The Regional Economics of Green Industrial Policy: Evidence from the Inflation Reduction Act

Enock Bulime

Kansas State University

bulimeenock@ksu.edu

October 30, 2025

Abstract

This paper examines how place-based energy subsidies affect regional employment and economic growth during the energy transition. The Inflation Reduction Act allocates over two hundred billion dollars toward clean energy manufacturing, with investments concentrated in specific counties. I will exploit this spatial heterogeneity to identify the causal effects of green industrial policy. The theoretical framework models how concentrated investments generate regional effects through direct employment, agglomeration economies, and spatial spillovers to neighboring regions. Empirically, I will employ five identification strategies: staggered difference-in-differences, synthetic control methods, regression discontinuity at Energy Community thresholds, instrumental variables using renewable energy potential, and spatial econometrics. The analysis will construct a county-quarter panel linking facility-level investments to administrative employment data. The findings will illuminate how public expenditures in energy markets affect regional economies and how policy design influences the spatial distribution of benefits

during energy transitions.

JEL Classification: Q43, Q48, R11, R12, E24, H71

Keywords: Energy subsidies, Industrial policy, Regional economics, Spatial spillovers, Inflation Reduction Act, Energy transition

1 Introduction

The transition from fossil fuels to low-carbon energy sources represents one of the most significant structural transformations facing modern economies. Understanding the regional economic consequences of this transition has become increasingly urgent as governments worldwide deploy substantial public expenditures to accelerate the shift toward clean energy. The United States Inflation Reduction Act of 2022 exemplifies this policy approach, allocating over four hundred billion dollars through tax credits and subsidies targeting clean energy manufacturing, renewable electricity generation, and electric vehicle production [Bistline et al., 2023]. Unlike economy-wide carbon pricing mechanisms that distribute costs and benefits broadly across regions, the place-based industrial policy embedded in the IRA concentrates investments geographically, with individual counties receiving facilities valued at billions of dollars while neighboring counties receive minimal direct investment. This spatial heterogeneity raises fundamental questions about how such policies affect regional employment, investment patterns, and economic growth, and whether benefits extend beyond directly treated locations through agglomeration economies and spatial spillovers.

This paper will develop and test a spatial economic framework for understanding how concentrated green energy investments affect regional economies during the energy transition. The research addresses three central questions. First, what are the direct employment and investment effects in counties receiving major clean energy facilities, and how do these effects vary with investment magnitude and baseline economic conditions? Second, do concentrated green investments generate agglomeration economies that amplify direct effects through supply chain development and productivity spillovers? Third, do benefits propagate spatially to neighboring regions through input-output linkages and labor market integration, and if so, how rapidly do these spillovers decay with distance?

These questions assume particular importance given ongoing debates about optimal energy transition policy design. If place-based energy subsidies generate substantial agglomeration economies and spatial spillovers, this suggests that concentrated investment strategies

may prove more cost-effective than dispersed allocation. Conversely, if benefits remain narrowly confined to direct recipients without generating broader regional development, this would argue for alternative policy approaches. Moreover, understanding the spatial distribution of benefits proves essential for assessing equity implications and political sustainability of energy transition policies. The geographic concentration of fossil fuel employment in specific regions has generated concerns about transition costs falling disproportionately on vulnerable communities. Whether place-based green subsidies successfully channel benefits to these communities or instead flow to regions already experiencing economic growth represents a critical policy question.

The theoretical model will incorporate three mechanisms through which place-based energy subsidies generate regional economic effects. First, green investments create direct employment through construction labor demand during facility development and operational manufacturing jobs once facilities become operational. The magnitude of these direct effects will depend on facility scale, technology type, and local labor market conditions. Second, concentrated investments may generate agglomeration economies as initial facilities attract supplier networks, complementary investments, and specialized labor pools. Following Kline and Moretti [2014], the model will formalize how manufacturing clusters create positive externalities that amplify the productivity and employment effects of initial investments. These agglomeration mechanisms prove particularly relevant for clean energy manufacturing, which requires specialized supply chains for components such as battery cells, solar panels, and electric vehicle assemblies. Third, concentrated benefits may propagate spatially to neighboring regions through several channels including commuting patterns allowing workers in adjacent counties to access new employment opportunities, input-output linkages as facilities source materials and services from regional suppliers, and demonstration effects as successful investments attract additional private capital to the broader region.

The model will generate testable predictions about heterogeneity in treatment effects across different settings. Counties with stronger existing manufacturing bases should expe-

rience larger employment multipliers from green investments due to complementarities with existing supply chains and specialized labor. Counties in closer proximity to other investment locations should exhibit enhanced effects through agglomeration spillovers. Energy communities with significant fossil fuel employment history may respond differently than other locations, with larger effects if green investments provide valued alternatives to declining fossil fuel employment or smaller effects if structural mismatches exist between fossil fuel workers and clean energy job requirements.

The empirical analysis will exploit quasi-experimental variation in IRA investment allocation across United States counties from August 2022 through December 2024 and potentially extending through 2026 as additional facilities become operational. We will construct a comprehensive dataset linking facility-level investment announcements to county-quarter economic outcomes. The primary investment data source will be the Clean Investment Monitor database maintained by Rhodium Group and the MIT Center for Energy and Environmental Policy Research, which tracks over twenty thousand clean energy facilities including manufacturing plants, utility-scale renewable projects, and related investments [Larsen et al., 2023, 2024]. This database provides facility-level information including precise geographic location, investment amounts, technology type, announcement and operational dates, and expected employment for both construction and operational phases.

We will merge this investment data with administrative employment records from the Bureau of Labor Statistics Quarterly Census of Employment and Wages, which covers over ninety-five percent of employment based on unemployment insurance records. The QCEW provides county-quarter employment and wage data with detailed industry breakdowns, enabling analysis of both aggregate employment effects and sectoral reallocation. Additional data sources will include establishment dynamics from County Business Patterns, county-level economic indicators from the Bureau of Economic Analysis Regional Economic Accounts, and geographic characteristics including renewable energy potential from the National Renewable Energy Laboratory. This comprehensive data infrastructure will enable

granular analysis of how specific facilities affect local labor markets, establishment formation, wages, and broader economic indicators.

The identification strategy will combine five complementary approaches that rely on different assumptions and exploit different sources of variation, providing robust evidence across multiple specifications. First, the primary specification will implement staggered difference-in-differences estimation using recent methodological advances that address treatment effect heterogeneity across cohorts and time periods [Callaway and Sant'Anna, 2021, Sun and Abraham, 2021]. The staggered rollout of IRA investments across counties and quarters provides quasi-experimental variation in treatment timing. We will estimate group-time average treatment effects for cohorts of counties receiving investment at different periods, then aggregate these estimates to construct event study plots showing dynamic treatment effects and overall average treatment effects on the treated. This approach addresses bias in standard two-way fixed effects estimation when treatment effects are heterogeneous [Goodman-Bacon, 2021].

Second, for counties receiving exceptionally large investments such as battery gigafactories or major solar manufacturing facilities, we will complement difference-in-differences with synthetic control analysis [Abadie, 2021]. The synthetic control method constructs data-driven counterfactuals by taking weighted averages of control counties that closely match treated counties on pre-treatment characteristics and outcome trajectories. This approach provides transparent comparisons for high-profile facilities where treatment is particularly salient. We will implement inference through in-space placebo tests that apply the synthetic control method to each potential control county, generating distributions of placebo effects for assessing statistical significance.

