UNIVERSITY OF CALIFORNIA

Los Angeles

Power and Adaptation to Climate Change

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Environment and Sustainability

by

Paul Grether Stainier

2024

ABSTRACT OF THE DISSERTATION

Power and Adaptation to
Climate Change

by

Paul Grether Stainier
Doctor of Philosophy in Environment and Sustainability
University of California, Los Angeles, 2024
Professor Alan Irwin Barreca, Co-Chair
Professor Michael L. Prelip, Co-Chair

The extent to which climate change will exacerbate already growing inequality, both across and within locations, is an open question. The power that people have, or lack, to adapt to a warming climate is central to this issue. This dissertation includes three studies on potential climate adaptations. In the first, my coauthors and I study whether heat-induced use of energy-intensive adaptation technologies, such as air conditioning, can lead to financial distress. We test for this possibility with data from California on electricity use and disconnections, a consequence of utility bill non-payment. We find that hot weather increases electricity expenses in the current billing period and the relative risk of disconnection 51 to 75 days later. In the second study, my coauthors and I examine the impact of temperature-related crop losses on household diets in rural India, a setting with a high prevalence of small family farms. High temperatures during the growing season reduce crop yields, but it is unclear how these losses affect household diets. While we find no significant impact of heat on average calorie or iron consumption in the subsequent year, the number of extremely malnourished households increases. We also find suggestive evidence that

households adapt to heat-induced losses of home-grown calories by purchasing more food, which helps to explain the lack of aggregate impacts. In the third study, I analyze worker responses to climate change. Past work in this area has largely focused on the on-the-job effects of extreme weather on environmentally exposed workers and potential short-term adaptations. Little is known about long-term adaptations, such as changing occupations. Well-documented challenges to occupational mobility, especially between occupations with different task requirements, suggest that this adaptation strategy may be prohibitively costly for many workers. Using individual-level panel data from France, I find that historically, inter-exposure mobility rates are low. The task composition of high exposure jobs provides a partial, but incomplete, explanation for this labor market segmentation.

The dissertation of Paul Grether Stainier is approved.

Karen McKinnon

R. Jisung Park

Michael L. Prelip, Committee Co-Chair

Alan Irwin Barreca, Committee Co-Chair

University of California, Los Angeles

2024

To David M. Grether

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank my dissertation chairs, Alan Barreca and Michael Prelip. I remain in awe of how you move through academic spaces with humility and joie de vivre, but without sacrificing rigor. I thank the members of my dissertation committee, Jisung Park and Karen McKinnon, for their advice and guidance. I also thank Manisha Shah, an excellent coauthor and mentor.

Chapter 2 is coauthored with Professors Alan Barreca and R. Jisung Park and has been published in the journal *Nature Energy*. The citation is: Barreca, A.; Park, R.J.; Stainier, P. (2022). "High temperatures and electricity disconnections for low-income homes in California." *Nature Energy*, 1-13. We all designed the research, interpreted the data, and wrote the paper. Professor Barreca and I analyzed the data. Chapter 3 (excluding Section 3.6 "Extension: Climate Projections")is coauthored with Professors Manisha Shah and Alan Barreca. We have submitted it for review to *American Economic Review: Insights*. We all designed the research, interpreted the data, and wrote the paper. I analyzed the data.

I acknowledge financial support from the Dissertation Year Fellowship.

This dissertation also benefits from advice, research, data, and code from many others. I thank Rachel Sheinberg, Will Krantz, May Wang, Saanchi Shah, Lilly Nhan, Roch Nianogo, Linghui Jiang, Anna Stansbury, Thomas Le Barbanchon, Gio Righi, Ina Drouven, Teevrat Garg, Vis Taraz, Maulik Jagnani, Urs Beyerle, Jan Sedlacek, and Duo Chan.

I thank IoES students, especially my cohort: Viraj, Dani, Alexandria, Rob, Regina, and Maggie. I thank the IoES staff. I am especially indebted to Harrison Levy, our fearless leader, and Joan Koyama, Robert Kunst, and Vanessa De La Rosa who helped me access the CASD data.

Thanks to my friends who have kept Los Angeles fresh: Julien, Akul, Dylan, Richard, Ciana, Allie, Josh, Amit, Angarika. Keep it fresh.

Thanks to William, Megan, Didier, and Gran for listening to countless project ideas, no matter how preliminary. Merci à Mamy d'avoir corrigé mon français pour la demande de

données du CASD. Thank you to Kamton and Edna for being my home base in West LA.

Finally, thank you to Alix Lee Joe for keeping my feet firmly planted on the ground. And thank you to Miley for not caring at all about my dissertation. Philadelphia awaits.

VITA

2020-2024	Teaching Fellow/Associate, UCLA Institute of the Environment and Sustainability
2019-2020	Graduate Student Researcher, UCLA Institute of the Environment and Sustainability
2018-2019	Research Assistant, UCLA Institute of the Environment and Sustainability
2018	B.A. in Applied Mathematics and Food Systems, Harvard College

PUBLICATIONS

Barreca, A.; Park, R.J.; Stainier, P. (2022). High temperatures and electricity disconnections for low-income homes in California. *Nature Energy*, 1-13.

Callejas, I.A.; Huang, L.; Cira, M.; Croze, B.; Lee, C.M.; Cason, T.; Schiffler, E.; Soos, C.; Stainier, P.; Wang, Z.; Shaked, S.; McClellan, M.; Hung, W.; Jay, J. (2023) Use of Google Earth Engine for Teaching Coding and Monitoring of Environmental Change: A Case Study among STEM and Non-STEM Students. *Sustainability* 15, 11995.

Eshel, G.; Stainier, P.; Swaminathan, A. (2019). Environmentally optimal, nutritionally sound, protein and energy conserving plant based alternatives to US meat. *Scientific reports*, 9(1), 1-11.

1

Introduction

In the past two decades, economists have become increasingly interested in understanding the impacts of extreme weather events. While rooted in a broader study of the relationship between natural resources and the economy, the pace of this work has increased due to impending climate change (Carleton and Hsiang, 2016). Two important developments have made this research possible. First, advances in computational power have allowed researchers to use large datasets, important for identifying signals in noisy real-world data.

Second, applied economics has become increasingly interested in, and adept at, identifying causal relationships. The "credibility revolution" refers to the field's heightened emphasis on causal inference based on careful research design. At the core of this methodology is a desire to emulate, as closely as possible, the randomized experiment framework used in the natural sciences. This approach entails either running experiments or mimicking them where administering one would be impossible or unethical (Angrist and Pischke, 2010). In the latter case, economists exploit "natural quasi-experiments," such as a minimum drinking age. For example, Carpenter and Dobkin (2009) measures the effect of drinking on mortality in young adults. The authors exploit the fact that in the US, right after turning 21, people drink more alcohol. By comparing mortality rates immediately before and after this birthday, they estimate the impact of this increased drinking on mortality. The assumption here is that, within a narrow enough age range (e.g., 1 month on either side of 21), people are the same other than having received the "treatment" of being legally allowed to drink.

The economics literature on the causal impact of heat on workplace injuries typifies the strengths of this methodology. Characterizing an injury as "caused" by the heat, usually an

indirect contributing factor, is a difficult task. For example, heat-related dehydration can decrease worker attentiveness, resulting in a fall from a ladder and a fractured wrist. The direct cause of this injury, and the only one listed on an official report, would be "fall from ladder." So, counting the number of injuries where heat is listed as the direct cause, such as heat stroke, would vastly undercount the actual number of injuries caused by heat (Park et al., 2021). However, counting the number of injuries that happen on hot days would lead to an overestimate, as the worker may have fallen even if it were cold.

To tackle this question, applied microeconomists use a combination of big administrative datasets and causal inference. Big datasets, such as workers' compensation records as in Park et al. (2021) or Dillender (2021), allow researchers to isolate relatively small signals from surrounding noise. Workplace injuries are a relatively rare occurrence: the mean number number of injuries per zip code-day in California is 1, with a median of 0 (Park et al., 2021). Therefore, a 10 percent increase in injury risk due to a day above 100°F, as found in Park et al. (2021), represents just a single additional injury every 10 days. Identifying this effect with qualitative methods, or even with a smaller dataset, is difficult. However, a dataset that includes millions of injury records, with spatial and temporal granularity (e.g., zip code by day) and diversity (e.g., 21 years and all of California), is well-suited to this task.

To identify causality, environmental economists exploit the fact that, holding climate constant, weather is effectively "randomly assigned" to a given location and date (Deschênes and Greenstone, 2007). By controlling for average levels of heat and injuries across locations and over time, researchers isolate these random heat fluctuations and corresponding fluctuations in injuries. Estimating the relationship between these two sets of fluctuations identifies the causal effect of heat on injuries. Using this methodology, the field has estimated the causal effects of heat, cold, and precipitation on a vast a array of outcomes, such as mortality (Deschênes and Greenstone, 2011), learning proficiency (Park et al., 2020), and energy use (Auffhammer and Mansur, 2014).

These methods can be particularly useful for policy evaluation. We may intuitively

know that heat makes one more likely to fall off a ladder, and a qualitative account can provide more than enough anecdotes to substantiate that intuition. The role of applied microeconomics is to establish the size of this relationship. Understanding how much injury risk increases due to a day above 100°F allows one to calculate the costs of heat-attributable injuries, including workers' compensation payments, lost productivity, and lost quality of life for workers. With this estimate, a policymaker can then balance these costs against the potential costs (primarily to firms but also to workers, for example through reduced wages) of a policy such as California's 2006 Heat Illness Prevention Standard, which mandates breaks and access to water when the temperatures rise above 80°F (Riley et al., 2012).

One critique of the credibility revolution is that its insistence on causal inference limits its ability to study "potentially important but slow-acting mechanisms" (Deaton, 2024). Climate change is exactly one such mechanism. Clean causal inference often relies on short-run weather variation, thereby limiting applied microeconomists' ability to research the potential effects of slow changes in the climate. The most commonly accepted approaches in this field make it difficult to account for long-term adaptations to climate change. For example, as the climate gets warmer and standards of living rise, historically cooler or poorer areas might increasingly adopt air conditioning (AC) technologies. This increased adoption could mitigate the impacts of heat on social outcomes in the future.

The costs of adaptation make anticipating the extent to which people will adapt difficult. Installing or running the AC costs money. How individuals balance these costs with the benefits of thermal comfort will determine how much they protect themselves from the heat. Importantly, both the need and the ability to adapt to extreme heat vary substantially across people within the same location. For example, past work shows that workers in climate-exposed industries decrease the amount of time they work on hot days. In contrast, workers in non-exposed industries, such as law, do not need to, as they work in climate-controlled offices. However, not all exposed workers can necessarily adapt in this way, as some may not have sufficient control over their schedules (Graff Zivin and Neidell, 2014). Understanding these multiple determinants of adaptation, both across and within geographies, is important

to anticipating the total costs of climate change and designing policies to protect those most vulnerable to extreme weather events.

Several complementary approaches attempt to bridge the gap between causal estimates based on short-run weather variation and projections of climate change-induced damages. One such strategy studies specific technological or behavioral adaptations. For example, Barreca et al. (2016) finds that the effect of heat above 80°F on mortality in the US declined by 75 percent over the course of the 20$^{\text{th}}$ century, almost exclusively due to increased AC use. This finding, combined with projections of future levels of AC use, can be used to predict potential changes in the temperature-mortality relationship in the future. Another approach estimates the temperature-mortality relationship as a function of local wealth and climate (Carleton et al., 2022). In doing so, the authors proxy for any adaptations that might result from slowly shifting climates or rising incomes. They combine their results with sub-country projections of future wealth and climate to calculate climate change-induced mortality. However, they must make the relatively strong assumption that the temperature-mortality relationship is a deterministic function of wealth and climate.

With the "long differences" approach, researchers most directly mimic climate change by estimating the effects of long-term changes in the climate and long-term changes in outcomes. Researchers have used this method for outcomes such as crop yields (Burke and Emerick, 2016) and worker sectoral allocation (Liu et al., 2023). Burke and Emerick (2016) calculate 5-year average temperatures and yields in counties in the American Midwest for two time periods centered around 1980 and 2000. They take the differences in the average temperatures and crop yields between these two periods, then estimate the relationship between these differences. Causal interpretation of results using this approach requires stronger assumptions than when exploiting short-run variation in the weather. Here, a key assumption is that, within each state, there are no time-varying unobservables affecting crop yields that are correlated with long-term changes in average temperature. While the authors substantiate this assumption for their case, it may not always hold true for other outcomes of interest.

These approaches all suffer from a common limitation of applied microeconomics research: a reliance on quantitative methods that limits questions to those where data exist. For example, Barreca et al. (2016) uses mortality records dating back to 1900, a luxury that exists for few outcomes of interest. In addition, Carleton et al. (2022) includes no data from the poorest parts of the world, including the entire continent of Africa. Importantly, there are many areas that are projected to be poorer in 2100 than the poorest region included in their study. Therefore, for their future climate damage projections, they need to extrapolate the temperature-mortality relationship as a function of wealth and climate to these regions. This limitation motivates turning to qualitative disciplines to supplement one's understanding of the potential costs to climate adaptation. While the methods I employ in this dissertation are firmly rooted in applied microeconomics, the questions I ask are inspired by ethnographic studies that highlight potential barriers to adaptation (Blanchette, 2020; Hatton, 2020).

This dissertation includes three studies that contribute to the literature on the costs of climate adaptation, and how they might differ within the same location. Each study focuses on a different geographical context and social outcome: energy use and financial distress in California, food availability in rural India, and worker welfare in France. In all three studies, I use applied microeconomics methodologies to answer questions relevant to the distributional costs of climate change.

Study 1

The first study examines how low-income households in California change their energy use in response to extreme heat. AC protects people from the worst effects of heat, such as increased mortality (Barreca et al., 2016) and decreased learning ability (Park et al., 2020). However, the costs of increased AC use during heat spells may cause financial distress. To test this possibility, my coauthors and I match over 13 million bills from 300,000 low-income households with contemporaneous weather data at the zip code level. In line with past work (Auffhammer and Mansur, 2014), we find that households consume more electricity when it

is unusually hot. This increased energy use results in higher energy bills, providing a cost of this adaptation in dollar terms. For example, a day with a maximum temperature above 100°F increases a household's monthly energy bill by around $3.

The above story, however, misses the fact that households may differ in their power to afford these unusually high energy bills. Many people would hardly notice a $3 increase per day above 100°F. However, for low-income households, these additional costs can be enough to cause financial hardships. We use disconnection events, which refer to households being disconnected from their utility for non-payment, as a measure of financial distress. We find that heat during a particular billing period increases disconnections when the corresponding bill is due. A day above 100°F leads to 6.2 additional disconnections per 100,000 bills, an increase of 2.5% relative to the baseline risk of disconnection.

