Distributional Equity in the Employment and Wage Impacts of Energy Transitions

Ben Gilbert, Ben Hoen, Hannah Gagarin

Abstract: We use restricted-access, geocoded data on the near-universe of workers in 23 US states to quantify the impact of wind energy development on local earnings and employment, by race, ethnicity, sex, and educational attainment. We find significant impacts that persist for several years beyond the project construction phase. Our estimates are larger than those from previous studies but still small relative to typical economic multipliers for fiscal spending or investment in other industries. We find the largest percentage increases for black workers and workers who either do not have a high school diploma or who have a college degree. We also find the economic gains for men to be much larger than those for women. Finally, we find estimates from data aggregated to the county level to be significantly lower than our worker-level estimates. We suggest a number of areas for further study building off the justice implications of our findings.

JEL Codes: J4, J21, Q4, R11

Keywords: wind power, equity, energy transition, employment, income

CONSIDERING THE CURRENT BIDEN ADMINISTRATION'S climate goals, which include a net-zero carbon pollution–free power sector by 2030 and a net-zero carbon pollution–free economy by 2050, the United States is poised to undergo an energy

Ben Gilbert is in the Department of Economics and Business and Faculty fellow at Payne Institute for Public Policy, Colorado School of Mines (bgilbert@mines.edu). Ben Hoen is at Lawrence Berkeley National Laboratory, Berkeley, CA (bhoen@lbl.gov). Hannah Gagarin is at Sandia National Laboratory, Albuquerque, NM (hrgagar@sandia.gov). Any views expressed are those of the authors and not those of the US Census Bureau. The Census Bureau's Disclosure

Dataverse data: https://doi.org/10.7910/DVN/UKISED

Received August 15, 2023; Accepted June 12, 2024.

Correction: This article was reposted on November 4, 2024, to change the Dataverse data link.

Journal of the Association of Environmental and Resource Economists, volume 11, number S1, November 2024.

https://doi.org/10.1086/732186

transition away from fossil fuels and toward renewable energy technologies. This has the potential to change the landscape of local labor markets as large-scale renewable energy technologies are deployed in communities across the United States. Renewable energy development may demand different skills, and occur in different places, than legacy energy sources. Renewable energy can also change the composition of the local economy through rent and royalty payments to landowners, tax payments to local jurisdictions and provision of public services, or local demand from manufacturing facilities for renewable energy components (Kim 2019; Brunner, Hoen, and Hyman 2022). Some scholars and policymakers raise concerns, however, that these investments still occur within the same political and cultural contexts that historically create an uneven distribution of economic welfare, along racial, ethnic, gender, educational, or other socioeconomic dimensions, which has the potential to continue marginalizing vulnerable or underrepresented populations (Mueller and Brooks 2020; Mejía-Montero et al. 2021). The concept of incorporating a justice element into existing and future US energy policy is often referred to as a "just transition" and lies at the heart of recent social science research on environmental and energy justice (EEJ).[1]

In this study, we use worker- and county-level data to investigate how large are the local employment and earnings gains for workers from wind energy development and to whom these gains accrue in terms of different worker subpopulations by race,

---

Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC project no. 1947 (CBDRB-FY22-P1947-R9663). This analysis was funded by the Wind Energy Technologies Office (WETO) of the US Department of Energy (DOE) under contract no. DE-AC02-05CH11231. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation grants, SES-9978093, SES-0339191, and ITR-0427889; National Institute on Aging gGrant AG018854; and grants from the Alfred P. Sloan Foundation. We also thank the Rocky Mountain Research Data Center staff and in particular Philip Pendergast, Kas McLean, and Carlos Becerra for their support. We are grateful for detailed and helpful comments from Johannes Schmieder and all of the participants at the National Bureau of Economic Research conference and preconference Workshop on Distributional Impacts of New Energy Policies, which benefited from funding from the Alfred P. Sloan Foundation. We could not have performed this work without the support of the DOE WETO, specifically Patrick Gilman, Rin Ball, and Maggie Yancey.

1. Environmental justice and energy justice are distinct but overlapping concepts that together provide conceptual, analytical, and decision-making frameworks for pursuing a sustainable energy system that redistributes welfare to avoid undue burdens on marginalized communities (McCauley et al. 2019; Carley and Konisky 2020). We use the combined acronym "EEJ" because of the acute interest in how energy transition policy will affect both environmental and equity outcomes.

ethnicity, sex, and educational attainment. We estimate impacts separately for black, American Indian/Native Alaskan, and white workers; for Hispanic workers; for male and female workers; and for workers in four educational attainment brackets. We view this as an important contribution to both economics and EEJ because such estimates do not exist in the literature at the granularity we are able to provide. Although EEJ concepts are much broader than jobs and income, distribution is an important pillar of EEJ, and, further, energy transition policies are likely to have distributional consequences through labor markets. As a secondary research question, we also ask: is there exclusivity in the research community regarding the ability to answer these questions? Specifically, we illustrate differences in the magnitude and precision of estimates using county-level data, such as are publicly available, compared to estimates using geocoded worker-level data such as are only available in restricted-access settings for researchers with the institutional capacity or connections required to obtain such data.

In order to answer these questions, we combine geocoded data on the near-universe of workers in 23 US states from 2000 to 2020 with geocoded data on all US wind projects. We estimate the causal effect of the arrival of utility-scale wind capacity within 20 miles of a worker's residence on that worker's earnings and employment status, controlling for wind capacity in place at greater distances. Specifically, we use restricted-access worker-level data provided by the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) dataset for 23 participating states, which contains worker residence coordinates, employment status, and earnings as well as age, sex, race, ethnicity, and educational attainment for all workers who paid into, or received benefits from, their state's unemployment insurance program—more than 96% of workers. We combine this with the US Wind Turbine Database, which contains the coordinates, arrival year, and capacity rating of wind plants.

In order to deal with the computational needs for a dataset of this size, we employ two separate empirical approaches. We first implement a unique shift-share instrumental variables (IV) design within a spatial lag model using a subset of the data. This approach estimates the effect of wind capacity in increasing 20-mile donut-distances from each residence on that worker's outcomes. Wind development near an individual may be correlated with local economic shocks that are not picked up in worker-by-county or state-by-year fixed effects, such as changes in zoning or land-use rules, expansion/contraction in other local industries, or changes in agricultural productivity and land rents.[2] We instrument for capacities in each donut using predicted capacities that come from a shift-share design. In this design, we predict wind capacity in localized hexagonal grid cells using local average wind speeds interacted with national aggregate wind capacity expansion trends and global prices of metals and energy commodities. We then aggregate these localized predictions to donut-distances around each worker's

2. We thank an anonymous referee for pointing out the potential agricultural markets interactions.

residence to construct the IVs for actual capacity in each donut. Results from this exercise indicate that the majority of the employment and earnings impacts are captured by wind development within the first 20-mile ring, with much smaller coefficients and larger standard errors at greater distances. However, this approach is very computationally intensive on a large geocoded dataset.

In order to reduce computation time and draw inference from the full dataset, we then use a brand new difference-in-differences (diff-in-diff) estimator: local projections difference-in-differences (LPDID) (Dube et al. 2023). The LPDID estimator is a regression-based approach that avoids the biases associated with two-way fixed effects in the presence of staggered treatment and heterogeneous, dynamic treatment effects and provides computational advantages over recent alternative estimators (Cengiz et al. 2019; Callaway and Sant'Anna 2021; Sun and Abraham 2021) while allowing for continuous treatment variables. Having established in the initial shift-share approach that most impacts occur within 20 miles, we use the 20-mile capacity as our "treatment" variable in the diff-in-diff approach and control for capacity at greater distances as additional regressors. Despite the computational advantages of LPDID, estimation on the full dataset is still not feasible. Because we have the near-population of workers, however, we repeatedly randomly sample 1 million unique workers (with replacement) and reestimate our models 100 times on each random sample in order to report the "true" variation in coefficient estimates rather than using analytical standard errors calculated from a single sample.

We find that wind power development within 20 miles of a given worker causes a statistically and economically significant sustained increase in earnings and employment. These impacts vary meaningfully across subpopulations. Black workers experience the largest proportional marginal impact despite there being few black workers within 20 miles of a wind project. Men enjoy larger gains than women. Additionally, workers without high school completion have the largest proportional gains among the four educational categories, followed by workers with a college degree.

With causal estimates of the effect of wind capacity on each subpopulation at the geocoded worker level in hand, we then aggregate the worker-level data to county-level averages and reestimate analogous models on the aggregate data.[3] We find that estimates using county-level aggregates are dramatically attenuated compared to the analogous estimates from worker-level data, especially for earnings. While earnings impacts are statistically and economically significant for all subpopulations in the worker-level data, our estimates are effectively zero using the same data aggregated to the county level. The degree of attenuation is also not uniform across subpopulations, nor does it vary in predictable ways. For example, the employment impact for workers with some

3. Aggregating to the county level vs. aggregating to the county plus any other county with a centroid within 20 miles produces very similar effect estimates, so we report county-level estimates here as a comparison to the worker level plus 20 miles results.

college coursework is 16% as large using county-level versus worker-level data, whereas the impacts for workers with a bachelor's degree are 60% as large, and the impact estimates for Hispanic/Latino and American Indian/Native Alaskan workers is actually larger using county-level data.[4] We further discuss potential explanations for these differences in the paper. These differences suggest that researchers from lower-resourced institutions, without access to such data resources, face additional barriers to fully understanding the employment impacts of new energy investments that may be important to the populations that their institutions serve.

Our findings contribute to a growing literature on the local economic and equity implications of large-scale energy transitions. The impacts that we estimate are far larger than can be explained by maintenance workers at the wind plant alone. Our findings confirm other recent studies showing larger local impacts (Brunner and Schwegman 2022; Gilbert et al. 2023) than earlier studies (Brown et al. 2012). However, while our employment estimates are in line with these recent studies, our earnings estimates are much larger and our evidence suggests that using county-level data understates earnings impacts by a far wider margin than employment impacts. The magnitudes of these effects suggest that the majority of local economic impacts occur through indirect channels. First, wind projects pay royalties to local landowners that can either be fixed annual payments or shares of the annual wind revenue. If these landowners reside locally and spend money in the community, this generates a local economic impact multiplier. Second, wind installations contribute to the local tax base, raising revenue for school districts (Brunner, Hoen, and Hyman 2022), counties (Castleberry and Greene 2017), and other local jurisdictions and community services (Shoeib et al. 2021). Expenditures by these jurisdictions on local services also generate economic impact multipliers. Third, the infusion of capital expenditures during the construction phase of a wind project may stimulate the opening of new establishments, or investments in existing establishments, that may persist when the construction phase ends, such as manufacturers of energy components (Kim 2019). Although our coefficients are larger than previous wind energy literature, when we convert these coefficients to local economic multipliers they are still modest in size compared to central estimates in the literature on earnings and employment multipliers for fiscal stimulus or industry investment from sectors other than wind energy. Back-of-the-envelope calculations suggest an employment multiplier of 0.51 jobs per million dollars of wind capacity investment, whereas employment multipliers in most other sectors are more than an order of magnitude larger. Our calculated earnings multiplier is approximately 0.16 dollars of local worker earnings per dollar of wind capacity investment, which is about one-fifth or one-sixth as large as central estimates of fiscal stimulus multipliers. Because direct employment at operating wind projects is very small, these

4. This is in line with recent work by Colmer et al. (2021), who find that county-level estimates of air pollution impacts on health may obscure some important within-county heterogeneity in responses.

are likely to be almost entirely indirect multiplier effects. The size of these multipliers will also depend on what share of wind project development costs are capital expenditures and what share is spent locally.[5]

We also contribute to the burgeoning field of EEJ. EEJ examines how systemic injustices embedded in societal and cultural norms can persist in energy transitions and addresses how to approach solving this from a policy perspective, to ensure a more "just" energy transition. Access to good-paying jobs is an important component of this. However, there is currently a knowledge gap in the EEJ literature in terms of being able to carefully measure impacts of energy development on disadvantaged subpopulations at a granular level. This study attempts to alleviate this gap by obtaining causal estimates from geocoded worker-level data on disadvantaged subpopulations as well as illustrating the challenges that researchers typically face in measuring impacts on specific groups using county-level data.