Third, the IRA's designation of Energy Communities receiving enhanced tax credits creates discontinuities in treatment probability at statutory thresholds, enabling regression discontinuity identification [Lee and Lemieux, 2010]. Census tracts qualify as Energy Communities if fossil fuel employment exceeded specific thresholds during the period from 2010 through 2021, or if they contain coal mines or coal-fired power plants that closed after 1999.

For counties where fossil fuel employment shares were close to qualification thresholds, we will implement sharp regression discontinuity designs comparing outcomes for counties just above versus just below the cutoff. We will also implement fuzzy regression discontinuity designs treating Energy Community designation as an instrument for actual investment received, since the relationship between eligibility and investment may not be deterministic.

Fourth, counties differ substantially in renewable energy potential due to exogenous geographic characteristics including solar irradiance, wind resources, and proximity to transmission infrastructure. This variation provides the foundation for instrumental variables estimation using shift-share designs [Goldsmith-Pinkham et al., 2020, Borusyak et al., 2022]. We will construct instruments combining location-specific renewable energy potential (the share component) with nationwide shifts in investment toward different technologies following IRA enactment (the shift component). The exclusion restriction requires that conditional on controls, renewable energy potential affects regional economic outcomes only through its effect on attracting IRA investments. We will assess instrument validity through first-stage strength tests, balance tests on pre-treatment characteristics, and sensitivity analyses examining how results change under various violations of the exclusion restriction.

Fifth, spatial econometric specifications will explicitly model spillovers to neighboring counties [LeSage and Pace, 2009]. We will estimate spatial Durbin models that incorporate both spatial lags of the dependent variable and spatially lagged covariates, allowing us to decompose total treatment effects into direct effects in treated counties and indirect effects in neighboring counties. We will specify spatial weight matrices based on both geographic contiguity and distance decay functions, examining how spillover magnitudes vary with distance from investment locations. Following LeSage and Pace [2009], we will compute direct effects, indirect effects, and total effects that properly account for spatial feedback loops where impacts propagate through networks of connected locations.

This research will contribute to three distinct literatures. First, it will advance understanding of how public expenditures in energy markets affect regional and national economic

aggregates. The literature on energy shocks and macroeconomic outcomes has focused primarily on oil price shocks and their effects on aggregate employment, inflation, and growth [Hamilton, 1983, Kilian, 2009]. More recent work examines how the shale gas revolution affected local labor markets and manufacturing activity [Feyrer et al., 2017]. However, relatively little research analyzes the regional economic consequences of place-based energy subsidies designed to accelerate the clean energy transition. Understanding these effects proves essential for assessing whether subsidy-based approaches can effectively complement or substitute for carbon pricing mechanisms.

Second, the paper will contribute to the literature on place-based policies and regional development. Seminal work by Kline and Moretti [2014] demonstrates that large-scale place-based investments can generate persistent effects through agglomeration economies, with manufacturing employment gains intensifying after federal subsidies ended. More recent research examines enterprise zones, infrastructure investments, and targeted business incentives, finding mixed evidence on effectiveness [Neumark and Simpson, 2015]. This paper will extend this literature by examining whether clean energy manufacturing exhibits agglomeration dynamics similar to traditional manufacturing, and whether such dynamics justify concentrated investment strategies during energy transitions.

Third, the findings will inform policy debates about optimal energy transition design. Theoretical work on directed technical change emphasizes that environmental policy can redirect innovation toward clean technologies when inputs are sufficiently substitutable [Acemoglu et al., 2012]. However, this literature typically abstracts from spatial considerations and implementation details. Recent empirical work documents that energy transitions create geographically concentrated disruption in fossil-fuel-dependent regions [Hanson, 2023], raising questions about whether place-based policies can address these distributional concerns. The spatial analysis will reveal whether IRA subsidies successfully target vulnerable communities and whether benefits extend beyond narrow geographic boundaries through regional spillovers.

The remainder of this proposal proceeds as follows. Section 2 develops the theoretical framework and derives testable predictions about how green investments affect regional economies through direct effects, agglomeration economies, and spatial spillovers. Section 3 describes the institutional context of the Inflation Reduction Act and patterns of investment allocation across regions. Section 4 presents the five empirical specifications and identification strategies in detail. Section 5 describes data sources, variable construction, and sample characteristics. Section 6 outlines the expected contributions and policy implications. Section 7 discusses the research timeline and concludes.

2 Theoretical Framework

This section develops a spatial economic model incorporating green investment, employment dynamics, agglomeration externalities, and regional spillovers. The model generates predictions about how investment magnitude, spatial concentration, and baseline economic conditions affect regional employment and income.

2.1 Economic Environment

Consider a spatial economy with a continuum of locations indexed by $\ell \in \mathcal{L}$, where $\mathcal{L}$ represents the set of all counties in the United States. Time is discrete and indexed by $t = 0, 1, 2, \ldots, T$. Each location has a fixed population of workers normalized to unity for notational convenience.

2.1.1 Production Technology

Output in location ℓ at time t is produced using labor $L_{\ell t}$ and capital according to a production function that combines dirty inputs $D_{\ell t}$ and clean inputs $C_{\ell t}$. Following Acemoglu

et al. [2012], we specify production as a constant elasticity of substitution aggregate:

$$Y_{\ell t} = A_{\ell t} \left[\gamma_t D_{\ell t}^{\frac{\epsilon-1}{\epsilon}} + (1-\gamma_t) C_{\ell t}^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}} \tag{1}$$

where $A_{\ell t}$ represents total factor productivity in location ℓ at time t, the parameter $\epsilon > 0$ is the elasticity of substitution between clean and dirty inputs, and $\gamma_t \in (0,1)$ represents the relative productivity of dirty technology at time t. This nested structure captures the essential trade-off between production methods during the energy transition, with the elasticity of substitution determining the ease with which clean inputs can substitute for dirty inputs.

The clean input aggregate $C_{\ell t}$ is produced using location-specific labor according to:

$$C_{\ell t} = B_{\ell t} L_{\ell t}^C \tag{2}$$

where $L_{\ell t}^C$ represents labor employed in clean energy production and $B_{\ell t}$ is the productivity of clean energy production in location ℓ. Similarly, the dirty input is produced as $D_{\ell t} = D_{\ell}^{\text{fixed}}$, representing fossil fuel inputs sourced from outside the region.

2.1.2 Green Investment and Capital Accumulation

Green investment $G_{\ell t}$ in location ℓ augments the productivity of clean energy production according to:

$$B_{\ell,t+1} = (1-\delta) B_{\ell t} + \phi(G_{\ell t}, \bar{G}_{\ell t}) \tag{3}$$

where $\delta \in (0,1)$ is the depreciation rate and $\phi(\cdot)$ is an investment transformation function. The key innovation in this formulation is that investment effectiveness depends on the aggregate investment level in nearby locations $\bar{G}_{\ell t}$, capturing agglomeration economies. Following Kline and Moretti [2014], we model this spillover effect through a logarithmic specification:

$$\phi(G_{\ell t}, \bar{G}_{\ell t}) = G_{\ell t} \left(1 + \alpha \log \left(\frac{\bar{G}_{\ell t}}{\bar{G}_0} \right) \right) \tag{4}$$

where $\alpha > 0$ measures the strength of agglomeration externalities and $\bar{G}_0$ is a normalization constant. The aggregate nearby investment is defined as a distance-weighted average:

$$\bar{G}_{\ell t} = \sum_{\ell' \in \mathcal{N}(\ell)} w_{\ell\ell'} G_{\ell' t} \tag{5}$$

where $\mathcal{N}(\ell)$ denotes the set of neighboring locations and $w_{\ell\ell'}$ are spatial weights that decline with distance. This specification embeds the empirical finding that manufacturing agglomeration economies decay with geographic distance [Rosenthal and Strange, 2004].

