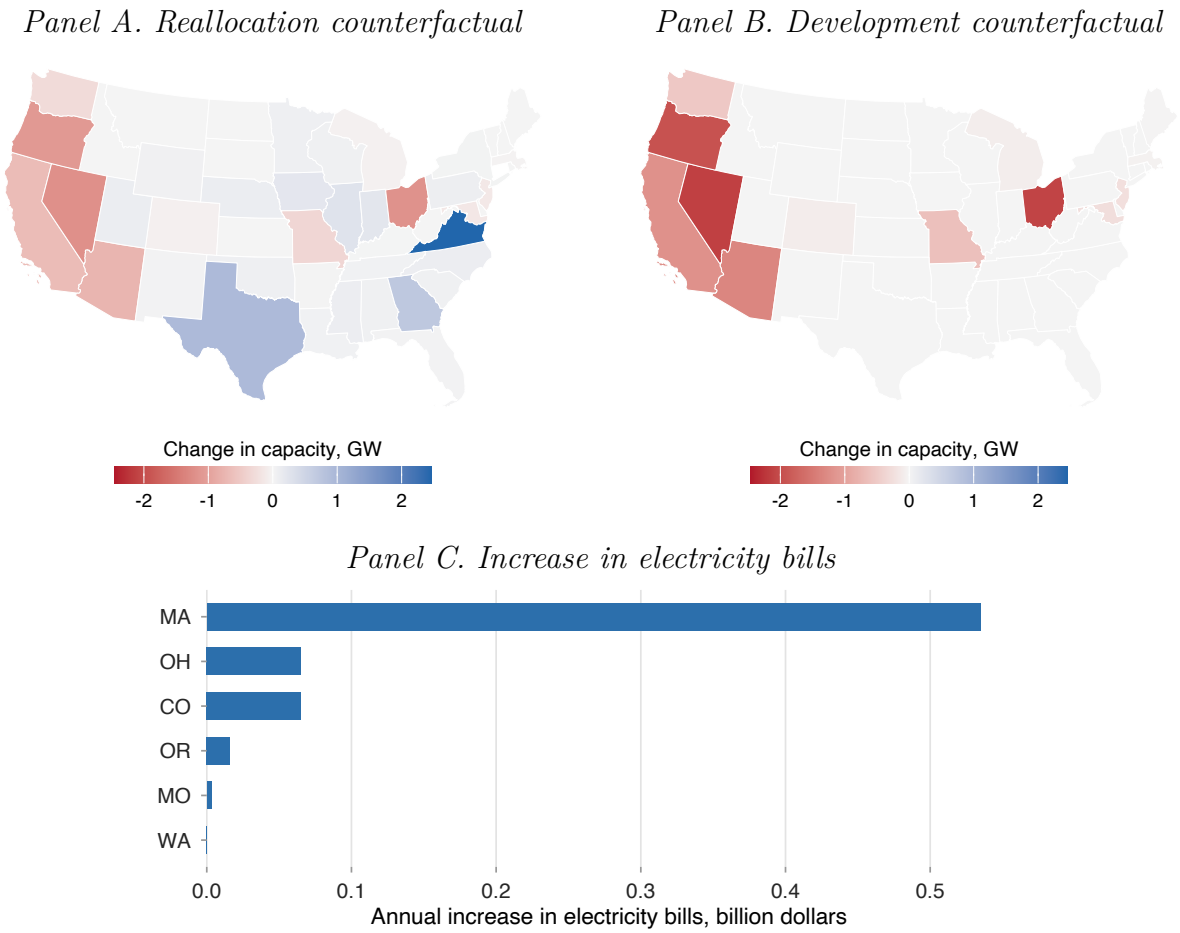


Panel C. Water-utility revenue



Notes: Each panel plots the coefficients on *Pass* interacted with indicators for each year relative to the ballot, from eight years before to eight years after the ballot, for the outcomes of Appendix Table C30. Panels A and B use the inverse hyperbolic sine of deliveries in gallons per day from USGS estimates for 2000-2020. Panel C uses the inverse hyperbolic sine of the water-utility revenue of local governments in thousands of nominal dollars from the Census of Governments and the Annual Survey of State and Local Government Finances. The year before the ballot is the omitted reference period. Bars show 90 and 95 percent confidence intervals based on standard errors clustered at the ballot level, and the estimation sample is restricted to $|R_{sa}| \leq 0.10$.

Figure B5. Changes in data center capacity and electricity bills, by state



Notes: In Panels A and B, states are shaded by the change in 2024 active data center capacity that the model attributes to legalization by ballot, measured as modeled capacity with the legalization shift minus modeled capacity without it, so the changes have the opposite sign of those in Table 9. Panel A shows the reallocation counterfactual, which holds the 35.9 gigawatt total fixed. Panel B shows the development counterfactual, in which the capacity that legalizing states lose is not built elsewhere, leaving capacity in other states unchanged. Both panels use the estimates in Column (2) of Table 8 and share one color scale. Blue shades mark states whose capacity is higher because of legalization by ballot elsewhere, the destinations of the reallocation, and red shades mark states that legalized by ballot, whose capacity the legalization shift lowers. Panel C shows the annual increase in captive customers' electricity bills by state. For each county, we apply the retail-price increase implied by the cultivation-intensity coefficient in Column (2) of Table 2 to 2024 electricity consumption net of data center load, and state totals include only counties whose ballots fall within the estimation bandwidth $|R_{sa}| \leq 0.10$. The maps show the continental U.S. and omit Alaska and Hawaii. Values are the authors' calculations from the Aterio data center database, EIA Form 861, and detected cultivation.