This finding highlights the strength of applied microeconomics. Even knowing that heat increases energy bills, it might be hard for many to believe that this effect is large enough to cause financial harm. With over 13 million energy bills, we harness enough statistical power to identify a meaningful but quantitatively small effect: 6.2 more disconnections per 100,000 bills. This increase would likely be difficult to identify using qualitative methods.

This study shows the importance of within-geography analyses of climate adaptation. For high-income Californians, the effect of temperature on energy bills provides an estimate of how much people are willing to pay for thermal comfort. In contrast, for low-income households, such an analysis would understate their true willingness to pay for protection from the heat. First, increased energy bills alone do not account for the costs of a disconnection, which can cause significant emotional and financial distress (Harrison and Popke, 2013). Second, households may be under-utilizing their AC, relative to their true thermal preferences, to avoid increased costs. Third, households may be making other tradeoffs, such as shrinking their spending on food, to meet their energy demands (Bhattacharya et al., 2003).

Study 2

In the second study, my coauthors and I estimate the impacts of extreme heat during the growing season (June-December) on diet quality in rural India. Relative to the low-income households in California, people here have even less power to adapt to extreme heat. In 2012, air-conditioning ownership in India was only 12 percent, with even lower rates in rural areas (Pavanello et al., 2021). In addition, the prevalence of small family farms means that hot weather, which decreases crop yields, is a direct threat to their sources of food and income (Taraz, 2018). Despite these challenges, households may still have options to respond to heat shocks, for example by using savings to purchase foods. The magnitude of these responses is an important determinant of the impacts of heat on welfare in rural India.

We study this question using 300,000 rural households' responses to India's National Sample Survey (NSS) from 2003 to 2012. We pair these responses with weather data at the district-year-season level. Despite its negative impacts on crop yields, we find no effect of extreme heat during the growing season on aggregate calorie or iron consumption in the subsequent year. We also find evidence to suggest that, in response to a heat-induced decrease in home-grown food consumption, households purchase more foods than usual. This response helps explain why we see no effects of heat on aggregate measures of diet quality. However, not all households have the power to adapt in this way: while extreme heat does not increase the number of households experiencing undernourishment, it does increase the number of households experiencing extreme undernourishment[1].

This study, which benefited from the advice of public health and climate science researchers, illustrates the strengths of interdisciplinary work. My co-advisor, Professor Prelip, a public health expert with a background in nutrition, helped design our initial research question. He suggested that we include iron in our analysis due to high rates of iron deficiency in India (Sharma et al., 2020a). Iron is largely absent in past economics work on

[1]We define a household as experiencing undernourishment if it consumes below 100 percent of its recommended levels of dietary intake, and extreme undernourishment as below 80 percent.

diet quality in rural India, which usually focuses on calories or macronutrients (Deaton and Drèze, 2009).

This absence exemplifies how applied microeconomics can prioritize questions based on data accessibility rather than theoretical importance. The NSS provides conversion tables to translate food items into calorie values, but these tables omit iron. Instead, motivated by Professor Prelip's advice, I manually added iron values for 150 food items to these tables. If I had drawn inspiration only from past economics work, this extra effort would not have seemed worthwhile. Including iron allowed for a more nuanced analysis of household responses to heat. For example, we find suggestive evidence that, although households increase food purchases after a hot growing season, the foods they purchase are low in iron.

A second way that this chapter benefits from interdisciplinary engagement is in its treatment of climate projections. The first two studies in my dissertation both include projections of how increased heat due to climate change will affect the outcomes of interest. In the first study, we use an ensemble of 22 climate models to generate projections of how the temperature distribution in 2080 and 2099 will differ from the baseline period of 2000 to 2019. Then, assuming that the temperature-disconnections relationship will remain the same in the future, we calculate an estimate of climate change-induced increases in energy disconnections.

Despite its many limitations, the above process represents a relatively sophisticated version of how applied microeconomists incorporate climate projections into their work. As recently as 2020, papers have used a single model to calculate climate projections (Garg et al., 2020b). As an example of a more involved approach, the Climate Impact Lab, an interdisciplinary group led by economists, uses 21 models and estimates "surrogate models" to fill in the under-represented tails of the climate projection distribution (Carleton et al., 2022). However, this method's computational intensity and multiple complicated assumptions decrease the likelihood of its widespread adoption.

My methods in study 2 are more sophisticated than usual for economics but still relatively easy to implement. I create a projection ensemble of 190 members from 26 models to calculate

a more comprehensive picture of the potential spread of climate damages across, and within, climate models. In addition, I focus on mid-century projections (2030 to 2049) as opposed to late-century ones (2080 to 2099). I do so because mid-century projections are more likely to motivate relevant adaptation policy today, in part because end-of-century ones include far more uncertainty. For near-future projections, however, within-model variability is a more substantial portion of the total uncertainty, heightening the importance of using multiple members of the same model (Schwarzwald and Lenssen, 2022).

The difference in the treatment of climate projections between the two studies is thanks to advice from two climate scientists: Professor Karen McKinnon and fellow IoES PhD student Will Krantz. Professor McKinnon helped me grasp the importance of within-model projection variability. For example, CanESM5 is one of the 26 Coupled Model Intercomparison Project Phase 6 (CMIP6) models I use to calculate the projected increase in the percentage of rural Indian households experiencing extreme iron undernutrition in 2030 to 2049. Within this single model, which includes 50 members, the variance in this increase is 0.11, a third of the size of the variance across the full 190-member ensemble. Will Krantz shared code and technical knowledge that allowed me to quickly implement the ideas from my discussions with Professor McKinnon. As I continue to work in this space, I hope to push economists to include larger climate model ensembles in their work.

Study 3

Labor Market Power and Climate Adaptation

My third study concerns a setting where the power to adapt can take on a more oppositional nature: the labor market. In the first two studies, the key determinant of adaptation potential has been access to resources to decrease one's vulnerability to heat. If a household can afford to run its AC, it does so and pays the bill. In the labor market, however, there is a disconnect between these costs and benefits. While employers bear the costs of adapta-

tions, such as running the AC, workers reap the most direct benefits in terms of health and thermal comfort. In theory, employers may also benefit by mitigating heat's effect on worker productivity (Cachon et al., 2012). However, business pushback to California's proposed heat illness prevention regulations for indoor workers suggests that firms may not see these productivity benefits as worth the costs of climate control (Kuang, 2024b). The realized level of adaptation is therefore likely the result of negotiations between employer and employee.

Competitive market theories of economics are poorly equipped to analyze adaptation decisions in settings involving power dynamics (Deaton, 2024). In the theoretical competitive market, employers must pay "compensating differentials" (e.g., higher wages) to make up for workplace disamenities such injury risk or exposure to the elements (Rosen, 1986). Without a compensating differential for a disamenity, workers would choose work in other occupations or firms with lower levels of that disamenity. From this perspective, identifying differences in realized levels of adaptation across firms is of relatively limited interest, as firms with low levels of adaptation must be compensating for it in some other (perhaps unobserved) way. For example, if a firm does not pay for AC in its factory, then it might pass on those savings to its employees who are working in more uncomfortable conditions.

Firms investing in sub-optimal (from the worker's perspective) levels of adaptation due to high labor market power, however, could motivate policy interventions to protect vulnerable workers. Economic theory suggests that firms with labor market power can set compensation at levels lower than worker productivity, resulting in suboptimal levels of employment and output, hurting both employees and consumers (Berger et al., 2022). In addition, firms may not fully realize the benefits of climate control, such as increased productivity and reduced workers' compensation payments. Therefore, firms with labor market power may be investing in suboptimal levels of adaptation, even from their own perspective.

Policymakers have different options to address this issue. If one believes that employees better understand the mutual benefits of climate control, then one might adopt policies that help equalize bargaining power between them and their employers (e.g., by supporting the

formation of unions). Another option is to adopt mandatory heat prevention protocols in the workplace, as currently exist in several US states (OSHA, 2024). For some stakeholders, simply identifying disparities in adaptation ability might be enough to justify a redistributive policy intervention. However, many economists and policymakers think in terms of aggregate welfare. In order to get broader buy-in for policy making from this group, quantifying labor market power, and the degree to which it constrains climate adaptation, is important.

Empirical Difficulties of Studying Labor Market Power

In *Porkopolis*, Alex Blanchette describes working at "Dover Foods," a meatpacking company in the Great Plains, alongside "Raul," a Cuban immigrant whose family lives in Miami. Raul was in construction, but after the industry's collapse in 2008, he could not find work in Miami, an issue exacerbated by his limited English proficiency. He was recruited by "Grensome Meats," another meatpacking plant in the Great Plains, at a job fair in Miami. After his stint there ended, he found work at Dover Foods (Blanchette, 2020). Raul's case exemplifies the difficulty of grasping and measuring "his" labor market. Doing so is an important first step in studying the balance of power between him and his prospective employers. Measuring the labor market power of local employers in the Great Plains misses the fact that his labor market should include meatpacking jobs in Miami. In addition, measuring the labor market power that he and his fellow meatpacking workers face omits his ability to work in other occupations.

In quantitative studies, the first challenge is defining the contours of a particular labor market. In economics, many recent studies have done so at the occupation-location level (Azar et al., 2019). In practice, even this deceptively simple definition requires some difficult decisions. Standard geographical units (e.g., state or county) tend to poorly represent labor markets. For example, Kansas City straddles the Kansas-Missouri border. Instead, researchers typically use "commuting zones," which more likely represent local economies (Marinescu et al., 2021). These zones are difficult to delineate, especially given how many

people commute large distances to their jobs. Choosing the granularity of occupations to use can be equally challenging. Take too coarse a definition, and one runs the risk of pooling together workers who do vastly different tasks (e.g. "Sales" could mean telemarketers or highly technical sales representatives for a software firm). But overly granular definitions risk separating largely overlapping groups of workers (e.g., sales associates vs. sales representatives).

This occupation-location definition of the labor market offers several relatively straightforward options to quantitatively measure labor market power. One option, employer market share, measures the proportion of employees hired by a single firm in an occupation-location. For example, if Walmart hires half of the cashiers in a commuting zone, it has a market share of 50 percent. Often, in order to match it with available data on wages, researchers calculate an aggregate measure of employer market share at the occupation-location level. This measure, called the Herfindahl-Hirschman index (HHI), is the sum of the squares of the market shares of all firms in that labor market. Many recent studies have documented the negative effects of HHI on wages, providing evidence that is consistent with the existence of significant market power (Marinescu et al., 2021).

Recent work has also begun to use the concept of "outside occupation options," which refers to the jobs that workers might take outside of their current occupation. This approach addresses some issues with defining the labor market at the occupation-location level. For most occupations (other than, say, brain surgeons or professional athletes), it is highly unlikely that workers are limiting their job search to a single occupation. Researchers have measured this idea in multiple complementary ways, including with task composition (i.e., which occupations have similar tasks to a worker's current occupation?) (Macaluso, 2023), worker characteristics (i.e., which occupations do similar employees work in?) (Caldwell and Danieli, 2024), and worker transitions (i.e., which occupations are workers in this current occupation likely to transition into?) (Schubert et al., 2022). This concept approaches the reality of a local, individually-relevant labor market that can still be measured quantitatively and at scale. However, even this more sophisticated approach to defining labor markets

misses much of the nuance from an anthropology text like *Porkopolis*.

After arriving at a relatively satisfactory definition of a labor market, an applied microeconomist can only study labor market power at that level if the necessary data exist. For example, Rode et al. (2023) studies how workers worldwide respond to extreme heat by decreasing their time at work. These wage losses from this decreased work represent a "willingness to pay" (WTP) to avoid extreme heat in the workplace. The authors use the WTP, alongside socioeconomic and climate projections, to calculate climate change's cost to worker well-being in dollar terms. An important assumption in this study, that the authors note is likely untrue, is that workers have the power to set their own schedules on a daily basis (Rode et al., 2023). However, incorporating measures of workers' control over their schedules would require data that do not exist. This paper exemplifies the difficulties of incorporating a thorough analysis of power into quantitative climate adaptation work.

Sometimes bridging the gap between important questions and existing data entails leveraging advances in computational power and software to create one's own dataset. For example, Schubert et al. (2022) defines outside occupation options using job transitions by collecting data from 16 million online resumes. The authors combine these occupational transition data with occupation-by-location wage data to create a measure of the local quality of outside occupation options. Then, using an instrumental variables regression, a causal inference tool commonly used in applied microeconomics, they find that a 1% increase in the wages in a worker's outside occupation options causes a 0.1% increase in their current wages (Schubert et al., 2022). This result highlights the importance of underlying labor market conditions in determining worker compensation. It also showcases the promise of combining careful causal inference with high computational power.

As I describe above, the usual challenges of quantitative research are accentuated when one seeks to study labor market power. These difficulties result in a gap in the economics literature that motivate turning to other disciplines to better understand power dynamics between workers and employers. In *Porkopolis*, anthropologist Alex Blanchette undertakes an

ethnographic study of workers in meatpacking plants. He describes how meatpacking firms recruit laborers with low employment prospects, including refugees, formerly incarcerated people, and undocumented immigrants. By immersing himself in the plant, he learns that many of his fellow workers, such as Raul, lack competitive alternative employment opportunities. The resulting power imbalance between them and their employer means that they face harsh working conditions without adequate pay (Blanchette, 2020). Texts and concepts such as these, combined with the economics literature on climate adaptation, inform the questions I ask in this third study.

The Study

In my third study, I test the potential for workers to respond to climate change by switching occupations, from ones exposed to the elements ("high exposure") to ones that are not ("low exposure"). This study fills a gap in the literature on potential worker responses to climate change, which is largely limited to short-term responses, such as working less on hot days (Graff Zivin and Neidell, 2014). It also contributes to long-standing debates on the degree of segmentation that exists in our labor markets (Reich et al., 1973; Eichhorst et al., 2017). Labor market segmentation, which refers to the division of workers and jobs into distinct "segments," reduces the set of jobs a worker can access, thereby enhancing employer power. Qualitative work can document instances of segmentation, such as Blanchette's description of meatpacking workers having few other employment prospects (Blanchette, 2020). However, quantitative measures of labor market segmentation, while complicated by the challenges outlined above, are important for designing appropriate policy interventions.

Segmentation between high and low exposure occupations will contribute to the overall consequences of climate change on inequality, in both its wage and non-wage forms. On the one hand, if high exposure workers are unable to transition to low exposure occupations, then climate change will likely widen already-existing inequality between these sets of workers. On the other hand, if such transitions are common, then low exposure workers, while insulated

from the direct effects of extreme weather, may still see climate-induced wage decreases as more workers seek to move to these jobs. These spillover consequences are more likely to affect lower-wage low exposure occupations, whose workers are more similar to those in high exposure work (Maestas et al., 2017).