2.1.3 Employment Determination and Regional Labor Markets

Total labor demand in location ℓ combines employment in clean energy production with employment in other sectors:

$$L_{\ell t} = L_{\ell t}^{C} + L_{\ell t}^{\text{Other}} \tag{6}$$

Employment in clean energy production is determined by the marginal product condition equating the wage rate w_t to the value marginal product of labor:

$$w_t = p_t^C \cdot \frac{\partial Y_{\ell t}}{\partial C_{\ell t}} \cdot \frac{\partial C_{\ell t}}{\partial L_{\ell t}^C} = p_t^C \cdot MPL_{\ell t}^C \tag{7}$$

where p_t^C is the price of clean inputs and $MPL_{\ell t}^C$ is the marginal physical product of labor in clean production.

The employment effect of green investment operates through both direct channels and indirect agglomeration channels. The total employment response can be decomposed as:

$$\frac{\partial L_{\ell t}}{\partial G_{\ell t}} = \underbrace{\frac{\partial L_{\ell t}^C}{\partial B_{\ell t}} \cdot \frac{\partial B_{\ell t}}{\partial G_{\ell t}}}_{\text{Direct Effect}} + \underbrace{\frac{\partial L_{\ell t}^C}{\partial B_{\ell t}} \cdot \frac{\partial B_{\ell t}}{\partial \bar{G}_{\ell t}} \cdot \frac{\partial \bar{G}_{\ell t}}{\partial G_{\ell' t}}}_{\text{Agglomeration Effect}} \tag{8}$$

The direct effect captures how investment in location ℓ increases productivity and hence labor demand in that location. The agglomeration effect captures how investment in lo-

cation ℓ increases investment in neighboring locations $\ell' \in \mathcal{N}(\ell)$, which in turn increases productivity in location ℓ through the spatial spillover term $\bar{G}_{\ell t}$.

2.1.4 Spatial Spillovers and Regional Integration

Employment effects propagate across space through several mechanisms. First, workers may commute across county boundaries, allowing residents of location ℓ' to access employment opportunities created in neighboring location ℓ. We model this through a commuting probability $\pi_{\ell\ell'}$ that declines with distance. Second, facilities in location ℓ may source intermediate inputs from suppliers in neighboring locations, generating derived labor demand through input-output linkages. Third, successful investments may attract additional private capital to the broader region through demonstration effects and improved infrastructure.

The total employment effect in location ℓ' from investment in location ℓ can be expressed as:

$$\frac{\partial L_{\ell' t}}{\partial G_{\ell t}} = \pi_{\ell'\ell} \cdot \frac{\partial L_{\ell t}}{\partial G_{\ell t}} + \sum_k \theta_{\ell' k} \cdot \frac{\partial I_{\ell k t}}{\partial G_{\ell t}} \tag{9}$$

where $\pi_{\ell'\ell}$ is the commuting share, $\theta_{\ell' k}$ represents input-output coefficients linking location ℓ' to sector k in location ℓ, and $I_{\ell k t}$ represents intermediate input demand.

2.2 Equilibrium Characterization and Predictions

Definition 1 (Regional Economic Equilibrium)**.** A regional economic equilibrium consists of allocations $\{L_{\ell t}, L^C_{\ell t}, C_{\ell t}, Y_{\ell t}\}$ for all locations ℓ and periods t, prices $\{w_t, p^C_t\}$, and investment levels $\{G_{\ell t}\}$ such that firms maximize profits choosing optimal labor and input combinations, labor markets clear with total employment equal to population in each location, and productivity evolves according to investment and depreciation.

The model generates several testable predictions about the regional economic effects of green investments.

Prediction 1 (Direct Employment Effects)**.** *Counties receiving larger green investments will*

experience greater employment growth. The magnitude of employment effects will increase with investment scale and baseline manufacturing productivity:

$$\frac{\partial L_{\ell t}}{\partial G_{\ell t}} > 0 \quad \textit{and} \quad \frac{\partial^2 L_{\ell t}}{\partial G_{\ell t} \partial A_{\ell,t-1}} > 0 \tag{10}$$

This prediction follows directly from the production structure and will be tested using difference-in-differences and synthetic control specifications comparing employment growth in counties receiving investments to control counties.

Prediction 2 (Agglomeration Amplification)**.** *Employment effects will be larger in counties with proximity to other investment locations. The employment multiplier increases with aggregate nearby investment:*

$$\frac{\partial}{\partial \bar{G}_{\ell t}} \left(\frac{\partial L_{\ell t}}{\partial G_{\ell t}} \right) = \frac{\alpha}{G_{\ell t}} \cdot \frac{\partial L^C_{\ell t}}{\partial B_{\ell t}} > 0 \tag{11}$$

This prediction captures the strategic complementarity in green investment allocation. We will test this through heterogeneity analysis examining whether treatment effects are larger in counties located near other investment sites.

Prediction 3 (Spatial Spillovers with Distance Decay)**.** *Counties neighboring high-investment counties will experience positive employment spillovers that decline with distance:*

$$\frac{\partial L_{\ell' t}}{\partial G_{\ell t}} > 0 \textit{ for } \ell' \in \mathcal{N}(\ell), \quad \frac{\partial}{\partial d_{\ell\ell'}} \left(\frac{\partial L_{\ell' t}}{\partial G_{\ell t}} \right) < 0 \tag{12}$$

where $d_{\ell\ell'}$ represents the distance between locations.

This prediction will be tested using spatial econometric specifications that explicitly model spillovers as a function of distance, allowing us to estimate how rapidly spillover effects decay with geographic separation.

Prediction 4 (Heterogeneity by Baseline Conditions)**.** *Employment effects will be larger in*

counties with stronger baseline manufacturing presence and smaller in counties with structural mismatches between existing workforce skills and clean energy job requirements. Effects will also vary by Energy Community status, with ambiguous predictions depending on whether green jobs provide valued alternatives to declining fossil fuel employment versus whether skills mismatches dominate.

We will test this prediction through extensive heterogeneity analysis interacting treatment variables with pre-treatment county characteristics including manufacturing employment share, educational attainment, fossil fuel employment concentration, and Energy Community designation.

3 Institutional Context: The Inflation Reduction Act

The Inflation Reduction Act, enacted on August 16, 2022, represents the most significant climate and energy legislation in United States history. This section describes the key provisions affecting investment location decisions and patterns of investment allocation across regions during the initial implementation period.

3.1 Key Provisions and Tax Credit Structure

The IRA allocates approximately three hundred and seventy billion dollars over ten years through tax credits and direct subsidies targeting clean energy technologies. The legislation employs several mechanisms to influence investment location decisions. The Advanced Manufacturing Production Credit under Section 45X provides tax credits for domestic manufacturing of clean energy components including solar photovoltaic cells and wafers, battery cells and modules, wind turbine nacelles and blades, and inverters and other power electronics. Credit values vary by component type, with battery cells receiving thirty-five dollars per kilowatt-hour of capacity and solar cells receiving four cents per watt of capacity.