C Additional Tables

Table C1. Clients lobbying on the Ratepayer Protection Act, 2026

Client type	Clients	Examples
Technology and data center firms	17	Data Center Coalition, Microsoft, Google, Meta, Amazon, Oracle, Equinix, CoreWeave, Anthropic
Electric utilities and energy	28	Edison Electric Institute, National Rural Electric Cooperative Association, American Electric Power, Duke Energy, Southern Company, Entergy, Xcel Energy, Constellation
Other organizations	46	Cleveland-Cliffs, Nucor, American Iron and Steel Institute, Dow, American Chemistry Council, International Paper, TSMC Arizona, Sierra Club
All	91	

Notes: The table counts the clients whose LD-2 reports under the Lobbying Disclosure Act name the Ratepayer Protection Act, H.R. 9340 in the House or S. 5028 in the Senate, in the first three quarters of 2026. Each client is counted once across its 2026 reports. Clients are grouped by keyword rules on client names, as in Appendix Figure B1. Two clients whose reports cite unrelated bills with similar numbers are excluded. The bill would require state utility regulators to consider making large new loads pay the full cost of the generation and transmission upgrades built to serve them. Its first version covered every project with a peak demand of 100 megawatts or more, and committee leaders narrowed it to data centers before the House Energy and Commerce Committee approved it in July 2026. Data come from the Lobbying Disclosure Act database as of October 22, 2026.

Table C2. The data center stock effect across the three designs

Outcome:	Stock		
Single-ballot cross-section: Pass	-1.355		
	(0.8572)		
Pooled cross-section: Pass		-1.464**	
		(0.7014)	
Panel difference-in-discontinuities: Pass×Post			-0.7019***
			(0.2194)
Fixed effects			
County × ballot			Yes
Calendar year			Yes
Observations	690	1,098	16,870
R^2	0.01339	0.01256	0.92833
Mean of dep. var.	0.462	0.436	0.446

Notes: The table reports the effect of recreational cannabis legalization on the active data center stock under three designs, using the 2000-2024 county-year data center panel. The single-ballot cross-section keeps each state's most recent ballot and has one observation per county. The pooled cross-section stacks every ballot event and has one observation per county and ballot. Both cross-sections measure the outcome one year after the ballot and include no fixed effects. The panel difference-in-discontinuities is estimated at the county-ballot-year level over a window of plus or minus eight years around the ballot and includes county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C3. Single-ballot cross-sectional estimates by data center outcome

Outcome:	Stock	Entry	Stock (MW)	Entry (MW)	Without on-site generation	With on-site generation	Pipeline
$\hat{\beta}$: Pass	-1.355	-0.1084*	-19.97	-4.532	-0.7566	-0.5985*	-0.6321
	(0.8572)	(0.0556)	(12.38)	(3.642)	(0.6516)	(0.3193)	(0.8170)
Observations	690	690	690	690	690	690	690
R^2	0.01339	0.00695	0.00988	0.00586	0.01253	0.00879	0.00303
Mean of dep. var.	0.462	0.029	6.876	1.239	0.309	0.154	0.213

Notes: The table reports single-ballot cross-sectional regression discontinuity estimates of the effect of recreational cannabis legalization on data center outcomes, with one observation per county. Each state enters with its most recent ballot, and outcomes are measured one year after the ballot, using the 2000-2024 county-year data center panel. The columns report the active stock, new entries, operating and entrant capacity in megawatts, the active stock without and with on-site generation, and the pipeline stock. The specification includes no fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 13 clusters, and are reported in parentheses.

Table C4. Pooled cross-sectional estimates by data center outcome

Outcome:	Stock	Entry	Stock (MW)	Entry (MW)	Without on-site generation	With on-site generation	Pipeline
$\hat{\beta}$: Pass	-1.464**	-0.1109**	-18.97**	-2.545	-0.9570*	-0.5071**	-0.2812
	(0.7014)	(0.0529)	(8.854)	(1.561)	(0.5140)	(0.2123)	(0.3343)
Observations	1,098	1,098	1,098	1,098	1,098	1,098	1,098
R^2	0.01256	0.00555	0.01033	0.00723	0.01159	0.00946	0.00465
Mean of dep. var.	0.436	0.027	6.033	0.903	0.301	0.135	0.144

Notes: The table reports pooled cross-sectional regression discontinuity estimates of the effect of recreational cannabis legalization on data center outcomes. The pooled cross-section stacks every ballot event and has one observation per county and ballot. Outcomes are measured one year after the ballot, using the 2000-2024 county-year data center panel. The columns report the active stock, new entries, operating and entrant capacity in megawatts, the active stock without and with on-site generation, and the pipeline stock. The specification includes no fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 23 clusters, and are reported in parentheses.

Table C5. Panel difference-in-discontinuities estimates by data center outcome, levels

Outcome:	Stock	Entry	Stock (MW)	Entry (MW)	Without on-site generation	With on-site generation	Pipeline
$\hat{\beta}$: Pass×Post	-0.7019***	-0.0886*	-16.19**	-4.394**	-0.2567***	-0.4452**	-0.3656
	(0.2194)	(0.0447)	(7.620)	(2.063)	(0.0847)	(0.1626)	(0.2935)
Fixed effects							
County × ballot	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,870	16,870	16,870	16,870	16,870	16,870	16,870
R^2	0.92833	0.35427	0.70371	0.27831	0.96300	0.80055	0.29634
Mean of dep. var.	0.446	0.023	5.714	0.634	0.320	0.127	0.058

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on data center outcomes, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. The columns report the active stock, new entries, operating and entrant capacity in megawatts, the active stock without and with on-site generation, and the pipeline stock. All specifications include county-by-ballot and calendar-year fixed effects and control for $R_{sa} \times \text{Post}$ and $R_{sa} \times \text{Pass} \times \text{Post}$. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 26 clusters, and are reported in parentheses.