The remainder of this paper is organized as follows. Section 1 reviews the literature on regional economic impacts from energy in relation to the EEJ literature. Section 2 describes the data used in the analyses, which inform the empirical methods we use described in section 3. Section 4 presents and discusses the results, while section 5 concludes.

1. LITERATURE REVIEW

Often EEJ focuses on the ways in which underrepresented communities are harmed in the context of energy systems. Energy policy and investment may limit or exacerbate these harms and/or redistribute costs and benefits of energy systems. Within the context of a US transition to a low-carbon-energy sector, renewable energy can often be portrayed as a more socially just form of energy due to decarbonization (Mejía-Montero et al. 2021); however, Mueller and Brooks (2020) point out that there can still be an uneven distribution along social dimensions such as race/ethnicity and other marginalized populations. While an energy system might transition, the cultural and societal practices and norms that existed in the "old" energy system (i.e., fossil based) still remain ingrained, maintaining the continued marginalization and underrepresentation of communities in the "new" energy system. This could be true along many dimensions of equity and justice, including hiring practices and wage setting. Understanding how renewable energy investments have altered local labor markets and, in particular, the position of disadvantaged subpopulations within local labor markets, is therefore important for evaluating the EEJ dimensions of the energy transition.

There is a wealth of literature studying the impact of energy development on economic outcomes in local communities. Focusing on wind project development and economic outcomes alone, both Brown et al. (2012) and Varela-Vázquez and Sanchez-Carreira

5. We thank an anonymous referee for pointing this out.

(2015) estimate positive economic impacts to local communities. However, these works focus on average impacts without delving into distribution.

Studies on wind energy development in the EEJ literature tend to focus on public acceptance of wind energy or what is referred to as a "social gap": where there is widespread public support but localized opposition (Jones and Eiser 2010; Fergen and Jacquet 2016; Hoen et al. 2019; Mills et al. 2019; Acheampong, Erdiaw-Kwasie, and Abunyewah 2021; Shoeib et al. 2021). There are some works that focus not just on the overall acceptance and public perception of wind projects but on the outcomes of underrepresented communities located near wind projects. Some authors argue that siting wind projects in areas where residents have lower social capital, financial capital, and labor force participation imposes costs and burdens on historically marginalized communities and exacerbates energy injustice (Mueller and Brooks 2020; Acheampong, Dzator, and Shahbaz 2021). To the extent that wind development generates local employment, income, tax revenue, and community services, however, the benefits may exceed the costs. Shoeib et al. (2021), for example, find that wind development in rural areas increases public services provision without increasing the cost of living or the overall demand on public services.

Shoeib et al. (2022), however, argue that larger, more urban counties may be better able to internalize the potential economic benefits from energy development, in that they are able to provide specialized services such as banking and legal and secondary inputs such as skilled workers and raw inputs. Shoeib et al. (2022) also find that wind development only increases farm income in rural areas, whereas earnings, employment, and poverty improve on average when studying all counties. Mauritzen (2020) supports this idea of uneven impacts between rural and urban areas, showing that rural areas characterized by low incomes are least likely to capitalize on positive economic benefits of wind power development, compared to more urban, metro areas. Indeed, Pedden (2006) also demonstrates this, with results showing that smaller communities see more leakage into nearby areas that are better able to provide more services. Mauritzen (2020) also finds, however, that wages rise in rural areas, arguing that the likely mechanism is from landowner lease payments and local tax revenue. While the rural/urban divide in local economic benefits of renewable energy has been studied extensively, there has been less attention paid to heterogeneous impacts along other dimensions of historical disadvantage, which our study provides.

A puzzle for studies such as ours that find sizable employment and earnings impacts is by what mechanism does the wind installation generate these local benefits? Hoen et al. (2015) find that home prices near wind projects are not negatively impacted, while Brunner and Schwegman (2022) find that wind development increases nearby home values. More recent studies find evidence for preconstruction proximal impacts that later fade after operation begins, however (Dong et al. 2023; Brunner et al. 2024). Brunner and Schwegman (2022) also estimate sizable positive impacts from wind using county-level data, including increases in GDP per capita, income per capita, and median household income, with employment shifts into construction and manufacturing jobs and out-of-farm employment. Castleberry and Greene (2017) and Shoeib et al. (2021) find that wind projects increase local tax revenue and availability of community services, which can generate and sustain employment. Brunner, Hoen, and Hyman (2022) find sizable increases in school-district-level revenue and expenditures after wind project installations.

2. DATA SOURCES AND CHARACTERISTICS

We combined several geocoded datasets in order to estimate the models described in section 3. We gathered latitude and longitude, year of operation, and nameplate capacity for all wind turbines in the United States over the study period from the US Wind Turbine Database (Rand et al. 2020; Hoen et al. 2021). Lawrence Berkeley National Laboratory provided project-specific ID numbers to group turbines into wind projects (collections of turbines analogous to a power plant).

We use restricted-access granular data on workers from the US Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) database, which includes quarterly data on earnings and employment status for individual workers and their geocoded residences.[6] The LEHD is released in multiyear "snapshots." This study uses the 2014 snapshot, with data from 2000 through 2014, as well as the recently released 2021 snapshot with data from 2000 through 2020. We worked with the US Census Bureau's Federal Statistical Research Data Center (FSRDC) program to access the LEHD infrastructure files from within a secure Census data center facility.

The LEHD is built on quarterly state-level unemployment insurance rolls, which cover more than 96% of workers who reside in the United States. We aggregate to the annual frequency for analysis. Because the LEHD is constructed from state-level unemployment insurance programs, state-by-state approval to access the data is required. The 23 states shown in light gray in figure 1 agreed to participate in our approved project. As this figure indicates, although we do not have data for Texas or Minnesota, we do have access to states with a significant share of wind capacity. For this study, we used the LEHD's Individual Characteristics File, which indicates race, ethnicity, sex, and educational attainment of individual workers. The data do not indicate whether a worker identifies with multiple races or ethnicities or is gender nonbinary. This is a limitation of our study. We also used the Employment History Files, which includes a worker's quarterly earnings from all jobs held in that quarter. We aggregated earnings to the annual level. We calculated our employment variable as the number of quarters in each year in which the worker had positive earnings, divided by 4. We also

6. The LEHD infrastructure provides the inputs required to produce the quarterly workforce indicators (QWI) (Abowd et al. 2009). Researchers with either paid access or institutional access to a Federal Statistical Research Data Center can propose research questions using the underlying restricted-access individual-worker-level and establishment-level data.

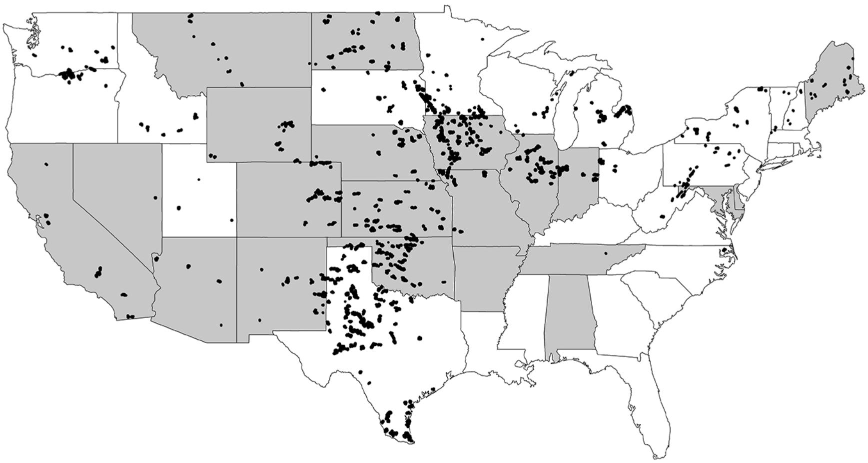

Figure 1. US states with wind projects. States highlighted in light gray are the states examined in this study; the black dots represent US wind project installations. Source: Graphic created by author B. Gilbert based on wind turbine location data provided by Hoen et al. (2021).

used the ICF LEHD Residence Files, which contain latitude and longitude coordinates for the worker's residence in each year.[7]

For this study, we report shift-share IV results using the 2014 LEHD snapshot with data from 2000 to 2014, and diff-in-diff results using the recently released 2021 LEHD snapshot, which includes employment data through 2021, but residence data through only 2020. Our preferred results are the diff-in-diff estimates from the 2021 snapshot, but we use our shift-share IV results from the earlier snapshot in order to guide our estimation approach as we discuss in section 3. Because of migration that occurred during the COVID-19 pandemic, we were not confident in imputing residence locations for 2021. We therefore limit our analysis to the years 2000 through 2020.

In each of these datasets we use georeferenced locations (i.e., latitude and longitude) of both worker residences or centroids of wind projects in order to find the aggregate wind capacity at various distance bands from each worker residence. This produced an annual worker-level panel with employment and earnings outcomes, demographic characteristics, and wind capacity exposure at various distances. Summary statistics

7. To the extent that workers live in one place and periodically travel to distant work sites with energy installations, our approach would not pick these people up, although our sense is that this is more common in the oil and gas industry than in the wind industry. One could possibly look for workers in the LEHD whose establishments of employment are farther from their residence than a reasonable commuting distance, but this is not straightfoward and is beyond the scope of this study.

for this combined panel are given in table 1 of the appendix (appendix is available online).

Any statistics or model estimates that are publicly released from the LEHD must undergo strict disclosure avoidance procedures by the US Census Bureau in order to protect privacy. As part of this process, all statistics, including summary statistics, coefficients, standard errors, and so forth., conform to the US Census Bureau's rounding rules for disclosure avoidance.

3. METHODS

Worker-level residential location choice and labor supply decisions may be endogenous with respect to determinants of local wind energy development because of location preferences, local economic shocks, or state and local policies affecting local labor markets. For example, municipal or county changes in zoning or land-use policies, local business development in related sectors, and changes in agricultural productivity or land rents may all affect the local attractiveness of wind development and may not be picked up in worker-by-county or state-by-year fixed effects. If these changes in local economic activity make wind development more or less attractive, our wind capacity coefficients would be picking up these effects rather than isolating the effect of wind projects. We therefore need a careful identification strategy to recover the causal effect of wind development on local labor outcomes.