The Investment Tax Credit under Section 48 offers a baseline thirty percent credit for

qualified renewable electricity generation facilities including solar, wind, geothermal, and energy storage projects. Projects meeting domestic content requirements receive additional ten percentage point bonus credits. The Clean Vehicle Credit under Section 30D provides up to seven thousand five hundred dollars for qualifying electric vehicles, with strict requirements that final assembly occur in North America and that substantial portions of battery components and critical minerals be sourced domestically or from free trade agreement partners.

3.2 Energy Communities and Geographic Targeting

The IRA designates certain areas as Energy Communities eligible for enhanced tax credits. Specifically, projects located in Energy Communities receive ten percentage point bonus credits above baseline rates, creating strong incentives for firms to locate facilities in these areas. Census tracts qualify as Energy Communities under three criteria. First, metropolitan statistical area census tracts qualify if fossil fuel employment exceeded five percent of total employment in any year from 2010 through 2021, while non-metropolitan tracts qualify at a 0.17 percent threshold. Second, census tracts qualify if they contain coal mines or coal-fired power plants that closed after 1999 and before the tax credit placed-in-service date. Third, brownfield site designations under federal environmental programs confer Energy Community status.

This geographic targeting reflects explicit policy goals of directing clean energy investment benefits toward communities historically dependent on fossil fuel industries and facing economic transitions as those industries decline. Analysis by Raimi and Pesek [2022] indicates that approximately eight percent of United States census tracts qualify as Energy Communities under these criteria, with substantial concentration in Appalachian coal regions, Western coal and natural gas producing states, and Gulf Coast oil and gas regions.

3.3 Investment Patterns and Spatial Distribution

Investment announcements accelerated dramatically following IRA enactment. According to tracking by the Clean Investment Monitor [Larsen et al., 2023, 2024], firms announced over two hundred and thirteen billion dollars in clean energy investments during the first two years post-enactment through August 2024, representing a one hundred and sixty-five percent increase from the five-year pre-IRA average annual investment rate. Manufacturing investment grew particularly rapidly, with battery manufacturing capacity announcements reaching one hundred and thirty-six gigawatt-hours annually compared to less than twenty gigawatt-hours prior to the IRA.

Investment exhibits substantial geographic concentration. Battery manufacturing facilities concentrate heavily in Southeastern states, with major announcements in Georgia including Hyundai Motor Group's seven point six billion dollar electric vehicle and battery complex in Bryan County, Rivian Automotive's five billion dollar electric vehicle plant in Morgan and Walton Counties, and multiple battery cell manufacturing facilities by SK Innovation, Panasonic, and other suppliers. Michigan received substantial electric vehicle and battery investment building on existing automotive manufacturing clusters. Solar manufacturing capacity concentrates in Southeastern states with existing industrial infrastructure and lower labor costs.

Energy Communities received substantial but not exclusive investment flows. Preliminary analysis indicates that approximately thirty-five percent of announced manufacturing investment by dollar value located in counties with Energy Community designation, compared to their eight percent population share. However, many large facilities also located in non-Energy Community counties with strong manufacturing infrastructure, renewable energy resources, or proximity to major markets. This variation in investment intensity conditional on Energy Community status will enable regression discontinuity identification around eligibility thresholds.

The geographic patterns suggest several factors beyond tax credits influence location

decisions including proximity to existing automotive manufacturing clusters for electric vehicle assembly and battery production, access to renewable energy resources for electricity-intensive manufacturing processes such as battery cell production, availability of large industrial sites with electrical infrastructure capacity, proximity to ports for international component imports and exports, and state-level incentive packages supplementing federal tax credits. This heterogeneity in location determinants motivates the instrumental variables approach using renewable energy potential as an exogenous determinant of investment attraction.

4 Empirical Strategy

This section presents five complementary identification strategies that will provide robust evidence on the regional economic effects of IRA investments under varying assumptions.

4.1 Staggered Difference-in-Differences with Heterogeneous Treatment Effects

The primary empirical specification will exploit the staggered rollout of IRA investments across counties and time periods. We will define treatment timing based on when counties first receive substantial green investment, operationalized as cumulative announced investment exceeding one hundred million dollars or expected employment exceeding five hundred jobs. These thresholds correspond to major manufacturing facilities or utility-scale renewable generation projects that plausibly affect local labor markets.

4.1.1 Event Study Framework

For county ℓ, let G_ℓ denote the quarter when treatment first occurs, with $G_\ell = \infty$ for never-treated counties. The event study specification is:

$$Y_{\ell t} = \alpha_\ell + \gamma_t + \sum_{k=-K}^{-2} \beta_k D_{\ell t}^k + \sum_{k=0}^{K} \beta_k D_{\ell t}^k + X_{\ell t}'\delta + \varepsilon_{\ell t} \tag{13}$$

where $Y_{\ell t}$ represents the outcome of interest including log employment, log average wages, or establishment counts, α_ℓ are county fixed effects absorbing time-invariant location characteristics, γ_t are time fixed effects capturing common aggregate shocks, $X_{\ell t}$ are time-varying control variables, and $D_{\ell t}^k = \mathbb{1}\{t - G_\ell = k\}$ are indicators for event time relative to initial treatment. The coefficient β_k captures the average treatment effect k periods relative to treatment, with $k = -1$ normalized to zero for identification.

However, recent methodological advances demonstrate that standard two-way fixed effects estimation of this specification produces biased estimates when treatment effects vary across cohorts or calendar time [Goodman-Bacon, 2021, de Chaisemartin and D'Haultfœuille, 2020]. The bias arises because later-treated units serve as controls for earlier-treated units, and comparison of earlier-treated units to later-treated units can generate negative weights on some treatment effects even when all underlying effects are positive.

4.1.2 Callaway-Sant'Anna Estimator

To address these concerns, we will implement the Callaway and Sant'Anna [2021] approach that estimates treatment effects separately for each combination of treatment cohort g and time period t. The group-time average treatment effect is defined as:

$$ATT(g,t) = \mathbb{E}[Y_{\ell t}(g) - Y_{\ell t}(\infty) | G_\ell = g] \tag{14}$$

where $Y_{\ell t}(g)$ represents the potential outcome when first treated in period g and $Y_{\ell t}(\infty)$ represents the never-treated potential outcome.

We will estimate this using a doubly-robust procedure combining outcome regression and propensity score weighting:

$$\begin{aligned}\widehat{ATT}(g,t) =& \frac{1}{n_g} \sum_{\ell: G_\ell = g} \left[Y_{\ell t} - \hat{m}(X_{\ell, g-1}, t) \right] \\ & - \frac{1}{n_g} \sum_{\ell: G_\ell = g} \left[\frac{\sum_{\ell' \in C(g,t)} \hat{p}_{\ell'} (Y_{\ell' t} - \hat{m}(X_{\ell', g-1}, t))}{\sum_{\ell' \in C(g,t)} \hat{p}_{\ell'}} \right]\end{aligned} \tag{15}$$

where n_g is the number of counties first treated in period g, $C(g,t)$ denotes the comparison group (either never-treated counties or not-yet-treated counties in period t), $\hat{p}_{\ell'}$ are propensity score weights estimated via logistic regression of treatment status on covariates, and $\hat{m}(X_{\ell,g-1}, t)$ are outcome regression predictions. This doubly-robust approach remains consistent if either the propensity score model or the outcome regression model is correctly specified.