I test how often workers transition between high exposure and low exposure occupations using administrative data from France from 2011 to 2019. This dataset is particularly well-suited to this analysis. First, it allows me to track individual workers, and their occupational transitions, over time. Second, as it is compiled from responses to a mandatory employer survey, it is highly representative of the whole labor market. Other datasets that could be used to quantify labor market segmentation, such as the online resumes used in Schubert et al. (2022), underrepresent many workers who often engage in high exposure work, such as those without a bachelor's degree (Maestas et al., 2017). Third, this dataset is a roughly 1 in 12 (all workers born in October) sample of all employees for the vast majority of occupations in France, providing me with a very large sample. I observe a median of 2,100 workers leaving each of the 368 occupations I study, allowing me to test for within-occupation differences in occupational mobility.

I find a high degree of segmentation between high and low exposure labor markets. By my preferred definition, high exposure workers only make up 9% of the jobs in France. However, 49% of worker transitions out of one high exposure occupation result in a move to another high exposure occupation. Skill requirement differences cannot fully explain this segmentation. For example, I show that occupations involving more manual tasks are more likely to be high exposure. However, the segmentation between high and low exposure occupations persists even when controlling for occupations' degree of manual intensity.

Finally, I analyze how the ability to transition to low exposure work differs between workers within the same high exposure occupation. I find that, among workers leaving a high exposure occupation, relatively high earners (after controlling for characteristics such as age, gender, and experience) are 23% more likely to transition to a low exposure occu-

pation than are low earners. This finding suggests that, as the planet warms, high-paid high exposure workers will more likely be able to switch to low exposure work than their lower-paid counterparts. It also emphasizes the importance of future labor market research at the sub-occupation level.

Sometimes, correlational analyses are key to broaching topics where data availability limits causal inference. One methodological inspiration of my third study is Chetty et al. (2014), which documents intergenerational mobility in the US. The authors show strong heterogeneity in mobility across the country and summarize the characteristics of high-mobility areas. They answer a hugely important question despite noting that their "descriptive analysis does not identify the causal mechanisms that determine upward mobility" (Chetty et al., 2014). Indeed, their primary goal seems to be to thoroughly document an existing issue in order to spur future, possibly causal, research on the topic (Chetty et al., 2014). My motivation for this third study is similar to theirs. I cannot, due to data constraints, conduct a causal analysis of how workers will react to a changing climate. Instead, I describe historical segmentation between high and low exposure occupations as a first step in anticipating potential worker responses in the future. In addition, I hope to pursue and inspire future research on factors that cause this segmentation.

2

High Temperatures and Electricity Disconnections for Low-Income Homes in California[1]

[1]This chapter is coauthored with Professors Alan Barreca and R. Jisung Park and has been published in the journal *Nature Energy*. The citation is: Barreca, A.; Park, R.J.; Stainier, P. (2022). High temperatures and electricity disconnections for low-income homes in California. *Nature Energy*, 1-13.

2.1 Introduction

Poor households face a unique set of tradeoffs when adapting to extreme heat events. They may need to temporarily incur higher energy expenses by using home cooling technologies, like air conditioning, to protect their health (Anderson et al., 2013; White, 2017; Barreca et al., 2016; Isen et al., 2017; Deschênes and Greenstone, 2011). Past studies have investigated whether households reduce consumption of necessities, like food, to pay for high energy bills (Snell et al., 2018; Bhattacharya et al., 2003; Cullen et al., 2004; Beatty et al., 2014; Frank et al., 2006; Nord and Kantor, 2006). Less is known about whether the financial distress causes households to miss paying important bills altogether. In particular, nonpayment of an energy bill can lead to a household being disconnected by their utility provider. Sociological interviews suggest that the cost of disconnection includes significant material hardship in the short term as well as any associated costs of reconnecting service (Harrison and Popke, 2013; Hernández, 2016; Heflin et al., 2011). Defaulting on energy debt may also cause longer-term harm, like limiting access to credit or adding to the cumulative stresses of poverty (Mullainathan and Shafir, 2013; Kishiyama et al., 2009; Evans, 2016; Schofield and Venkataramani, 2021).

Here, we estimate the relationship between temperature, electricity expense, and electricity disconnections using administrative data from California from 2012 through 2017. Our data come from Southern California Edison (SCE), a major utility that serves over 15 million people in a 50,000 square-mile area of central, coastal, and southern California. The overwhelming bulk of the accounts in our sample was enrolled at some point in the California Alternate Rates for Energy (CARE) program, which provides a subsidy of around 30 percent for qualifying low-income households (Southern California Edison, 2017a, 2012a, 2017b, 2012b). SCE employs a tiered pricing plan, but these tiers have varied historically with minor differences for CARE customers (Southern California Edison, 2017a, 2012a, 2017b, 2012b; California Public Utilities Commission, 2009). Time of use pricing is relatively rare in our sample period since SCE had only begun piloting an opt-out time of use pricing plan

in 2016 (Southern California Edison, 2019).

This research innovates on existing work in two key ways. First, this disconnection analysis adds to the broader energy insecurity literature (Jessel et al., 2019; Brown et al., 2020; Hernández and Bird, 2010; Bednar and Reames, 2020; Agbim et al., 2020), but more specifically complements the sociological and economic studies that investigate bill juggling strategies and difficult consumption choices that households face with higher energy bills. National survey data suggests that 14 percent of U.S. households receive a shut off or delivery stop notice at some point during the year, but information on the frequency of actual disconnections is typically not collected (Energy Information Administration, 2018). Qualitative studies featuring in-depth interviews of low-income households find that fear of disconnection is a significant psychological stressor (Harrison and Popke, 2013; Hernández, 2016). Quantitative analyses find that high energy bills lead to less food consumption, something referred to as the heat or eat tradeoff (Snell et al., 2018; Bhattacharya et al., 2003; Cullen et al., 2004; Beatty et al., 2014; Frank et al., 2006; Nord and Kantor, 2006). Analysis of disconnections is largely absent from such studies, with the exception of one study from Australia, likely due to the dearth of available disconnection data (Longden et al., 2022; National Consumer Law Center, 2020).

Second, our estimates help reveal the timing between extreme heat events and disconnection risk. To the extent that temperature affects disconnections through energy expenses, disconnections are likely to be delayed since bills are typically not considered overdue until a few weeks after the close of a billing period (Southern California Edison, 2013a). Furthermore, existing policies - including in our study area - often prevent utilities from disconnecting homes when the day's temperature is above 100°F (Southern California Edison, 2020). Higher expenses in one billing period could increase the cumulative risk of default in subsequent months or simply cause intertemporal shifts that expedite an impending disconnection. Temperature could also affect disconnection risk for other reasons than energy expenses, such as lost labor income or health-related costs (Barreca et al., 2016; Graff Zivin and Neidell, 2014).

Our estimates show that hot weather in one billing period leads to higher disconnections later on. In particular, each additional day with a maximum temperature of 95°F in one billing period causes the relative risk of disconnection to increase by 1.2 percent 51-75 days after the close of that billing period. Higher electricity expenses themselves appear to be a likely contributor. We find electricity expenses increase by 1.6 percent in the current billing period with each additional 95°F day. The disconnection timing matches the energy provider's policy to disconnect no sooner than 53 days after a bill is issued, which is usually just after the close of the billing period (Southern California Edison, 2013a). We further document the importance of energy expenses as a determinant of disconnections. We show that both expenses and disconnections increase dramatically after 26 periods on the CARE program, right when a portion of households are disenrolled from the CARE subsidy as part of the two-year eligibility verification process (Southern California Edison, 2014a).

2.2 Estimates of the Temperature-Disconnection Relationship

Table 2.1 presents the means and standard deviations of our key variables broken down by season of year. Energy consumption is strongly seasonal. The average bill amount is nearly $50 higher in the summer than in the spring months. This suggests that higher temperatures are strongly predictive of electricity use in our low-income population, which is generally consistent with existing studies' findings on the temperature-energy relationship (Deschênes and Greenstone, 2011; Auffhammer and Mansur, 2014; Auffhammer et al., 2017; Barreca, 2012; Aroonruengsawat and Auffhammer, 2011). The mean of the disconnection variable is 246 disconnections per 100,000 bills or about 0.25 percent on each bill. Disconnections are at their lowest in the summer, highlighting the need to account for lags between when energy expenses are incurred and when bills are delinquent. Temperatures are relatively high in our sample area, with an average daily maximum temperature of 77°F.

Table 2.1: Summary Statistics: SCE and PRISM

	Full	Season			
Means	Sample	Winter	Spring	Summer	Fall
Bill Amount	72.79	68.29	55.34	101.64	65.47
Electricity used (kWh per cycle)	531.55	476.24	451.56	688.64	508.40
Disconnections per 100,000 bills	245.93	245.67	275.20	211.24	250.95
Bills on CARE	0.74	0.74	0.74	0.74	0.74
Bills on FERA	0.01	0.01	0.01	0.01	0.01
Accounts Ever Disconnected	0.043				
Accounts Disconnected 2 or More Times	0.021				
Disconnections for Disconnected Sample	2.54				
Daily Maximum Temperature (F)	77.25	67.87	76.22	87.35	77.31
Days with Max Temperature <60F	2.06	6.16	0.68	0.00	1.51
Days with Max Temperature 60F-70F	6.68	12.01	7.48	0.41	7.01
Days with Max Temperature 70F-80F	9.51	9.10	12.42	6.52	9.91
Days with Max Temperature 80F-90F	7.45	3.14	6.85	11.57	8.26
Days with Max Temperature 90F-100F	3.85	0.14	2.36	9.18	3.62
Days with Max Temperature >100F	0.88	0.00	0.34	2.57	0.55
Daily Precipitation (inches)	0.03	0.06	0.01	0.00	0.02
Standard Deviations					
Bill Amount ($)	85.00	71.13	70.03	108.53	75.9
Electricity used (kWh per cycle)	397.07	325.02	316.59	513.28	349.69
Disconnections per 100,000 bills	4,953.08	4,950.38	5,238.71	4,591.27	5,003.22
Bills on CARE	0.44	0.44	0.44	0.44	0.44
Bills on FERA	0.12	0.12	0.12	0.12	0.12
Accounts Ever Disconnected	0.20				
Accounts Disconnected 2 or More Times	0.14				
Disconnections for Disconnected Sample	2.85				
Daily Maximum Temperature (F)	9.63	5.59	5.75	7.20	7.93
Days with Max Temperature <60F	4.32	6.15	1.61	0.03	3.42
Days with Max Temperature 60F-70F	6.63	5.06	5.89	1.92	6.66
Days with Max Temperature 70F-80F	6.45	5.13	5.58	8.14	4.67
Days with Max Temperature 80F-90F	6.00	3.44	4.81	6.87	4.90
Days with Max Temperature 90F-100F	5.81	0.58	3.37	7.54	4.50
Days with Max Temperature >100F	2.88	0.00	1.49	4.87	1.55
Daily Precipitation (inches)	0.05	0.07	0.02	0.01	0.03
Number of Bills	13,329,955	3,339,091	3,618,488	3,449,100	2,923,276

Notes: This table has the means of our key variables for 2012-2018. Winter includes December-February, Spring March-May, Summer June-August, and Fall September-November. There are a total of 299,225 accounts in our sample. Accounts were selected based on whether they ever enrolled in the California Alternate Rates for Energy (CARE) or Family Electric Rate Assistance Program (FERA) low-income subsidy programs.

Before examining impacts on disconnections, we first show that hot weather causes higher electricity use for this low-income sample. Figure 2.1A illustrates estimates from a model where the natural log of electricity usage is the outcome. The model controls for the current bill period's temperature as well as a set of additional control variables to address confounders. The model's estimates should be interpreted as the impact of one additional day at a given temperature, relative to a day at 75°F, in the current billing period on the current billing period's usage. Our model uses a cubic spline that allows the temperature-electricity relationship to vary flexibly along the temperature distribution because the spline has multiple pieces (Smith, 1979). We preserve the variation in daily temperatures by calculating the daily value of each spline piece before summing up to the billing period. By controlling for zip code by calendar-week fixed effects and billing period date fixed effects, our key estimates are driven by unpredictable year-to-year variation in temperatures for a given zip code and time of year. In addition, the model includes household fixed effects to address time-invariant differences in household electricity consumption behaviors.

We find that our sample of low-income households increase energy use when there are more hot days in a given billing period. For example, each additional 95°F day causes electricity use to increase by 1.2 percent that billing period, or about 8.26 kWh based on the summer average of 688.64 kWh. These findings are consistent with previous studies which find a U-shaped temperature-energy relationship in California's general population (Aroonruengsawat and Auffhammer, 2011) and the U.S. more broadly (Deschênes and Greenstone, 2011; Auffhammer et al., 2017; Barreca, 2012). Differences in cooling efficiency (e.g. window vs. central air conditioning) could mask important disparities in household welfare due to indoor thermal comfort. In a global analysis, one study finds that the relationship between temperature and electricity is generally flat for low-to-middle income countries, and other energy sources exhibit an L-shaped relationship with higher demand at colder temperatures only (Rode et al., 2021). One study examines the effect of temperature on electricity consumption in Mexico (Davis and Gertler, 2015), finding a similar pattern to the U.S. studies including ours here. These studies illustrate that the relationship between temperature and

energy use varies by different climates, housing characteristics, appliance penetration, and local infrastructure.

Figure 2.1B estimates a similar model with bill amount in current dollars as the outcome. Each additional 95°F day causes bill amounts to increase by $1.60 on a baseline average of $101.64 during summer months, or about 1.6 percent. The relative increase in expenses is slightly higher than the relative increase in energy consumption. This finding is likely due to SCE's tiered pricing plan, where rates increase when total consumption over a billing period exceeds the baseline level for that area (California Public Utilities Commission, 2009).

Figure 2.1C uses a binned temperature model, as opposed to the spline, for log electricity usage. These binned models are relatively popular because of their ease of interpretation and certain flexibility in functional form (Dell et al., 2014). The findings are qualitatively similar to the spline model. One additional day with temperatures between 90 and 100°F causes electricity use to increase by 1.0 percent that billing period. One additional day above 100°F causes a 2.0 percent increase in electricity use. Figure 2.1D presents the binned model for bill amount and produces similar estimates to the spline. Each day above 100°F causes electricity expenses to increase by $2.92.