In addition to these potential endogeneity concerns, our empirical approach needs to overcome two more important challenges. First is the computational limits on performing many geospatial calculations of individual worker distances to wind projects and estimating multiple regression models with over a billion observations. Second, it is not obvious who belongs in the "treatment" versus "control" groups given previously documented spatial spillovers of energy development on labor markets (Feyrer et al. 2017; James and Smith 2020).

It is not straightforward to classify worker residences as "wind" versus "non-wind" households as other individual-level studies have done (Jacobsen et al. 2023) because of this second concern. As figure 1 indicates, wind projects are spatially clustered. Many workers may live within a short distance of more than one wind project, so assigning them as "treated" by their closest wind project will understate exposure. This is one reason we aggregate capacity in distances around worker residences, despite the computational intensity of doing so.[8] One option to determine treatment status might be

8. Another benefit of aggregating capacity around residences, using the residence as the unit of analysis, rather than using the wind project as the unit of analysis and including all nearby households in the "treatment group" is that we can include capacity that may be just across state borders in states for which we do not have worker data but which may still impact workers that we do observe. For example, plants in southern Minnesota may affect workers from northern Iowa who are in our sample.

to select a certain distance cutoff beyond which we assume exposure is zero. In the context of oil and gas, previous studies have found economically and statistically significant impacts as far as 60–100 miles away (Feyrer et al. 2017; James and Smith 2020). A 100-mile buffer around all wind projects leaves very few areas from which to select a control group (as illustrated in appendix fig. 1). In order to deal with both the endogeneity issues and the spillover concerns with treatment assignment, we first estimate a spatial lag model à la James and Smith (2020), but with a shift-share IV at each spatial lag that is akin to Feyrer et al. (2017). This approach allows us to establish the extent of spillover as a guide to determining treatment status. However, the approach is incredibly computationally intensive, and we are only able to implement it on a small random subset of the full dataset.

As we will show, this exercise suggests that most impacts are captured within the first 20 miles. Impacts at greater distances are much smaller, but not zero. We therefore proceed with a more computationally feasible estimation approach that uses capacity within 20 miles as the main treatment variable, while still controlling for capacity at greater distances and that leverages the entire dataset.

The solution that we implement, and the source of our preferred results, is the local projections difference-in-differences (LPDID) framework recently proposed by Dube et al. (2023), described in detail below. In order to deal with remaining computational constraints and capture the sampling uncertainty in our estimates, we reestimate our LPDID models 100 times using random draws of worker IDs from the near-population in our dataset. Finally, we aggregate the full dataset, as well as relevant subpopulations, to the county level and reestimate the LPDID coefficients for comparison.

3.1. Shift-Share IV

We estimate our spatial lag regressions with shift-share instrumental variables on a 0.1% random sample of the 2014 LEHD snapshot dataset. The “share” component of the IV, or the exogenous cross-sectional variation, comes from average local wind speeds, whereas the “shift” component, or time series shock, comes from national aggregate trends in expansion of wind capacity, number of turbines, and number of wind plants, as well as commodity prices of crude oil, natural gas, aluminum, and a rare earth metals composite index. In a shift-share estimator, identification can come from either exogeneity of the “shares” (Goldsmith-Pinkham et al. 2020) or from many exogenous “shocks” or “shifts” (Borusyak et al. 2022). In our case, we argue that identification comes from both sources. Exogenous spatial variation in wind resource has been argued in Brown et al. (2012) and Brunner, Hoen, and Hyman (2022) as well as for other resources like oil and gas deposits in Feyrer et al. (2017) and Maniloff and Mastromonaco (2017). We augment this with exogenous time series shocks in the form of national or global trends in commodity prices that might affect local wind development and national wind development trends that capture the national policy and investment climate in a given year.

Specifically, we estimate the following spatial lag model:

$$Y_{ict} = \sum_{d \in \{20,40,60,80,100\}} \gamma_d D_{ict}^d + \alpha_{ic} + \mu_{st} + \epsilon_{ict}, \tag{1}$$

where D_{ict}^d is the amount of wind capacity in donut distance d from person i's residence in year t, with d ranging from 0 to 20 miles, 20 to 40, and so forth through 80 to 100 miles. We include worker-by-county fixed effects α_{ic} to isolate impacts on workers within a spell of living in a given community rather than confounding impacts with migration. We also include state-by-year fixed effects μ_{st} to control for state-level macroeconomic trends that influence employment and earnings but may be correlated with energy development in the region. The worker-by-county fixed effects also help reduce recent concerns about using individual fixed effects in long panels if unobserved characteristics may change over time (Millimet and Bellemare 2023). Working-age adults move about every five years on average (US Census Bureau 2021), which reduces the time span over which we must assume characteristics are fixed within a worker-county spell. Although many people move less often or move within their county, our shift-share IV recovers variation in the wind capacity treatment due to spatial variation in average wind speeds and temporal variation in global commodity prices and national wind energy expansion that is likely independent of slowly changing worker-level unobserved attributes.[9]

If secular changes in local economic activity like agricultural productivity and/or land rents are affecting where wind projects get built and how big they are, then identification relies on the assumption that, conditional on worker-by-county and state-by-year fixed effects, (*a*) the preexisting wind resource is uncorrelated with these localized (substate-level) secular agricultural market changes (and any other local economic shocks like changes to zoning rules that might affect whether/how much wind gets built) or that (*b*) global oil and metals prices and national wind capacity expansion are uncorrelated with these localized changes agricultural markets or other local economic shocks that move independently of state-by-year trends. These assumptions are plausible; variation in global prices and national trends is already captured in state-by-year fixed effects, and variation in wind speeds over space is fairly stable over time and should not affect sudden changes in land rents or agricultural productivity.

Alternatively, wind project development could affect agricultural productivity and land rents by affecting irrigation, planting, or crop growth, which in turn affect local jobs and earnings. That would be just fine in our framework, as our coefficients would measure the net effect of these indirect impacts along with all the other indirect impacts of wind development that we do not directly measure such as landowner royalty

9. In a replication analysis of James (2015), Millimet and Bellemare (2023) find that their IV results are also robust to potential temporal changes in supposedly fixed attributes.

payments leading to increased local spending, or local tax payments leading to increased government expenditures on public services.

We instrument for each D_{ict}^{d} using predicted capacity in each donut/ring around each person's residence. We construct these predicted capacity instruments using the following approach. We segment the lower 48 US states into approximately 216,000 evenly sized hexagons and gather data on the average wind speed within each hexagon. These hexagons are about the size of the average census block group. We used hexagons rather than other grid tessellations like squares or triangles because we wanted to aggregate the predicted wind capacity (our IV) at fine-scale spatial units into donuts around each worker residence in order to most closely match the second-stage spatial lag model (eq. [1]), which uses actual wind capacity in donuts around each residence. The donut circles are less likely to cut the hexagons at awkward angles than a shape with less obtuse angles like a square or a triangle. Each hexagon is also equidistant from all of its neighbors, unlike squares or triangles, making it preferable for aggregating among nearest neighbors (Birch et al. 2007).

Using the geocoded US Wind Turbine Database, we then calculate the total number and capacity of turbines in each hexagon in each year. The vast majority of hexagons have zero wind capacity, so we use a fixed effects Poisson quasi-maximum likelihood regression (Wooldridge 1999) to predict total wind capacity and number of turbines in each hexagon in each year.[10] The right-hand side of the fixed effects Poisson regression includes hexagon fixed effects, state-by-year fixed effects, and a cubic polynomial in average wind speed interacted with the total national US wind capacity in each year, the total number of wind turbines in the United States in that year, the number of US wind plants in that year, and the vector of commodity prices. We use a cubic function of wind speed in order to capture the fact that wind development is not feasible at either very low or very high average speeds.[11] The variation in predicted hexagon-year wind development is then driven by unobserved hexagon-specific effects such as suitability for transmission access, secular state-by-year macroeconomic or policy trends such as state-level renewable energy incentives or mandates, and a nonlinear shift-share component driven by the interaction of wind speed with national aggregate investment behavior, essentially allocating nationwide investment in each year to the locations that are most suitable for development according to wind speed.

A problem with the fixed effects Poisson approach with hexagon fixed effects is that the predicted value in any hexagon that never has wind capacity—the vast majority of them—is zero. As such, the predicted values are extremely highly correlated with the

10. This estimator is also known as the Poisson pseudo-maximum likelihood estimator (Cameron and Trivedi 1986).

11. The physics and engineering literature also uses a cubic function of wind speed to model wind power output (Manwell et al. 2010; Pryor et al. 2020). In the economics literature, Cicala (2022) uses a cubic function of wind speed to impute missing wind generation observations.

actual values. Aggregating these predicted values around worker residences in order to construct instruments for equation (1) thus produces instruments that essentially reproduce the ordinary least squares (OLS) estimates.

In order to deal with this issue, we further aggregate both actual wind capacity from the US Wind Turbine Database and predicted wind capacity from the Poisson regression on each hexagon, to the county-year level. We also calculate average wind speed in each county. We then predict capacity in each county-year using a linear OLS-FE regression with county fixed effects (FE), state-by-year fixed effects, and interactions between the county-average wind speeds and the sum of hexagon-level predictions within each county. This procedure generates county-year variation in predicted wind capacity. Finally, we use county centroids to aggregate the county-year predicted capacity values in donut distances around each worker's residence. This produces a nonlinear transformation of unique fixed county and hexagon characteristics, and state-level macroeconomic trends, and shift-share variation from local wind speed and national aggregate wind development trends. These county-level predictions aggregated to 20-mile donut-distances from worker residences to surrounding county centroids become our IVs for equation (1) in a standard just-identified linear IV framework.[12]

We perform this shift-share estimation approach for equation (1) on both a 0.1% random sample and on county-aggregated data for comparison.

3.2. Shift-Share IV Results

Here we report results from the estimation of equation (1) using a 0.1% sample of the worker data from 2000 through 2014 (i.e., using the 2014 snapshot of the LEHD). We cluster standard errors at the worker level in worker-level regressions in order to adjust for possible autocorrelation in worker outcomes that may persist even when workers move across counties. We cluster at the county level in county-level regressions in order to adjust for possible autocorrelation within counties in aggregate worker outcomes.

In table 1 we can see that at the worker level (cols. 1 and 3) the magnitude of the point estimate for capacity declines steeply at distances greater than 20 miles. This provides confidence that results for the 20-mile ring reported are capturing most of the relevant impacts. By comparing the first row across all four columns of table 1 we also see that within 20 miles the difference between the county-level estimate and the worker-level estimate is much greater for log earnings than for employment. We will see a similar general pattern when we compare individual and county-level estimates using the LPDID estimator on the data through 2020.

Although these are not our preferred estimates, the magnitudes are also comparable to the LPDID results we will see in section 4, despite the different sample and

12. It may be possible to estimate this in a nonlinear approach in one step, but computational limitations make the linear IV framework more feasible to implement.