For presentation, we will aggregate the group-time effects to construct event study plots showing treatment effects by periods relative to treatment:

$$\widehat{ATT}^{ES}(e) = \sum_g \frac{n_g}{\sum_{g'} n_{g'}} \widehat{ATT}(g, g+e) \tag{16}$$

where $e = t - g$ represents event time. We will also compute an overall average treatment effect:

$$\widehat{ATT}^{\text{overall}} = \sum_g \sum_{t \geq g} \frac{n_g \cdot \mathbb{1}\{t \geq g\}}{\sum_{g'} \sum_{t' \geq g'} n_{g'} \cdot \mathbb{1}\{t' \geq g'\}} \widehat{ATT}(g,t) \tag{17}$$

4.1.3 Parallel Trends Assessment

The key identifying assumption is parallel trends: in the absence of treatment, treated and control counties would have followed parallel outcome trajectories. We will assess this assumption through several approaches. First, we will examine pre-treatment event study

coefficients β_k for $k < 0$, testing whether they are jointly equal to zero. Rejection would indicate pre-existing differential trends between treated and control counties. Second, we will compare treated and control counties on baseline characteristics and pre-treatment outcome trends using formal balance tests. Third, following Roth [2022], we will implement sensitivity analysis examining how our estimates change under various violations of parallel trends with different magnitudes and functional forms.

4.1.4 Control Variables and Fixed Effects

All specifications will include rich sets of controls to address concerns about differential trends across regions with different economic structures. The baseline specification includes county fixed effects α_ℓ, year-quarter fixed effects γ_t, and state-specific linear time trends. Augmented specifications will include time-varying county characteristics interacted with time fixed effects including log county population, log median household income, and baseline manufacturing employment share. The most saturated specifications will include state-by-year fixed effects α_{st}, absorbing all variation in outcomes across states and time including state-level policy changes, business cycle dynamics, or other factors affecting states differentially.

4.2 Synthetic Control Method for Major Facilities

For counties receiving exceptionally large investments, we will complement the difference-in-differences analysis with synthetic control methods [Abadie, 2021]. The synthetic control approach constructs counterfactual outcomes for treated units by taking weighted averages of control units that closely match the treated unit on pre-treatment characteristics.

4.2.1 Synthetic Control Construction

For a treated county ℓ_1 receiving major investment in period T_0, we construct a synthetic control by solving:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \|\mathbf{X}_1 - \mathbf{X}_0\mathbf{w}\|_{\mathbf{V}} = \arg\min_{\mathbf{w}} (\mathbf{X}_1 - \mathbf{X}_0\mathbf{w})'\mathbf{V}(\mathbf{X}_1 - \mathbf{X}_0\mathbf{w}) \tag{18}$$

subject to $w_j \geq 0$ for all j and $\sum_{j=2}^{J+1} w_j = 1$, where $\mathbf{X}_1$ is a $(k \times 1)$ vector of pre-treatment characteristics for the treated county, $\mathbf{X}_0$ is a $(k \times J)$ matrix of the same characteristics for J potential control counties, and $\mathbf{V}$ is a $(k \times k)$ positive semidefinite matrix assigning weights to the importance of matching on different characteristics.

Following Abadie [2021], we will choose $\mathbf{V}$ to minimize the mean squared prediction error of the outcome variable during the pre-treatment period:

$$\mathbf{V}^* = \arg\min_{\mathbf{V}} \sum_{t=1}^{T_0-1} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^*(\mathbf{V}) Y_{jt} \right)^2 \tag{19}$$

The predictor variables in $\mathbf{X}$ will include several lags of the pre-treatment outcome variable at different time points, baseline county characteristics including population, median income, educational attainment, manufacturing employment share, and pre-treatment economic trends including employment growth and wage growth from 2018 through 2021. Including multiple outcome lags often achieves better pre-treatment fit than including only covariates [Kaul et al., 2015].

4.2.2 Treatment Effect Estimation and Inference

The estimated treatment effect in post-treatment period t is:

$$\hat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \quad \text{for } t \geq T_0 \tag{20}$$

The average post-treatment effect over the observation period is:

$$\hat{\tau}_1 = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^{T} \hat{\tau}_{1t} \tag{21}$$

We will conduct inference using in-space placebo tests as developed by Abadie et al. [2010]. We will iteratively apply the synthetic control method to each control county, pretending it was treated at the same time as the actual treated county. This generates a distribution of placebo effects $\{\hat{\tau}_{jt}\}_{j=2}^{J+1}$ for each post-treatment period. The permutation-based p-value for period t is:

$$p_t = \frac{1 + \sum_{j=2}^{J+1} \mathbb{1}\{|\hat{\tau}_{jt}| \geq |\hat{\tau}_{1t}|\}}{J + 1} \tag{22}$$

Following Abadie et al. [2015], we will restrict the inference sample to placebo units with pre-treatment root mean squared prediction error no more than twice that of the treated unit, avoiding poorly-matched placebos that inflate p-values.

4.2.3 Robustness Checks

We will implement several robustness exercises. First, leave-one-out analysis will iteratively exclude each control county with positive weight and re-estimate the synthetic control, assessing whether results depend critically on any single control unit. Second, we will conduct backdating exercises that artificially move the intervention date earlier to pseudo-treatment periods that are actually pre-treatment, verifying that the synthetic control tracks the treated unit well during these pseudo-post periods. Third, we will vary the set of predictor variables to assess sensitivity to predictor choice. Fourth, we will implement the augmented synthetic control method of Ben-Michael et al. [2021], which addresses interpolation bias when no convex combination of control units closely approximates the treated unit by allowing for regularization.

4.3 Regression Discontinuity at Energy Community Thresholds

The IRA's designation of Energy Communities receiving enhanced tax credits creates discontinuities in treatment probability at statutory thresholds. We will exploit these discontinuities for identification using regression discontinuity designs.

4.3.1 Sharp RD: Fossil Fuel Employment Threshold

Census tracts qualify as Energy Communities if fossil fuel employment as a share of total employment exceeded statutory thresholds during 2010 through 2021. For metropolitan statistical areas, the threshold is five percent, while for non-metropolitan areas, the threshold is 0.17 percent. For counties where the maximum fossil fuel employment share during this period fell close to the relevant threshold, assignment to Energy Community status is effectively deterministic conditional on the running variable.

The sharp regression discontinuity specification is:

$$Y_\ell = \alpha + \tau \cdot \mathbb{1}\{R_\ell \geq c\} + f(R_\ell - c) + X_\ell'\beta + u_\ell \tag{23}$$

where R_ℓ is county ℓ's fossil fuel employment share measured as the maximum over 2010-2021 (the running variable), c is the relevant threshold (five percent for metropolitan areas, 0.17 percent for non-metropolitan), $\mathbb{1}\{R_\ell \geq c\}$ indicates treatment (Energy Community status), and $f(\cdot)$ is a flexible polynomial or spline function in the centered running variable allowing different slopes on either side of the threshold.

Following best practices [Lee and Lemieux, 2010], we will estimate local linear regression specifications:

$$Y_\ell = \alpha_0 + \alpha_1(R_\ell - c) + \tau \cdot \mathbb{1}\{R_\ell \geq c\} + \beta_1 \mathbb{1}\{R_\ell \geq c\}(R_\ell - c) + u_\ell \tag{24}$$

using observations within bandwidth h of the cutoff, where $|R_\ell - c| \leq h$. We will select

the bandwidth using the Calonico et al. [2014] data-driven optimal bandwidth selector, which trades off bias and variance to minimize mean squared error. We will also present results using alternative bandwidths including half and twice the optimal bandwidth to assess sensitivity.