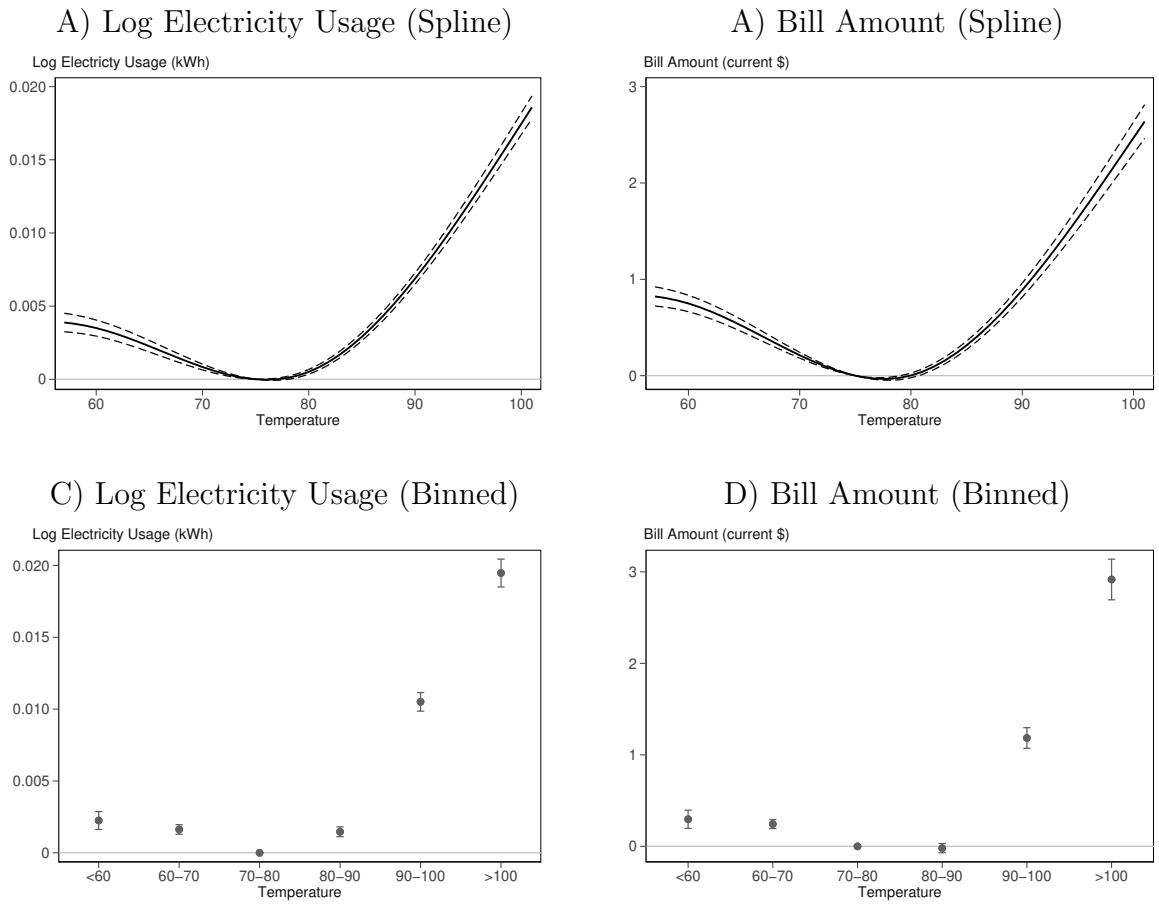


Figure 2.1: Effect of Temperature During Current Billing Cycle on Electricity Usage and Bill Amount

Panel A shows the change in the log electricity usage for each additional day at a given temperature relative to 75°F in the current billing cycle with a spline temperature model. Panel B shows the change in the bill amount (current $) for each additional day at a given temperature relative to 75°F in the current billing cycle with a spline temperature model. Dashed lines are 95% confidence intervals centered on the estimated coefficient. Panel C shows the change in the log electricity usage for each additional day at a given temperature in the current billing cycle with a binned temperature model with days between 70 and 80°F as the omitted category. Panel D shows the change in the bill amount (current $) for each additional day at a given temperature in the current billing cycle with a binned temperature model with days between 70 and 80°F as the omitted category. Brackets are 95% confidence intervals centered on the estimated coefficient. All models include billing start date, zip by calendar week, age of account, and individual account fixed effects, as well as precipitation controls. Standard errors are clustered at the zipcode level. Panels A and C include 13,311,787 observations, and B and D include 13,329,871 observations. There are fewer observations in a and c because we drop zeros and negative values for energy consumption when we convert to logs. There are fewer observations in A and C because we drop zeros and negative values for energy consumption when we convert to logs.

We next quantify the relationship between temperature in the current billing cycle and subsequent disconnection probabilities. Figure 2.2A presents the estimates from several temperature spline models where the outcomes are disconnection indicators at 25-day intervals into the future from the close of the current billing period (1-25 days, 26-50 days, ..., 126-150 days into future). We also estimate impacts of temperature on disconnection during the current bill period. Across the different model outcomes, the largest relationship exists between high temperatures in the current period and disconnections 51-75 days later. One additional 95°F day causes an additional 3.0 disconnections per 100,000 bills. Given an average disconnection rate of 246 per 100,000, this effect size represents a 1.2 percent increase in the relative risk of disconnection.

Figure 2.2A illustrates that colder temperatures appear to have little effect on disconnection risk, which could be explained by the facts that very few homes use electricity for heating in our study area and there are relatively few cold days. Data from the 2019 California Residential Appliance Saturation Study suggests that only 13 percent of households in our setting use electric central furnace space heating, and 4 percent use heat pumps(Energy Insights USA, 2020). In contrast, 68 percent of households in SCE territory have central air conditioners, and 18 percent have room air conditioners, suggesting that electricity use for cooling is higher than for heating (Energy Insights USA, 2020). Appendix Figures A1 and A2 illustrate each estimate separately with confidence intervals, including one placebo check where we estimate the relationship between the current billing period's temperature and disconnections 1-25 days prior to the start of the billing period. There is no discernable relationship observed in the placebo test, suggesting our model appropriately controls for confounders tied to climatic differences by zip code.

Figure 2.2B illustrates the effect of a 95°F day across time periods, but with confidence intervals. By construction, the effect size at 51-75 days is the same as the estimated effect at 95°F presented in Figure 2.2A. We do not observe any statistically meaningful effects outside of the 51-75 day window indicating that hot weather causes an increase in cumulative risk as opposed to simply shifting disconnections forward or backward in time.

The large effect observed 51-75 days after the close of the billing period is consistent with homes failing to pay the bill in which the hot weather occurred. In the case of non-payment on a bill, SCE's policy is to disconnect approximately 53 days from the date the bill is presented (Southern California Edison, 2013a). In practice, customers can appeal the bill amount, arrange for partial payment and/or enroll in balanced payment plans, which might delay disconnections beyond the 53 days. Appendix Figure A3 presents an analysis where we break out exposure into 10-day windows that supports our main findings by showing that the effects are concentrated in the 51-60, 61-70, and 71-80 day windows.

Figure 2.2C also plots the estimated effects of temperature on disconnections 51-75 days later for the spline model with confidence intervals. Figure 2.2D shows the estimates are similar when we use a binned temperature model that captures the number of days in a given temperature range during the current billing period. For example, one additional day with max temperature between 90-100°F in the current billing period causes disconnection risk to increase by 2.9 disconnections per 100,000 bills 51-75 days later. Each day above 100°F causes an additional 6.2 disconnections per 100,000 bills, equivalent to a 2.5 percent increase on the mean risk in the full sample.

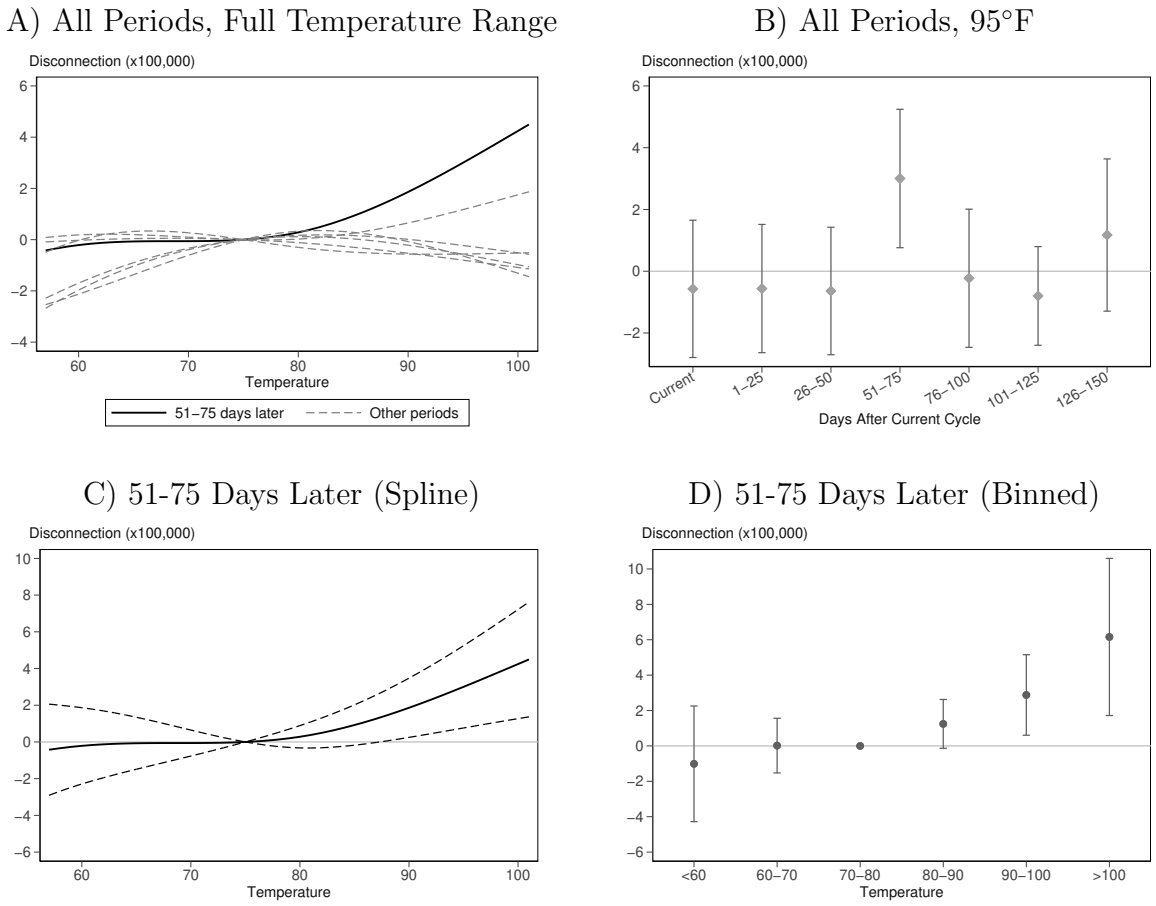


Figure 2.2: Effect of Temperature on Subsequent Disconnection Risk

Panel A shows the temperature response function from several different models, each quantifying the relationship between temperatures in the current billing cycle on disconnections at various time periods in the future. The disconnection time periods include 1-25, 26-50, 51-75, 76-100, 101-125, and 126-150 days after the current billing cycle closes as well disconnections in the current billing period. We omit confidence intervals for clarity. Panel B shows the change in disconnections due to one day at 95°F relative to 75°F for different time periods. Brackets are 95% confidence intervals centered on the estimated coefficient. Panel C shows the temperature-response function from a single model correlating temperature in the current billing cycle with disconnections 51-75 days after the close of the billing cycle with a spline temperature model. Dashed lines are 95% confidence intervals centered on the estimated coefficient. Panel D shows he temperature-response function from a single model correlating temperature in the current billing cycle with disconnections 51-75 days after the close of the billing cycle with a binned temperature model with days between 70 and 80°F as the omitted category. Brackets are 95% confidence intervals centered on the estimated coefficient. All models include billing start date, zip by calendar week, age of account, and individual account fixed effects, as well as precipitation controls. Standard errors are clustered at the zipcode level. In Panel B, the models include (in order from current cycle to 126-150 days after current cycle) 13,329,871, 13,237,428, 12,993,175, 12,758,335, 12,522,895, 12,393,882, and 12,170,748 observations. For C and D the sample size is 12,758,335 billing-period observations.

We can also put the magnitude of the effect into the context of our sample, using our binned estimates for 90-100°F and ≥100°F days in Figure 2.2D for ease of exposition. The average billing period experiences 3.85 days 90-100°F and 0.88 days above 100°F (see Table 2.1). Based on our binned estimates for 90-100°F (2.880) and ≥100°F days (6.157), hot weather accounts for an additional 0.00016506 disconnections in the average account-cycle ((3.85*2.880 + 0.88*6.157)/100,000) or about 2,200 disconnections for our sample of 13,329,955 bills. Only 4.3 percent of accounts in our sample ever experienced a disconnection, or 12,931 out of the 299,225 accounts. There were a total of 32,783 disconnections, with each of these accounts experiencing 2.54 disconnections on average. As such, our estimates suggest that hot weather, categorized as days above 90°F, explains 6.7 percent (2,200/32,783) of all disconnections in our sample.

Appendix Figure A4 tests the sensitivity of our results to changes in our model. We drop the household fixed effects in case these controls are oversaturing the model. We add one lag and one lead of our temperature variables to account for potential serial correlation in weather. We use daily mean temperature in place of daily maximum temperature to better account for diurnal temperature swings. We drop billing cycles with unusual start and end dates in case these observations are adding measurement error to the temperature variables. In all tests, estimates are nearly identical to our preferred model.

2.3 Change in Disconnection Risk After Losing the CARE Subsidy

Energy expenses are one potential channel underlying the relationship between temperature and disconnection risk, though we cannot rule out other channels, like lost labor income or added health costs. To further assess the relevance of the energy expense channel, we consider whether disconnection risk changes after accounts disenroll from the CARE program for plausibly idiosyncratic reasons. We exploit the fact that the utility company provides existing CARE customers with a notice to recertify their eligibility after two years on the

program, and customers have 45 days from the notice to recertify (Southern California Edison, 2014a). This verification process leads to an acute drop in CARE enrollment after the 26th bill, something we quantify in a regression discontinuity framework.

Before estimating the regression discontinuity Equation 2.3, we first illustrate the basic changes in CARE enrollment, electricity use, expenses, and disconnection risk around the 26th bill. As illustrated in Figure 2.3A, there is a discontinuous drop in CARE enrollment after the 26th bill. Figure 2.3B shows there is little observable change in energy consumption after the 26th bill outside of the regular cyclicality. Figure 2.3C illustrates that energy expenses also follow a similar seasonality to energy consumption. However, there is a jump in energy expenses after the 26th bill due to the fact that electricity prices increased (from losing the CARE subsidy) and consumption changed little.

Figure 2.3D illustrates the dramatic increase in disconnection risk as of the 28th bill. There is little change in disconnection risk at the 27th cycle since the accounts have yet to be presented with their first non-CARE bill. The increase in disconnections in the 28th bill occurs before the 27th bill is even past due, suggesting that the increased expense might expedite the disconnection process for those households already carrying an overdue balance and on the margin of missing a payment. This contrasts with our temperature-disconnection analysis (Figure 2.2) where we observe a 51-75 day delay, or about two to three billing periods delay, from when the bill period closes to when households are disconnected.