Table 1. Spatial Lag Coefficients (GW within Each 20 Mile Donut)

	Outcome			
	Employment		Log Earnings	
	Worker Level (1)	County Level (2)	Worker Level (3)	County Level (4)
0–20 miles	2.6	3.8	20	4.1
	(2.1)	(3.3)	(12)	(2.3)
20–40 miles	−.66	1.8	1.7	2.2
	(.37)	(1.7)	(2.4)	(.84)
40–60 miles	−.33	2.1	.12	1.2
	(.15)	(1.0)	(.96)	(.45)
60–80 miles	−.14	1.4	−.27	.75
	(.086)	(.71)	(.57)	(.31)
80–100 miles	−.05	1.0	−.075	.51
	(.052)	(.53)	(.34)	(.22)
R^2	.20	.22	.18	.90
K-P Wald F-statistic	265	2,110	265	2,110
$N \cdot T$	1,438,000	29,750	1,438,000	29,750

Note. This table reports the spatial lag coefficients from eq. (1) using the shift-share IV procedure described in sec. 3.1. The dependent variables are the percentage of the year employed (quarters with nonzero earnings divided by 4) and the log of earnings. Earnings regressions limit the sample to employed workers (those with at least two quarters of nonzero earnings in a year). The treatment variables are gigawatts (GW) of capacity within each 20-mile donut distance of a worker's residence. We used a 0.1% random sample of workers from the 2014 snapshot of the LEHD, which includes the years 2000–2014. Standard errors are clustered at the worker level in cols. 1 and 3 and at the county level in cols. 2 and 4.

estimation method. The first row in column 1 of table 1 indicates that an additional 100 megawatts (MW) of wind capacity within 20 miles causes employment of the average worker to increase by 0.26 percentage points. The comparable effect using LPDID on the 2021 LEHD snapshot is 0.13 percentage points from a 100 MW increase, taken from the first column of panel A in appendix table 2. Similarly, the first row of column 3 of table 1 shows that an additional 100 MW of nearby wind capacity increases earnings by 2%. The comparable estimate of the continuous treatment effect using LPDID on the more recent data is in the first column of panel A of appendix table 3, which corresponds to a 1.2% increase from 100 MW of wind.

3.3. Local Projections Diff-in-Diff

It is not computationally feasible to estimate the shift-share IV approach described in section 3.1 on the full 2021 LEHD snapshot or even to downsample and perform the geospatial calculations repeatedly on large random subsamples. We therefore turn to the LPDID method for our preferred results.

Under LPDID, the estimation of a diff-in-diff event study is analogous to the estimation of an impulse response function by local projections in a time series analysis. One estimates a sequence of long-difference regressions with increasingly long differences in the outcome from the treatment date. Treatment coefficients from each regression in the sequence make up each element of the event study relative to the treatment date. In each regression in the sequence, one can obtain "clean controls" (control units that have not previously been treated) by limiting the sample to those units whose treatment status changed in the current period and those that have never been treated during the time horizon of the given long difference.

For this approach, we assume that workers living within 20 miles of utility-scale wind capacity are "treated," and all other workers are not. We include capacity exposure at greater distances as control variables in our regressions. It is possible that some workers with a utility-scale wind plant 19 miles away may be in our treatment group while others with utility-scale wind 21 miles away are not, despite very similar exposure levels. However, this is a limitation of any study that chooses a hard spatial cutoff for treatment exposure. We deal with this by using the regression to explicitly control for the exposure that any worker who is just outside the edge of our treatment group might experience.

We pull from the entire dataset outside 20 miles for our control group. While some authors argue that the control group should first be narrowed using propensity score matching or inverse propensity weighting, we are comfortable with our approach for several reasons. First, most workers have at least some level of exposure to wind even if they do not live within the 20-mile band, as we can see in table 1 and appendix figure 1, even if the exposure is small. Second, wind projects have a mix of urban, suburban, and rural exposure as can be seen in figure 1 with wind projects near Chicago, Los Angeles, and Denver as well as small-to-medium-sized cities like Des Moine and Omaha. Third, the observable characteristics between treated and control groups are not incredibly different even in the near-population (appendix table 1). Finally, although our randomly selected control workers might have systematically different characteristics in levels from the treated workers, the parallel trends assumption is strongly satisfied in all but a few subsamples, as we will see in the results section.

Specifically, our model for outcomes is

$$Y_{ict} = \gamma D_{ict} + X'_{ict}\beta + \alpha_{ic} + \mu_{st} + \epsilon_{ict}, \tag{2}$$

where Y_{ict} is either the fraction of the year that person i is employed in year t and county c, or the log of earnings. In logged earnings regressions we restrict the sample to workers having nonzero earnings in at least two quarters of the year, so coefficients are impacts on earnings conditional on being employed. In this case, γ is a semi-elasticity approximating the percentage change in earnings for each unit change in the treatment variable. The term D_{ict} is a treatment variable defined as either an indicator for the presence of at least 10 MW of wind capacity in year t within a 20-mile radius of person i's residence or as a continuous variable measuring the total capacity in gigawatts within 20 miles that year. Ten megawatts is roughly equivalent to the 5th percentile of project size during the study period and can be considered a minimum size threshold for a commercial wind project (Hoen et al. 2021; Brunner, Hoen, and Hyman 2022). The average exposure to wind capacity for treated workers within our sample is approximately 300 MW, so in regressions with continuous treatment we can interpret $\gamma \cdot 0.3$ as the semi-elasticity at the mean, from the addition of one average-sized wind project. The term X_{ict} is a vector of control variables. In our case, if wind development is spatially autocorrelated and there are regional economic spillovers, we would need to control for wind capacity at greater distances than 20 miles. The term X_{ict} is the spatial lag of wind capacity in increasing 20-mile donut distance bins from person i's residence in year t, for example, capacity 20–40 miles away, 40–60, and so forth, out to 100 miles. As it is challenging to causally identify multiple spatial lags within the same diff-in-diff framework, we do not report these coefficients or give them a causal interpretation but rather control for them in order to causally identify γ. The term α_{ic} is a worker-by-county fixed effect, capturing time-invariant characteristics of the worker-place combination, such as the worker's productivity within a particular set of local work opportunities. Finally, μ_{st} is a state-by-year fixed effect capturing regional macroeconomic trends at the state level that may be correlated with both worker outcomes and wind energy development. With both sets of fixed effects, variation in the treatment variable comes from changes in wind development near a person's residence, during spells between major cross-county or cross-state moves, that occur independently of state-level macroeconomic trends.

In order to estimate equation (2) by LPDID, we take successively long differences of (2) and estimate them one at a time. Specifically, using the full near-population dataset and each subpopulation of interest, we estimate

$$Y_{ic,t+h} - Y_{ict} = \delta_h \Delta D_{ict} + \Delta X'_{ict}\beta^h + \Delta\mu^h_{st} + \Delta\epsilon^h_{ict} \tag{3}$$

sequentially for different values of h, limiting the sample each time to newly treated ($\Delta D_{ict} > 0$) and not-yet-treated ($D_{ic,t+h} = 0$). The treatment coefficient δ_h in each regression is the event-study estimate for event time h. Averaging δ_h over the posttreatment periods gives an estimate of the average treatment effect, γ, from equation (2). Examining δ_h for negative values of h can help evaluate the parallel trends assumption, such as by plotting the individual coefficients on an event-study graph and/or testing the significance of the cumulative sum of δ_h over pretreatment periods. We report both approaches. Including temporal lags of Y_{ict} on the right-hand side of the regression can help control away preexisting differences in trends in order to obtain conditional parallel trends, while also controlling for potentially endogenous selection into treatment based on pretreatment outcomes. Parallel trends tests are satisfied in almost all of our worker-level specifications, so we do not include temporal lags of Y_{ict} in results reported here. Note that we have differenced out the worker-by-county fixed effect α_{ic}, which improves computation time, but this also means that each long difference is within a spell in which a worker lives in a particular county. Using separate worker fixed effects and county fixed effects, and differencing within worker (but not necessarily within county if a worker migrates) produces very similar estimates that we omit for the sake of brevity. Another advantage of our LPDID implementation is that the time horizon over which the α_{ic} must remain fixed is relatively short. Millimet and Bellemare (2023) propose a set of rolling difference estimators to address the problem of individual characteristics slowly changing in long panels. The successive "long" differences in each of our regressions are between one and seven years.[13]

This approach economizes on computational resources, making estimation on very large datasets feasible, for several reasons. First, a very large set of worker-by-county fixed effects is removed by differencing before estimation. Second, because each regression in the sequence limits the sample to clean controls, the "stacking" approach of Cengiz et al. (2019) is not required. Stacking involves appending panels of treated units with panels of units that are not treated within the same time window, such that each treated unit has a set of clean controls. This approach significantly magnifies the size of the analysis dataset, requiring additional memory and RAM. In our setting with 1.4 billion observations on the near-universe of workers in 23 states observed over 20 years, this approach is not feasible. Third, because this is a regression-based estimator, the computationally intensive task of averaging thousands of individual 2×2 pre-/posttreatment comparisons as in Callaway and Sant'Anna (2021) or estimating shares of cohort weights as in Sun and Abraham (2021) is not necessary. In fact, Dube et al. (2023) demonstrate that computation time for LPDID is comparable to two-way fixed effects estimation and much faster compared to Callaway and Sant'Anna (2021) and Sun and Abraham (2021), for which computation time is more than two orders of magnitude greater. This is a crucial benefit in our case using a massive restricted-access dataset. Further, LPDID achieves the same reduction in treatment effect bias as these alternative new estimators when treatment arrival is exogenous, while achieving even less biased results when selection into treatment depends on past outcomes.

We estimate the regressions in equation (3) using the entire dataset, then estimate them separately for black, American Indian/Native Alaskan, white, Hispanic, male, and female workers, as well as those without a high school education, with a high school education, some college coursework or an associates degree, and those with a bachelors degree or higher.

3.4. Monte Carlo Estimation

Even with the computational advantages of LPDID, estimating equation (3) on the full LEHD dataset is not computationally feasible. In order to solve this problem, we

13. A reader who is skeptical that characteristics remain fixed for seven years could focus on the event-study coefficients that are closer to the event date (which use shorter differences).

repeatedly randomly sampled (with replacement) 1 million unique worker IDs. For each random 1-million-worker sample, we estimated the sequence of LPDID regressions in (3) for all available time periods and retained the δ_h coefficients. We repeated this exercise 100 times and calculated the mean and standard deviation of coefficient estimates across the 100 draws. These coefficient means and standard deviations across draws are what we report in the results section below, rather than using analytical standard errors from a particular draw.

Our sampling approach for this Monte Carlo procedure requires some explanation. In order to deal with the fact that the full population and each subpopulation are unevenly geographically distributed across the United States, we randomly sampled the same number of people from each county such that the total number of unique workers would sum to 1 million. If too few workers in a given subpopulation resided in a given county, we took all workers from that subpopulation in that county in each draw and reallocated the remaining workers to other counties in order to reach 1 million workers in each draw. This means that subpopulations and counties with fewer workers are oversampled relative to their population share. This is by design.