4.3.2 Fuzzy RD: Actual Investment as Outcome of Energy Community Status

Energy Community designation may affect investment through the enhanced tax credits but the relationship between eligibility and actual investment received is not deterministic. Some Energy Communities receive no investment while some non-Energy Communities receive substantial investment based on other location advantages. This motivates a fuzzy regression discontinuity design treating Energy Community status as an instrument for investment received.

The first-stage relationship is:

$$\text{Invest}_\ell = \pi_0 + \pi_1 \cdot \mathbb{1}\{R_\ell \geq c\} + g(R_\ell - c) + X_\ell'\gamma + v_\ell \tag{25}$$

The reduced-form relationship is:

$$Y_\ell = \rho_0 + \rho_1 \cdot \mathbb{1}\{R_\ell \geq c\} + g(R_\ell - c) + X_\ell'\delta + \epsilon_\ell \tag{26}$$

The fuzzy RD estimator is the ratio of reduced-form to first-stage discontinuities:

$$\hat{\tau}^{FRD} = \frac{\hat{\rho}_1}{\hat{\pi}_1} \tag{27}$$

This estimates the local average treatment effect for compliers, defined as counties whose investment status is affected by crossing the Energy Community eligibility threshold.

4.3.3 Validity Tests

We will assess the validity of the RD design through several diagnostic tests. First, we will implement the McCrary [2008] density test examining whether the density of the running variable is continuous at the threshold. A discontinuity in the density would suggest manipulation of the running variable around the cutoff, violating the identification assumptions. Second, we will conduct balance tests estimating RD specifications with pre-treatment covariates as outcome variables. If covariates exhibit discontinuities at the threshold, this suggests that units just above and below the cutoff differ systematically on characteristics that may independently affect outcomes. Third, we will implement placebo tests examining whether outcomes measured in pre-treatment periods exhibit discontinuities at the threshold. Any such discontinuities would indicate confounding factors correlated with threshold crossing rather than causal effects of Energy Community status. Fourth, we will implement donut RD specifications excluding observations very close to the threshold [Barreca et al., 2016], assessing whether results are driven by potential sorting or measurement error near the cutoff.

4.4 Instrumental Variables Using Renewable Energy Potential

Counties differ substantially in their potential for renewable energy generation due to exogenous geographic characteristics. We will exploit this variation using instrumental variables strategies based on shift-share designs.

4.4.1 Shift-Share Instrument Construction

The instrument combines location-specific exposure measures (shares) with common aggregate shocks (shifts). For county ℓ, we construct:

$$Z_\ell = \sum_{k \in \mathcal{K}} s_{\ell k} \cdot g_k \tag{28}$$

where k indexes renewable energy technology types (solar photovoltaic, onshore wind, offshore wind, geothermal), $s_{\ell k}$ represents county ℓ's exposure or suitability for technology k (the share), and g_k represents the nationwide shift in investment toward technology k following IRA enactment.

We will construct the shares using data from the National Renewable Energy Laboratory on county-level renewable energy technical potential and capacity factors. Specifically:

$$s_{\ell k} = \frac{\text{CapacityFactor}_{\ell k}}{\sum_{\ell' \in \mathcal{L}} \text{CapacityFactor}_{\ell' k}} \tag{29}$$

where $\text{CapacityFactor}_{\ell k}$ represents the average capacity factor (ratio of actual to maximum potential generation) for technology k in county ℓ based on solar irradiance, wind speeds at relevant hub heights, or geothermal gradients. These capacity factors depend on geological and meteorological conditions that are plausibly exogenous to economic outcomes.

The shifts are defined as:

$$g_k = \frac{\text{Investment}_k^{\text{post-IRA}} - \text{Investment}_k^{\text{pre-IRA}}}{\text{Investment}_k^{\text{pre-IRA}}} \tag{30}$$

measuring the percentage increase in nationwide investment in technology k comparing periods after and before IRA enactment.

4.4.2 Identification Approaches

Following Borusyak et al. [2022], we will implement the quasi-experimental shift-share approach where identification relies on exogeneity of the shifts g_k. The key assumption is that conditional on controls, nationwide policy-driven investment shifts are orthogonal to unobserved local determinants of economic outcomes. This assumption is plausible if IRA passage and the resulting investment shifts were driven by national political economy considerations rather than economic conditions in specific locations.

Alternatively, following Goldsmith-Pinkham et al. [2020], we can implement the expo-

sure shares approach where identification relies on exogeneity of the shares $s_{\ell k}$. The key assumption is that conditional on controls, counties with different renewable energy potential are comparable on unobserved determinants of economic outcomes. This assumption is plausible if geology and meteorology determining renewable potential are independent of local economic factors.

4.4.3 Empirical Specifications

The first-stage relationship is:

$$\text{Invest}_{\ell t} = \delta_0 + \delta_1 Z_\ell \times \text{Post}_t + \alpha_\ell + \gamma_t + X'_{\ell t}\pi + v_{\ell t} \tag{31}$$

where $\text{Post}_t = \mathbb{1}\{t \geq 2022Q3\}$ indicates periods after IRA enactment. The reduced-form relationship is:

$$Y_{\ell t} = \rho_0 + \rho_1 Z_\ell \times \text{Post}_t + \alpha_\ell + \gamma_t + X'_{\ell t}\theta + u_{\ell t} \tag{32}$$

The two-stage least squares estimator is:

$$\hat{\beta}^{2SLS} = \frac{\hat{\rho}_1}{\hat{\delta}_1} \tag{33}$$

We will assess instrument strength using the first-stage F-statistic, with the Stock and Yogo [2005] rule of thumb that F-statistics exceeding ten provide adequate protection against weak instrument bias. For the shift-share instrument with clustered standard errors, we will use the Kleibergen-Paap rk Wald F-statistic.

4.4.4 Validity Assessment

We will assess the exclusion restriction through several approaches. First, we will conduct balance tests examining whether counties with higher instrument values differ systematically on pre-treatment observable characteristics after conditioning on the included controls. Sec-

ond, we will implement Rotemberg weight analysis [Goldsmith-Pinkham et al., 2020], which decomposes the IV estimate into a weighted average of individual share-specific coefficients and examines which shares drive the overall estimate. Shares with implausibly large weights or implausible sign patterns suggest violations of the exclusion restriction. Third, when using different technology types as separate instruments, we will implement overidentification tests. Fourth, we will conduct sensitivity analysis following Conley et al. [2012], examining how results change under various magnitudes of exclusion restriction violations.

4.5 Spatial Econometric Methods

The theoretical model predicts spatial spillovers operating through agglomeration economies and input-output linkages. We will test these predictions using spatial econometric specifications that explicitly model dependence across neighboring locations.

4.5.1 Spatial Durbin Model

The spatial Durbin model (SDM) incorporates both a spatial lag of the dependent variable and spatially lagged explanatory variables:

$$\begin{aligned} Y_{\ell t} = \rho \sum_{\ell' \neq \ell} w_{\ell\ell'} Y_{\ell' t} + \beta_1 \text{Invest}_{\ell t} + \beta_2 \sum_{\ell' \neq \ell} w_{\ell\ell'} \text{Invest}_{\ell' t} \\ + X_{\ell t}' \theta_1 + \left(\sum_{\ell' \neq \ell} w_{\ell\ell'} X_{\ell' t} \right)' \theta_2 + \alpha_\ell + \gamma_t + u_{\ell t} \end{aligned} \tag{34}$$

where $w_{\ell\ell'}$ are elements of a row-normalized spatial weight matrix $\mathbf{W}$ defining the neighborhood structure, ρ is the spatial autoregressive parameter capturing the strength of spatial dependence in outcomes, β_1 measures direct effects of own-county investment, and β_2 captures effects of investment in neighboring counties.