Tables 2.2 (CARE enrollment and energy use) and 2.3 (bill amount and disconnection) estimate the regression discontinuity Equation 2.3, which captures the shift in outcome after the 26th billing period and controls for a rich set of fixed effects to address the seasonality observed in Figure 2.3. For this regression discontinuity analysis, we restrict the sample to the 12 bills before and after the 26th bill to avoid capturing longer-term changes in household behavior that might confound our estimates. In our core model, we find that CARE enrollment declines by 13.63 percentage points after the 26th cycle. Electricity usage declines by only 0.3 percent. While energy use changes little, bill amounts are $2.33 higher

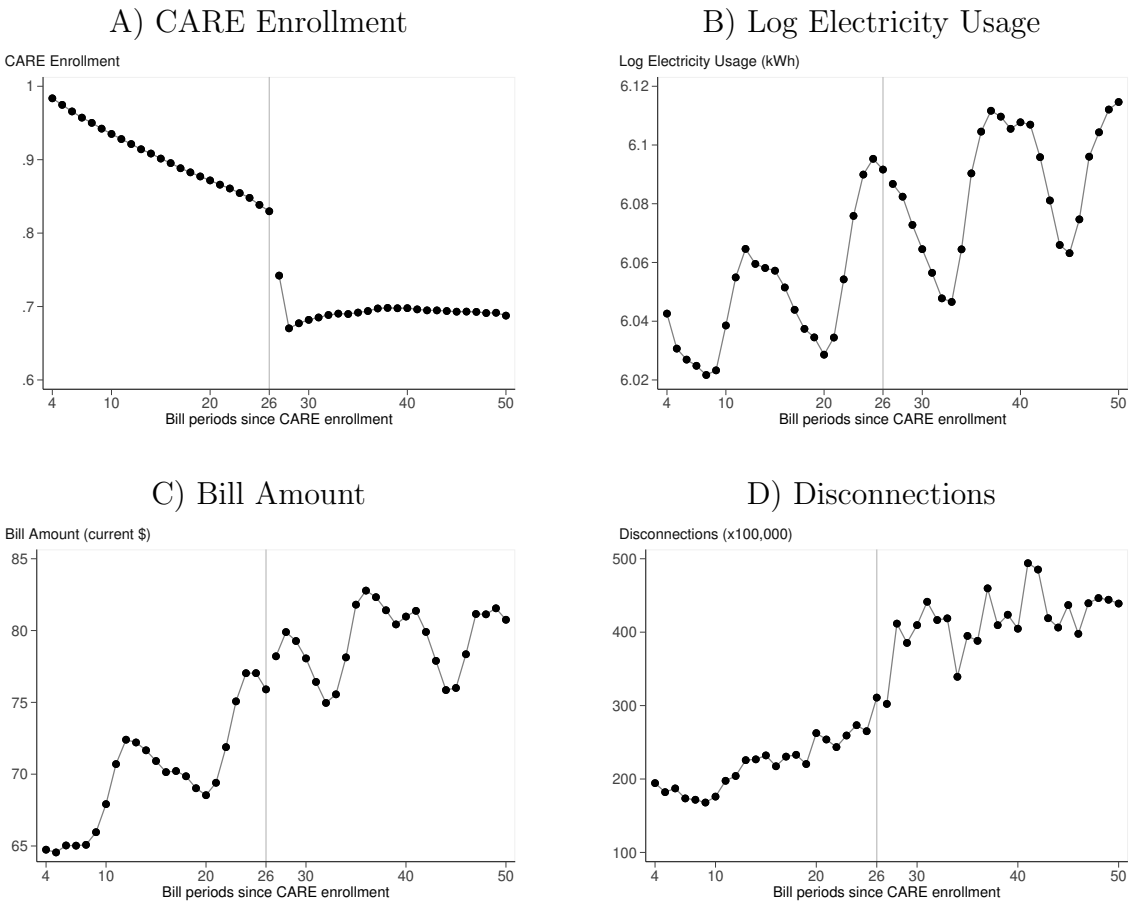


Figure 2.3: Means in Outcomes Since Starting CARE, RD Sample

Panel A shows mean CARE enrollment by number of bills since first enrolled in CARE. Panel B shows mean log electricity usage by number of bills since first enrolled in CARE. Panel C shows mean energy bill (current $) by number of bills since first enrolled in CARE. Panel D shows mean disconnection (x100,000) by number of bills since first enrolled in CARE. The vertical line is drawn at the 26th billing period. We top code the billing periods at 50, since relatively few accounts exceed this amount of time in our sample (Jan 2012-August 2017). We omit the first three periods on CARE to aid in the illustration.

after the 26th period due to the increase in electricity prices. There are two possible, and mutually inclusive, explanations as to the unresponsiveness of consumption to such a large price change. First, energy consumption is a basic necessity for which households likely exhibit limited price sensitivity in the short run, as suggested by existing studies (Reiss and White, 2005; Burke and Abayasekara, 2018; Deryugina et al., 2020). Second, households may have been relatively unaware of being dropped from the CARE program. This seems plausible given a number of accounts re-enroll in CARE after billing period 28 as illustrated in Figure 2.3A.

Disconnections increase by 72.4 per 100,000 bills after the 26th cycle, which is a much larger effect compared to our temperature-disconnection findings. Above, we found that each 95F day caused energy expenses to increase by $1.60 and the disconnection risk to increase by 3.0 accounts per 100,000 bills 51-75 days later. One possible explanation is that disconnection risk is increasing at an increasing rate in the size of the shock to expenses. With hot weather, the measured increase in expenses may have been spread over more households with varying levels of baseline financial distress. In the regression discontinuity analysis, the increased expenses of $2.33 only impacted the 14 percent of households that were disenrolled from CARE, suggesting an average increase of $16.64 for those households losing the subsidy ($2.33/0.14).

Tables 2.2 and 2.3 also presents a series of robustness checks. We assess sensitivity to choice of temporal bandwidth by shrinking the bandwidth to 9 billing periods (column 2) and 6 billing periods (column 3) around the 26th billing period. We assess whether the results are sensitive to dropping the 27th billing period, something that is more relevant for the disconnection analysis given the non-CARE bills had not been presented yet (column 4). We restrict the analysis to accounts that never closed, to assess whether disenrolling from CARE affected the sample population (column 5). We provide estimates without the household fixed effects (column 6). And, weather controls are added to our core model in one test (column 7). In nearly all cases, the qualitative conclusions are unchanged. When we shrink the bandwidth to 6, the effect size for disconnections decreases, the standard error

increases, and the estimate is no longer statistically significant. However, this is likely due to the fact that smaller bandwidths put more weight on the 27th bill. The effect size is larger for disconnections when we drop the 27th billing period. Appendix Figure A5 tests whether households are more vulnerable to higher temperatures after the 26th bill, however the large confidence intervals around the estimates preclude learning meaningful lessons.

Table 2.2: 26 Months After CARE Enrollment Using Different Controls (1)

	Core model (1)	BW = 9 (2)	BW = 6 (3)	Drop 27th bill (4)	Never closed (5)	Drop Account FE (6)	Add Weather (7)
CARE Enrollment							
Post 26 Dummy	-0.1363	-0.1301	-0.1166	-0.1619	-0.1457	-0.1376	-0.1363
	(0.0023)	(0.0022)	(0.0020)	(0.0028)	(0.0027)	(0.0023)	(0.0023)
Observations	2,101,474	1,560,000	1,027,922	2,018,433	1,259,677	2,106,434	2,101,474
Log Electricity Usage (kWh)							
Post 26 Dummy	-0.003	-0.006	-0.003	-0.003	-0.008	-0.007	-0.003
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	2,098,351	1,557,659	1,026,367	2,015,435	1,257,923	2,103,307	2,098,351

This table shows the intercept shift in CARE enrollment and electricity usage after 26th billing cycle since CARE enrollment. All specifications have a linear term for time since enrolled in CARE that is allowed to vary across the 26th cycle. The running variable is the number of bills since first enrolled in CARE. Each set of rows represents a separate model estimated for a given outcome. The model has zip by calendar week, billing start date, and individual account fixed effects. Standard errors (in parentheses) are clustered at the zip code level. Column 4 drops the 27th cycle. Column 2 and 3 altar the bandwidth. Column 5 only uses accounts that did not close during the entire bandwidth period. Column 6 drops account fixed effects. Column 7 adds weather from the current billing period: a cubic spline function of maximum temperature, the number of days with precipitation above 0 and below 0.5 inches, and the number of days with precipitation above 0.5 inches. There are fewer observations with log electricity usage because some accounts have non-positives values and are dropped from the analysis due to the log transformation.

Table 2.3: 26 Months After CARE Enrollment Using Different Controls (2)

	Core model (1)	BW = 9 (2)	BW = 6 (3)	Drop 27th bill (4)	Never closed (5)	Drop Account FE (6)	Add Weather (7)
Bill Amount (Current $)							
Post 26 Dummy	2.33	1.85	2.50	2.95	2.53	2.35	2.32
	(0.23)	(0.24)	(0.21)	(0.28)	(0.28)	(0.23)	(0.23)
Observations	2,101,474	1,560,000	1,027,922	2,018,433	1,259,677	2,106,434	2,101,474
Disconnection (x100,000)							
Post 26 Dummy	72.43	70.66	28.03	111.50	73.84	79.67	72.41
	(16.65)	(19.18)	(24.20)	(19.04)	(22.37)	(16.94)	(16.65)
Observations	2,101,474	1,560,000	1,027,922	2,018,433	1,259,677	2,106,434	2,101,474

This table shows the intercept shift in bill amounts and disconnections after 26th billing cycle since CARE enrollment. All specifications have a linear term for time since enrolled in CARE that is allowed to vary across the 26th cycle. The running variable is the number of bills since first enrolled in CARE. Each set of rows represents a separate model estimated for a given outcome. The model has zip by calendar week, billing start date, and individual account fixed effects. Standard errors (in parentheses) are clustered at the zip code level. Column 4 drops the 27th cycle. Column 2 and 3 altar the bandwidth. Column 5 only uses accounts that did not close during the entire bandwidth period. Column 6 drops account fixed effects. Column 7 adds weather from the current billing period: a cubic spline function of maximum temperature, the number of days with precipitation above 0 and below 0.5 inches, and the number of days with precipitation above 0.5 inches.

2.4 Change in Disconnections Using 2080-2099 Weather Projections

We provide a back-of-the-envelope calculation of possible changes in electricity usage and disconnection risk that would occur if today's weather resembled projected weather for the 2080-2099 period. We use RCP scenario 8.5 output from 22 climate models from Climate Model Intercomparison Project phase 5 (CMIP5) (Taylor et al., 2012) and calculate the within-model changes in the distribution of daily temperatures for our sample area from a base period of 2000-2019. We preserve the daily variation in temperature by calculating the spline values at the daily level before aggregating to the monthly and annual level.

These calculations cannot be used to definitively project the effects of climate change for several reasons. Household energy appliance penetration rates, energy efficiency, incomes, and relative prices of essential goods and services might change. The climate change models themselves include uncertainty. Some households might adjust their behavior due to new expectations regarding the frequency of hot weather, potentially saving more for summer months or improving insulation in their homes. Households could become more vulnerable in the future with greater climatic variability or compounding hazards (e.g. wildfires coinciding with heat waves). Nonetheless, these calculations provide a stylized illustration of the potential financial strain on low-income households caused by weather-related changes holding all else constant.

Figure 2.4A shows the annual changes in log electricity usage for every CMIP5 model (gray lines) and the average across all models for each year (black line). On the average, we project a 9.1 percent increase in electricity usage between the 2000-2019 and 2080-2099 time periods. We project an increase in disconnections of 30.25 per 100,000 on average each billing cycle (2.4B). This represents a 12 percent increase for our low-income population by 2080-2099 relative to our sample average (30.25/246). Given there are approximately 1.2 million accounts on CARE in SCE territory (Southern California Edison, 2017d), this translates into 363 more disconnections each month or 4,356 more disconnections each year

in this population alone.

The year-to-year variation within each model (gray lines in Figure 2.4A) highlights the importance of considering peak demand when assessing the impact of climate change on energy systems. This finding complements existing analyses that show that energy supply might be constrained during heat waves, at the same time that energy demand peaks (Auffhammer et al., 2017; Sathaye et al., 2013; Burillo et al., 2017; Ke et al., 2016; Brockway and Dunn, 2020). However, our study focuses on a low-income population whereas existing studies focus on the general population.

These annualized projections mask important seasonality in usage and disconnections. Figure 2.4B projects changes for each calendar month assuming that billing periods correspond to the calendar month and the disconnections occur three calendar months later (as opposed to 51-75 days later). For this figure, we take the difference in the projections between the 2000-2019 and 2080-2099 periods for each model. We see that usage increases by about 20 percent on average during the summer on average across the 22 models. Disconnections increase by over 45 per 100,000 in the late summer and early fall on average, or an 18 percent increase over the baseline risk.

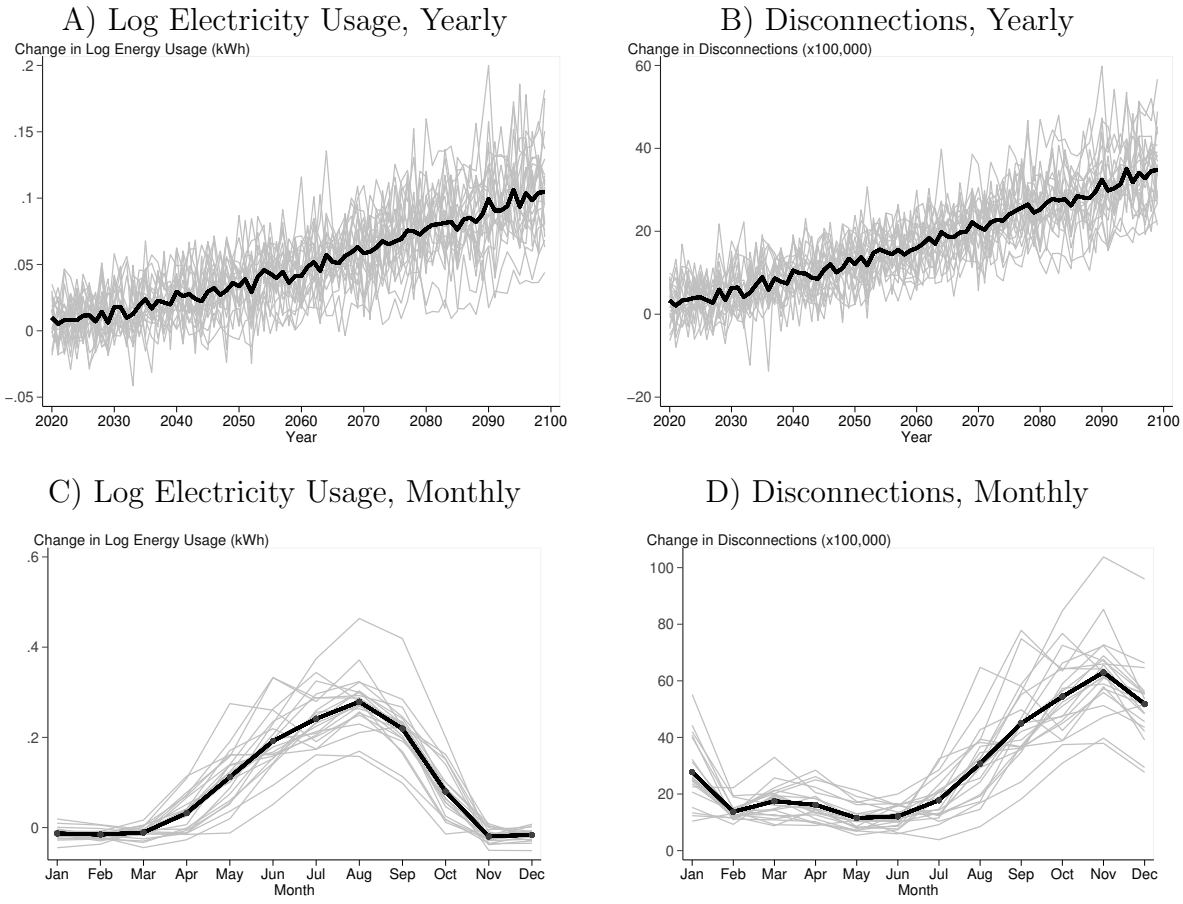


Figure 2.4: Climate Change Projections by Month and Year

Panel A shows the projected change in log electricity usage by year compared to 2000-2019. Panel B shows the projected change in disconnections (x100,000) by year compared to 2000-2019. Panel C shows the projected change in log electricity usage for each month in 2080-2099 compared to 2000-2019. Panel D shows the projected change in disconnections (x100,000) for each month in 2080-2099 compared to 2000-2019. The data from these projections comes from 22 CMIP5 models (RCP 8.5), restricted to the two CMIP model grid points in Southern California Edison's service area. The changes in daily temperatures are calculated using each model's own output for the 2000-2019 period as the baseline. We then multiply the change in daily temperatures by the estimates in Figures 2.1 and 2.2. The solid black line is the average across all models and years and the gray lines represent the average for each model.

