Nice Work if You Can Get It? The Distribution of Employment and Earnings during the Early Years of the Clean Energy Transition*

Jonathan Colmer
University of Virginia

Eva Lyubich
U.S. Census Bureau

John Voorheis
U.S. Census Bureau

Abstract

The transition to clean energy represents a fundamental and important shift in economic activity. We present new facts about workers in clean and legacy energy sectors between 2005 and 2019 using linked, administrative employer-employee data for all W-2 workers in the United States. We show that both clean and legacy energy establishments hire a disproportionate share of non-Hispanic White and male workers compared to the working population, that workers rarely move from legacy to clean firms, and that, conditional on education, workers do not earn more in clean firms than in legacy firms. The occupational categories of jobs at clean firms differ notably from occupations at legacy firms and, on average, tend to be performed by workers with higher levels of education. Regional overlap in employment opportunities is not sufficient to facilitate worker transitions from legacy to clean firms. Substantially lower earnings outside of the energy sector combined with low mobility between legacy and clean firms suggests that the costs of the clean transition on workers in legacy fossil fuel sectors may be substantial. At the same time workers moving into clean activities from outside of the energy sector experience significant increases in earnings and greater job stability, suggesting that clean jobs are "good jobs" for those who can access them.

*This Version: November 2023. First Version: June 2022. Correspondence: j.colmer@virginia.edu. Affiliations: Colmer—Department of Economics, University of Virginia, USA; Lyubich—U.S. Census Bureau, USA; Voorheis—U.S. Census Bureau, USA; Addison—Department of Economics, University of Virginia, USA. We thank Arik Levinson, Catie Hausman, and Gilbert Metcalf for many helpful thoughts, comments, and discussions. We are also grateful to seminar participants at the NBER Distributional Consequences of New Energy Technologies Workshop and AERE for helpful comments and suggestions. We thank Kyle Addison for outstanding research assistance. This project was supported by the Sloan Foundation. All errors and omissions remain our own. Any opinions and conclusions expressed herein are those of the authors and do not reflect the views of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7505723, Disclosure Review Board (DRB) approval numbers: CBDRB-FY22-CES014-032 and CBDRB-FY22-CES014-034).

1 Introduction

The United States is undergoing a historic shift away from carbon-intensive energy sources and toward lower carbon and clean energy sources. This process, which is driven by a combination of market forces and active industrial policy, will result in a reallocation of capital and labor from legacy fossil fuel activities to new clean energy and production. In the presence of market failures, most notably the negative externalities associated with extracting and burning fossil fuels, a change in the composition of economic activity will increase aggregate welfare because the value marginal product of labor and capital in clean activities is higher than the value marginal product of labor and capital in legacy activities. This does not, however, imply that reallocation is without cost. Even when new jobs are not replaced by capital, they may be located in different markets or require different skills. Frictions may result in mismatch, leading to a prolonged adjustment process. Understanding the extent to which disruptions affect workers is important for evaluating the distributional consequences of the clean energy transition even if the long-run aggregate consequences deliver much greater benefits to society as a whole.

In this paper we present new facts about the labor market consequences of the clean energy transition as experienced to date, evaluate how these effects are distributed across individuals, and consider the broader implications for economic opportunity and inequality in the United States. To answer these questions it is necessary to identify the universe of jobs destroyed and created, the characteristics of lost and gained jobs, and which workers are affected. We do this by constructing a new linked employer-employee data set from U.S. Census Bureau microdata, Internal Revenue Service (IRS) tax data and other administrative records, which contains information on residential histories, earnings histories, and employment histories, as well as employer characteristics, for the population of workers in the United States between 2005-2019. A challenge is that many clean energy activities are relatively nascent and are therefore not yet separately identified in the industrial classifications used in the construction of economic statistics and business-level microdata. We overcome this by combining industry information with a text analysis approach using keywords applied to firm names, followed by a detailed clerical review. Our approach yields the most comprehensive set of establishments engaging in legacy and clean energy activities, and captures approximately twice as many establishments and workers compared to the narrow industry-based classification approach.

We first present new facts about the aggregate trends in employment in clean and legacy energy establishments. Both the number of clean establishments and total employment in clean establishments have grown dramatically since 2005. Nevertheless, while the number of legacy establishments began to decline in 2018, employment in legacy establishments remains an order of magnitude higher than employment in clean establishments and the number of new legacy jobs continues to outpace the number of new clean jobs.

Second, we document that both clean and legacy jobs are disproportionately made up of non-Hispanic White, male workers compared to the working population as a whole. Workers in clean establishments are more likely to be college-educated than workers in legacy establishments, but no more so than the overall working population.

Third, we document that while unconditionally, workers in clean establishments have longer employment spells and receive higher earnings, gaps between clean and legacy establishments can be almost entirely explained by differences in educational attainment. The outside options available to workers with experience in clean and legacy establishments, however, are much worse. Even after controlling for individual fixed effects we estimate a 30% wage gap between energy and non-energy jobs. Decomposing this effect, we find that, on average, workers who move into clean jobs from outside of the energy sector earn 40% more and workers that are unable to transition from legacy to clean establishment earn 25% less.

Given the cost of not being able to transition from a legacy establishment to a clean establishment, we calculate the likelihood that workers make this transition. We show that both the likelihood of moving to a clean job conditional on leaving a legacy job, $p(clean_t|legacy_{t-1})$, and the likelihood of coming from a legacy job conditional on currently working in a clean job, $p(legacy_{t-1}|clean_t)$ are both incredibly small, even after accounting for indirect transitions, i.e. transitions where a worker is either at a non-energy job or does not have a W-2 for some period between legacy and clean jobs. Conditional on legacy-to-clean transitions arising, the workers who do move are disproportionately likely to be non-Hispanic white and college educated. The cost to workers after separating from legacy establishments have fallen disproportionately on less educated and non-white workers.

These facts suggest that new clean jobs may have created economic opportunity for workers who were previously outside of the energy sector, but that the transition has been particularly costly to separated workers in legacy sectors (Colmer et al., 2023). Less clear is why workers who have separated from legacy establishments have been unable to access jobs in clean establishments or why workers from outside of the energy sector did not take advantage of higher-paying jobs in legacy establishments prior to clean jobs becoming available. One barrier to transition may be skill mismatch, whereby workers in legacy establishments lack the skills required to work in new clean jobs. Mobility costs, search costs, and information frictions could also impede reallocation. However, it is also possible that fully informed workers may select out of higher-paying opportunities due to compensating differentials associated with working conditions or the geographic locations of jobs leading to average wage gaps between sectors.

Our initial analysis suggests that skill mismatch is likely to be a more important explanation for why workers have historically not reallocated from legacy to clean energy jobs. In addition to the fact that clean jobs are more likely to be occupied by college-educated workers, we find that the rank ordering of top occupations is quite different between sectors, despite overlap in the top occupations for workers at clean and legacy establishments. We also observe that while legacy jobs are more geographically concentrated, clean jobs are distributed more evenly across space. We estimate low transition rates even after accounting for the existence of clean jobs in a local area, suggesting that spatial mismatch is unlikely to be a first-order concern. The overlapping geography of clean and legacy jobs also suggests that information frictions, search frictions, and selection out of clean jobs due to place-based compensating differentials are unlikely to be first-order explanations.

Our main contribution is to shift from a place-based to a worker-based understanding of the clean transition's impact on labor markets. Existing work has either used aggregate data or highly selected samples of individual data. These data constraints have limited research on the distributional consequences of the clean transition to date by impeding serious study of heterogeneity (Carley and Konisky, 2020). In contrast, our comprehensive individual-level data with national coverage allows us to provide systematic and comprehensive evidence about the distributional consequences of the clean energy transition as experienced to date. Overall, the energy sector accounts for a very small share of employment and activity and so has little direct scope to fundamentally affect aggregate patterns of inequality. However, as the energy sector is a major source of greenhouse gas emissions and other local pollutants the reallocation of activity away from the production of electricity using fossil fuels towards clean electricity production has the potential to deliver broader, but more diffuse benefits in the form of climate mitigation.