4.5.2 Spatial Weight Matrix Specifications

We will specify spatial weight matrices using three approaches. First, contiguity-based weights where $w_{\ell\ell'} = 1$ if counties ℓ and ℓ' share a border and zero otherwise, followed by row normalization so that each row sums to one. Second, distance-based weights where $w_{\ell\ell'} = 1/d_{\ell\ell'}^{\alpha}$ for counties within radius r kilometers of each other, where $d_{\ell\ell'}$ is the distance between county centroids and the decay parameter α will be set to one or two. Counties beyond radius r receive zero weight. Third, economic distance weights based on similarity of industrial structures or input-output linkages between counties.

4.5.3 Direct, Indirect, and Total Effects

In spatial models, marginal effects must account for feedback loops where changes in one location affect neighbors, which in turn feed back to the original location. Following LeSage and Pace [2009], we will compute three summary measures. The average direct effect captures the impact of changing the explanatory variable in location ℓ on the outcome in location ℓ:

$$\bar{M}_D = n^{-1}\text{tr}\left[(\mathbf{I}_n - \rho\mathbf{W})^{-1}(\beta_1\mathbf{I}_n + \beta_2\mathbf{W})\right] \tag{35}$$

The average indirect effect captures the impact of changing the explanatory variable in all other locations $\ell' \neq \ell$ on the outcome in location ℓ:

$$\bar{M}_I = n^{-1}\mathbf{i}_n'\left[(\mathbf{I}_n - \rho\mathbf{W})^{-1}(\beta_1\mathbf{I}_n + \beta_2\mathbf{W}) - \text{diag}\left((\mathbf{I}_n - \rho\mathbf{W})^{-1}(\beta_1\mathbf{I}_n + \beta_2\mathbf{W})\right)\right]\mathbf{i}_n \tag{36}$$

The average total effect is the sum:

$$\bar{M}_T = \bar{M}_D + \bar{M}_I \tag{37}$$

These measures properly account for spatial feedback and provide interpretable summaries of direct and spillover effects.

4.5.4 Estimation and Inference

We will estimate spatial panel models with fixed effects using maximum likelihood on transformed data that eliminates the fixed effects. To address incidental parameters bias arising in short panels, we will implement the bias-corrected estimator of Lee and Yu [2010]. For inference, we will compute standard errors that are robust to heteroskedasticity and spatial correlation using the Conley [1999] approach, which allows for spatial correlation declining with distance up to some cutoff.

5 Data Sources and Construction

We will construct a comprehensive county-quarter panel dataset covering all United States counties from the first quarter of 2018 through the fourth quarter of 2024, with potential extension through 2026 as additional data become available.

5.1 Green Investment Data

The primary investment data source will be the Clean Investment Monitor database maintained by Rhodium Group in partnership with the MIT Center for Energy and Environmental Policy Research [Larsen et al., 2023, 2024]. This database tracks over twenty thousand clean energy facilities nationwide including manufacturing plants, utility-scale renewable generation projects, energy storage installations, and related investments. We will request academic data access from the research team to obtain facility-level records including geographic coordinates, county FIPS codes, congressional district identifiers, investment amounts in millions of dollars, technology classifications, announcement dates, planned and actual operational dates, expected employment during construction and operational phases, funding sources distinguishing federal tax credits from state incentives and private capital, and company ownership information.

We will supplement the Clean Investment Monitor with data from the Atlas Public Policy

Clean Economy Tracker, which provides weekly updates on manufacturing facilities and includes explicit Energy Community eligibility flags. We will also incorporate Department of Energy Loan Programs Office data on federal financing for major projects. After cleaning and deduplication across sources, we will geocode all facilities to county FIPS codes and aggregate to county-quarter level, creating variables for cumulative investment amounts, facility counts by technology type, expected employment, and treatment indicators based on substantial investment thresholds.

5.2 Employment and Economic Outcomes

We will obtain county-level employment and wage data from the Bureau of Labor Statistics Quarterly Census of Employment and Wages, which provides comprehensive coverage based on state unemployment insurance records. The QCEW includes employment counts, total wages, average weekly wages, and establishment counts with detailed industry breakdowns at the NAICS two-digit through six-digit level. We will extract quarterly data for all counties covering total employment, manufacturing employment by detailed subsector, clean energy manufacturing including solar panel manufacturing (NAICS 334413), battery manufacturing (NAICS 335912), and motor vehicle manufacturing (NAICS 3361), and fossil fuel employment including coal mining (NAICS 2121), oil and gas extraction (NAICS 211), and fossil fuel electric power generation (NAICS 221112).

Additional economic data will come from the American Community Survey five-year estimates providing county-level population, demographic composition, educational attainment, median household income, poverty rates, and industry of employment. We will use the Census Bureau County Business Patterns for annual establishment counts and employment distributions by firm size class. County-level gross domestic product, personal income, and industry contributions will come from the Bureau of Economic Analysis Regional Economic Accounts.

5.3 Geographic and Infrastructure Characteristics

We will obtain renewable energy potential data from the National Renewable Energy Laboratory including the National Solar Radiation Database providing county-level solar irradiance and capacity factors, the Wind Integration National Dataset providing wind speeds and capacity factors at various hub heights, and technical potential estimates for utility-scale solar, rooftop solar, onshore wind, offshore wind, and geothermal generation. We will process these geospatial datasets using GIS software to calculate county-level averages.

Additional geographic data will include existing energy infrastructure from the Energy Information Administration Form 860 database of all power plants and generators, transmission line locations from Homeland Infrastructure Foundation Level Data, port and freight rail access from the Bureau of Transportation Statistics National Transportation Atlas, and Energy Community designation boundaries from the Internal Revenue Service interactive map. We will calculate distances from county centroids to nearest major infrastructure and create spatial weight matrices for the econometric analysis.

5.4 Panel Construction Procedure

We will construct the final analysis dataset through the following steps. First, we will process the facility-level investment data to create county-quarter aggregates of investment amounts, facility counts, and expected employment. We will define treatment variables including the quarter of first substantial investment and cumulative investment through each quarter. Second, we will merge Bureau of Labor Statistics employment data creating a balanced county-quarter panel and calculate growth rates and industry shares. Third, we will merge annual data from the American Community Survey and Bureau of Economic Analysis, forward-filling quarterly observations within each year. Fourth, we will merge time-invariant geographic characteristics including renewable energy potential, existing infrastructure, and Energy Community designation. Fifth, we will calculate spatial weight matrices and spatially lagged variables for the spatial econometric analysis. Finally, we will implement data quality

checks, verify merge accuracy, and export the final dataset in multiple formats for analysis.

6 Expected Contributions and Policy Implications

This research will make several contributions to understanding the regional economics of energy transitions and place-based industrial policy.