2.5 Discussion

The results from this study show that hot weather causes financial strain on low income households, as evidenced by an increase in electricity expenses and subsequent electricity disconnections. The theoretical link between temperature and disconnections operates through multiple channels, though we focus our attention on electricity expenses. Financial strain may also be exacerbated by reduced labor income or added morbidity risk due to hot weather (Barreca et al., 2016; Graff Zivin and Neidell, 2014). We present additional evidence to support the expense channel. We show that households experience higher expenses and far more disconnections due to abruptly losing the CARE subsidy after an automatic eligibility verification at two years on the CARE program.

With regards to climate change, the projected increase in summer temperatures at the end of the century could increase electricity expenses and disconnection risk for our low-income population. Our back-of-the-envelope calculation suggests a 12 percent increase in disconnections if today's weather resembled the projected weather for 2080-2099, though this calculation should be interpreted with caution since we do not account for future changes in energy efficiency, among other things. In addition to an expected increase in summer temperature, summer variability from year-to-year might also impact low income households to the extent that financial distress is increasing in the size of the shock to expenses. Analyzing interannual variability in summer temperature projections, as done in one recent study,55 represents an area where advances in climate science would help assess the financial stressors of climate change. Year-to-year variability is also important due to issues related to peak energy demand (Auffhammer et al., 2017; Sathaye et al., 2013; Burillo et al., 2017; Ke et al., 2016; Brockway and Dunn, 2020).

Our findings are consistent with emerging evidence that some sub-populations may be adversely affected by hot weather in ways that are masked in more aggregated analyses (Park et al., 2020; Kim et al., 2021; Heilmann et al., 2021). This appears to be true even in highly developed economies such as the U.S. We find that currently routine heat events

cause significant financial distress for poorer households that manifests in electricity disconnections. The immediate costs of these disconnections include significant material hardship, one potential reason why disconnection durations themselves are generally short (Harrison and Popke, 2013; Hernández, 2016; Heflin et al., 2011; Southern California Edison, 2014a). The experience of electricity disconnection itself, even if brief, may have lasting psychological costs (Harrison and Popke, 2013; Hernández, 2016). The heightened salience of disconnection risk may lead low-income households to reduce their energy consumption during hot weather events in the future, which could lead to higher morbidity risk.

There are two broad categories of policy response that could be employed to reduce energy burdens related to extreme heat events. One policy response may be to reduce the price of energy, including subsidies like California's CARE program. A concern with such subsidies is that they might lead to more energy consumption, which could hinder efforts to curb aggregate greenhouse gas emissions. While theoretically possible, existing research suggests that energy consumption responds little to price changes in the general population (Reiss and White, 2005; Burke and Abayasekara, 2018; Deryugina et al., 2020; Labandeira et al., 2017; Zhu et al., 2018), and could be even less so for the low-income population. Time-of-use rates could be used to reduce financial distress while also encouraging intraday demand shifting to when solar power is more plentiful (Dutta and Mitra, 2017). Another policy response is to incentivize the adoption of energy efficiency technologies or better insulation, such as intended with California's 1977 Title 24 building codes (Novan et al., 2022; Chuang et al., 2022; Kotchen, 2017). This may be especially relevant in Los Angeles County where people tend to live in older homes that are less energy efficient (Fournier et al., 2019). Principal-agent problems due to high rates of rentership in low-income populations may call for a more nuanced policy response that targets landlords (Myers, 2020; Energy Insights USA, 2020).

Policymakers may have dual goals of improved indoor thermal comfort and reduced financial burdens, which could require multifaceted policies. Technological solutions that improve cooling efficiency or provide access to air conditioning could improve thermal comfort but the impact on energy expenses is unclear since households might use the cooling technology

considerably more (Greening et al., 2000). Energy subsidies, though they could reduce expenses, might do little to improve thermal comfort to the extent that low-income households possess inferior cooling technology or lack access to air conditioning altogether (Chen et al., 2019).

2.6 Methods

Electricity Data

The electricity data come from Southern California Edison (SCE), which serves approximately 15 million people in Southern California, including much of the greater Los Angeles area. SCE provided a sample of accounts from 300 zip codes with the most low-income households in absolute numbers. Appendix Figure A6 illustrates the map of the sample zip codes. This is a climatically diverse area which helps with statistical power when controlling for time trends. The California Energy Commission defines 16 Climate Zones across California based on temperature, weather, and energy use patterns (California Energy Commission, 2022). The zip codes in our sample cover eight of these 16 zones, highlighting the large degree of spatial variation in temperature across SCE territory which spans coastal, inland, and desert climates. Appendix Figure A7 illustrates the daily max temperature for four different zip codes in our sample area for only August of 2015. For example, the average temperature in the city of Palm Springs (located inland in the desert) is over 20 degrees higher than in the city of Santa Monica (located on the coast).

SCE provided us with a random sample of households that enrolled in the California Alternate Rates for Energy (CARE) or the Family Electric Rate Assistance (FERA) subsidy programs between 2012 and 2018. The sampling procedure was restricted to those accounts who were present for one year to increase the sample size. The original sample included up to 1,000 accounts per zip code for the 2012-2018 time period for a total of 299,799 accounts. Though we sampled accounts present for one year, we dropped 574 accounts with fewer than

12 bills (due to irregular billing period lengths) to help facilitate the use of account fixed effects in our core empirical model. Our final sample includes 299,225 accounts. SCE was unable to provide disconnection data after August of 2017, so we also dropped bills that started after that point.

We have information on the electricity usage in kilowatt-hours (kWh) and expense amount in current dollars for each billing period. The data have specific disconnection dates for each account. We also have information on which billing periods the accounts were enrolled in CARE or FERA, though CARE is by far more common. Accounts in our low-income sample are on FERA for fewer than 1 percent of all bills (Table 2.1). We observe the zip code of residence and billing period start and end dates, which we use to match to daily weather data.

Sampling on CARE accounts is a good approach to identifying low-income households. Unlike other social insurance programs, the CARE program has an excellent takeup rate among the target population (Currie, 2006). In 2017, an estimated 84 percent of eligible households enrolled in CARE (Southern California Edison, 2017d). The household income cutoff for CARE is set at 200 percent of the poverty line. For example, a four-person household with two children earning $49,200 or less (in 2017) would qualify for CARE (Southern California Edison, 2017a). Households who earn one dollar over the cutoff would no longer qualify for CARE, but they would qualify for the FERA program with incomes up to 250 percent of the poverty line (Southern California Edison, 2017a). The CARE subsidy is generally close to 30 percent every year of our sample, though the exact amount varies depending on the consumption tier, e.g. baseline vs. 101-400 percent of baseline, which is itself a function of the Climate Zone (Southern California Edison, 2012b,a, 2013c,b, 2014c,b, 2015b,a, 2016b,a, 2017b,a). The FERA subsidy was either much smaller during our sample period, i.e. 12 percent in years 2015-2017, or only covered certain tiers of consumption (Southern California Edison, 2012c, 2013d, 2014d, 2015c, 2016c, 2017c). Accounts in our sample were much more likely to be on CARE than FERA. The accounts in our sample were on CARE for 74 percent of bills, but only on FERA for 1 percent of bills (see Table 2.1).

SCE attempts to keep billing periods to approximately one month, though the length varies because meters are not typically read on the weekends. The billing periods last between 27 and 34 days in 98.7 percent of bills. In rare cases where the bill dates start on a weekend or holiday, we reassign the start date to the previous Friday or non-holiday weekday if Friday is a holiday (e.g. Thanksgiving), which affects 1.1 percent of bills. Most accounts with the same billing start date have the same end date. In order to mitigate potential confounders tied to unusual bill lengths, we assign approximately 2.6 percent of bills with unusual lengths to the modal bill length for a given billing start date.

We drop the first 3 bills after an account opens since SCE's disconnection policy effectively precludes disconnection so early on. SCE's policy is to disconnect no sooner than 53 days from the date the bill is presented (Southern California Edison, 2013a). We do not have data on reconnection dates, but households that are disconnected are usually reconnected very quickly. According to an SCE report on August 2018 disconnections, 85 percent of all customers disconnected were reconnected within 24 hours (Southern California Edison, 2018).

Weather Data

We use PRISM data that covers SCE's service area (as well as the rest of the contiguous U.S.) with a 4x4 kilometer grid of model-predicted daily weather conditions. The key variables include maximum temperature and precipitation. We assign a PRISM point to a zip code if the point falls within the zip code boundaries. The average area of zip codes in our sample is 95.9 square kilometers, though half the zip codes are less than 22.5 square kilometers. There are 62 zip codes that do not contain a PRISM point, so we assign the nearest PRISM point to the zip code's centroid. The PRISM data are better than weather station observations for this analysis because of the excellent spatial coverage, which has been shown to closely mirror weather station observations when there are overlap between PRISM points and stations (Walton and Hall, 2018; Auffhammer et al., 2013). Our empirical model includes

date fixed effects, causing the estimates to be powered by differences in temperature across coverage points on the same day. There are over 1,600 PRISM points for the 300 zip codes in our sample area. Even drawing from weather stations in and around SCE's service area, there are only some 60 official weather stations.

We aggregate the days' temperature over the duration of the billing period for a given zip code. We use a nonlinear spline function of the day's maximum temperature in our core empirical model, though we also show results using binned categorical temperatures (e.g. days with maximum temperatures above 100°F). We calculate the nonlinear daily function value of the spline (or bin) at the PRISM point prior to aggregating to the zip-billing period level to preserve the daily variation in temperature. For a given billing period, we assign the weather from the bill start date to the day before the bill's end date, which is necessary to avoid double counting on the days billing periods change.

Contemporaneous impact of temperature on electricity use

Our empirical model relies on unpredictable weather variation for a given zip code and time of year consistent with the emerging climate-economics literature (Dell et al., 2014). Before assessing impacts on disconnection, we first estimate the following empirical model with electricity use and bill amount as outcomes:

$$Y_{hzt} = \beta f(TMAX_{zt}) + X_{hzt} + \omega_h + \alpha_{zw} + \rho_t + \epsilon_{hzt} \tag{2.1}$$

where Y is either the natural log of electricity use (kWh) or the bill amount in current dollars for household h in zip code z for the billing period starting on year-month-day t. $f(TMAX)$ is a cubic spline function in the daily temperatures in zip code z and billing period with start date t. The cubic spline has separate pieces determined by knots at 40, 60, 80, and 100°F (Smith, 1979). We construct the value of each piece at the daily level before summing up to the billing cycle to ensure our estimates capture daily variation in temperatures. In our analysis, we set 75°F as the reference category. Specifically, we

calculate the difference between a given temperature and 75°F for each piece of the spline and then multiply that difference by the piece's parameter value. Because there are multiple pieces, this spline model is more flexible than the popular Cooling Degree Days approach, which imposes a linear relationship between temperature and electricity consumption above a given baseline temperature and a null relationship below that baseline. We favor the spline model as a means to improve statistical power compared to the binned model that is popular in the weather-economics literature (Dell et al., 2014), though we estimate the latter as a robustness check.

X controls for account age fixed effects in billing periods with a topcode of 24 periods (4, 5, 6, ..., 24, >24 periods). As mentioned in the Data section, we drop the first three bills for each account. X also controls for rainfall (days with 0.01-0.50 inches of rain and over 0.5 inches of rain, respectively), though rainfall is relatively infrequent in this area (Table 2.1). Rainfall is included as a control to account for the possibility that people's behaviors change on rainy days in ways that might be tied to electricity consumption.

We have household fixed effects (ω) to account for households entering and exiting our sample in ways that might be tied to the households' electricity consumption. Zip-code-by-calendar-week fixed effects (α) mitigate potential confounding effects from differences across zip codes, like climate type and housing characteristics, that might be spuriously correlated with electricity consumption patterns. Zip-code-by-calendar-week fixed effects also address seasonal changes in energy consumption that might be spuriously correlated with other behavioral factors or baseline characteristics. For example, zip codes with a larger proportion of school-age children at home during summer vacation may exhibit different seasonality. These fixed effects, which are common in the climate-economics literature, are each unique indicator variables that absorb any potential confounder, whether measurable or unmeasurable, that is fixed or constant at the given dimensions (Dell et al., 2014). We assign the first seven days of the year to week 1, the next seven to week 2, and so on up until week 52; any remaining days are included in week 52. We control for bill start date fixed effects (ρ) to help improve precision and avoid spurious time-series correlation between

energy outcomes and temperature. Standard errors are clustered at the zip code level to account for serial correlation within zip codes.