Table C6. Other state policies around cannabis ballots

Outcome:	RPS (binding)	RGGI member	Right-to-work	Retail choice	Corporate tax
Pass×Post	-0.0160	0.0618*	-0.0157	0.0000	-0.8835
	(0.0910)	(0.0351)	(0.0576)	(0.0000)	(0.8341)
Fixed effects					
County × ballot	Yes	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	16,870	16,870	16,870	16,870	16,870

Notes: Each column reports the Pass×Post coefficient from the baseline panel difference-in-discontinuities specification applied to the state-year regulatory covariate listed in the header, merged by ballot state onto the stacked county-ballot-year panel over 2000-2024. A cannabis ballot may signal voter willingness to adopt other regulation that also affects data center siting. If so, these covariates would jump at the cutoff after passage. RPS indicates that the state has a binding mandatory Renewable Portfolio Standard in that year, from DSIRE and NCSL. RGGI indicates that the state is a current participating member of the Regional Greenhouse Gas Initiative. Right-to-work indicates that the state has a right-to-work statute, from NCSL. Retail choice indicates that the state has full retail electricity choice, from EIA. Corporate tax is the top statutory state corporate income tax rate in percent, from the ITEP Historical State Tax Rate Data compiled from FTA. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C7. Not-yet-treated control exclusion

Outcome	Active stock		Active entry	
Specification	Baseline	Excluding legalized	Baseline	Excluding legalized
$\hat{\beta}$: Pass×Post	-0.7019***	-0.4582**	-0.0886*	-0.0357
	(0.2194)	(0.1654)	(0.0447)	(0.0242)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Observations	16,870	16,369	16,870	16,369
R^2	0.92833	0.94347	0.35427	0.38702
Mean of dep. var.	0.446	0.413	0.023	0.020

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on the active data center stock and on new entries, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. The excluding-legalized sample drops each control state's observations from the year the state itself legalizes onward, which removes California from 2016 on, Oregon from 2014 on, and Arizona from 2020 on. South Dakota's overturned 2020 initiative and Florida's 2024 measure, which fell short of the required 60 percent supermajority, remain passing ballots and are not treated as legalization. All specifications include county-by-ballot and calendar-year fixed effects and control for $R_{sa} \times \text{Post}$ and $R_{sa} \times \text{Pass} \times \text{Post}$. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 26 clusters, and are reported in parentheses.

Table C8. The data center stock effect across event windows

	±3 yr	±6 yr	±8 yr
$\hat{\beta}$: Pass×Post	-0.1961**	-0.4963***	-0.7019***
	(0.07175)	(0.1585)	(0.2194)
Fixed effects			
County × ballot	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes
Observations	7,930	13,538	16,870
R^2	0.98131	0.95786	0.92833
Mean of dep. var.	0.402	0.427	0.446

Notes: The table reports panel difference-in-discontinuities estimates of the Pass×Post coefficient on the active data center stock at three event windows, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot and restricted to the event window around each ballot. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C9. The data center stock and entry effects across bandwidths

Bandwidth $\lvert R_{sa}\rvert$	G_-	G_+	$\hat{\beta}$ Stock	$\hat{\beta}$ Entry
0.04	5	4	-1.193***	-0.244***
			(0.344)	(0.068)
0.05	6	6	-1.035**	-0.177*
			(0.393)	(0.085)
0.06	9	10	-0.787**	-0.117
			(0.312)	(0.068)
0.07	10	12	-0.741**	-0.113*
			(0.293)	(0.061)
0.08	11	14	-0.839***	-0.142**
			(0.285)	(0.061)
0.10	12	14	-0.702***	-0.089*
			(0.219)	(0.045)
0.12	13	15	-0.722***	-0.114**
			(0.232)	(0.052)
0.15	14	15	-0.642***	-0.088*
			(0.206)	(0.049)

Notes: The table reports panel difference-in-discontinuities estimates of the Pass×Post coefficient on the active data center stock and on new entries across estimation bandwidths, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot. G_- and G_+ are the ballot events below and above the cutoff within each bandwidth. The event window is plus or minus eight years around the ballot, and all specifications include county-by-ballot and calendar-year fixed effects. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C10. The data center stock effect across fixed-effect structures

	County	County×ballot	County×ballot + year
$\hat{\beta}$: Pass×Post	-0.6839**	-0.7188***	-0.7019***
	(0.2667)	(0.2485)	(0.2194)
Fixed effects			
County	Yes		
County × ballot		Yes	Yes
Calendar year			Yes
Observations	16,870	16,870	16,870
R^2	0.90647	0.92740	0.92833
Mean of dep. var.	0.446	0.446	0.446

Notes: The table reports the panel difference-in-discontinuities Pass×Post coefficient on the active data center stock under three fixed-effect structures, each including the ballot-specific post indicator, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot. The event window is plus or minus eight years around the ballot, and the estimation bandwidth is $\lvert R_{sa}\rvert \leq 0.10$. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C11. The data center effect by ballot-year cohort

Ballot calendar year	Passing ballots	Treated counties	Data center stock
2012	2	103	-0.9174***
			(0.2159)
2014	2	65	-0.6592**
			(0.2750)
2016	4	105	-0.6157**
			(0.2388)
2018	1	83	-0.9744***
			(0.2503)
2020	2	122	-0.9325***
			(0.2543)
2022	1	115	-0.6804***
			(0.2162)
2023	1	88	-0.6877**
			(0.2558)
2024	1	67	-0.8867***
			(0.2395)

Notes: Each row reports the effect of passage for the ballots held in that year, estimated jointly in one regression in which Pass×Post is interacted with cohort indicators, so every cohort is measured against the same pooled control group of failing-ballot counties. A separate regression per cohort is not possible, because a within-year comparison needs both a passing and a failing ballot in that year, and 2014, 2020, and 2023 have no failing ballot while the 2002-2010 ballots include no passing one. The sample is the 2000-2024 county-year data center panel, stacked by ballot at the county-ballot-year level and restricted to a window of plus or minus eight years around the ballot. The regression includes county-by-ballot and calendar-year fixed effects, and the estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C12. Running-variable polynomial order

Outcome	Active stock		Active entry	
Specification	Baseline	Quadratic R_{sa}	Baseline	Quadratic R_{sa}
$\hat{\beta}$: Pass×Post	-0.7019***	-0.9930**	-0.0886*	-0.2000**
	(0.2194)	(0.3962)	(0.0447)	(0.0862)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Observations	16,870	16,870	16,870	16,870
R^2	0.92833	0.92846	0.35427	0.35478
Mean of dep. var.	0.446	0.446	0.023	0.023

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on the active data center stock and on new entries, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. All specifications include county-by-ballot and calendar-year fixed effects and control for $R_{sa} \times \text{Post}$ and $R_{sa} \times \text{Pass} \times \text{Post}$. The quadratic specification adds $R_{sa}^2 \times \text{Post}$ and $R_{sa}^2 \times \text{Pass} \times \text{Post}$ to the running-variable polynomial. The levels of R_{sa} and R_{sa}^2 are constant within a ballot and are absorbed by the county-by-ballot fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 26 clusters, and are reported in parentheses.