Our second contribution is to the classification of clean and legacy activities. Classifying clean jobs is a non-trivial and subjective task because, in many cases, new activities are emerging that do not yet have industry or occupational classifications. Existing work has tended to categorize clean jobs by concentrating on narrow occupational definitions based on direct involvement in clean activities, which miss supporting activities that facilitate the generation of clean energy, or by identifying "green" or "clean" occupations(Bowen et al., 2018; Consoli et al., 2016; Vona et al., 2018, 2019; Rutzer et al., 2020; Curtis and Marinescu, 2023; Park et al., 2023). We approach this measurement task from the employer side, and use an establishment-based definition instead, which allows us to capture all jobs created and destroyed by establishment turnover in this transition. We apply keyword identification to firm names so that we are not limited to existing sector classifications and do extensive background checks to rule out false positives.

Finally, we contribute to the literature on the labor market consequences of structural and technological change (Tinbergen, 1974; Katz and Murphy, 1992; Goldin and Katz, 1998, 2008; Katz and Autor, 1999; Acemoglu, 2002; Autor et al., 2006; Acemoglu and Autor, 2011; Lin, 2011; Autor et al., 2013; Deming, 2017; Korinek and Stiglitz, 2018; Acemoglu and Restrepo, 2019, 2022; Autor, 2022; Emanuel et al., 2023; Emanuel and Harrington, 2023; Barrero et al., 2023). Our context differs from existing work exploring the consequences of automation and globalization, and therefore provides different insights into how workers and firms respond to structural change. Unlike automation, the clean energy transition could in principle create more jobs, not less, as it reflects a switch in the type of employment opportunities available rather than a replacement of workers with capital. Unlike the process of globalization, where jobs are being outsourced, clean jobs are locally available and in many cases cannot be outsourced. Unlike working from home, where the boundary of the workplace changes, the clean energy transition reflects a change in employer. Given differences in the geographic concentration of clean and legacy activities as well as potential differences in skill requirements, our context provides an opportunity to better understand the extent to which workers in legacy sectors or in the non-energy jobs are able to transition to available clean jobs, and if not why. We provide new insights into the role of skill and spatial mismatch in shaping labor market responses to structural change.

2 Data

To systematically evaluate the distribution of employment and earnings in the clean and legacy energy sectors, we have to first identify establishments engaged in clean energy related activities (as well as incumbent legacy energy establishments). From there, we identify which employers own these establishments, which workers are employed by those employers, and finally the characteristics of those identified workers.

Our approach to this set of tasks is to make use of administrative tax data, large nationally representative household surveys, and other administrative records that can be combined in the Census Bureau's data linkage infrastructure (Voorheis et al., 2023b). We first use a combination of industry classification codes, text analysis, and clerical review to identify clean and legacy establishments in Census business microdata (the County Business Patterns Business Register). We then use employer-establishment links in the business microdata to identify all employers (indexed by IRS Employer Identification Numbers or EINs) that own these establishments. After this, in the universe of IRS form W-2 recipients, we identify all workers who have ever worked for a clean and legacy employer and create a balanced panel of their employment histories. We finally link basic demographic information (age, gender, race, and ethnicity) from various sources for this entire population, as well as more detailed sociodemographic information (including educational attainment and occupation) from the American Community Survey for a separate, nationally representative sample of these workers. We provide more detail on these data steps below.

2.1 Classifying Clean and Legacy Energy Firms

Identifying clean and legacy businesses, workers, or jobs is a complicated task. Our focus in this paper is the set of employers and workers whose primary business activities are associated with energy production, transmission, and distribution.[1] Our approach utilizes North American Industry Classification System (NAICS) codes, a well-defined set of industry codes used to categorize economic activity. These codes are well-suited to identify obviously clean and legacy economic activity. However, they are less suited to easily identify companies that indirectly support clean or legacy energy or companies that engage in new clean activities. We account for this with a thorough review of broader NAICS codes covering notable indirect energy activity. Our approach excludes industrial activities that produce high levels of direct emissions but are not involved in energy production, such as carbon-intensive steel or cement production. While changes to the operations of these activities play a significant role in decarbonization, we cannot measure their contributions to the clean energy transition in the same way that we can for businesses directly involved in energy production. To increase comparability in the types of activity employers engage in, we opt for a more narrow definition of clean/legacy employers and explicitly restrict ourselves to the energy sector.

We identify establishments from the County Business Patterns Business Register (CBP-BR) using their listed business names and their EIN to identify unique observations. The Census Business Register County Business Patterns version is the underlying microdataset used to create the Census Bureau's County Business Patterns data products. The CBP-BR is a harmonized version of the Census Business Register, which is derived primarily from Internal Revenue Service corporate and individual tax filings. We use the same set of NAICS codes each year to create separate lists of clean and legacy establishments while allowing for entry and exit. As incumbents, legacy energy sector NAICS codes are well classified. The sector codes used to classify establishments engaged in legacy energy activities are listed in Columns 1 and 2 of Table 1. "Crude Petroleum Extraction", "Fossil Fuel Electric Power

[1]We will use establishment level information to identify clean and legacy activities but will primarily use employers, as identified by their EIN, as our relevant business unit of analysis. Most employers are made up of one or more establishments, though a small share of them contain a very large number of establishments. Firms are an additional business unit that will be used; they are composed of one or more employers and identified by their firm id from the Longitudinal Business Database (LBD).

Generation", "Bituminous Coal Mining", "Natural Gas Distribution", and a variety of other legacy energy activities are among them. We classify all establishments in each of these industries as legacy.

Identifying clean energy sectors is more difficult. We start with ten NAICS codes that clearly represent clean energy activities (Columns 3 and 4 of Table 1). These codes represent various types of renewable energy generation as well as "Storage Battery Manufacturing." We consider all establishments in each of these sectors to be clean. These NAICS codes, however, only include activities directly involved in electricity production and not broader support activities. Because clean energy related activities are still nascent, important clean energy activities are often categorized under broader NAICS codes alongside other economic activities. Solar cell manufacturing, for example, is included in the NAICS code "334413 – Semiconductor and Related Device Manufacturing." Other examples include solar panel installation, which is classified under NAICS code "238210 – Electrical Contractors and Other Wiring Installation Contractors"; wind turbine installation, which is classified under NAICS code "237130 – Power and Communication Line and Related Structures Construction"; and wind turbine manufacturing, which is classified under NAICS code "333611 – Turbine and Turbine Generator Set Unit".

To identify establishments participating in these supporting clean energy activities, we first identify the set of NAICS codes they could be grouped under (Columns 5 and 6 of Table 1). After this, we pull all establishments categorized under these codes, filter by matching business names to clean energy keywords, and clerically review the results. A thorough manual review provides an additional level of certainty in determining whether establishments are legitimately engaged in clean activities and allows us to rule out false positives. The keywords chosen to signal likely clean energy activity were limited to `solar`, `sun`, `wind`, `renewable`, and `energ`[2]. Clerical review of the string-matched establishments took an extensive amount of time but proved worthwhile due to the number of non-energy observations that were removed from the sample. For example, the strings "solar" and "sun" were commonly identified in a wide variety of business names in high sun states such as Florida, Arizona, and Nevada. Manual review allowed us to retain relevant matches, such as "Southern Solar Installers", and remove irrelevant activities, such as "Sunny Days Pool Cleaners", that would otherwise have remained in our sample. This method will still miss a subset of establishments engaged in clean energy activities with more ambiguous names but is much less likely to include false positives. Observations categorized as clean after the

[2]The keyword "energ" was chosen in order to pick up establishments whose businesses have variations of the word "energy" in their name. Nonexistent but plausible examples are "Energized Futures Electric" or "Energetic Power Alternatives"

clerical review are combined with the known clean results from our established NAICS codes to make up our full collection of clean establishments. The addition of clean activities from these additional sectors substantially expanded the number of clean energy establishments.

Finally, because many firms have multiple establishments engaged in different activities and may have different NAICS codes, we identify their modal activity (defined using year end employee headcount). Because the majority of clean energy economic activity occurs in newly formed entities (there are approximately four times as many clean firms in 2019 as there were in 2005), our narrower classification approach is more likely to capture entrepreneurial activity focused on new clean energy technologies rather than investments in these sectors by existing multi-establishment firms.[3]

2.2 Employer-Employee Linkages

Having defined a set of clean and legacy establishments, we assemble a list of all workers who have ever worked for the employers that own these establishments. This is accomplished by linking the population of W-2 tax forms to the set of EINs from our employers. This set includes those in direct and supporting energy roles as well as a broader selection of those likely not working in energy but still for an EIN associated with an energy establishment. For each of these workers, we then pull all of their W-2 forms to get a full record of their employment over the 2005-2019 period.