Impulse response function: Impact on disconnections

To assess impacts on disconnection risk, we rely on an impulse response function that traces out disconnection risk over time after the billing period is over (Jordà, 2005). We estimate:

$$Y_{hzt} = \beta^p f(TMAX_{zt}) + X^p_{hzt} + \omega^p_h + \alpha^p_{zw} + \rho^p_t + \epsilon^p_{hzt}, \quad \text{where } p = 0, 1, 2, ..., 6 \tag{2.2}$$

where Y is an indicator variable (equal to 0 or 1) if household h was disconnected in time period t+p, which is p periods after the close of the billing period t. The model controls are identical to (1), with temperature representing exposure during the billing period t. The disconnection outcome periods (p) are broken out into 25-day period intervals after the close of the current billing period up to 150 days later, i.e. days 1-25, 26-50, 51-75, ..., 126-150 days after the billing period t has ended. Appendix Figure A3 estimates an impulse response function using a 10-day window as a robustness check.

We expect the effects to be largest 51-75 days later since SCE's policy is to disconnect approximately 53 days from the date the bill is presented (Southern California Edison, 2013a). In practice, customers can appeal the bill amount, arrange for partial payment or enroll in balanced payment plans, which might delay disconnections beyond the 53 days. To the extent that high energy bills expedite disconnections for households already carrying large overdue balances, we might see impacts sooner than 50 days.

Equation 2.2 is similar to a linear projection model, which is a popular approach for estimating effects of a shock over time (Jordà, 2005). However, we do not control for lagged weather in our core model in the interest of improving power (Choi and Chudik, 2019). We test the robustness of this modeling approach by controlling for one previous and one future

billing periods' weather in a robustness check (Appendix Figure A4).

We estimate the contemporaneous impacts of temperature on disconnection risk in the current billing cycle as well. SCE's policy is to not disconnect homes when the day's temperature is above 100F (Southern California Edison, 2020). As a result, hot weather could simply decrease disconnections in the current period and displace them into subsequent periods. We also estimate a placebo test to verify that temperatures in the current billing period are unrelated to disconnections 1-25 days prior to the temperature realization.

Regression discontinuity model: Disenrollment from CARE after 26 bill periods

With Equation 2.2, we cannot isolate whether temperature influences disconnection risk through energy expenses or some other channel, like lost labor income. To isolate the effects of energy expense, we exploit a feature of the CARE program. The utility company provides existing CARE customers with a notice to recertify their eligibility after two years on the program, and customers have 45 days from the notice to recertify (Southern California Edison, 2014a). This verification process leads to an acute drop in CARE enrollment probability after the 26th bill, which is unlikely to coincide with a drop in income that might independently increase disconnection risk. If anything, we might expect households to increase their labor hours to offset the increase in prices, which would understate the importance of the energy expense channel.

To quantify how the change in CARE status might affect disconnection risk, we estimate a regression discontinuity design of the following form:

$$\begin{aligned} Y_{hzt} = \psi BILL_{hzt} + \lambda BILL_{hzt}(BILL_{hzt} > 26) + \theta(BILL_{hzt} > 26) \\ + \omega_h^p + \alpha_{zw}^p + \rho_t^p + \epsilon_{hzt}^p, \end{aligned} \tag{2.3}$$

where $BILL$ is the running variable measuring the number of periods from the current billing periods to the 26th billing period. $(BILL_{hzt} > 26)$ is an indicator for whether the

current billing period is greater than 26, when the chance of CARE disenrollment goes up. Similar to Equation 2.1, we include zip-calender-week, billing date, and household fixed effects. We do not control for account age fixed effects in Equation 2.3, however, since account age is strongly correlated with billing periods on CARE. Equation 2.3 does not control for weather, though we add temperature as a control in a robustness check. Appendix Figure A5 takes this a step further and interacts temperature with the ($BILL_{hzt} > 26$) dummy.

We focus on four main outcomes (Y) for this analysis, all measured in the cycle t for household h. First, we estimate Equation 2.3 on the probability of CARE enrollment to validate the assumption that CARE enrollment drops discontinuously after the 26th billing period. Second, we estimate Equation 2.3 on log electricity usage to explore whether households might reduce energy consumption in a way that would dampen disconnection risk. Third, we quantify impacts on energy expense in order to assess the magnitude of the added expense. Finally, we estimate effects on disconnection probability. We focus on contemporaneous disconnection risk for ease of interpretation, though we do not expect disconnection risk to increase immediately after the 26th cycle due to the delay between usage and invoicing. Unlike with Equation 2.2, the regression discontinuity approach is not ideally suited for estimating an impulse response function since there is only one break in time.

For this analysis, we restrict the sample to the 15th through 38th bill after CARE enrollment, or a bandwidth of 12 bills around the 26th bill. We drop accounts that are enrolled in CARE as of January 2012, the first month in our sample, since we cannot determine their CARE enrollment date. These sampling restrictions leave us with 130,076 accounts and 2,106,662 bills. The regressions naturally drop a small number of observations due to multicollinearity, depending on the controls included in the model.

Figure A5 tests whether households are more vulnerable to higher temperatures after the 26th bill. We present the main effect of temperature as well as the modified effects after the 26th bill on the probability of disconnection 51-75 days into the future. The hypothesis is that temperature shocks after the 26th bill are more expensive, and therefore, should

increase disconnection risk further. The estimates are inconclusive due to large confidence intervals. The main effect of temperature is similar to our core estimates in Figure 2.2C, though the standard errors are larger. There is no differential effect of temperature after the 26th bill, however these estimates have large confidence intervals. Future work should test this hypothesis over a longer time period and broader geography for more statistical power.

3

The Effects of Hot Weather on Rural Indian Diets[1]

3.1 Introduction

Rural households in developing countries, like India, are vulnerable to extreme heat due to their reliance on home-grown foods for both sustenance and income (Taraz, 2018). Existing research documents that temperatures above 85°F are particularly harmful to crop yields in India (Garg et al., 2020b). Crop losses could impact household welfare across several dimensions, though household nutrition is a primary outcome of interest given the downstream costs. Poor nutrition can have an impact on long-term health (Sharma et al., 2020b), human capital accumulation (Georgieff, 2020), psychological well being (Pailler and Tsaneva, 2018), and labor productivity (Thomas et al., 2006). Crop losses are likely to harm household nutrition, which is determined by both the quantity of calories as well as the nutrient and mineral composition of these calories. However, directly measuring the impacts of temperature extremes on household nutrition is an important endeavor because the magnitude of the effect could be mitigated by adaptation. Households could decrease the amount of calories they eat or consume cheaper, less nutritious calories. Households might also spend down savings or decrease non-food purchases to avoid cutting food expenditures. In order to add insights into the magnitude of these adaptation strategies, we use survey responses from over 300,000 households in rural India to estimate the impact of extreme heat on food consumption choices and diet quality.

[1]This chapter is coauthored with Professors Manisha Shah and Alan Barreca. The coauthored paper does not include climate projections, which I conducted alone.

We measure food consumption using India's National Sample Survey (NSS), a nationally representative repeated cross-section of Indian households, from 2003 to 2012 (the last year that includes detailed food consumption). The NSS includes questions on the consumption of roughly 150 food items, which we convert into household-level calorie and iron consumption. We use calories because they feature, along with macronutrients, as the traditional measure of nutrition in past studies, which also use the NSS (Carpena, 2019; Deaton and Drèze, 2009). We use iron because iron deficiency is endemic to rural India. India's National Family Health Survey 4 (NFHS-4), run in 2015 and 2016, found that 59 percent of children below 5, 58 percent of breastfeeding women, and 53 percent of all women were anemic, caused in part by iron deficiency (IIPS and ICF, 2017).[2] Iron is particularly important during pregnancy, as pregnant women need more iron than usual, in part because some of it is diverted to their fetus. Iron deficiency during pregnancy is linked with increased maternal mortality, preterm births, low birthweight, and decreased human capital accumulation in children.[3] In adults, anemia lowers energy levels, labor productivity, and mental health (Marcus et al., 2021; Thomas et al., 2006).

We estimate the impact of temperature during the growing season on food consumption in the subsequent year, which would presumably be when households either consume or sell their production. Also, by focusing on nutritional outcomes in the subsequent year, our estimates abstract from heat's impact via more contemporaneous channels such as labor productivity. Our treatment variable is the frequency of days in a given temperature bin, e.g., days between 90 and 100°F, for the growing season (June-December). The unit of observation for this variable is the district and year. To address potential confounders, our key control variables include district and year fixed effects. The core estimates are identified

[2]Anemia, when a person lacks healthy red blood cells for the transport of oxygen, is a common clinical proxy measure for iron deficiency. Anemia has multiple causes, including iron deficiency, genetic factors, gastrointestinal diseases such as hookworm, postpartum hemorrhage, and folate or B12 deficiencies. This multitude of factors, and their frequent overlap, makes it difficult to assess exactly how much anemia is caused by iron deficiency (Chaparro and Suchdev, 2019).

[3]Breastmilk is low in iron, so newborn infants rely on iron stored during the fetal stage for the proper development of organs, including the brain (Georgieff, 2020).

by unpredictable temperature swings across growing seasons for a given district. We also control for rainfall shocks, which can have a meaningful effect on crop yields and food consumption (Carpena, 2019). Our preferred specification includes state specific time trends to control for correlated differences in warming and nutrition over time between states. We show that the results are robust to including district specific time trends as well.

There are two important findings from our research. First, although extreme heat decreases crop yields, we find no statistically significant impact of extreme heat on total household caloric or iron consumption, and we can rule out large effect sizes. However, we find important distributional impacts, with an increase in the number of households that qualify as extremely malnourished, in terms of both calories and iron. For example, one day above 110°F increases the fraction of households consuming less than 80 percent of the recommended caloric intake by 0.36 percentage points, an increase equivalent to 3.1 million people. On average, households in our sample experienced 0.6 days above 110°F per growing season, with 10 percent of households experiencing over 2 such days.

Second, we find suggestive evidence that households respond to crop losses by purchasing more food, though these purchased foods may be relatively low in iron. We estimate our model separately for home-grown calories, purchased calories, home-grown iron, and purchased iron. We find that heat during the growing season causes home-grown calorie consumption to decline but leads to a commensurate increase in purchased calories, explaining why we find no impact of heat on total calories. However, we find that iron from purchased foods does not increase alongside calories. This finding implies that households respond to heat shocks by sacrificing the quality of their diet to maintain the quantity of calories consumed.

This paper makes three main contributions. First, we add to a growing literature on heat's impact on food consumption by considering adaptations. Heat's impacts on yields are well documented (Taraz, 2018), but the crop losses are difficult to translate into societal welfare losses without understanding how households respond. We consider the impacts

of weather shocks on home-grown and purchased foods separately, something that is largely absent from the literature but important in rural contexts (Jayasinghe et al., 2017). We show that households purchase more foods in response to heat shocks. In doing so, households may be cutting into savings or sacrificing non-food purchases or investments (e.g. human capital).

Second, we capture important distributional effects. Past papers on nutrition tend to focus on aggregate measures, such as total calorie consumption or diet composition (e.g., percentage of calories from fruits vs. grains). By using low-consumption thresholds (e.g., consumption below 80 percent of recommended levels of calories), we capture important distributional impacts that aggregate measures miss.

Third, we make methodological contributions to measuring diet quality. Our work emphasizes the need to include micronutrients and minerals when estimating climate change's impact on diet quality. Although it is particularly important in rural India, iron rarely features in past economics studies on diet quality in the area (Carpena, 2019; Deaton and Drèze, 2009). The NSS provides tables to convert food quantities into calories and macronutrients such as protein and fat. In this project, we use the raw nutritional data that the NSS tables reference, allowing us to directly measure iron consumption (Gopalan et al., 1989). We add to previous work that shows that calorie consumption and diet composition are affected by droughts in Mexico (Hou, 2010) and India (Carpena, 2019) and rainfall and temperature shocks in Nigeria (Dillon et al., 2015) and Tanzania (Randell et al., 2022).

3.2 Data

National Sample Survey (NSS)

The National Sample Survey (NSS) is a nationally representative repeated cross-section of Indian households. We use Schedule 1 of the NSS, which includes detailed household-level consumption data. Our dataset consists of 8 rounds of the data (rounds 59-64, 66, and 68),

spanning 2003 to 2012 (National Sample Survey Office, 2012).[4] Each household response details the consumption records of roughly 150 food items, separated into home-grown and purchased foods, over the previous 30 days. It also includes the number of meals that 1) household members eat outside of the home and that 2) non-household members eat in the home. Following past work, we use these numbers to adjust our food consumption figures (NSS, 2014). Note that the survey does not measure exactly how much food was eaten, as some could be lost during the cooking process or wasted after preparation. It also does not indicate who, within the household, consumed specific foods. These consumption measures therefore indicate household-level food availability for intake, rather than intake itself. The survey also includes all household members' basic demographic information. The dataset includes 314,425 households, with a total of 1,599,551 people, across 575 district-areas that we construct to be geographically consistent over the entire time frame.

Nutritional Outcomes

In order to generate the main outcomes of interest, we translate the detailed food consumption data from the NSS into household-level caloric and iron consumption. We obtain each food item's caloric value from nutritional value tables provided by the NSS. The NSS-provided tables do not include the iron content of foods, resulting in iron being excluded from previous analyses (Deaton and Drèze, 2009; Carpena, 2019). These NSS-provided tables, however, are based on a dataset from *Nutritive Value of Indian Foods* (Gopalan et al., 1989), which does include iron. We use tables from this original source to create two variables: daily per capita calories consumption and daily per capita iron consumption. These variables are both right-skewed and roughly log normal, so we take their log.

We generate four additional dependent variables to estimate the distributional impacts of extreme heat on food consumption. We match the NSS data with individual-level dietary requirements from *Nutrient Requirements and Recommended Dietary Allowances For*

[4]Rounds 65 and 67 do not include household food consumption questions.

Indians, and then aggregate requirements to the household-level (ICMR, 2010). We calculate calories and iron adequacy percentages by dividing each household's consumption by its recommended intake levels. Using these percentages, we create dummy variables indicating whether household-level caloric or iron consumption is below 100 percent or 80 percent of the recommended levels. In addition, we create analogous dummy variables for thresholds ranging from 50 percent to 150 percent.

We use these threshold dummy variables to assess how extreme heat affects households at different points in the nutrient adequacy distribution. In our preferred definition, we label a household below the 100 percent threshold as experiencing "undernourishment," and one below the 80 percent threshold as experiencing "extreme undernourishment." Our results are robust to defining extreme undernourishment with thresholds between 50 and 80 percent adequacy. An increase in the percentage of households below the 100 percent threshold indicates households being pushed from adequate dietary availability to undernourishment, and an increase in the percentage of households below the 80 percent threshold indicates households being pushed from undernourishment to extreme undernourishment. Because each household only answers the survey once, after treatment, we are unable to investigate distributional effects by separating households into groups by baseline food consumption levels. In addition, as we only have household-level food consumption we are unable to tell exactly who, within the household, is consuming sub-adequate levels of calories or iron.