Table C13. The data center effect conditioning on pre-ballot generation capacity

	Estimate	95% CI	N
No supply control (baseline)	-0.7019***	[-1.154, -0.250]	16,870
	(0.2194)		
+ existing and renewable capacity × Post	-0.6158***	[-0.994, -0.237]	16,870
	(0.1838)		
+ existing and renewable capacity × year	-0.5922***	[-0.975, -0.209]	16,870
	(0.1860)		
+ queued capacity as well, × Post	-0.5681***	[-0.927, -0.209]	16,870
	(0.1742)		
+ queued capacity as well, × year	-0.5365***	[-0.888, -0.185]	16,870
	(0.1708)		

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on the active data center stock, at the county-ballot-year level, without and with controls for pre-ballot generation capacity. The sample is the 2000-2024 county-year data center panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. Existing and renewable generation capacity are the utility-scale nameplate megawatts sited in the county in the year before its own ballot, from EIA Form 860M. Queued capacity is the megawatts of interconnection requests filed in the county before the ballot, from the Lawrence Berkeley National Laboratory queue data. All three are predetermined and enter as the inverse hyperbolic sine of megawatts, since 209 of the 891 bandwidth counties host no utility-scale generator at all. Conditioning on capacity as measured *at* the outcome year would be conditioning on a response to the same shock, since interconnection requests are themselves an outcome in Section 6.3. Interacting pre-ballot capacity with Post allows the post-ballot change to vary with initial capacity. Interacting it with calendar-year indicators allows this relationship to vary by year. Form 860 covers generators of one megawatt and above, so this is utility-scale capacity. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C14. Inference for the headline data center stock effect

Inference procedure	G	SE	p	95% CI
Analytical, cluster by ballot	26	0.219	0.004	[−1.154, −0.250]
Analytical, cluster by state	17	0.236	0.009	[−1.201, −0.202]
Wild bootstrap (unrestricted), ballot	26	0.219	0.028	[−1.317, −0.086]
Wild bootstrap (null-imposed), ballot	26	0.219	0.010	[−1.226, −0.177]
Wild bootstrap (null-imposed), state	17	0.236	0.023	[−1.307, −0.096]

Notes: The table reports inference for the panel difference-in-discontinuities effect of recreational cannabis legalization on the active data center stock, at the county-ballot-year level. The sample is the 2000-2024 county-year data center panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. The specification includes county-by-ballot and calendar-year fixed effects, and the estimation bandwidth is $|R_{sa}| \leq 0.10$. Every row uses the same point estimate, and only the inference procedure varies. Ballot clustering, with 26 clusters, is the primary specification. Clustering by state groups repeated ballots within the same state, yielding 17 clusters. Analytical intervals are cluster-robust t intervals with $G-1$ degrees of freedom, where G is the number of clusters. Bootstrap intervals are symmetric studentized percentile-t intervals from $B = 4{,}999$ Rademacher draws (Cameron et al., 2008). "Unrestricted" draws residuals from the fitted model. "Null-imposed" draws them from the model with the treatment effect set to zero, the variant recommended by MacKinnon and Webb (2018).

Table C15. Electricity price effect: full sample, detection-window sample, and cultivation-intensity interaction

	Average			Commercial			Industrial		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\hat{\beta}_1$: Pass×Post	0.1243	0.1199	0.1197	0.1178	0.1823	0.1823	0.1035	0.1449	0.1444
	(0.1414)	(0.1064)	(0.1051)	(0.1282)	(0.1341)	(0.1331)	(0.1337)	(0.1195)	(0.1179)
$\hat{\beta}_2$: km^2×Pass×Post			0.0219***			0.0231***			0.0264*
			(0.0072)			(0.0066)			(0.0136)
Observations	11,563	6,868	6,868	11,563	6,868	6,868	11,355	6,841	6,841
Sample	Full	Restricted	Restricted	Full	Restricted	Restricted	Full	Restricted	Restricted
km^2 interaction	No	No	Yes	No	No	Yes	No	No	Yes

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on log electricity prices, at the county-ballot-year level. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus five years around the ballot. The full sample uses all ballot-state counties in the bandwidth. The restricted sample covers 2017-2024, the period of satellite detection coverage, and retains 19 of the 26 ballots. Four ballots have no post-2017 window, and three retain a single calendar year and are absorbed by the county-by-ballot fixed effects. Columns (3), (6), and (9) include the interaction between detected cultivation area in km^2 and Pass×Post and report its coefficient, $\hat{\beta}_2$. The interaction is defined only on the restricted sample. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C16. Electricity prices and detected cultivation in levels

	Average (1)	Commercial (2)	Industrial (3)
$\hat{\gamma}$: Detected cultivation area (km^2)	0.0035*	0.0019	0.0064***
	(0.0017)	(0.0021)	(0.0018)
Fixed effects			
County × ballot	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes
Fit statistics			
Observations	6,868	6,868	6,841
R^2	0.84657	0.85500	0.89337
Within R^2	0.00072	0.00021	0.00112

Notes: The table reports within-county estimates of the relationship between detected cultivation and electricity prices, at the county-ballot-year level. The outcome in every column is the natural log of the population-weighted retail electricity price in dollars per kWh for a customer class, namely the all-customer average in Column (1), commercial in Column (2), and industrial in Column (3). The regressor is the level in km^2 of detected cultivation area from the three-signal satellite measure. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the sample is restricted to 2017-2024. Standard errors are clustered at the ballot level, with 26 or fewer clusters, and are reported in parentheses.