Since the W-2 tax forms only tell us the employer of each worker, we use residential data from the Environmental Impacts Frame (EIF) to determine the likeliest establishment for each worker every single year that a W-2 is available. Longitude and latitude values are available for a majority of establishments and worker residences; for those missing exact locations in both sets, we impute estimates with the best available data. For establishments, the following data is layered in sequentially where available: publicly available Census Bureau crosswalks, time-invariant averages from all CBP-BR establishments by zip code, pooled EIF estimates by zip code, and time-invariant averages from all CBP-BR establishments by county. For workers, means of EIF residential lat/lon values by zip code + 4 are used.[4] Complete spatial estimates enable us to calculate distances between all establishments under a worker's employer and determine the closest and thus likeliest place of work. We then keep the top paying job each individual in a given year, while also retaining their total earnings

[3]Exploring the extent to which employment at clean establishments in otherwise non-clean firms (for example, if General Motors opens a battery manufacturing plant) may differ from employment in clean firms under our baseline definition is an important future avenue for research, but it requires more precise matching between employees and establishments than we are able to do here.

[4]There are approximately 56 million zip + 4 locations in the U.S. compared to approximately 44,000 zip codes.

across all jobs. In the final steps, we merge in individuals' 1040 earnings, and establishment-level and firm-level characteristics from the LBD. We create a nationally representative sample of comparison workers by following a similar procedure to create a linked panel from the set of workers we observe in the American Community Survey (ACS) from 2005-2019, removing those who have worked for an energy establishment to avoid duplicating observations.

Finally, we attach sociodemographic information to all workers in our panel by leveraging the Census Bureau's linkage infrastructure, which allows us to link individuals across datasets by unique "Person Identification Keys", or PIKs (Wagner and Layne 2014). Our sociodemographic information comes from two data sources. First, we derive information on race and ethnicity from a composite Census Bureau dataset –the Title 13 Best Race and Ethnicity File – which draws information from a combination of Decennial Census, survey and administrative data. We define five main racial groups of interest: Hispanic of any race, Non-Hispanic White, Non-Hispanic Black, Non-Hispanic AIAN and Non-Hispanic Other (combining Asian, Native Hawaiian/Pacific Islander, Other Race and multi-racial individuals). Second, we attach information on date of birth and sex from the Demographic spine of the Environmental Impacts Frame (Voorheis et al., 2023a).

3 Results

3.1 Aggregate Patterns and Trends

Between 2005 and 2019, each year there were an average of 2,587 clean establishments, 43,667 legacy establishments, 110,055 clean jobs, and 1,132,900 legacy jobs, according to our classification. They account for 0.04 and 0.6 percent of all establishments and 0.08 and 0.8 percent of all employment in the United States. While a small overall share of employment and activity, electricity is a critical input to all activity in the United States. In addition, the energy sector is a major source of greenhouse gas emissions and other local pollutants. As such, the reallocation of activity away from the production of electricity using fossil fuels towards clean electricity production has the potential to deliver far broader social benefits through climate change mitigation and improvements in local environmental quality.

Figure 1 explores aggregate trends in the number of establishments and aggregate employment in these two sectors between 2005 and 2019. Panel (a) shows an annual time series of the percent change in clean and legacy establishment counts relative to 2005. Between 2005 and 2019, the number of clean establishments increased by over 200 percent, from 1,200 in 2005 to 3,800 in 2019 (Figure 1a).[5] By contrast, the number of establishments in legacy energy sectors decreased slightly by 1.2 percent between 2005 and 2019. The rate of growth has slowed, albeit from a much higher base; there were 41,500 legacy establishments in 2005 and 41,000 legacy establishments in 2019 (Figure 1b).

Panels (c) and (d) presents trends in aggregate employment. As with the number of establishments, we see that the number of workers in both clean and legacy establishments increased significantly between 2005 and 2019. Clean establishments employed approximately 79,110 people in 2005 and 143,000 people in 2019, an 81 percent increase (Figure 1d). For comparison Curtis and Marinescu (2023) identify 66,000 wind and solar jobs postings in 2019 using Burning Glass/Lightcast data. Clean employment growth has been slower than clean establishment growth, which is consistent with new clean establishments being relatively small.

For legacy activities, the opposite is true: while the number of establishments shrunk in 2019 relative to 2005, aggregate employment has grown. The number of workers in legacy establishments grew 33 percent, from 925,800 in 2005 to 1.18 million in 2019 (Figure 1d). We are not yet in the phase of the clean energy transition where aggregate employment in legacy energy activities is declining.

3.2 The Geography of Clean and Legacy Energy Jobs

Having documented aggregate patterns and trends we next evaluate the spatial distribution of legacy and clean jobs in 2019 (Figure2). This is important given concerns that geographic mismatch could be a constraint to labor reallocation from legacy to clean activities.

Legacy energy employment is more spatially concentrated, with many states having close to no legacy employment and others having employment shares as high as 7.8% (Figure 2a). By contrast, clean employment shares are much more evenly distributed across states, and make up for less than one percent of employment even in the most concentrated market (Figure 2b).

3.3 Who Works in Clean and Legacy Firms?

Public discussion surrounding the clean transition frequently emphasizes its potential to create new clean jobs that will provide opportunities for people living in disadvantaged and underserved communities. A relevant baseline for considering the extent to which clean firms may provide opportunities in the future is to explore how emerging clean firms have

[5] All numbers are rounded in line with US Census Bureau disclosure requirements.

historically provided opportunities across demographic groups. We present an overview of the demographic composition of workers in clean and legacy firms between 2005 and 2019.

The demographic composition of workers in clean and legacy jobs over time is reported in Figure 3. The grey line represents the demographic composition of a random sample of non-energy workers. Non-Hispanic white workers are over-represented in both clean and legacy energy activities compared to workers in non-energy activities, although the gap between the white share in energy and non-energy activities has narrowed over time. Non-Hispanic Black workers are substantially under-represented in both clean and legacy activities, though the Black share of clean jobs shows to have grown faster in the last decade than for legacy jobs or non-energy jobs. Hispanic workers remain under-represented in clean activities, although the gap is narrowing over time. Compared to non-energy activities Hispanic workers are over-represented in legacy energy activities. Female workers are also extremely under-represented in both clean and legacy energy jobs, making up less than 20% of employment in both sectors. Finally, we explore how employment varies with education. College educated workers appear just as likely to be employed in clean energy jobs as in jobs outside of the energy sector. By contrast, workers in legacy energy activities are less likely to have a college degree. Workers in legacy energy activities are also less likely to have a high school diploma than workers in clean energy activities.

These findings suggest that clean energy jobs may require a more educated workforce than legacy energy, pointing to the potential for skill mismatch in impeding access to these new opportunities. In addition they suggest that the transitional costs to those working in legacy energy activities may fall disproportionately on Non-Hispanic white, Hispanic, Male, and less educated workers.

3.4 The Characteristics of Clean and Legacy Energy Jobs

Wages

First and foremost, we want to understand the extent to which there are differences in the returns to working in clean and legacy activities. Do workers in clean firms get paid more, and if so, why? To explore this we estimate the growth rate of earnings from working in a clean firm compared to a legacy firm. We estimate variations on the following specification,

$$logWages_{ijt} = \alpha + \beta_1\ Clean_{ijt} + \beta_2\ NonEnergy_{ijt} + \phi_t + \epsilon_{ijt}$$

where $logWages_{ijt}$ are the log of real earnings ($2019) from worker i's W-2 tax filing while working for establishment j in year t, $Clean_{ijt}$ is a binary indicator equal to 1 if worker i is employed at a clean establishment j in year t, and $NonEnergy_{ijt}$ is a binary indicator equal to 1 if worker i is employed at a non-energy establishment j in year t, and ϕ_t is a vector of year fixed effects, meaning that all comparisons are made across workers within a given year.

The parameters of interest are β_1 and β_2. In the primary specification, these reflect average unconditional wage differences compared to employment in legacy establishments. In the second specification, we include additional controls for education, experience, race, and sex. In the least parsimonious specification, we also include individual fixed effects, absorbing all time-invariant variation in earnings and individual characteristics, such as race, education, sex, and other unobservable characteristics such as ability and motivation. This final specification only exploits variation from workers who move between activities (i.e., from clean to non-energy, or non-energy to legacy, legacy to clean, etc.) and so while more internally valid, reflects moves from a more selected sample. Because workers have some choice in where they they work, whether they switch activities, and the timing in which they switch, the conditional wage difference between clean and legacy establishments $\hat{\beta}_1$ only represents the causal effect of working in clean firms under very stringent conditions (Mincer, 1958, 1974; Becker, 1964; Card, 1999; Heckman et al., 2006). Our goal is to simply document the unconditional and residual differences in earnings between clean, legacy, and non-energy activities to form clearer expectations about earnings opportunities and outside options.