We want to measure the effect of the average wind farm on nearby people rather than the effect of a wind farm on the average person. This is a key distinction. The reason for using approximately equal sample sizes from each county in each draw is that the "average" person in the United States lives in a city and does not have any wind capacity within 20 miles of their home. Using a true random sample (or even a random sample stratified in proportion to county populations) produces a sample that is heavily weighted toward urban dwellers who only live near wind capacity in very unusual situations that are not representative of the typical "treatment" of a local wind project. That sample would be representative of the US population, but it would not be representative of our treated population. Our treatment group would then be heavily skewed toward nonrepresentative urban-adjacent wind farms, rather than capturing typical effects of expansion in the wind industry. By sampling equal numbers from each county for the worker-level regressions, the counties are equally represented in terms of sample proportion in each draw. This is as close as we can get to an apples-to-apples comparison with the county-level regressions. We do not weight the county-level regressions with population shares, so that each location has an equal weight rather than giving more weight to more populous counties, and we similarly do not weight the worker-level regressions with sample weights for the same reason. This is also as close as possible to the thought experiment of what would happen if a wind developer could randomly allocate a wind project to a particular location rather than randomly drawing a person from the population and putting them near a wind project (this is also, in effect, the desired mechanism behind relying on the exogenous spatial distribution of natural resources as an IV).

In making these weighting decisions, we follow Solon et al. (2015), who thoroughly review the econometric theory literature in order to provide guidance on weighting for practitioners. Specifically, Solon et al. (2015) show that if the regression controls for characteristics that are also the basis for calculating the weights, then including weights will not improve consistency and may harm precision. Our individual-level and county-level models include county fixed effects (or worker-by-county fixed effects for the worker-level estimates), so according to Solon et al. (2015) there is no need to add weights that vary by county, such as county-level population or sample weights.

Solon et al. (2015) further argue that if the goal of weighting is to capture parameter heterogeneity (e.g., different effects for small vs. large counties) then using weights is a very constraining functional form to get at that heterogeneity. Practitioners should directly model the heterogeneity rather than using the constraining functional form of putting a weight in the denominator. Within the restricted-access US Census data center, we have estimated a large number of models, including those using pure and stratified random sampling, population and sample weights, county population interactions, and subsamples that drop the most populous counties from the analysis at various population thresholds. We chose to obtain Census approval for public release of, and to report here, the subset of results that are most representative of the effect of the average wind farm on nearby people, rather than the effect of a wind farm on the average person. Our findings are robust across a wide variety of specifications, however. Code to run all of our alternative specifications is available to anyone with LEHD access upon request from the authors.

3.5. County-Level Aggregation

We also wish to illustrate how the use of data aggregated to arbitrary and differently sized administrative boundaries such as counties can produce different impact estimates than individual data and investigate the extent of these differences for different subpopulations. In order to do so, we aggregate the individual-level outcomes within a county and reestimate the sequence of regressions from equation (3). In order to define analogous outcome variables at the county level, we calculate the average fraction of the year employed for all workers in each county-year, and we calculate the log of average earnings per worker who had nonzero earnings in at least two quarters of the year. We perform this calculation for the entire dataset and for each subpopulation of interest. We then modify equation (2) by using county fixed effects (rather than worker-by-county fixed effects). We define the treatment variable as either a binary indicator for whether or not a given county has at least 10 MW of wind capacity or the continuous number of gigawatts of wind capacity in the county.[14] We then estimate the regressions from equation (3) as well as the shift-share exercise for equation (1) on this

14. To be consistent with the 20-mile ring used in the individual-level regressions, we also estimated regressions in which we defined treatment by aggregating capacity in all counties with centroids within 20 miles of a given county's centroid. Results are not noticeably different—most counties do not have another county centroid within 20 miles of their own centroid.

county-level aggregate data. As discussed at length in the previous subsection, these county-level regressions are not weighted by population shares.

County-level results may differ from individual-level results for several reasons. The ecological fallacy literature suggests that aggregate and individual-level statistics are not in general the same (Robinson 1950). The composition of county populations changes over time. Covariances between county residents are captured in the aggregate coefficients but not in the individual-level coefficients. These issues could lead to differences in the aggregate and individual-level coefficients even if we were aggregating to a consistent group size and shape.

In addition, the irregular shapes and sizes of counties introduces (potentially nonclassical) measurement error in treatment exposure that is correlated with where wind resources exist and where wind energy is developed. Our shift-share IV based on spatial variation in average wind speeds is then less helpful with county-aggregate data because wind speed may be correlated with the location of irregular county shapes and sizes. This is closely related to an issue covered more extensively in the field of geography: the modifiable areal unit problem (MAUP). The choice of an administrative boundary (e.g., county border) for data aggregation is a "modifiable areal unit." The size and shape of the unit affects the correlation in aggregate data between units and therefore influences how much the aggregate relationships differ from individual relationships. Larger units will appear more correlated, and differently shaped units will lead to different correlations based on the same underlying individual data depending on how the unit borders are drawn. Unfortunately it is very hard to predict in advance for particular cases or make general statements about how aggregate estimates will differ from individual estimates (Openshaw 1984; Fotheringham and Wong 1991). The extent to which MAUP drives differences between individual and aggregate estimates depends on complex interactions between the underlying spatial distributions of individual data, the covariances of the specific variables being used, and specific sizes and shapes of the areal units being used.

4. RESULTS

We first report results for the average treatment effects at the individual worker and county aggregate levels. These are averages of the δ_b coefficients from equation (3) in the posttreatment period, which are estimates of the γ parameter from equation (2). We then discuss comparisons of these treatment effects between subpopulations and illustrate the dynamics of these effects through event-study graphs.

4.1. Average Effects

Tables 2 and 3 show the average treatment effects (average of event-study coefficients in the posttreatment period) for binary treatment on employment and log earnings, respectively. Appendix tables 2 and 3 report an analogous set of results for continuous wind capacity treatment. Panel A in each table shows the effects from the same model

Table 2. Average Treatment Effects (Binary > 10 MW within 20 Miles): Employment

	All	Black	Am. Ind	White	Hispanic	Female	Male	No High	High Sch.	Some Col.l	College
						A. Worker Level					
ATE	.42	.64	.40	.36	.45	.33	.46	.57	.31	.36	.48
	(.17)	(.38)	(.47)	(.16)	(.21)	(.15)	(.15)	(.20)	(.14)	(.14)	(.15)
Jobs	231	25	4	160	34	88	131	45	47	65	67
Cumulative pretrend	.0037	−.33	−1.2	−.00089	.32	.12	−.011	−.091	.16	.041	.019
	(.35)	(.46)	(.46)	(.30)	(.37)	(.32)	(.25)	(.31)	(.26)	(.30)	(.26)
						B. County Level					
ATE	.15	−.05	.49	.16	.45	.14	.15	.19	.14	.05	.28
	(.12)	(.99)	(.93)	(.12)	(.52)	(.13)	(.15)	(.24)	(.15)	(.14)	(.16)
	[−.09,.39]	[−2.0,1.9]	[−1.3,2.3]	[−.07,.38]	[−.57,1.5]	[−.11,.39]	[−.15,.46]	[−.28,.66]	[−.15,.43]	[−.22,.32]	[−.04,.60]
Cumulative pretrend	.22	−.62	−.63	.25	.66	.15	.29	.44	.14	.070	.63
	(.36)	(3.5)	(2.8)	(.34)	(1.7)	(.38)	(.43)	(.71)	(.44)	(.43)	(.58)

Note. This table reports the average of event-study coefficients (δ_h) in the posttreatment period as "average treatment effects" (ATE) for both worker-level and county-level regressions. These are estimates of γ from eq. (2). The dependent variable is the percentage of the year in which a worker had nonzero earnings. The treatment is a dummy variable equal to 1 if the worker had at least 10 megawatts (MW) of capacity within 20 miles of their residence or if the county had at least 10 MW of capacity in county-level regressions. Aggregating to 20 miles around county centroids produced similar results, which are omitted here. Worker-level estimates are parameter averages and standard deviations across 100 model estimates from repeated random draws from the near-population. County-level estimates use the full dataset aggregated to the county level, with standard errors clustered at the county level. The cumulative pretrends test reports the sum of event-study coefficients over the pretreatment period and its standard deviation across draws (worker level) or analytical standard error (county level). Brackets below the county-level average treatment effects contain 95% confidence intervals.

Table 3. Average Treatment Effects (Binary > 10 MW within 20 Miles): Log Earnings

	All	Black	Am. Ind	White	Hispanic	Female	Male	No High	High Sch.	Some Coll.	College
	A. Worker Level										
ATE	4.0	5.7	4.1	3.5	3.5	2.9	4.9	6.0	2.9	3.8	4.1
	(1.9)	(3.2)	(5.4)	(1.7)	(2.1)	(1.7)	(1.7)	(2.0)	(1.4)	(1.5)	(1.8)
Earnings ($)	1,270	1,330	768	1,110	883	710	1,900	1,170	706	1,130	2,020
Cumulative pretrend	−2.7	−.56	−13	−2.0	3.4	−1.7	−1.6	−2.5	.74	−1.9	−4.4
	(3.6)	(5.3)	(5.2)	(3.1)	(3.7)	(3.4)	(3.0)	(3.2)	(2.3)	(3.0)	(2.5)
	B. County Level										
ATE	−.13	−1.1	1.4	−.17	−.55	−.11	−.11	−.74	−.073	.014	−.79
	(.33)	(3.6)	(2.3)	(.33)	(1.2)	(.27)	(.42)	(.67)	(.41)	(.37)	(.44)
	[−.78,.51]	[−8.2,5.9]	[−3.1,5.8]	[−.82,.47]	[−3.0,1.9]	[−.65,.42]	[−.93,.70]	[−2.1,.57]	[−.89,.74]	[−.71,.74]	[−1.7,.08]
Cumulative pretrend	.44	4.1	6.6	.45	−3.3	.25	.63	−1.2	.72	−.21	1.2
	(1.0)	(12)	(7.2)	(.98)	(4.3)	(.82)	(1.3)	(1.8)	(1.4)	(1.1)	(1.3)

Note. This table reports the average of event-study coefficients (δ_h) in the posttreatment period as "average treatment effects" (ATE) for both worker-level and county-level regressions. These are estimates of γ from eq. (2). The dependent variable is the log of earnings, with the sample limited to employed workers (those with at least two quarters of nonzero earnings in a year). The treatment is a dummy variable equal to 1 if the worker had at least 10 megawatts (MW) of capacity within 20 miles of their residence or if the county had at least 10 MW of capacity in county-level regressions. Aggregating to 20 miles around county centroids produced similar results, which are omitted here. Worker-level estimates are parameter averages and standard deviations across 100 model estimates from repeated random draws from the near-population. County-level estimates use the full dataset aggregated to the county level, with standard errors clustered at the county level. The cumulative pretrends test reports the sum of event-study coefficients over the pretreatment period and its standard deviation across draws (worker level) or analytical standard error (county level). Brackets below the county-level average treatment effects contain 95% confidence intervals.

estimated on 100 Monte Carlo random draws of the geocoded worker-level data, reporting means and standard deviations of parameter estimates across draws. Panel B in each table shows the effects using the full data aggregated to the county level, with analytical standard errors clustered at the county level. The first column in each table reports these results for all workers, whereas each subsequent column reports results for each of the subpopulations that we study.