Table C17. Electricity prices and detected cultivation within county, eight-year window

	Average (1)	Commercial (2)	Industrial (3)
$\hat{\gamma}$: Inverse hyperbolic sine of detected cultivation area (km^2)	0.0510*** (0.0092)	0.0477*** (0.0132)	0.0670*** (0.0077)
Fixed effects			
County × ballot	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes
Fit statistics			
Observations	8,649	8,649	8,591
R^2	0.86565	0.86644	0.89844
Within R^2	0.00606	0.00503	0.00526

Notes: The table reports within-county estimates of the relationship between detected cultivation and electricity prices, at the county-ballot-year level. The outcome in every column is the natural log of the population-weighted retail electricity price in dollars per kWh for a customer class, namely the all-customer average in Column (1), commercial in Column (2), and industrial in Column (3). The regressor is the inverse hyperbolic sine of detected cultivation area in km^2 from the three-signal satellite measure. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, the event window is plus or minus eight years around the ballot, and the sample is restricted to 2017-2024. Standard errors are clustered at the ballot level, with 26 or fewer clusters, and are reported in parentheses.

Table C18. Cannabis cultivation, electricity prices, and consumption, eight-year window

	Total		Commercial		Industrial	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Electricity price						
$\hat{\beta}_1$: Pass×Post	0.1213	0.1209	0.1817	0.1816	0.1435	0.1427
	(0.1027)	(0.1020)	(0.1427)	(0.1423)	(0.1120)	(0.1112)
$\hat{\beta}_2$: km^2×Pass×Post		0.0110*		0.0146*		0.0116
		(0.0064)		(0.0073)		(0.0082)
Observations	8,649	8,649	8,649	8,649	8,591	8,591
Panel B. Electricity consumption						
$\hat{\beta}_1$: Pass×Post	0.0321	0.0325	-0.0647	-0.0645	-0.0046	-0.0042
	(0.3657)	(0.3655)	(0.3963)	(0.3962)	(0.2075)	(0.2074)
$\hat{\beta}_2$: km^2×Pass×Post		0.0067		0.0045		0.0015
		(0.0105)		(0.0138)		(0.0124)
Observations	8,649	8,649	8,649	8,649	8,649	8,649
km^2 interaction	No	Yes	No	Yes	No	Yes

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on electricity prices and consumption, at the county-ballot-year level. The specification follows Table 2 except for the event window. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Both panels use the sample restricted to 2017-2024, the period of satellite detection coverage, so the interaction columns are defined throughout. Standard errors are clustered at the ballot level and are reported in parentheses. The "Total" columns report the average retail price in Panel A and total consumption in Panel B. Price outcomes are in logs and consumption outcomes are in the inverse hyperbolic sine of MWh. Odd-numbered columns omit the interaction between detected cultivation area in km^2 and Pass×Post. Even-numbered columns include this interaction and report its coefficient, $\hat{\beta}_2$.

Table C19. Interconnection-queue response to cannabis cultivation: solar and gas generators

Outcome	Queue applications		Withdrawals	
	(1)	(2)	(3)	(4)
Variables				
$\hat{\beta}$: Pass×Post	0.2207	0.2621	-0.1835	-0.1736
	(0.2563)	(0.2584)	(0.1296)	(0.1221)
Total-grid cultivation×Pass×Post	-0.0415***		0.0116	
	(0.0123)		(0.0145)	
Own-county cultivation×Pass×Post		-0.1985**		-0.0190
		(0.0753)		(0.0440)
Rest-of-grid cultivation×Pass×Post		-0.0285**		0.0140
		(0.0128)		(0.0123)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	8,576	8,576	8,576	8,576

Notes: The table reports panel difference-in-discontinuities estimates on the stacked ballots, at the county-ballot-year level. The outcomes are county-year counts of solar and gas generator interconnection-queue applications in Columns (1)-(2) and dated withdrawals in Columns (3)-(4), from the LBNL Queued Up records. Columns (1) and (3) interact Pass×Post with total cannabis cultivation on the county's utility grid, the sum of own-county and rest-of-grid cultivation in the same state. Columns (2) and (4) decompose that grid dose into own-county and leave-one-out rest-of-grid cultivation, both interacted with Pass×Post in one regression, so the own-county coefficient is net of the rest-of-grid effect. Cultivation is satellite-detected area in km^2. The sample is the cultivation-dose subsample of 8,576 county-ballot-years in 16 states from 2017 onward. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C20. Electricity-price spillover under the two alternative grid-footprint rules

Footprint rule	Any shared utility		Bordering state	
Spillover	Pooled	Split	Pooled	Split
	(1)	(2)	(3)	(4)
Variables				
Pass×Post	0.0485	0.0678	0.0527	0.0653
	(0.0912)	(0.0853)	(0.0916)	(0.0880)
Own-county km^2×Pass×Post	0.0072*	0.0037	0.0102**	0.0082*
	(0.0039)	(0.0040)	(0.0036)	(0.0041)
Pooled spillover km^2×Pass×Post	0.0046*		0.0065**	
	(0.0024)		(0.0023)	
Within-state km^2×Pass×Post		0.0090***		0.0089**
		(0.0031)		(0.0032)
Cross-state km^2×Pass×Post		0.0014		0.0030
		(0.0023)		(0.0024)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	15,926	15,926	14,828	14,828

Notes: The table reports panel difference-in-discontinuities estimates of the spillover of cannabis cultivation onto the county electricity price under two alternative grid-footprint rules, at the county-ballot-year level. Columns (1)-(2) admit any out-of-state county sharing a utility with the balloting state, an alternative shared-utility rule that counts PacifiCorp's Utah counties as grid neighbors of California. Columns (3)-(4) instead require the county's state to border the balloting state, with borders derived from Census TIGER county boundaries. Table 4 reports the balancing-authority rule used in the text, and the outcome and dose terms are defined as there. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C21. Data center spillover under the two alternative grid-footprint rules