On average, we estimate that workers earn 16 percent more in clean establishments and 62 percent less in non-energy establishments compared to legacy establishments (Table 2, column 1). These unconditional wage gaps tell us that, on average, workers earn more in clean activities and less outside of the energy sector, but this does not imply that there is a causal effect of working for a clean establishment on earnings. These differences may simply reflect differences in the composition of workers. As documents above, we show that workers in clean activities are more likely to be male, non-Hispanic White, and have higher levels of educational attainment compared to workers in legacy activities and non-energy activities, for example. When we control for education, experience, race, and sex the conditional wage gaps are markedly smaller. We estimate that, on average, workers in clean establishments earn 3 percent more than workers in legacy establishments, suggesting that the composition of workers may play an important role in explaining income differences (Table 2, column 2). The conditional wage gap between workers in non-energy establishments and legacy establishments, however, remains largely unchanged – on average, workers in non-energy activities earn 58% less than workers in legacy establishments. This conditional gap is reduced substantially, however, when we further control for worker fixed effects, capturing all time-invariant individual heterogeneity, including unobservable components such as ability and motivation (Table 2, column 3). The conditional wage gap between workers in clean and legacy establishments remains relatively stable at (2.5 percent). By contrast, the conditional wage gap between workers in non-energy activities and legacy establishments shrinks to 31 percent.

We visualize wage differences between clean and legacy jobs by looking at workers' average wages over the course of a job spell, before and after they switch jobs, in Figure 4. Dots that are not connected by dotted lines are wages of people who move either into or out of not having a job in the other period. There are two main takeaways from this figure. First, on average, workers in clean and legacy activities receive higher earnings than other jobs. Second, there are important compositional differences in wages. People who switch between clean and legacy jobs have substantially higher wages than people who switch between either clean or legacy and non-energy jobs. Whether this is because of an energy-sector specific skill match or another source of selection, it is an important fact to keep in mind when considering the potential wages of workers in clean jobs as the number of clean jobs expands.

3.5 Job Transitions

In light of these findings, an important question is who receives these higher-paying clean jobs. Do workers move into clean activities from non-energy activities, potentially resulting in substantial increases in earnings, or do they move from legacy energy activities. If workers do not transition from legacy energy activities to clean activities their outside options in non-energy activities appear to results in much lower earnings. Answering these questions, has important implications for understanding the distributional consequences of the clean energy transition.

The extent to which workers are able to move from one activity to another depends on the geographic distribution of employment opportunities, i.e, the extent to which alternative employment opportunities exist in the same labor market or the extent to which workers are willing or able to relocate to new locations, as well as the extent to which workers have the necessary skills, training, education, and experience to move into new activities. Where geographic and skill mismatch exist, it is important to understand their relative importance so as to narrow the scope of possible inefficiencies such that policy can be targeted more efficiently to alleviate frictions.

Existing research based on the geographic overlap of clean and legacy activities has concluded optimistically that there is a lot of potential for workers to transition from legacy to clean sectors; however, geographic overlap is neither a sufficient nor even a necessary condition for such a transition to occur. At the worker level, we see that the likelihood of moving from a legacy to a clean firm is extremely low – the likelihood of working at a clean firm conditional on having worked for a legacy firm the previous year is 0.45 percent. This could be mechanically driven by the fact that there are very few clean jobs available, but conditioning on current job and looking at where workers came from, we see similarly low likelihoods of legacy-to-clean transitions.

The vast majority of workers in clean activities come from outside of the energy sector. On the one hand, this suggests that the increase in earnings for these workers is likely substantial. The pool of workers outside of the energy sector are more likely to be female, Black, or Hispanic, resulting in a greater unconditional potential to address historical economic disparities than if workers transition from legacy energy activities. At the same time, however, workers in legacy legacy activities are expected to earn substantially lower earnings outside of energy activities resulting in non-trivial transition costs. At present there are more legacy jobs than clean jobs. As such. the aggregate direct labor market effects of the transition will depend on the extent to which the number of clean jobs exceed the reduction in legacy jobs as well as the extent to which any increases in earnings for new workers in clean activities exceeds reductions in earnings for workers who transition from legacy activities to non-energy sectors. Given the limited conditional earnings differences between clean and legacy firms, the limited transition across these activities suggests that geographic or skill mismatch may be severe.

In figure 6, we explore who makes the rare moves between clean and legacy activities. Each bar is a ratio between the likelihood of someone in the relevant demographic group moving to someone not in the relevant demographic group moving. White workers are about 25% less likely than non-white workers to make any job switch, but they are *more* likely than non-white workers to switch from a legacy job to a clean job. We see the opposite trend for Hispanic workers; they are disproportionately likely to move between any job, but disproportionately less likely to move legacy to clean. Female workers and Black workers are also less likely to move from legacy to clean jobs. Most strikingly, college graduates are almost twice as likely as non-college graduates to move from a legacy to a clean job. These findings are robust to controlling for the commuting zones share of clean jobs, as a control for geographic opportunities, providing suggestive evidence that skill mismatch is a relatively more important constraint than spatial mismatch in the transition from legacy to clean activities. The overlapping geography of clean and legacy jobs also suggests that information frictions, search frictions, and selection out of clean jobs due to place-based compensating differentials are unlikely to be first-order explanations. In the following section, we look more closely at the distribution of occupations within clean and legacy establishments to further evaluate the extent to which skill mismatch could impede reallocation between legacy and clean activities.

Occupational Characteristics

Using 2-digit SOC classifications, we identify the top ten occupation groups in clean and legacy firms, by educational attainment.

Figure 7 shows occupation shares at clean and legacy establishments. For college educated workers (panel (a)), occupations in the top 10 most common categories for legacy jobs are also in the top 10 for clean jobs, and while there is some rank-switching (e.g. financial specialists are the third most common legacy occupation category but the 8th most common clean occupation category), there is also substantial overlap in the rank order (e.g. the top two most common occupations in both are management followed by architecture and engineering). For workers without a college education, there is not nearly as much occupational overlap. 8 out of 10 occupations are in the top 10 for both legacy and clean jobs, but "extraction" and "food" make it into the top 10 of legacy jobs and not clean jobs, while "architecture and engineering" and "protective services" make it into the top 10 of clean but not legacy jobs. There is also more switching in the rank order of overlapping occupations: Construction, Transportation, and Production rank 1, 2 and 3 in legacy jobs, but 2, 5, and 1 in clean jobs.

These results imply that while job switching is possible, it is likely to be easier for those with college degrees due to better overlap in the distribution of occupational characteristics of jobs. For non-college educated workers, the occupational mismatch might result in more difficulty switching from legacy to clean jobs.

Job Spells

Another possibility, is that workers perceive jobs in new clean activities as potentially less stable and so prefer to receive lower earnings outside of the energy sector in return for job stability – a compensating differential. To explore this we look at job spells, which reflect a combination of job satisfaction and job stability. We argue that activities that have, on average, longer job spells typically reflect a better quality of job.

We find that the average job spell outside of the energy sector is one year shorter than the average job spell in legacy activities and two years shorter than the average job spell in clean activities. This is despite the fact that we expect clean job spells to be mechanically biased downward since many clean jobs appear in later years in our data. These findings are largely inconsistent with the possibility that new clean jobs are less stable and if anything suggest that clean jobs are better quality jobs than non-energy and legacy-energy jobs.

Conclusion

Our findings add to our understanding of the labor market effects and distributional consequences of transitioning to a low-carbon economy. First, our findings suggest that the clean transition will have little impact on overall inequality in the United States. Workers in both the clean and legacy energy sectors make up a very small proportion of total employment, have similar demographic compositions, and earn comparable wages. As a result, the clean transition will most likely have a minor net effect on income distributions. Second, as the number of clean jobs increases, our findings indicate that they are more likely to be filled by workers outside of the energy sector. As a result, while the benefits of the transition will be diffuse, the economic costs may be more concentrated – workers who have held clean or legacy jobs earn significantly less outside of the energy sector. Individually, these costs have not historically been mitigated by the availability of new jobs in the expanding clean energy sector, and those who have transitioned from legacy to clean establishments have tended to be more educated and White. As a result, the costs of the clean transition may fall disproportionately on less educated and minority populations.