The first row and column of table 2, for example, shows that being exposed to utility-scale wind installations within 20 miles increases employment by 0.42 percentage points for the average worker. Given that approximately 55,000 workers live within 20 miles of a utility-scale wind project, this translates to an average effect of approximately 231 jobs per project. We use similar calculations to translate coefficients into jobs for each subpopulation in table 2. By contrast, previous studies find local impacts of approximately 50–90 jobs per project (Brown et al. 2012; Gilbert et al. 2023). In terms of an employment multiplier, this translates to approximately 0.51 jobs per million dollars in wind project investment.[15] This is a fairly modest employment multiplier compared to those in other industries or from federal stimulus that tend to be at least an order of magnitude larger (Chodorow-Reich 2019). Multipliers may be small because most of the expenditure in wind project development is capital expenditure rather than ongoing operations. However, the multiplier would be larger if only part of the project cost is spent locally. According to Stehly et al. (2023), about 40% of project costs are spent locally. This implies a multiplier on local dollars that is closer to 1.3 jobs per million dollars.

In the worker-level data, the cumulative pretrends are not statistically significant for any subpopulation except for American Indian/Native Alaskan workers. The other key feature to note from table 2 is that the county-level impact estimates are much smaller than the worker-level estimates. They are not statistically significant for any subpopulation or overall, and with the exception of Hispanic and American Indian workers, the point estimates are quite a bit smaller than in the worker-level regressions—coming in at about one-third of the worker-level estimates on average.

Using a continuous treatment variable in GW of capacity yields similar results. The first row and column of appendix table 2 indicates that each additional GW of capacity within 20 miles increases the average worker's employment by 1.3 percentage points. The average capacity within 20 miles of a treated worker is 0.3 GW, which yields a treatment effect at the mean that is very similar to the binary treatment coefficient. However, the average treatment effect estimates are somewhat noisier using the continuous variable.

---

15. We arrive at this multiplier by the following calculation: utility-scale wind installation costs are approximately \$1.5 million per MW (Wiser et al. 2023), and the average exposure of treated workers is 300 MW within 20 miles, or \$450 million of wind investment. Dividing 231 jobs by \$450 million gives the result.

Similar patterns arise when examining table 3 regarding impacts on log earnings (conditional on being employed). The first row and column of table 3 shows that earnings increase by 4.0% following the arrival of a nearby wind project. This is also somewhat larger than previous literature, which has fairly mixed results for earnings impacts ranging from no impact to about 3% (Brunner and Schwegman 2022). Considering that the average earnings of treated workers in our sample is $31,100, this translates to an impact of approximately $1,270 per employed worker per year. In terms of an earnings multiplier, this is equivalent to about 0.16 dollars in worker earnings per dollar invested in local wind capacity. This is also modest compared to similar infusions of spending such as government fiscal stimulus, which has earnings multipliers estimated to range from 0.3 to 2, but with most estimates falling between 0.6 and 1 (Ramey 2019). Again this may be small because most of the project costs are capital expenditures, but the multiplier on local dollars is likely larger. If 40% of project costs are spent locally (Stehly et al. 2023), this translates to a multiplier of about 0.4 dollars in worker earnings per dollar of wind investment spent locally, which is closer to the low end of fiscal multipliers.

Cumulative pretrends are again only significant for American Indian workers. When comparing earnings impacts at the worker level in panel A to the county level in panel B, however, we see even more dramatic attenuation at the county level for earnings than for employment. Almost all county-level estimates are close to zero and many are the opposite sign. None are statistically significant. They are especially lower for black and Hispanic workers, men, and workers either with a college degree or without a high school diploma.

As with employment, using a continuous treatment variable in GW of capacity yields similar results for earnings as the binary treatment variable. The first row and column of appendix table 3 indicates that each additional GW of capacity within 20 miles increases the average employed worker's earnings by 12 percentage points. With an average treated capacity of 0.3 GW the treatment effect at the mean is again very similar to the binary treatment coefficient. As with employment, the average treatment effect estimates for log earnings are somewhat noisier using the continuous variable.

4.2. Treatment Effect Comparisons

Table 2 shows that among race and ethnicity subgroups, the binary treatment effect on employment is proportionately largest for black workers at 0.64 percentage points as compared to white workers at 0.36 percentage points. However, because more white workers live near wind projects, the total jobs impact is still much larger for white workers (160 jobs) versus black workers (25 jobs). The impact is much larger for male than for female workers (0.46 vs. 0.33 percentage points increase in employment, translating to 131 vs. 88 jobs). The effect is also proportionately largest for workers with either very low skill (no high school education) or high skill (college education), while again the aggregate jobs impact is still greater for more-educated workers (roughly 46 jobs for people with high school education or below vs. 66 for people with at least some college).

Table 3 similarly shows that the binary treatment effects on log earnings are proportionately largest for black workers and workers without a high school education and are heavily imbalanced for men versus women. Unlike the employment impacts, these proportionate impacts are also reflected in the actual dollar magnitudes of earnings impacts. Black workers increase earnings by over $1,300 per year while workers without a high school education gain almost $1,200. However, the increase in male earnings is almost three times that of female earnings, and the largest earnings increase is enjoyed by college graduates at over $2,000 per year. These impacts are statistically significant for all subpopulations except for women, and black, Hispanic, and American Indian workers.

Table 4 summarizes the difference of estimated average treatment effects (in percentage point terms) between subpopulations, averaged across Monte Carlo draws

Table 4. Differences in Average Treatment Effects (Binary > 10 MW within 20 Miles)

Average Difference in ATE across Draws	Employment	Log Earnings
White - Black	−.28	−2.3
% positive	.23	.23
White - American Indian	−.036	−.65
% positive	.58	.57
White - Hispanic	−.085	−.025
% positive	.35	.49
Male - female	.13	2.0
% positive	.77	.84
No high school - high school	.26	3.1
% positive	.89	.90
No high school - some college	.21	2.2
% positive	.84	.85
High school - some college	−.049	−.91
% positive	.38	.30
College - no high school	−.088	−2.0
% positive	.35	.15
College - high school	.17	1.2
% positive	.84	.69
College - some college	.12	.25
% positive	.75	.50

Note. This table reports the average across Monte Carlo draws of the difference in average treatment effects (ATE) between different demographic groups from the individual worker-level regressions with a binary treatment variable. These are the differences between average treatment effects in tables 2 and 3, with slight differences due to US Census Bureau rounding rules for the release of individual estimates. For each comparison, this table also reports the percentage of draws in which the difference is positive as a way to describe the statistical significance of each comparison.

of worker-level data. The table also shows the percentage of draws for which the difference in treatment effects is positive. It should be noted that treatment effect differences between subpopulations are not statistically significant at conventional levels, with the percentage of draws with a positive difference always below 95%. However, the treatment effect differences are on average larger in percentage point terms and are most frequently larger across draws, for black workers compared to white workers, men compared to women, workers without a high school education compared to workers with either a high school diploma or some college, and workers with a college degree compared to those with a high school diploma. These comparisons reflect similar patterns that we can see by inspection of tables 2 and 3.

We conduct similar comparisons, and find similar patterns, using our continuous variable of GW of capacity in appendix table 4. In those results, however, the differences between black and white workers are more pronounced whereas the differences between workers with very high and very low education are somewhat dampened—both in terms of magnitude and frequency—relative to the binary treatment variable.

These differences are consistent with a variety of possible mechanisms, and we do not have enough evidence to isolate which ones are most important. It may be that landowners spend their royalties to hire more unskilled laborers who do not need a high school education. In addition, landowners and wind project construction workers may spend more money in local restaurants that employ workers without a high school diploma. Tax payments to local school districts may cause them to hire more teachers, which requires a college degree. Understanding these channels in more detail is a topic for future research.

Using county-level data, comparisons between the subpopulations can be made by inspecting the 95% confidence intervals in panel B of tables 2 and 3 (and analogously appendix tables 2, 3 for the continuous treatment variable). In almost all cases, the magnitude of differences in treatment effects is smaller at the county level, and there is considerable overlap in the confidence intervals. The exceptions are employment treatment effects for American Indian and Hispanic workers as compared to white workers. This is not surprising given that the county-level treatment effects for these are larger than the estimates using worker-level data whereas the county-level estimates are smaller than the worker-level estimates in most other subpopulations.

4.3. Event-Study Graphs

Figures 2 and 3 plot the LPDID event-study coefficients for employment and log of earnings, for worker-level and county-aggregate data. We plot the confidence intervals calculated from the analytical standard errors in the county-level event studies and from the standard deviation of point estimates across Monte Carlo draws in the worker-level event studies. We normalize to two years before any wind project is operational in order to capture potential construction effects in the year preceding the first year of operation.

Figure 2. Impact of wind capacity on employment: worker-level versus county-level. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2.

We can see several important findings in both figures. First, we can see from both figures that impacts are not concentrated around the construction phase of event years −1 and 0. Rather, impacts grow over time and persist as many as six years after the project becomes operational. This is consistent with the hypothesis that there are many indirect channels through which renewable energy can improve the local economy, not just related to direct employment at the renewable energy installation. Indirect impacts may occur, for example, because of local landowners spending royalty payments, additional community services stimulated by additional local tax payments, or other indirect channels. A similar time path of event-study coefficients was found by Brunner and Schwegman (2022) in their study of wind installations and local economic development.

Second, we can see from both figures that error bars on the worker-level estimates are relatively large compared to county-level error bars, despite the fact that worker-level

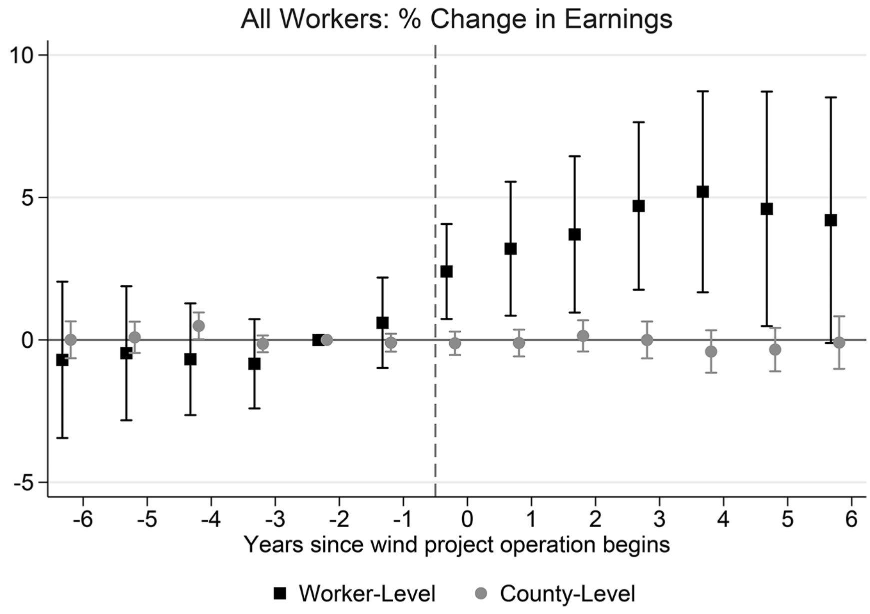

Figure 3. Impact of wind capacity on log earnings: worker-level versus county-level. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2. The dependent variable is logged earnings, so impact estimates are semi-elasticities, or approximately percentage changes in annual earnings.

sample sizes are in the millions while county-level sample sizes are in the thousands. This suggests that there is significant heterogeneity in impacts that is picked up through our repeated random sampling procedure.