Footprint rule	Any shared utility		Bordering state	
Spillover	Pooled	Split	Pooled	Split
	(1)	(2)	(3)	(4)
Variables				
In-state, cultivating	-0.2934*	-0.2988*	-0.3377*	-0.3278*
	(0.1584)	(0.1661)	(0.1833)	(0.1796)
Pooled spillover	-0.3714**		-0.4093**	
	(0.1403)		(0.1625)	
In-state, non-cultivating		-0.3468**		-0.3760**
		(0.1387)		(0.1576)
Other state, same grid		-0.4194**		-0.4496**
		(0.1741)		(0.1953)
Fixed effects				
County × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	29,518	29,518	25,898	25,898

Notes: The table reports panel difference-in-discontinuities estimates of the spillover of cannabis cultivation onto the county data center stock under two alternative grid-footprint rules, at the county-ballot-year level. The footprint rules are as in Table C20. Table 6 reports the balancing-authority rule used in the text, and the outcome and county groups are defined as there. The pooled spillover combines the in-state non-cultivating and other-state groups. All specifications include county-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C22. Allocation of utility prices to counties

Counties in the electricity panel	3,135
served by exactly one utility	256 (8.2%)
of those, share matching the reported utility price to within 0.1%	95.6%
Multi-county utility-state-years	7,117
share in which every county carries an identical price	5.8%
median within-cell coefficient of variation of price	0.039
R^2 of county log price on utility-state × year	0.935
R^2 of county log price on calendar year alone	0.342
In the spillover estimation sample	
county-ballot-year observations	14,630
distinct counties	1,125
distinct utility-states behind them	160
distinct ballots (clusters)	23

Notes: County electricity price is built by allocating each utility-state's revenue and its sales to counties with the same population weight and then dividing one by the other, so the weight cancels and a county served by a single utility inherits that utility's reported price exactly. The upper rows cover the counties of the electricity panel, and the lower rows cover the spillover estimation sample. Table C23 re-estimates the spillover at the utility-state level, with the allocation removed entirely.

Table C23. Cultivation spillover onto the utility-state electricity price

Footprint	Balancing authority			Any shared utility
Price	Average	Residential	Commercial	Average
	(1)	(2)	(3)	(4)
Variables				
Pass×Post	0.1067	0.1329	0.0629	0.0962
	(0.1810)	(0.1946)	(0.1712)	(0.1705)
Own utility-state km^2	0.0066***	0.0062***	0.0035**	0.0049*
	(0.0018)	(0.0018)	(0.0013)	(0.0024)
Same state, other utility km^2	-0.0003	-0.0004	-0.0004	0.0002
	(0.0017)	(0.0020)	(0.0018)	(0.0017)
Same utility, other state km^2	0.0022	0.0003	0.0020	0.0018
	(0.0030)	(0.0036)	(0.0034)	(0.0022)
Fixed effects				
Utility-state × ballot	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	3,421	3,421	3,421	3,724

Notes: The table reports panel difference-in-discontinuities estimates of the cultivation spillover onto the electricity price, with the utility-state as the unit of analysis. The outcome is the log price reported on EIA Form 861 for the utility-state-year. The average price is total revenue divided by total sales, and the residential and commercial prices are each class's own revenue divided by its own sales. The specification uses utility-state revenue and sales directly, without county-level allocation. Doses are cultivation in square kilometers, and each enters interacted with Pass×Post. The own dose covers the counties this utility-state serves. The same-state, other-utility dose covers a different utility in the same state, which shares neither revenue nor sales with this one. The same-utility, other-state dose covers the same utility across the state line. Counties are assigned to one utility-state by modal utility, so the three doses do not overlap. Columns (1)-(3) use the balancing-authority footprint rule of the main text, and Column (4) repeats the average-price specification under the alternative shared-utility rule. All specifications include utility-state-by-ballot and calendar-year fixed effects. The estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C24. Headline estimates with grid-connected control counties removed

Horizon (years)	Controls dropped	Share of controls	Data center stock	Log price
None	0	0%	-0.7019***	0.1832
			(0.2194)	(0.2076)
0	18	3%	-0.6693***	0.2005
			(0.2316)	(0.2064)
1	18	3%	-0.6693***	0.2005
			(0.2316)	(0.2064)
2	74	14%	-0.7135***	0.2150
			(0.2480)	(0.2095)
3	74	14%	-0.7135***	0.2150
			(0.2480)	(0.2095)
4	103	19%	-0.7355***	0.2023
			(0.2545)	(0.2105)
5	103	19%	-0.7355***	0.2023
			(0.2545)	(0.2105)
6	103	19%	-0.7355***	0.2023
			(0.2545)	(0.2105)
7	103	19%	-0.7355***	0.2023
			(0.2545)	(0.2105)
8	103	19%	-0.7355***	0.2023
			(0.2545)	(0.2105)

Notes: The table reports panel difference-in-discontinuities estimates of the headline specification after removing grid-connected control counties, at the county-ballot-year level. A control county lies in a state whose ballot failed. It is removed when a utility serving it also serves a *different* state that passed a ballot within the stated horizon of this county's own ballot. A horizon of zero counts only votes held in the same year, and eight counts the whole event window, so a larger horizon is a stricter check. All twelve failing ballots remain in the estimation sample at every horizon, so the set of comparison ballots is unchanged. The largest gap between any pair of grid-connected ballots in the sample is four years, which is why horizons above four remove nothing further. The specification follows Table C5, with county-by-ballot and calendar-year fixed effects, the estimation bandwidth $|R_{sa}| \leq 0.10$, and an event window of plus or minus eight years around the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C25. Industry employment and share of county labor

Industry (NAICS)	Mean emp. per county-year (workers) (1)	Mean industry share of county total emp. (%) (2)
Cannabis cultivation, indoor (NAICS 111419)	2.6	0.003%
Data centers (NAICS 518210)	110.9	0.042%
Information (NAICS 51)	868.1	0.784%
Construction (NAICS 23)	2,239.8	4.62%
Manufacturing (NAICS 31-33)	3,820.3	11.31%
Accommodation & food services (NAICS 72)	3,991.5	6.98%
Retail trade (NAICS 44-45)	4,754.2	11.43%
County total (NAICS 10)	44,974.2	100.00%

Notes: The table reports averages across 22,914 county-year observations from the BLS QCEW for 2017-2023. Column (1) reports mean employment in workers, using QCEW annual averages of monthly headcounts. County-years with no firm in the relevant NAICS code enter as zero. Column (2) reports the mean across county-years of each industry's share of total county employment, with shares calculated separately for each county-year.