References

Acemoglu, Daron, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 2002, *40* (1), 7–72.

_ **and David H. Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Vol. 4 2011, pp. 1043–1171.

_ **and Pascual Restrepo**, "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 2019, *33* (2), 3–30.

_ **and** _ , "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016.

Autor, David, "The Labor Market Impacts of Technological Change: From Unbridled Enthusiasm to Qualified Optimism to Vast Uncertainty," Working Paper 30074, National Bureau of Economic Research May 2022.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–2168.

_ , **Lawrence F. Katz, and Melissa S. Kearney**, "The Polarization of the U.S. Labor Market," *American Economic Review*, 2006, *96* (2), 189–194.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis, "The Evolution of Working from Home," 2023.

Becker, Gary, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, National Bureau of Economic Research, New York. Distributed by Columbia University Press, 1964.

Bowen, Alex, Karlygash Kuralbayeva, and Eileen Tipoe, "Characterising green employment: The impacts of 'greening' on workforce composition," *Energy Economics*, 2018, *72.*

Card, David, "The Causal Effects of Education on Earnings," *Handbook of Labor Economics*, 1999, *5.*

Carley, Sanya and David Konisky, "The Justice and Equity Implications of the Clean Energy Transition," *Nature Energy*, 2020, *5.*

Colmer, Jonathan, Eleanor Krause, Eva Lyubich, and John Voorheis, "Transitional Costs and the Decline of Coal:
Worker-Level Evidence," *Working Paper*, 2023.

Consoli, Davide, Giovanni Marin, Alberto Marzucchi, and Francesco Vona, "Do green jobs differ from non-green jobs in terms of skills and human capital?," *Research Policy*, 2016, *45* (5), 1046–1060.

Curtis, Mark and Ioanna Marinescu, "Green Energy Jobs in the United States: What Are They, and Where Are They?," *Environmental and Energy Policy and the Economy*, 2023, *4.*

Deming, David J., "The Growing Importance of Social Skills in the Labor Market," *The Quarterly Journal of Economics*, 2017, *132* (4), 1593–1640.

Emanuel, Natalia and Emma Harrington, "Working Remotely? Selection, Treatment, and the Market for Remote Work," FRB of New York Staff Report 1061, Federal Reserve Bank of New York 2023.

_ , _ , **and Amanda Pallais**, "The Power Of Proximity To Coworkers: Training for Tomorrow or Productivity Today," Working Paper 2023.

Goldin, Claudia and Lawrence F. Katz, "The Origins of Technology-Skill Complementarity," *The Quarterly Journal of Economics*, 1998, *113* (3), 693–732.

_ **and** _ , *The Race Between Education and Technology*, Harvard University Press, 2008.

Heckman, James, Lance Lochner, and Petra Todd, "Earnings Functions, Rates of Return and Treatment Effects," *Handbook of the Economics of Education*, 2006, *1.*

Katz, Lawrence F. and David H. Autor, "Changes in the Wage Structure and Earnings Inequality," *Handbook of Labor Economics*, 1999, *3*, 1463–1555.

_ **and Kevin M. Murphy**, "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Korinek, Anton and Joseph E. Stiglitz, "Artificial Intelligence and its Implications for Income Distribution and Unemployment," in "The Economics of Artificial Intelligence: An Agenda," University of Chicago Press, 2018, pp. 349–390.

Lin, Jeffrey, "Technological Adaptation, Cities, and New Work," *The Review of Economics and Statistics*, 2011, *93* (2), 554–574.

Mincer, Jacob, "Investment in human capital and personal income distribution," *Journal of Political Economy*, 1958, *66* (4), 281–302.

_ , *Schooling, Experience and Earnings*, National Bureau of Economic Research, New York. Distributed by Columbia University Press, 1974.

Park, J., M. Curtis, and L. O'Kane, "Workers and the Green-Energy Transition: Evidence from 300 million Job Transitions," *Working Paper*, 2023.

Rutzer, Christian, Matthias Niggli, and Rolf Weder, "Estimating the Green Potential of Occupations: A New Approach Applied to the U.S. Labor Market," *WWZ Working Paper*, 2020, (2020/03).

Tinbergen, Jan, "Substitution of Graduate by Other Labor," *Kyklos*, 1974, *27* (2), 217–226.

Vona, Francesco, Giovanni Marin, and Davide Consoli, "Measures, drivers and effects of green employment: evidence from US local labor markets, 2006-2014," *Journal of Economic Geography*, 2019, *19* (5), 1021–1048.

_ , _ , _ , **and David Popp**, "Environmental Regulation and Green Skills: An Empirical Exploration," *Journal of the Association of Environmental and Resource Economists*, 2018, *5* (4), 713–753.

Voorheis, J., J. Colmer, K. Houghton, E. Lyubich, C. Scalera, and J. Withrow, "Building the Prototype Census Environmental Impacts Frame," *NBER Working Paper No. 31189*, 2023.

_ , _ , _ , _ , **M. Munro, C. Scalera, and J. Withrow**, "Building the Prototype Census Environmental Impacts Frame," *NBER Working Paper No. 31189*, 2023.

Wagner, Deborah and Mary Layne, "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software," *CARRA Working Paper No. 2014-01, Center for Economic Studies, U.S. Census Bureau*, 2014.

Tables and Figures

Table 1: Legacy and Clean Sector Classification

Legacy Activities		Clean Activities		Potentially Clean Activities	
Industry	Description	Industry	Description	Industry	Description
211120	Crude Petroleum Extraction	221111	Hydroelectric Power Generation	236118	Residential Remodelers
211130	Natural Gas Extraction	221113	Nuclear Electric Power Generation	237130	Power and Communication Line and Related Structures Construction
212111	Bituminous Coal and Lignite Surface Mining	221114	Electric Power Generation, Solar	238160	Roofing Contractors
212112	Bituminous Coal Underground Mining	221115	Electric Power Generation, Wind	238210	Electrical Contractors and Other Wiring Installation Contractors
212113	Anthracite Mining	221116	Geothermal Electric Power Generation	238220	Plumbing, Heating, and Air-Conditioning Contractors
213111	Drilling Oil and Gas Wells	221117	Biomass Electric Power Generation	238990	All Other Specialty Trade Contractors
213112	Support Activities for Oil and Gas Operations	335911	Storage Battery Manufacturing	334413	Semiconductor and Related Device Manufacturing
213113	Support Activities for Coal Mining			333611	Turbine and Turbine Generator Set Units Manufacturing
221112	Fossil Fuel Electric Power Generation			423690	Other Electronic Parts and Equipment Merchant Wholesalers
221210	Natural Gas Distribution			423720	Plumbing and Heating Equipment and Supplies
237120	Oil and Gas Pipeline and Related Structures Construction				(Hydronics) Merchant Wholesalers
324110	Petroleum Refineries			423730	Warm Air Heating and Air-Conditioning Equipment
423520	Coal and Other Mineral and Ore Merchant Wholesalers				and Supplies Merchant Wholesalers
424710	Petroleum Bulk Stations and Terminals			444190	Other Building Material Dealers
424720	Petroleum and Petroleum Products Merchant Wholesalers			454390	Other Direct Selling Establishments
454310	Fuel Dealers			541330	Engineering Services
486110	Pipeline Transportation of Crude Oil			541690	Other Scientific and Technical Consulting Services
486210	Pipeline Transportation of Natural Gas			811310	Commercial and Industrial Machinery and Equipment
486910	Pipeline Transportation of Refined Petroleum Products				(except Automotive and Electronic) Repair and Maintenance

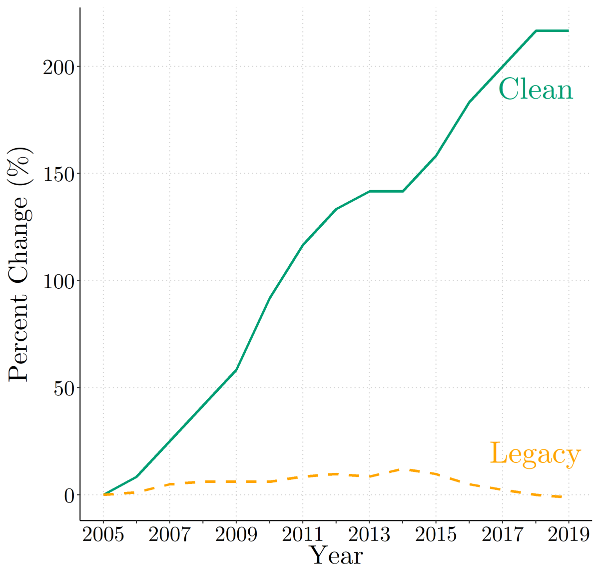
Notes: This table documents the full selection of North American Industry Classification System (NAICS) codes used in defining clean and legacy industry establishments.