Third, the figures help visualize the degree of attenuation from using county-level aggregates. Figure 2 shows that county-level estimates are smaller and still quite noisy; yet point estimates are on average positive and less than half of the worker-level estimates. By contrast, figure 3 shows that earnings estimates at using county-level data are, essentially, precisely estimated zeroes. This suggests that earnings impacts as reported in the literature are likely understating true earnings impacts by a wider margin than employment estimates.

4.3.1. Event Studies: Race and Ethnicity

Figures 4 and 5 also confirm the average treatment effects in the previous tables. The effect on black workers for both employment and earnings is both larger and more

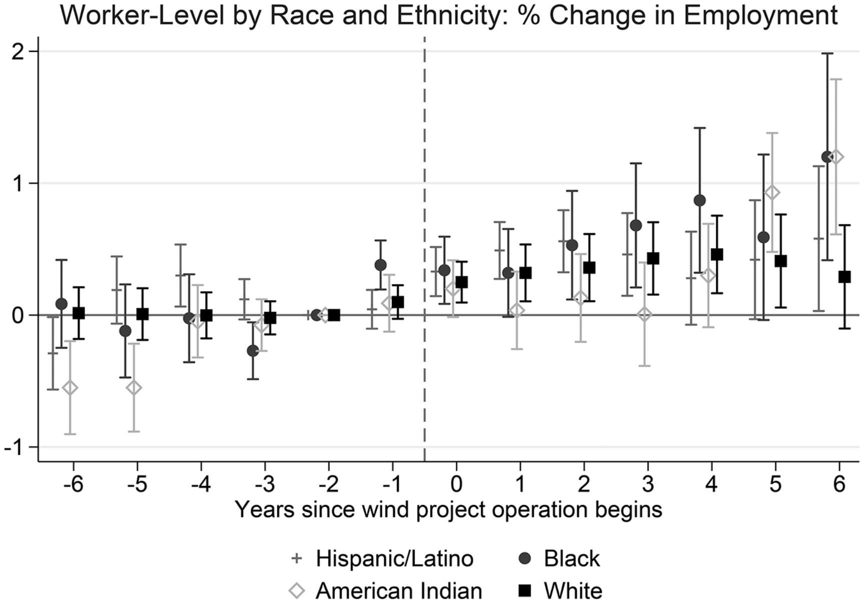

Figure 4. Impact of wind capacity on employment: worker-level by race and ethnicity. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2.

persistent than for other races and ethnicities; appendix table 1 shows that despite there being relatively few black workers with wind projects within 20 miles, the average "treated" black worker is also exposed to significantly more capacity. Yet appendix tables 2 and 3 show that the marginal impact of an additional gigawatt of wind capacity is also larger for black workers.

4.3.2. Event Studies: Sex

Figures 6 and 7 further show that impacts on employment and earnings are consistently larger for male workers than female workers. The impacts on men start sooner, are larger, and persist longer than for women. However, there is much overlap in the confidence intervals at each time step.

4.3.3. Event Studies: Educational Attainment

Finally, figures 8 and 9 again show persistent impacts that are largest for workers without a high school education or workers with a college degree. However, impacts for all

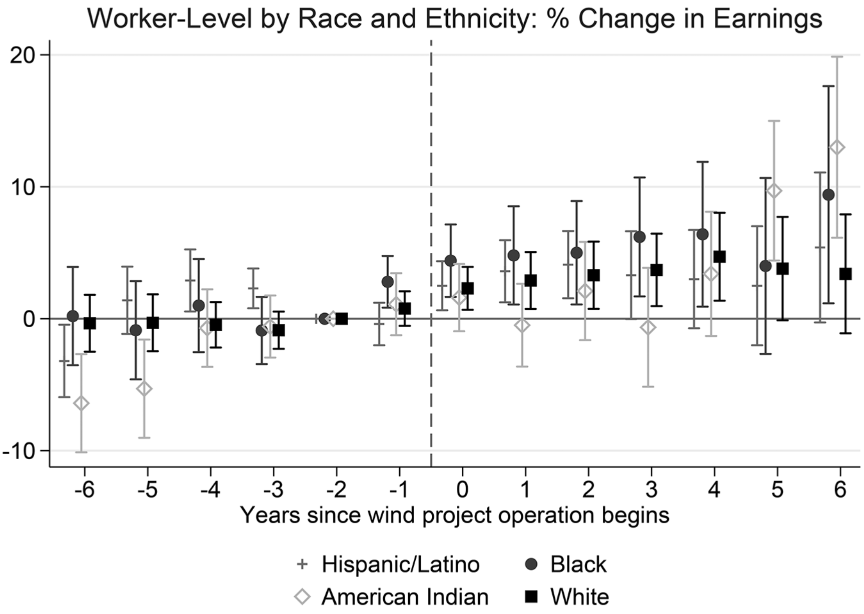

Figure 5. Impact of wind capacity on log earnings: worker-level by race and ethnicity. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2. The dependent variable is logged earnings, so impact estimates are semi-elasticities, or approximately percentage changes in annual earnings.

educational categories are statistically significant and persistent for many years following the arrival of a wind project.

5. CONCLUSION

As the United States continues to make unprecedented investments in renewable energy in order to meet carbon emissions goals, this will shift the demand for skilled and unskilled labor and generate new sources of income, tax revenues, and expenditures, possibly in places that have not previously been major energy-producing communities. These developments could either continue to allocate benefits to privileged groups while continuing to restrict access to disadvantaged groups or they could increase access to economic opportunity among vulnerable populations.

In this study, we use restricted-access geocoded data on the near-universe of workers in 23 US states in order to estimate the local earnings and employment impacts of wind energy development. We estimate these effects for all workers and separately for

Figure 6. Impact of wind capacity on employment: worker-level by sex. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2.

black, American Indian/Native Alaskan, white, and Hispanic workers, male versus female workers, and those without a high school diploma, with a high school diploma, some college coursework, and with a college degree. We then aggregate these data to the county level in order to compare our estimates with those we would have obtained with county-level aggregates using data such as is available in the public domain.

We find economically and statistically significant employment and earnings gains from wind development within 20 miles of a worker's residence. We also find that these impacts are relatively more pronounced for black workers, men, and very low skilled or high skilled workers. These impacts persist for years after the construction phase ends, suggesting that there may be multiple indirect channels through which wind capacity in place in a community can generate benefits.

We also find that impact estimates are dramatically lower when using county-level aggregate data and that the differences in the estimates vary across subpopulations in ways that are not obviously predictable. This likely arises because of the well-known

Figure 7. Impact of wind capacity on log earnings: worker-level by sex. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2. The dependent variable is logged earnings, so impact estimates are semi-elasticities, or approximately percentage changes in annual earnings.

issue of MAUP in the geography literature. This finding also suggests that there is an inequity within the research community in terms of which researchers have access to better data in order to generate a full understanding of impacts on the communities who may be served by their institutions.

Our study is limited in several respects. We were not able to access data on two major wind energy states: Texas and Minnesota. We were also not able to explore impacts of other energy sources, compare a variety of identification strategies, or evaluate additional outcome variables because of limitations on computation time using such a large dataset. Future research using the LEHD could explore a narrower selection of control groups using various matching and propensity weighting methods in order to more precisely measure comparisons between subpopulations or consider changes in earnings within a given spell at a specific employer. Further examination of the mechanisms driving the impact multipliers and subpopulation differences would also be valuable. For example, are multipliers low because expenditures travel through worker

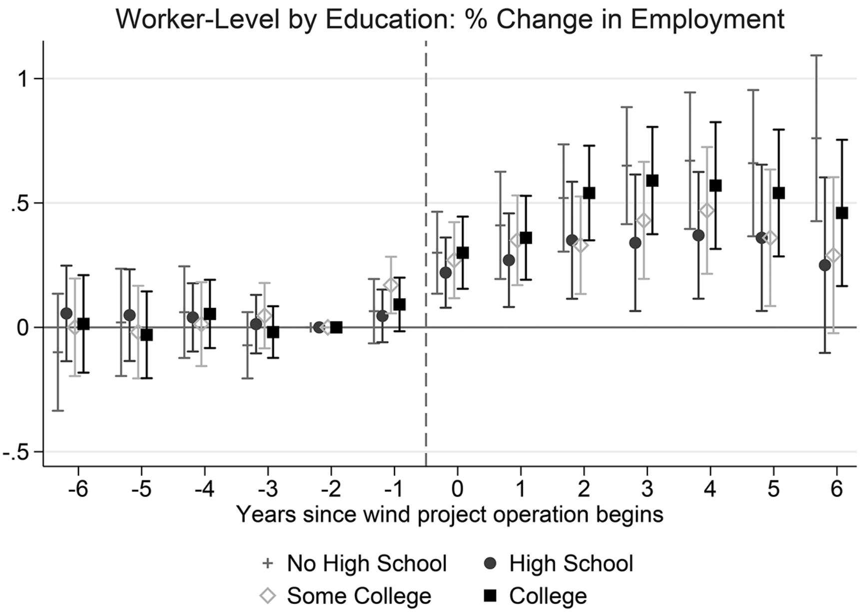

Figure 8. Impact of wind capacity on employment: worker-level by education. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2.

types with a lower marginal propensity to consume? What features of wind energy communities help men, black workers, and workers without a high school diploma gain the most from wind development? We were also not able to look at more intersectional outcomes, starting with more race and ethnicity categories but also including workers with more than one race or ethnicity, non-gender-binary workers, or more granularity in worker skill level. We further did not consider impacts on migration decisions or impacts on workers in specific industrial sectors. These are all important areas for future research. Similarly, understanding the "boundary" of a local community and its economy is also an interesting question for future research. While we have followed the common approach from the literature of using concentric circle distances, other methods such as isochrones for commuting time may be interesting to explore.[16]

16. We thank Justin Kirkpatrick for this suggestion.

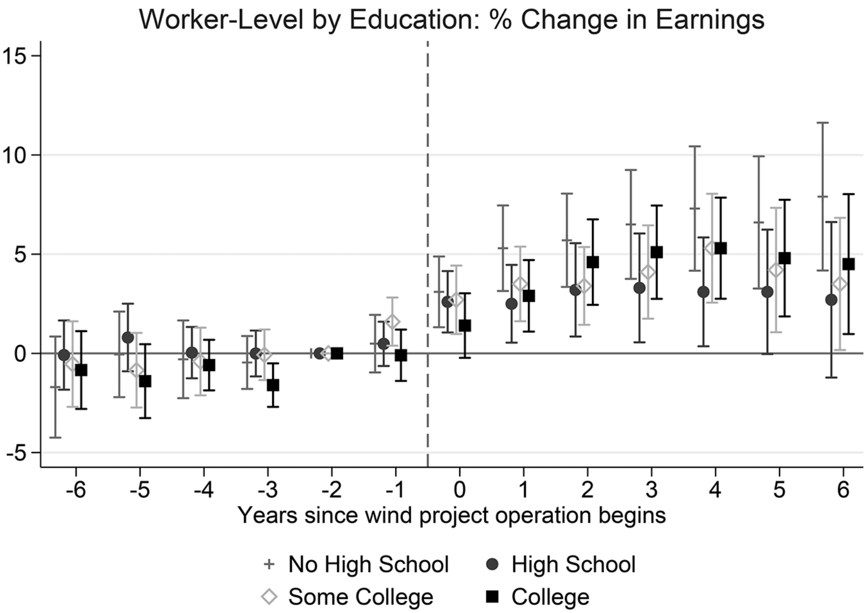

Figure 9. Impact of wind capacity on log earnings: worker-level by education. These are event-study estimates from equation (3) using worker-level data versus county-aggregate data. Treatment is a binary variable for whether at least 10 megawatts of wind has arrived within 20 miles of the worker's residence or arrived within the county. Event year 0 equates to the year wind project operations began. Wind project construction may have occurred in the preceding year, so we normalize to event year −2. The dependent variable is logged earnings, so impact estimates are semi-elasticities, or approximately percentage changes in annual earnings.