Table C26. Top occupations in data centers and cannabis cultivation

	Data centers NAICS 518200		Support Activities for Crop Production NAICS 115100		Agriculture, Forestry, Fishing and Hunting sector NAICS 11	
Rank	Occupation (SOC)	Share	Occupation (SOC)	Share	Occupation (SOC)	Share
1	Software Developers (15-1252)	16.1%	Farmworkers and Laborers, Crop, Nursery, and Greenhouse (45-2092)	55.4%	Farmworkers and Laborers, Crop, Nursery, and Greenhouse (45-2092)	45.7%
2	Sales Reps. of Services, Except Advertising, Insurance, Financial Services, and Travel (41-3091)	6.3%	Packers and Packagers, Hand (53-7064)	6.3%	Packers and Packagers, Hand (53-7064)	5.2%
3	Customer Service Representatives (43-4051)	5.4%	Agricultural Equipment Operators (45-2091)	4.5%	Logging Equipment Operators (45-4022)	4.0%
4	Computer and Information Systems Managers (11-3021)	5.0%	First-Line Supervisors of Farming, Fishing, and Forestry Workers (45-1011)	3.2%	Agricultural Equipment Operators (45-2091)	3.8%
5	Computer User Support Specialists (15-1232)	3.9%	Graders and Sorters, Agricultural Products (45-2041)	3.0%	First-Line Supervisors of Farming, Fishing, and Forestry Workers (45-1011)	3.6%
6	General and Operations Managers (11-1021)	3.9%	Laborers and Freight, Stock, and Material Movers, Hand (53-7062)	2.8%	Heavy and Tractor-Trailer Truck Drivers (53-3032)	3.2%
7	Market Research Analysts and Marketing Specialists (13-1161)	3.2%	Industrial Truck and Tractor Operators (53-7051)	2.3%	Farmworkers, Farm, Ranch, and Aquacultural Animals (45-2093)	3.0%

Continued on next page

(Table C26 continued)

	Data centers NAICS 518200		Support Activities for Crop Production NAICS 115100		Agriculture, Forestry, Fishing and Hunting sector NAICS 11	
Rank	Occupation (SOC)	Share	Occupation (SOC)	Share	Occupation (SOC)	Share
8	Computer Systems Analysts (15-1211)	2.8%	Heavy and Tractor-Trailer Truck Drivers (53-3032)	1.6%	Graders and Sorters, Agricultural Products (45-2041)	2.5%
9	Computer Occupations, All Other (15-1299)	2.8%	Packaging and Filling Machine Operators and Tenders (51-9111)	1.6%	Laborers and Freight, Stock, and Material Movers, Hand (53-7062)	2.4%
10	Project Management Specialists (13-1082)	2.5%	Office Clerks, General (43-9061)	1.3%	Industrial Truck and Tractor Operators (53-7051)	1.9%
11	Office Clerks, General (43-9061)	1.9%	General and Operations Managers (11-1021)	1.2%	Office Clerks, General (43-9061)	1.7%
12	Software QA Analysts and Testers (15-1253)	1.9%	Farm Equipment Mechanics and Service Technicians (49-3041)	0.9%	Animal Caretakers (39-2021)	1.5%
	Total emp.: 484,990		*Total emp.: 351,890*		*Total emp.: 427,250*	

Notes: The table reports the top 12 detailed-SOC occupations within each industry, ranked by employment, from the BLS OEWS May 2024 national release. "Share" is the percentage of the industry's total employment in that occupation, as published by OEWS. NAICS 115100, Support Activities for Crop Production, is the closest OEWS-published proxy for cannabis-cultivation labor. The direct crop-production codes NAICS 1114 and 1119, which contain the cannabis cultivation codes 111419 and 111998, are not published in OEWS. The NAICS 11 sector total in OEWS comprises only logging and support activities, and the direct agriculture production codes are not included in the sector total either.

Table C27. Weekly wage across broad sectors

Sector	$\hat{\beta}$: Pass×Post ($/wk)	Pre-period mean ($/wk)	Observations	Within R^2
Construction	-61.24*	860.0	14,850	0.009
	(33.26)			
Manufacturing	-66.22	891.1	13,760	0.004
	(46.16)			
Retail trade	-8.51	490.2	16,276	0.016
	(20.45)			
Information	-261.34**	886.4	12,678	0.008
	(111.03)			
Accommodation & food	5.53	295.1	13,109	0.035
	(23.99)			
Leisure & hospitality	9.14	294.5	16,380	0.045
	(26.15)			
Professional services	-51.33	850.1	15,445	0.002
	(70.10)			
County × ballot FE	Yes			
Year FE	Yes			

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on the intensive-margin weekly wage in seven broad sectors, at the county-ballot-year level, dropping county-years with a missing wage. All specifications include county-by-ballot and calendar-year fixed effects, and the estimation bandwidth is $|R_{sa}| \leq 0.10$ with an estimation window of eight years on either side of the ballot. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C28. Cannabis cultivation wage by indoor and outdoor segment

	Agg. 6-digit (in.+out.)	Indoor only 111419	Outdoor only 111998	Agg. 5-digit (in.+out.)	Indoor only 11141	Outdoor only 11199
	(1)	(2)	(3)	(4)	(5)	(6)
Pass × Post	95.6	619.0**	11.0	67.6	159.6	22.0
	(174.4)	(256.1)	(115.9)	(123.1)	(173.3)	(106.6)
Pre-period mean ($/wk)	553.5	549.6	552.5	575.0	550.7	570.8
County × ballot FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,307	562	1,998	2,691	978	2,263
R^2	0.673	0.834	0.643	0.674	0.783	0.650
Within R^2	0.002	0.056	0.001	0.003	0.009	0.001

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on cannabis-cultivation wages, at the county-ballot-year level. The outcome is the weekly wage in dollars per worker. County-years with a missing wage are dropped, so the sample is restricted to county-years where at least one firm exists in the relevant NAICS aggregate and BLS disclosed the cell. Columns (1) and (4) are employment-weighted averages across the indoor and outdoor component NAICS codes, while Columns (2), (3), (5), and (6) are single-NAICS measures. All specifications include county-by-ballot and calendar-year fixed effects, the estimation bandwidth is $|R_{sa}| \leq 0.10$, and the event window is plus or minus eight years around the ballot. The 5-digit indoor code NAICS 11141 is the residual not assigned to a 6-digit subcode, and the 5-digit outdoor code NAICS 11199 is the aggregate of other crop farming including outdoor cannabis. Standard errors are clustered at the ballot level and are reported in parentheses.