Table 2: Regression of Log Wages

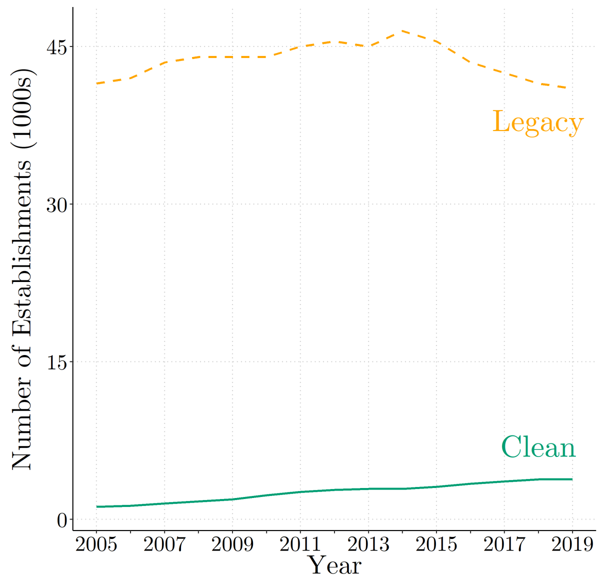
	(1)	(2)	(3)
Clean	0.162***	0.0296***	0.0247***
	(0.00405)	(0.00484)	(0.00448)
Neither	-0.617***	-0.584***	-0.309***
	(0.00172)	(0.00208)	(0.00140)
Observation	8,738,000	8,738,000	8,738,000
Year FEs	X	X	X
Individual Controls		X	X
Individual FEs			X

Notes: This table presents unconditional and conditional wage gaps between clean, legacy, and non-energy activities. We restrict our analysis to the sub-sample of workers who have ever responded to the ACS to ensure a common sample size across columns. We do this because we only have education information for these workers, which we show is important in explaining the wage gap for workers between clean and legacy activities. Results for column 1 and 3 are similar if we use the full set of workers. Year fixed effects are included in all specifications. In column 2 we include individual controls. These include indicators for high school completion and some college, linear and quadratic terms for experience, indicators for race, and and an indicator for sex. In column 3, we use individual fixed effects, identifying gaps off workers who move between clean, legacy, or non-energy activities. Significance levels are indicated as * 0.10 ** 0.05 *** 0.01. Standard errors are clustered at the worker level.

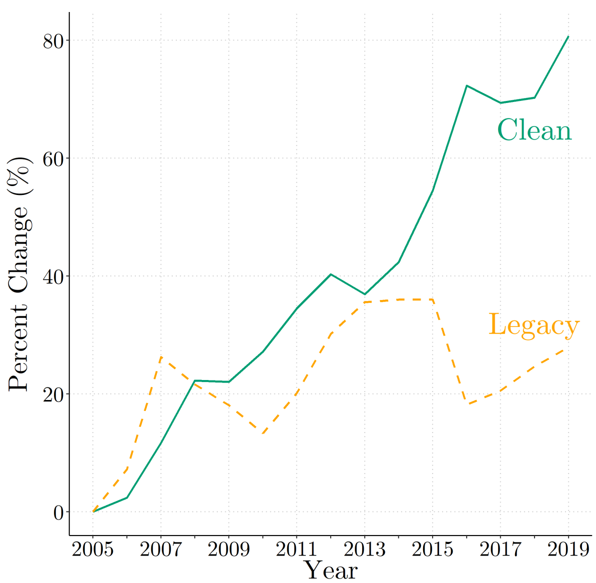
Figure 1: Aggregate Trends in Establishments and Jobs



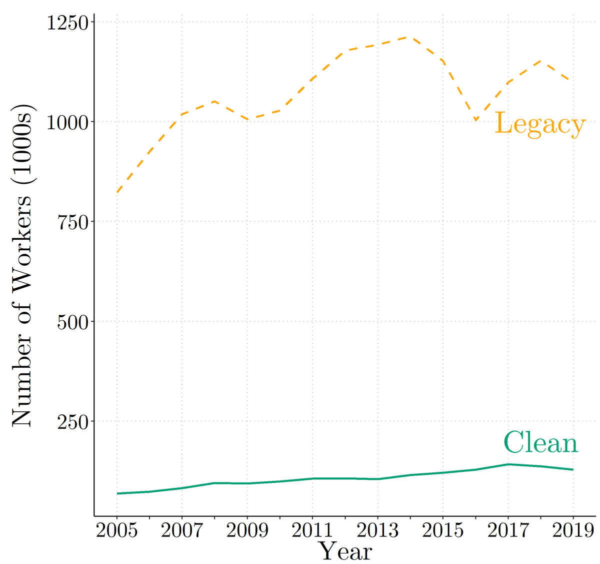
(a) Growth in Establishment Counts



(b) Aggregate Establishments



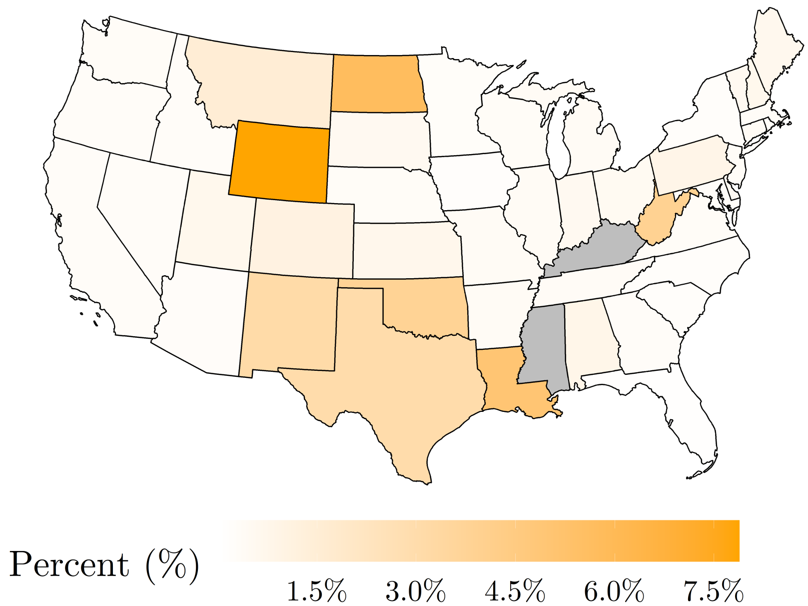
(c) Growth in Employment



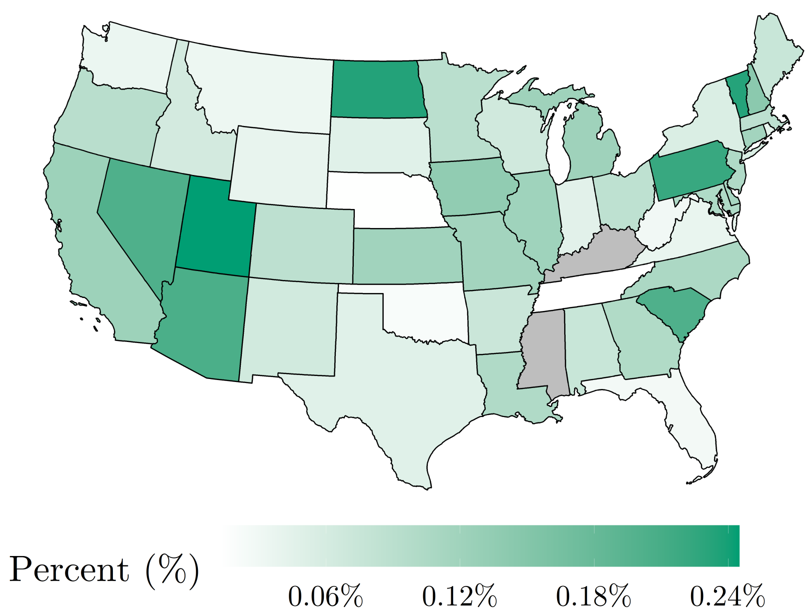
(d) Aggregate Employment

Notes: This figure shows aggregate trends in clean (solid green line) and legacy (dashed orange line) energy establishments and employment over 2005-2019. Panel a) presents the percent change in the number of establishments between 2005 and 2019. Panel b) presents the aggregate number of establihsments over time. Panel c) presents the percent change in number of workers between 2005 and 2019. Panel d) presents the aggregate number of workers over time.