6.1 Scientific Contributions

The research will advance three literatures. First, it will contribute to understanding how public expenditures in energy markets affect regional and national economic aggregates, a question of direct relevance given the magnitude of IRA subsidies and similar policies being implemented internationally. While substantial research examines aggregate macroeconomic effects of energy price shocks, relatively little analyzes the regional economic consequences of place-based energy subsidies. The spatial dimension proves critical for understanding both the magnitude and distribution of impacts.

Second, the paper will contribute to the literature on place-based policies and regional development by providing rigorous evidence on whether clean energy manufacturing exhibits agglomeration economies and generates regional spillovers. If concentrated green investments create self-reinforcing development similar to Kline and Moretti [2014]'s findings for the Tennessee Valley Authority, this has important implications for optimal subsidy allocation during energy transitions. The research will identify specific channels through which agglomeration operates and quantify how rapidly spillover effects decay with distance.

Third, the findings will inform understanding of energy transition dynamics and policy design. Theoretical models of directed technical change [Acemoglu et al., 2012] typically abstract from spatial considerations and implementation details. By examining how the geographic distribution of investments affects aggregate outcomes and distributional impacts, the research will provide evidence relevant for comparing subsidy-based approaches to carbon

pricing and for designing policies that balance efficiency, equity, and political sustainability concerns.

6.2 Policy Implications

The research will generate several policy-relevant insights. First, the magnitude of direct employment effects and their persistence will inform cost-effectiveness assessments of place-based energy subsidies. If employment multipliers are large and sustained, this strengthens the case for subsidy-based approaches. If effects are small or temporary, this suggests alternative policy designs may prove more efficient.

Second, the spatial analysis will reveal whether benefits concentrate narrowly in direct recipients or extend to neighboring regions. If spillovers are substantial, this implies that the effective coverage of place-based policies exceeds their direct geographic targeting. If spillovers are minimal, this suggests benefits accrue narrowly with implications for equity and political support.

Third, the Energy Community analysis will assess whether geographic targeting successfully channels benefits toward vulnerable fossil-fuel-dependent communities. If the regression discontinuity estimates show large effects at eligibility thresholds, this validates the targeting mechanism. If effects are small, this suggests additional policy tools may be needed to support these communities during transitions.

Fourth, the heterogeneity analysis will identify which baseline conditions moderate treatment effects, informing optimal allocation strategies. If effects are larger in regions with existing manufacturing presence, this suggests concentrating investments in industrial clusters. If effects are larger in economically distressed areas, this suggests prioritizing equity over efficiency may not sacrifice aggregate impacts.

7 Research Timeline

The research plan will proceed according to the following timeline. During the period from November 2025 through January 2026, we will finalize data acquisition from all sources, complete merging procedures, and conduct comprehensive data quality checks. During February through March 2026, we will implement the five empirical specifications, conducting estimation and inference for each identification strategy. During April through May 2026, we will complete robustness checks, heterogeneity analysis, and mechanism investigations. We will prepare a complete draft for the virtual pre-conference in May 2026. Following feedback from the pre-conference, we will revise and extend the analysis during summer 2026, incorporating any additional data that becomes available. The final paper will be completed for presentation at the December 2026 conference.

8 Conclusion

This research proposal outlines a comprehensive empirical investigation of how place-based energy subsidies affect regional employment, investment, and economic growth during the energy transition. The Inflation Reduction Act provides an unprecedented opportunity to examine these questions using quasi-experimental variation in investment allocation across United States counties. By combining spatial economic theory with five complementary identification strategies and comprehensive administrative data, the research will provide robust evidence on the regional economic effects of green industrial policy. The findings will advance scientific understanding of energy transitions, place-based policies, and agglomeration economies while informing ongoing debates about optimal climate policy design. The spatial analysis will reveal whether benefits extend beyond direct recipients through regional spillovers, with important implications for assessing the efficiency, equity, and political sustainability of subsidy-based approaches to accelerating clean energy deployment.

Acknowledgements

The author declares no sources of research support and no financial relationships or other potential conflicts of interest that apply to this research.

References

Alberto Abadie. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425, 2021.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505, 2010.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015.

Daron Acemoglu, Philippe Aghion, Leonardo Bursztyn, and David Hemous. The environment and directed technical change. *American Economic Review*, 102(1):131–166, 2012.

Alan I Barreca, Jason M Lindo, and Glen R Waddell. Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, 54(1):268–293, 2016.

Eli Ben-Michael, Avi Feller, and Jesse Rothstein. The augmented synthetic control method. *Journal of the American Statistical Association*, 116(536):1789–1803, 2021.

John E Bistline, Neil R Mehrotra, and Catherine Wolfram. Economic implications of the climate provisions of the inflation reduction act. *Brookings Papers on Economic Activity*, 2023(1):77–182, 2023.

Kirill Borusyak, Peter Hull, and Xavier Jaravel. Quasi-experimental shift-share research designs. *Review of Economic Studies*, 89(1):181–213, 2022.

Brantly Callaway and Pedro HC Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946, 2014.

Timothy G Conley. Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45, 1999.

Timothy G Conley, Christian B Hansen, and Peter E Rossi. Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272, 2012.

Clément de Chaisemartin and Xavier D'Haultfœuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996, 2020.

James Feyrer, Erin T Mansur, and Bruce Sacerdote. Us oil production and the macroeconomy: The incredible oil independence boom. *American Economic Journal: Macroeconomics*, 9(4):196–232, 2017.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624, 2020.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

James D Hamilton. Oil and the macroeconomy since world war ii. *Journal of Political Economy*, 91(2):228–248, 1983.

Gordon H Hanson. Local labor market impacts of the energy transition: Prospects and policies. NBER Working Paper 30871, National Bureau of Economic Research, Cambridge, MA, 2023.

Ashok Kaul, Stefan Klößner, Gregor Pfeifer, and Manuel Schieler. Synthetic control methods: Never use all pre-intervention outcomes together with covariates. *Working Paper*, 2015.

Lutz Kilian. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, 99(3):1053–1069, 2009.

Patrick Kline and Enrico Moretti. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, 129(1):275–331, 2014.

John Larsen, Hannah Pitt, Trevor Houser, Hannah Kolus, Whitney Herndon, and Peter Kolus. Clean investment monitor: Tracking clean energy investment in america. Technical report, Rhodium Group and MIT Center for Energy and Environmental Policy Research, 2023.

John Larsen, Hannah Pitt, Kate Larsen, Whitney Herndon, and Peter Kolus. Clean investment monitor: 2024 update. Technical report, Rhodium Group and MIT Center for Energy and Environmental Policy Research, 2024.

David S Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, 2010.

Lung-fei Lee and Jihai Yu. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154(2):165–185, 2010.

James P LeSage and R Kelley Pace. *Introduction to Spatial Econometrics*. Chapman and Hall/CRC, Boca Raton, FL, 2009.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

David Neumark and Helen Simpson. Place-based policies. In Gilles Duranton, J Vernon Henderson, and William C Strange, editors, *Handbook of Regional and Urban Economics*, volume 5, pages 1197–1287. Elsevier, 2015.

Daniel Raimi and Madison Pesek. Defining energy communities for the inflation reduction act. Issue Brief 22-16, Resources for the Future, Washington, DC, 2022.

Stuart S Rosenthal and William C Strange. Evidence on the nature and sources of agglomeration economies. *Handbook of Regional and Urban Economics*, 4:2119–2171, 2004.

Jonathan Roth. Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322, 2022.

James H Stock and Motohiro Yogo. Testing for weak instruments in linear iv regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108, 2005.

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.