Table C29. Arrests and business establishments

	Arrests per 1,000 residents				Establishments
	(1) Violent offenses	(2) Property offenses	(3) Non-drug offenses	(4) Non-cannabis drugs	(5) County total
$\hat{\beta}$: Pass×Post	0.183	-0.043	0.860	-0.388	0.060
	(0.261)	(0.503)	(2.881)	(0.531)	(0.099)
Fixed effects					
County × ballot	Yes	Yes	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes	Yes	Yes
Observations	15,543	15,543	15,543	15,543	16,870
R^2	0.688	0.782	0.726	0.651	0.979
Mean of dep. var.	1.268	2.753	24.614	2.120	4,289

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on county arrests and business establishments, at the county-ballot-year level. Columns (1)-(4) are arrests per 1,000 county residents, summed over the county's police agencies excluding tribal and state agencies, from the Uniform Crime Reporting arrest data for 2000-2024 (Kaplan, openICPSR #102263). Violent offenses are murder and nonnegligent manslaughter, rape, robbery, and aggravated assault. Property offenses are burglary, larceny-theft, motor vehicle theft, and arson. Non-drug offenses are all offenses except drug offenses, runaways, curfew and loitering, and suspicion. Non-cannabis drug arrests are for the possession or sale of opium or cocaine and their derivatives, synthetic narcotics, and other drugs, excluding marijuana. County-years with no reporting agency or no population count are excluded. Column (5) is the inverse hyperbolic sine of the county's total number of establishments in the Quarterly Census of Employment and Wages, with missing cells coded as zero, as in Table 7. The mean of the dependent variable is the sample mean of arrests per 1,000 residents in Columns (1)-(4) and the pre-period mean number of establishments in Column (5). The sample is the county-year panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. All specifications include county-by-ballot and calendar-year fixed effects and control for $R_{sa} \times \text{Post}$ and $R_{sa} \times \text{Pass} \times \text{Post}$. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 26 clusters, and are reported in parentheses.

Table C30. Water deliveries and water-utility revenue

Outcome:	Total public-supply deliveries	Deliveries to commercial, industrial, institutional, and irrigation customers	Water-utility revenue
	(1)	(2)	(3)
$\hat{\beta}$: Pass×Post	-0.012	0.043	-0.119
	(0.036)	(0.051)	(0.209)
Fixed effects			
County × ballot	Yes	Yes	Yes
Calendar year	Yes	Yes	Yes
Observations	12,037	12,037	12,556
R^2	0.999	0.998	0.922
Mean of dep. var.	17.431	4.379	47.632

Notes: The table reports panel difference-in-discontinuities estimates of the effect of recreational cannabis legalization on county water deliveries and water-utility revenue, at the county-ballot-year level. All outcomes are in inverse hyperbolic sine. Columns (1) and (2) are total public-supply deliveries and the deliveries to commercial, industrial, institutional, and irrigation customers, in gallons per day, from the public-supply water-use estimates of the U.S. Geological Survey for 2000-2020. The deliveries in Column (2) are part of those in Column (1). Column (3) is the water-utility revenue of local governments in thousands of nominal dollars, from the Census of Governments and the Annual Survey of State and Local Government Finances for 2000-2024, excluding county-years in which no government reported. The mean of the dependent variable is in raw units, million gallons per day in Columns (1) and (2) and millions of dollars in Column (3). The sample is the county-year panel, stacked by ballot and restricted to a window of plus or minus eight years around the ballot. All specifications include county-by-ballot and calendar-year fixed effects and control for $R_{sa} \times \text{Post}$ and $R_{sa} \times \text{Pass} \times \text{Post}$. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level, with 24 clusters in Columns (1) and (2) and 26 in Column (3), and are reported in parentheses.

Table C31. Cannabis legalization and state trade in AI-related products

Triple difference	Coefficient	SE	Ballots
Imports			
All AI-related	-0.034	(0.089)	18
Compute hardware	-0.433*	(0.237)	18
HS heading 8471	-0.354**	(0.158)	18
Exports			
All AI-related	0.061	(0.107)	18
Compute hardware	-0.129	(0.210)	18
HS heading 8471	-0.245	(0.154)	18

Notes: The outcome is the inverse hyperbolic sine of a state's annual trade. Each cell is the triple-difference coefficient on $Pass \times Post \times AI$, comparing trade in AI-related products against all other trade within the balloting state. The *All AI-related* rows include all products that Waugh (2026) classifies as having high relevance to AI data center infrastructure. The *Compute hardware* rows include only products in the classification's Compute Hardware category. State trade is published at HS6 and the classification is at HS10, so each HS6 subheading is split by the national share of its value in the relevant HS10 lines that year. The HS heading 8471 rows include only compute-hardware products under that heading, which covers automatic data processing machines and units. Fixed effects are ballot-by-AI, year-by-AI, and ballot-by-year. The sample covers 2013-2024 and the 18 ballots inside the bandwidth with pre- and post-ballot trade. The unit of analysis is the state-year-product-group cell, with separate groups for AI-related and all other trade. The estimation bandwidth is $|R_{sa}| \leq 0.10$. Standard errors are clustered at the ballot level and are reported in parentheses.