Figure 2: The Spatial Distribution of Employment (2019)



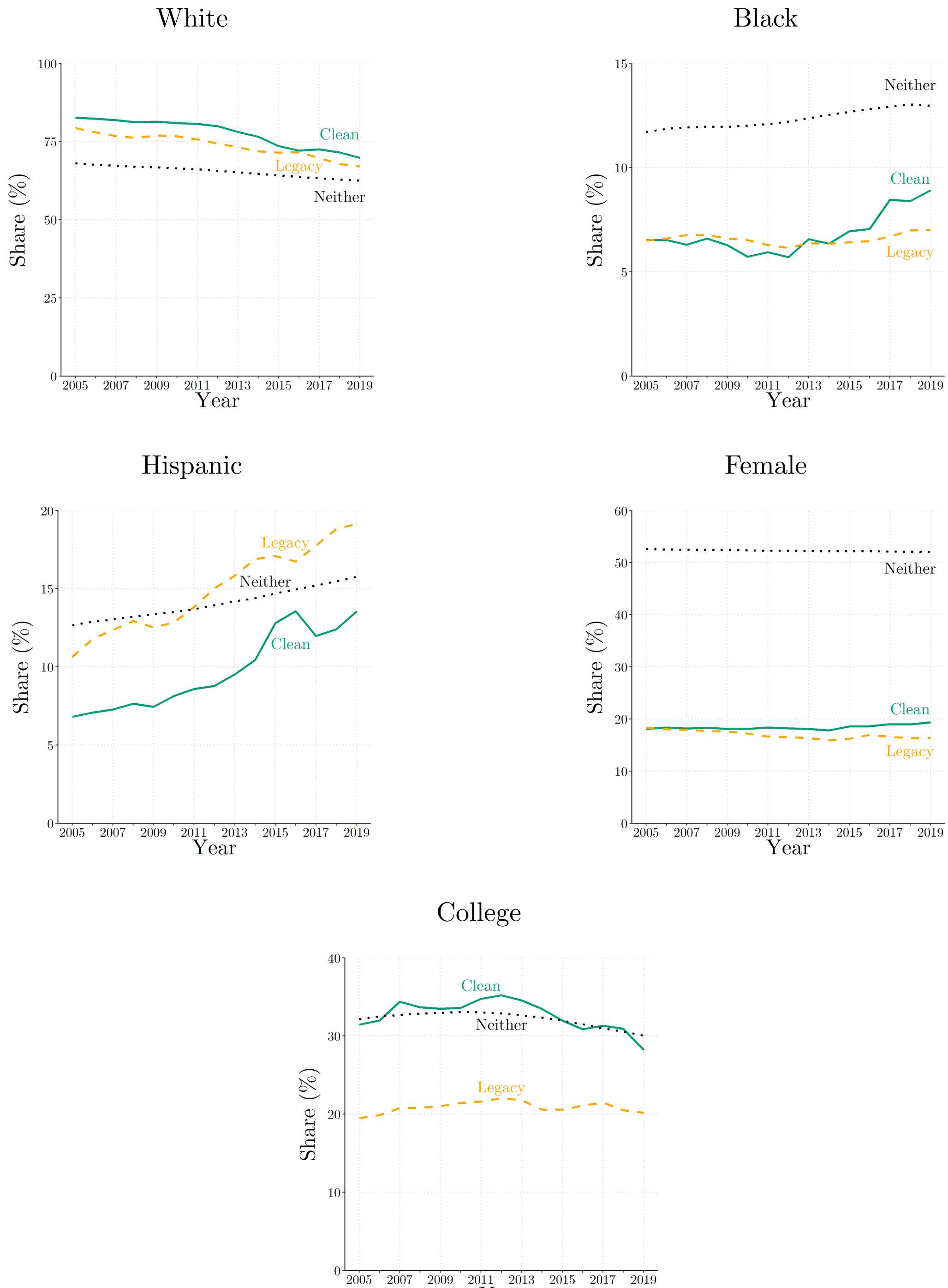
(a) Legacy Share of State Employment



(b) Clean Share of State Employment

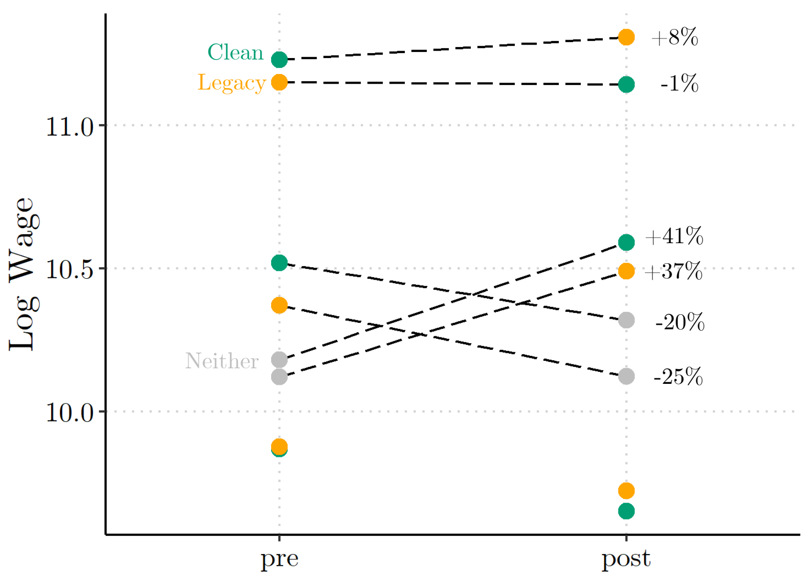
Notes: This figure presents the spatial distribution of employment across states. Panel a) presents the share of employment accounted for by workers in legacy energy activities. Panel b) presents the share of employment accounted for by workers in clean energy activities. The shares are calculated by taking the total number of employees of in each activity in each state and dividing this by the total number of employees across all within the state. Data for Kentucky and Mississippi are yet to be disclosed.

Figure 3: Demographic Shares



Notes: This figure shows the share of workers participating in clean energy (solid green line), legacy energy (dashed orange line), and non-energy (dotted black line) activities between 2005 and 2019 by demographic group.