REFERENCES

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock. 2009. The LEHD infrastructure files and the creation of the Quarterly Workforce Indicators. In *Producer dynamics: New evidence from micro data*, ed. Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts, 149–230. Chicago: University of Chicago Press.

Acheampong, Alex O., Janet Dzator, and Muhammad Shahbaz. 2021. Empowering the powerless: Does access to energy improve income inequality? *Energy Economics* 99:105288.

Acheampong, Alex O., Michael Odei Erdiaw-Kwasie, and Matthew Abunyewah. 2021. Does energy accessibility improve human development? Evidence from energy-poor regions. *Energy Economics* 96:105165.

Birch, Colin P. D., Sander P. Oom, and Jonathan A. Beecham. 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecological Modelling* 206 (3–4): 347–59.

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

Brown, Jason P., John Pender, Ryan Wiser, Eric Lantz, and Ben Hoen. 2012. Ex post analysis of economic impacts from wind power development in U.S. counties. *Energy Economics* 34 (6): 1743–54.

Brunner, Eric J., Ben Hoen, and Joshua Hyman. 2022. School district revenue shocks, resource allocations, and student achievement: Evidence from the universe of U.S. wind energy installations. *Journal of Public Economics* 206:104586.

Brunner, Eric J., Ben Hoen, Joe Rand, and David Schwegman. 2024. Commercial wind turbines and residential home values: New evidence from the universe of land-based wind projects in the United States. *Energy Policy* 185:113837.

Brunner, Eric J., and David J. Schwegman. 2022. Commercial wind energy installations and local economic development: Evidence from US counties. *Energy Policy* 165:112993.

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

Cameron, A. Colin, and Pravin K. Trivedi. 1986. Econometric models based on count data: Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics* 1 (1): 29–53.

Carley, Sanya, and David M. Konisky. 2020. The justice and equity implications of the clean energy transition. *Nature Energy* 5 (8): 569–77.

Castleberry, Becca, and J. Scott Greene. 2017. Impacts of wind power development on Oklahoma's public schools. *Energy, Sustainability and Society* 7 (1): 34.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. The effect of minimum wages on low-wage jobs. *Quarterly Journal of Economics* 134 (3): 1405–54.

Chodorow-Reich, Gabriel. 2019. Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy* 11 (2): 1–34.

Cicala, Steve. 2022. Imperfect markets versus imperfect regulation in US electricity generation. *American Economic Review* 112 (2): 409–41.

Colmer, Jonathan, John Voorheis, and Brennan Williams. 2021. Air pollution and economic opportunity in the United States. Unpublished manuscript, University of Virginia and US Census Bureau. https://drive.google.com/file/d/19zLlSTaSJgs1c3FSHo2_l1xUmgSBJuq2/view.

Dong, Luran, Vasundhara Gaur, and Corey Lang. 2023. Property value impacts of onshore wind energy in New England: The importance of spatial heterogeneity and temporal dynamics. *Energy Policy* 179:113643.

Dube, Arindrajit, Daniele Girardi, Oscar Jorda, and Alan M. Taylor. 2023. A local projections approach to difference-in-differences event studies. Technical report, National Bureau of Economic Research, Cambridge, MA.

Fergen, Joshua, and Jeffrey B. Jacquet. 2016. Beauty in motion: Expectations, attitudes, and values of wind energy development in the rural U.S. *Energy Research and Social Science* 11:133–41.

Feyrer, James, Erin T. Mansur, and Bruce Sacerdote. 2017. Geographic dispersion of economic shocks: Evidence from the fracking revolution. *American Economic Review* 107 (4): 1905–13.

Fotheringham, A. Stewart, and David W. S. Wong. 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A* 23 (7): 1025–44.

Gilbert, Ben, Hannah Gagarin, and Ben Hoen. 2023. Geographic spillovers of wind energy development on wages and employment. Technical report.

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

Hoen, B. D., J. E. Diffendorfer, J. T. Rand, L. A. Kramer, C. P. Garrity, and H. E. Hunt. 2021. United States Wind Turbine Database (ver. 4.0). US Geological Survey, American Wind Energy Association, and Lawrence Berkeley National Laboratory data release.

Hoen, Ben, Jason P. Brown, Thomas O. Jackson, Mark Thayer, Ryan Wiser, and Peter Cappers. 2015. Spatial hedonic analysis of the effects of US wind energy facilities on surrounding property values. *Journal of Real Estate Finance and Economics* 51 (1): 22–51.

Hoen, Ben, Jeremy Firestone, Joseph Rand, Debi Elliot, Gundula Hübner, Johannes Pohl, Ryan Wiser, Eric Lantz, T. Ryan Haac, and Ken Kaliski. 2019. Attitudes of US wind turbine neighbors: Analysis of a nationwide survey. *Energy Policy* 134:110981.

Jacobsen, Grant D., Dominic P. Parker, and Justin B. Winikoff. 2023. Are resource booms a blessing or a curse? Evidence from people (not places). *Journal of Human Resources* 58 (2): 393–420.

James, Alexander. 2015. US state fiscal policy and natural resources. *American Economic Journal: Economic Policy* 7 (3): 238–57.

James, Alexander G., and Brock Smith. 2020. Geographic dispersion of economic shocks: Evidence from the fracking revolution; Comment. *American Economic Review* 110 (6): 1905–13.

Jones, Christopher R., and J. Richard Eiser. 2010. Understanding "local" opposition to wind development in the UK: How big is a backyard? *Energy Policy* 38 (6): 3106–17.

Kim, Haeyeon. 2019. Three essays on mineral economics and economic geography. PhD thesis, Colorado School of Mines.

Maniloff, Peter, and Ralph Mastromonaco, 2017. The local employment impacts of fracking: A national study. *Resource and Energy Economics* 49:62–85.

Manwell, James F., Jon G. McGowan, and Anthony L. Rogers. 2010. *Wind energy explained: Theory, design and application*. Hoboken, NJ: Wiley.

Mauritzen, Johannes. 2020. Will the locals benefit? The effect of wind power investments on rural wages. *Energy Policy* 142:111489.

McCauley, Darren, Vasna Ramasar, Raphael J. Heffron, Benjamin K. Sovacool, Desta Mebratu, and Luis Mundaca. 2019. Energy justice in the transition to low carbon energy systems: Exploring key themes in interdisciplinary research. *Applied Energy* 233–34:916–21.

Mejía-Montero, Adolfo, Matthew Lane, Dan van Der Horst, and Kirsten E. H. Jenkins. 2021. Grounding the energy justice lifecycle framework: An exploration of utility-scale wind power in Oaxaca, Mexico. *Energy Research and Social Science* 75:102017.

Millimet, Daniel L., and Marc F. Bellemare. 2023. Fixed effects and causal inference. IZA Discussion paper 16202, IZA, Bonn.

Mills, Sarah Banas, Douglas Bessette, and Hannah Smith. 2019. Exploring landowners' post-construction changes in perceptions of wind energy in Michigan. *Land Use Policy* 82:754–62.

Mueller, J. Tom, and Matthew M. Brooks. 2020. Burdened by renewable energy? A multiscalar analysis of distributional justice and wind energy in the United States. *Energy Research and Social Science* 63:101406.

Openshaw, Stan. 1984. Ecological fallacies and the analysis of areal census data. *Environment and Planning A* 16 (1): 17–31.

Pedden, M. 2006. Analysis: Economic impacts of wind applications in rural communities. Technical report, NREL/SR–500–39099, National Renewable Energy Laboratory, Golden, CO.

Pryor, Sara C., Rebecca J. Barthelmie, Melissa S. Bukovsky, L. Ruby Leung, and Koichi Sakaguchi. 2020. Climate change impacts on wind power generation. *Nature Reviews Earth and Environment* 1 (12): 627–43.

Ramey, Valerie A. 2019. Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives* 33 (2): 89–114.

Rand, Joseph T., Louisa A. Kramer, Christopher P. Garrity, Ben D. Hoen, Jay E. Diffendorfer, Hannah E. Hunt, and Michael Spears. 2020. A continuously updated, geospatially rectified database of utility-scale wind turbines in the United States. *Scientific Data* 7 (1): 15.

Robinson, W. S. 1950. Ecological correlations and the behavior of individuals. *American Sociological Review* 15 (3): 351–57.

Shoeib, Eman Ahmed Hamed, Elisabeth Hamin Infield, and Henry C. Renski. 2021. Measuring the impacts of wind energy projects on U.S. rural counties' community services and cost of living. *Energy Policy* 153:112279.

Shoeib, Eman Ahmed Hamed, Henry C. Renski, and Elisabeth Hamin Infield. 2022. Who benefits from renewable electricity? The differential effect of wind power development on rural counties in the United States. *Energy Research and Social Science* 85:102398.

Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. 2015. What are we weighting for? *Journal of Human Resources* 50 (2): 301–16.

Stehly, Tyler, Patrick Duffy, and Daniel Mulas Hernando. 2023. 2022 cost of wind energy review. Technical report NREL/88335, National Renewable Energy Laboratory, Golden, CO.

Sun, Liyang, and Sarah Abraham. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225 (2): 175–99.

US Census Bureau. 2021. Calculating migration expectancy using ACS data. https://www.census.gov/topics/population/migration/guidance/calculating-migration-expectancy.html (accessed February 8, 2024).

Varela-Vazquez, Pedro, and Maria del Carmen Sanchez-Carreira. 2015. Socioeconomic impact of wind energy on peripheral regions. *Renewable and Sustainable Energy Reviews* 50:982–90.

Wiser, Ryan, Mark Bolinger, Ben Hoen, Dev Millstein, Joe Rand, Galen Barbose, Naïm Darghouth, et al. 2023. Land-based wind market report: 2023 edition. Technical report, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA.

Wooldridge, Jeffrey M. 1999. Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics* 90 (1): 77–97.