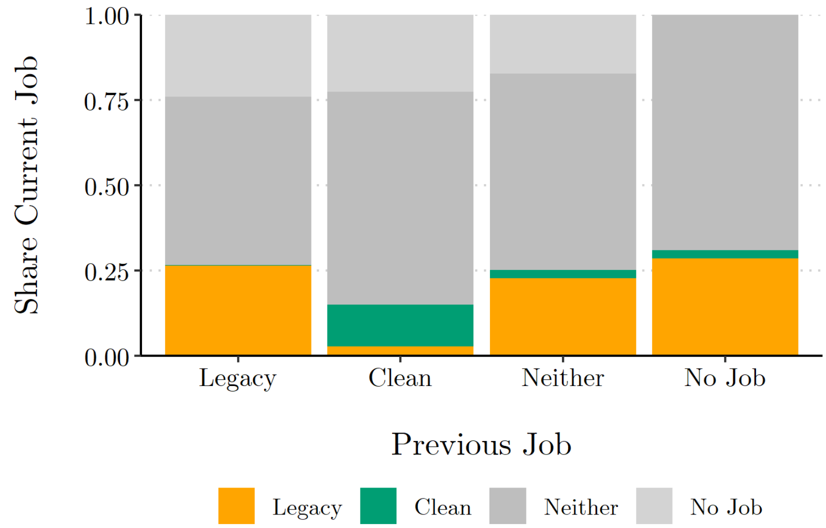
Figure 4: Wages



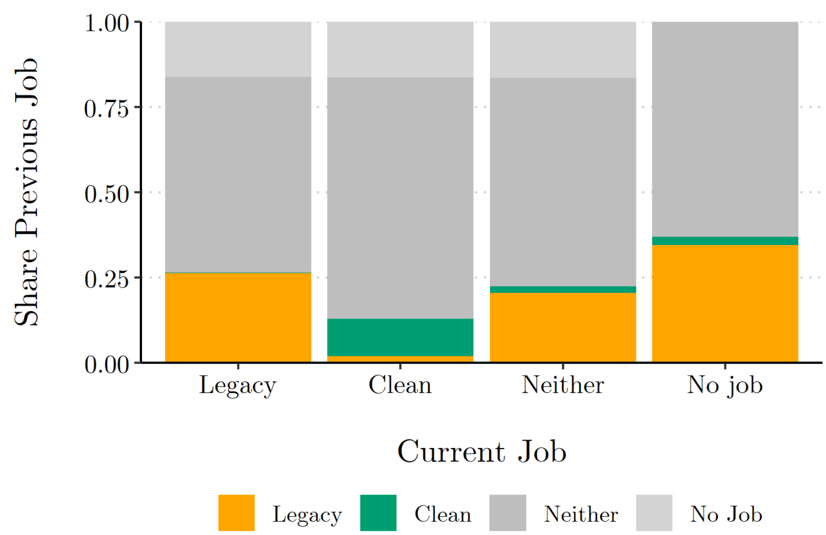
Notes: This figure shows average within-worker changes in wages before and after a transition between clean, legacy, and non-energy activities. Log wages for clean (green), legacy (orange), and non-energy (gray) activities are shown pre-transition and are connected to their respective post-transition values with percent changes labeled. Points that are not connected represent transitions into or out of non-employment.

Figure 5: Do Workers in Legacy Firms Move to Clean Firms?

(a) Transition Likelihoods by Previous Job

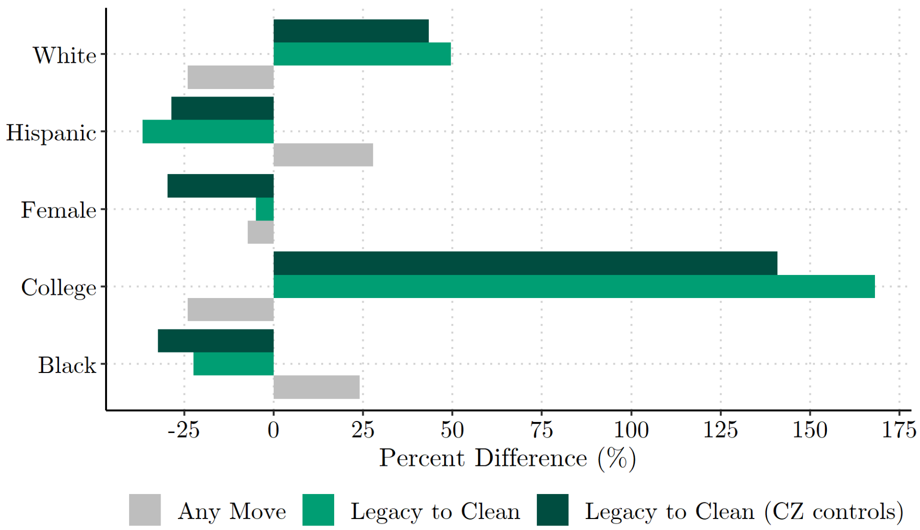


(b) Transition Likelihoods by Current Job



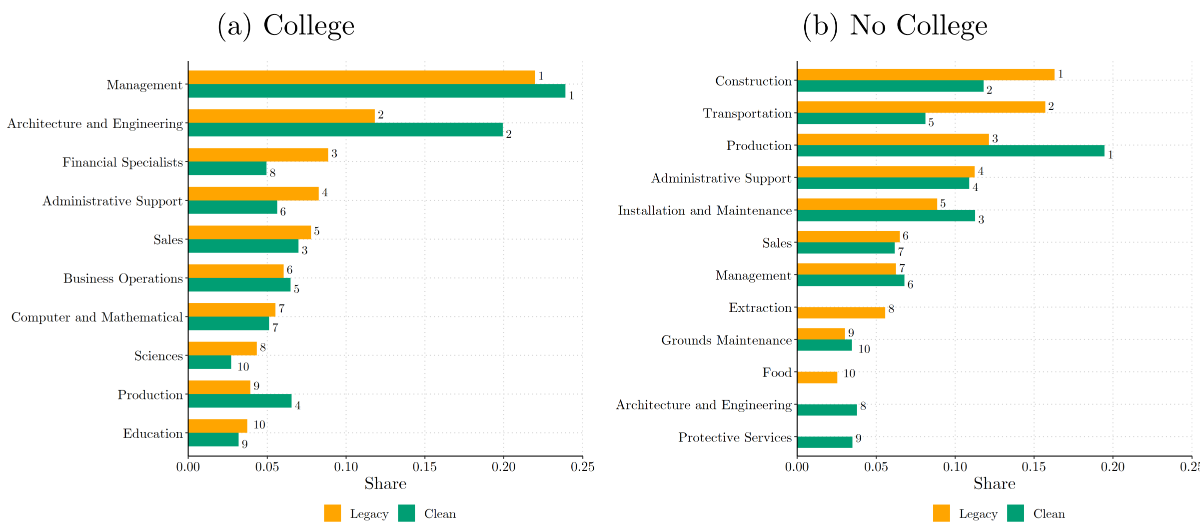
Notes: This figure shows transition probabilities across different activities. The likelihood of moving to any establishment is measured as a employer-employee separation. Panel a) reports the share of workers whose current position is in legacy energy activities (orange), clean energy activities (green), non-energy activities (dark gray), or not employed (light gray), separated by whether their previous position was legacy, clean, neither, or not employed. This tells us, for example, the likelihood of moving to a clean job conditional on leaving a legacy job, $p(clean_t|legacy_{tâ1})$. Panel b) reports the share of workers whose previous position was in legacy energy activities, clean energy activities, non-energy activities, or not employed, separated by whether their current position is legacy, clean, neither, or not employed. This tells us, for example, the likelihood of moving from a legacy job conditional on currently working in a clean job, $p(legacy_{tâ1}|clean_t)$.

Figure 6: Transition Probabilities by Demographics



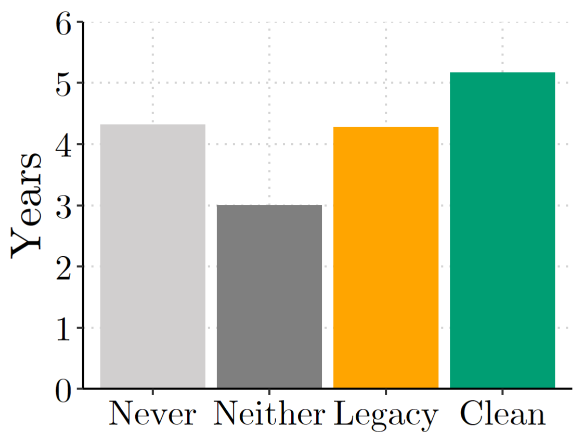
Notes: This figure shows the probability of a worker transition conditional on their demographic characteristics. The gray bar shows the likelihood of making any transition. The light green bar shows the likelihood of transitioning from legacy energy activities to clean energy activities. The dark green bar shows the likelihood of transitioning from legacy energy activities to clean energy activities conditional on commuting zone-level controls capturing the availability of clean jobs.

Figure 7: Top Occupations



Notes: This figure shows the top 10 occupation categories for workers in clean energy activities (green bars) and legacy energy activities (orange bars). Panel a) restricts the sample to college educated workers. Panel b) restricts the sample to workers without any college experience. The individual bars in each panel are labeled with the ranking of each occupation within each activity.

Figure 8: Job Spells



Notes: This figure shows the average length of time employees spend working for an establishment by type of activity. Workers from the "Never" (light gray) category come from our nationally representative American Communities Survey (ACS) sample. This category includes all economic activities that are not energy related. Workers from the "Neither" (dark gray), "Legacy" (orange), and "Clean" (green) categories come from our core sample of workers. "Neither" reflects the average job spell for workers in non-energy activities, conditional on them having worked in energy activities at some point during the sample. "Legacy" reflects the average job spell for workers within legacy energy activities. "Clean" reflects the average job spell for workers within clean energy activities.