Table 3.1 shows average levels of nutrient consumption across the households in the sample (for the full distribution of nutrient consumption, see Appendix Figure B2). The average per capita daily caloric consumption is around 2150 kCal/person-day, roughly the recommended amount of caloric intake, and the average calorie adequacy percentage is 103 percent. Fifty-one point five percent of households consume below 100 percent the recommended amount of calories, and 20.6 percent of households consume less than 80 percent of the recommended level of calories. Iron consumption, with an average of 16 mg/person-day, is often lower than needed. The average iron adequacy percentage is 86.9 percent. Sixty-seven point nine percent of households consume less than 100 percent of their recommended

level of iron, and 50.4 percent of households consume less than 80 percent of it. These statistics illustrate two main points. First, focusing on average levels of food consumption masks significant heterogeneity in diet quantity and quality across households. Second, focusing only on calories would lead one to miss how low iron consumption is in rural India.

The NSS separates foods into home-grown and purchased items. We use this distinction to construct four more dependent variables: iron and calories from home-grown foods versus from purchased foods. Around 65 percent of households consume some amount of home-grown foods, and about 27 percent of iron (4.39 mg/person-day) and 28 percent of calories (603 kCal/person-day) come from home-grown foods.

Weather Data

We use weather data from the ERA5 reanalysis model (Copernicus Climate Change Service, 2012), which includes hourly temperature and precipitation across a 31km by 31km grid covering the entire globe (Hersbach et al., 2020). For each day and gridpoint, we calculate the maximum temperature and total precipitation. Following Garg et al. (2020b), we separate weather into growing season (June to December) and non-growing season (March to May). We assign temperature to each district-year-season as follows. For each gridpoint-year-season, we count the number of days within distinct 10 degree temperature bins ranging from below 70°F to above 110°F. We then average these temperature bins across all the gridpoints within each district. For the 18 small districts that do not contain any ERA5 gridpoints, we assign them the nearest ERA5 grid-point to their respective centroid. Table 3.1 summarizes the average number of growing season days within each temperature range and the rainfall shock variable. It shows that, on average, India is hot: households in our sample experience an average of 109 days between 80 and 90°F, 51 days between 90 and 100°F, 8 days between 100 and 110°F, and 0.6 days above 110°F during the seven month growing season. Figure 3.1 (full year) and Appendix Figure B1A (growing season) map the number of days per year with maximum temperature above 100°F for each district. They

highlight the vast climatic diversity across India: the number of days per year with maximum temperature above 100°F ranges from 0 to above 140. About 25 percent of households in our sample reside in districts with no growing season days above 100°F over our study period. We also construct a "rainfall shock" control variable, which we describe in the Appendix, as is done in Shah and Millett Steinberg (2017).

Table 3.1: Summary Statistics: NSS and ERA5

	Mean	SD
Daily calories (kCal/capita)	2,149	571.34
Calories adequacy percentage (CAP)	103.20	30.06
CAP <100% (dummy)	51.51	49.98
CAP <80% (dummy)	20.64	40.48
Daily iron (mg/cap)	16.00	7.07
Iron adequacy percentage (IAP)	86.91	40.56
IAP <100% (dummy)	67.90	46.69
IAP <80% (dummy)	50.36	50.00
Daily calories from home (kCal/capita)	602.70	751.58
Daily calorie purchases (kCal/capita)	1,538.69	759.10
Daily iron from home (mg/capita)	4.39	6.49
Daily iron purchases (mg/capita)	11.53	6.95
Growing season days ≥110°F	0.59	1.71
Growing season days 100-110°F	7.65	11.00
Growing season days 90-100°F	50.61	34.23
Growing season days 80-90°F	109.26	44.01
Growing season days 70-80°F	33.23	26.98
Growing season days <70°F	12.66	35.26
Growing season rainfall (m)	1.24	0.65
Growing season rainfall shock	0.11	0.65
Number of Districts	575	
Number of Observations	314,425	

This table shows means and standard deviations of our outcome variables from 2003 to 2012 and weather variables from 2002 to 2011 (because we focus on the impacts of the previous year's weather). Diet quality variables are from the NSS and are winsorized within a survey round at the 1st and 99th percentiles. The adequacy percentages represent the percentage of recommended calories or iron that a household consumes. Weather variables are from ERA5, and this table only includes growing season (June to December) weather. The non-growing season (March to May) summary statistics are in Table B1. We do not include January and February in either season because they contain very few hot days. The rainfall shock takes on a value of 1 if rainfall in a given district-season-year exceeds the 80th percentile of its historical distribution (defined as 1979-2012) of rainfall for that district-season, and a value of -1 if rainfall is below the 20th percentile of the historical distribution (0 otherwise).

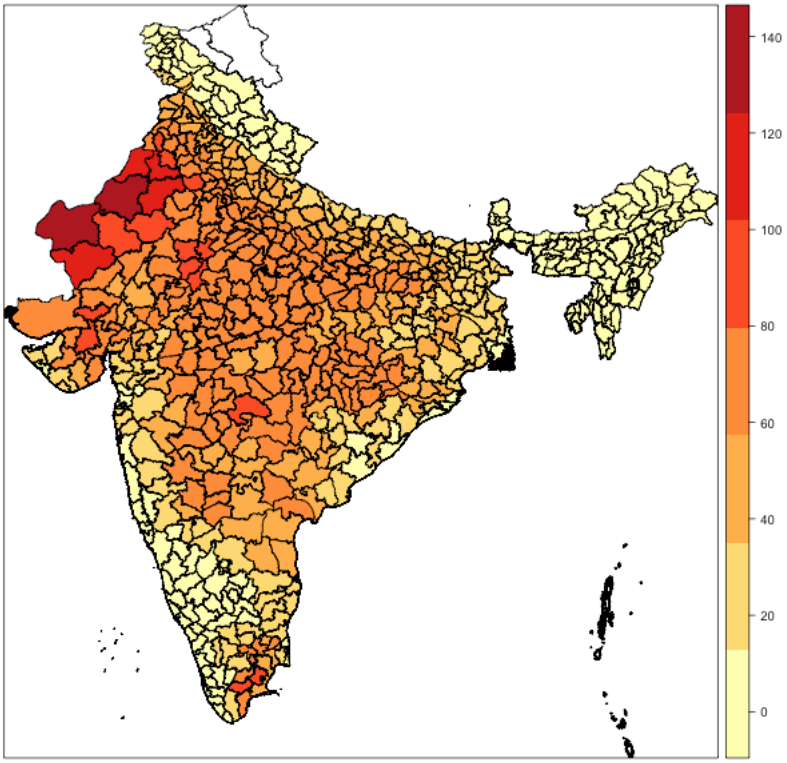


Figure 3.1: Yearly Days with Tmax $\geq$100°F , 2002-2011

This figure shows the average yearly number of days greater than or equal to 100°F in each district from 2002 to 2011. Temperature data are from ERA5.

3.3 Empirical Strategy

Our main empirical strategy is as follows:

$$N_{hdsmy} = \sum_{g=0}^{1}\sum_{t=0}^{1}\sum_{j=1}^{J}\beta_{y-t,g}^{j}T_{d,y-t,g}^{j} + \sum_{g=0}^{1}\sum_{t=0}^{1}\alpha_{y-t,g}P_{d,y-t,g} + \delta X_h + \gamma_d + \tau_y + \eta_m + \sigma_{sy} + \epsilon_{hdmy} \tag{3.1}$$

where N is a nutritional outcome for household h in district d in state s during month m of year y. $T_{d,y-t,g}^{j}$ is the number of days in 10 degree temperature bin j in district d during year $y-t$'s growing ($g=1$) and non-growing ($g=0$) seasons. P is the rainfall shock for this year and last year, separated by growing season. X_h is a set of household-level controls, including religion, caste, head of household education, and household composition by age group and gender. When the outcome variable is a nutrient adequacy threshold (e.g., below 100 percent of recommended levels of calories), we replace household composition with total household size. District fixed effects (γ_d) account for district-level differences in nutrition and temperature. For example, if wealthier districts are warmer, then without these fixed effects our estimates would pick up a spurious correlation that shows that heat is associated with better nutritional outcomes. Year fixed effects (τ_y) account for any yearly India-wide differences in temperature that may be correlated with differences in nutrition, and month fixed effects (η_m) improve precision by controlling for India-wide seasonality in nutrition. We include state-specific yearly trends (σ_{sy}) to control for any state-specific changes in nutrition that may be spuriously correlated with warming over time. Our results are robust to district-specific yearly trends as well. ϵ_{hdmy} is a household-specific error term. The standard errors are clustered at the district level to account for any correlation between the error terms of households in the same district. For example, all households in a district with a more iron-rich staple crop (such as wheat) might experience bigger swings in iron consumption due to heat compared to households in other districts.

In the Results section, we focus on the effect of last year's growing season temperature (the $\beta_{y-1,1}^{j}$) on this year's nutritional outcomes. We do so because weather during the most

recent growing period (June-December) likely has meaningful effects on a household's crop yields and income this year (Garg et al., 2020b). In turn, these effects could reduce food consumption. However, we include this year's weather and last year's non-growing season weather as controls.

3.4 Results

We estimate equation 3.1 to test how heat impacts total household-level caloric and iron consumption. Figure 3.2 plots the coefficient estimates from the previous year's growing season temperature, for the dependent variables log of daily calories per capita (Panel A), and log of daily iron per capita (Panel B). The coefficients document the effect of one day in the corresponding temperature bin relative to a day below 70°F. We multiply the dependent variables by 100 so that the coefficients can be interpreted as percentage changes. Despite heat's impact on yields[5], Panel A shows that heat in any range does not impact caloric consumption. Panel B shows a statistically significant decrease in iron consumption for all bins above 70°F, with the exception of the noisy above 110°F bin. These estimates range from 0.16 to 0.21 percent decreases in iron consumption.[6] The effects are small, as a 0.21 percent decrease corresponds to an average of around 0.03mg less iron per person per day, equivalent to the amount of iron in 3g of uncooked rice (where one serving size is 45g). In addition, the magnitudes do not increase with temperature.

Next, we consider whether extreme heat affects the number of households in relation to key nutritional adequacy thresholds. In Figure 3.3, Panel A we graph coefficient estimates for the effect of the previous year's growing season weather on the percent of households below 100 percent calories adequacy. In Panel B we plot them for the percent of households below 100 percent iron adequacy.[7] Panel A shows there is no effect of heat on the percentage

[5]See Appendix Figure B3 for details.

[6]See Table 3.2 columns 1 and 2 for the coefficient numbers.

[7]See Table 3.2 columns 3 and 5 for the coefficient numbers.

of households who consume 100 percent of their recommended level of calories. Panel B shows that temperatures between 80 and 100°F increase the percentage of households who consume below 100 percent of their recommended level of iron. However, the effect does not get stronger with higher temperatures. These figures suggest that extreme heat during the previous growing season does not push households from adequate dietary availability to undernourishment.

In contrast, we find that extreme heat does significantly increase the percentage of households in extreme undernourishment. In Figure 3.3 Panels C and D, we use the 80 percent adequacy threshold.[8] These graphs show that temperatures above 80°F increase the percentage of households who consume less than 80 percent of their recommended caloric or iron intake. A day above 110°F pushes 0.36 percent of households into extreme caloric undernourishment and 0.35 percent into extreme iron undernourishment. These percentage changes correspond to an increase of 3.1 million people and 3.0 million people per day above 110°F, respectively, out of rural India's 2012 population of 871 million (World Bank, 2024). These results suggest that extreme heat does exacerbate the diet quality of households already experiencing undernourishment, pushing many into extreme undernourishment.

Next, we show that this distributional pattern is consistent across multiple adequacy thresholds. We estimate equation 3.1 for the percentage of households below calorie and iron adequacy thresholds ranging from 50 percent to 150 percent. We plot the effect of a previous year's growing season day above 110°F on these outcomes in Figure 3.4. Panel A shows that extreme heat does not increase the percentage of households below calorie adequacy thresholds of 90 percent and above. However, the percentage of households below 80 percent, 70 percent, and 60 percent adequacy does increase due to extreme heat. While there is no statistically significant effect on the percentage of households below the 50 percent threshold, this is in large part due to the infrequency of this outcome: only 1.4 percent of

[8]See Table 3.2 columns 4 and 6 for the coefficient numbers.

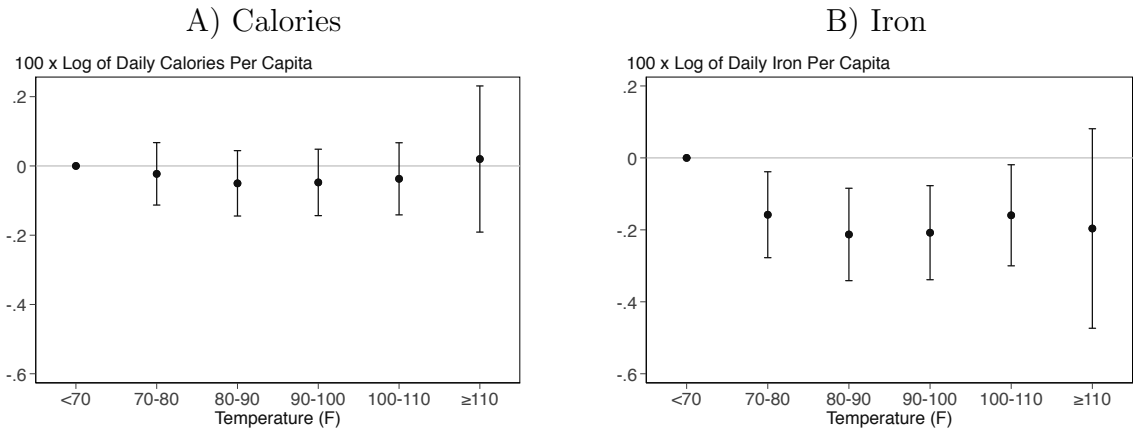


Figure 3.2: Calorie and Iron Availability

Panel A shows the change in 100 times the log of per capita daily consumption of calories for each additional day at a given temperature during last year's growing season relative to a day below 70°F. The coefficients can be interpreted as percentage changes. Panel B shows the same relationship for iron. We match household-year-month nutritional outcomes with weather data at the district-year-season level. All models include 10°F temperature bins from 70-80°F. to above 110°F for this year and last year's growing season (June to December) and non-growing season (March to May). They omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. The models include district, month, and year fixed effects, state time trends, rainfall shocks, demographic (religion, social group, and head of household education) and household composition controls. Standard errors (error bars represent the 95% confidence interval) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. The model in panel A includes 314,424 observations, and panel B includes 313,790 observations. These coefficients correspond to those in Table 3.2, columns 1-2.

households ever fall below this threshold.[9] Panel B graphs the corresponding coefficients for different iron adequacy thresholds. Similar to calories, extreme heat does not push households below thresholds ranging from 100 percent to 150 percent iron adequacy, but does increase the percentage of households below thresholds ranging from 50 percent to 90 percent.

[9]See Appendix Figure B2 and Appendix Table B2 for summary statistics.

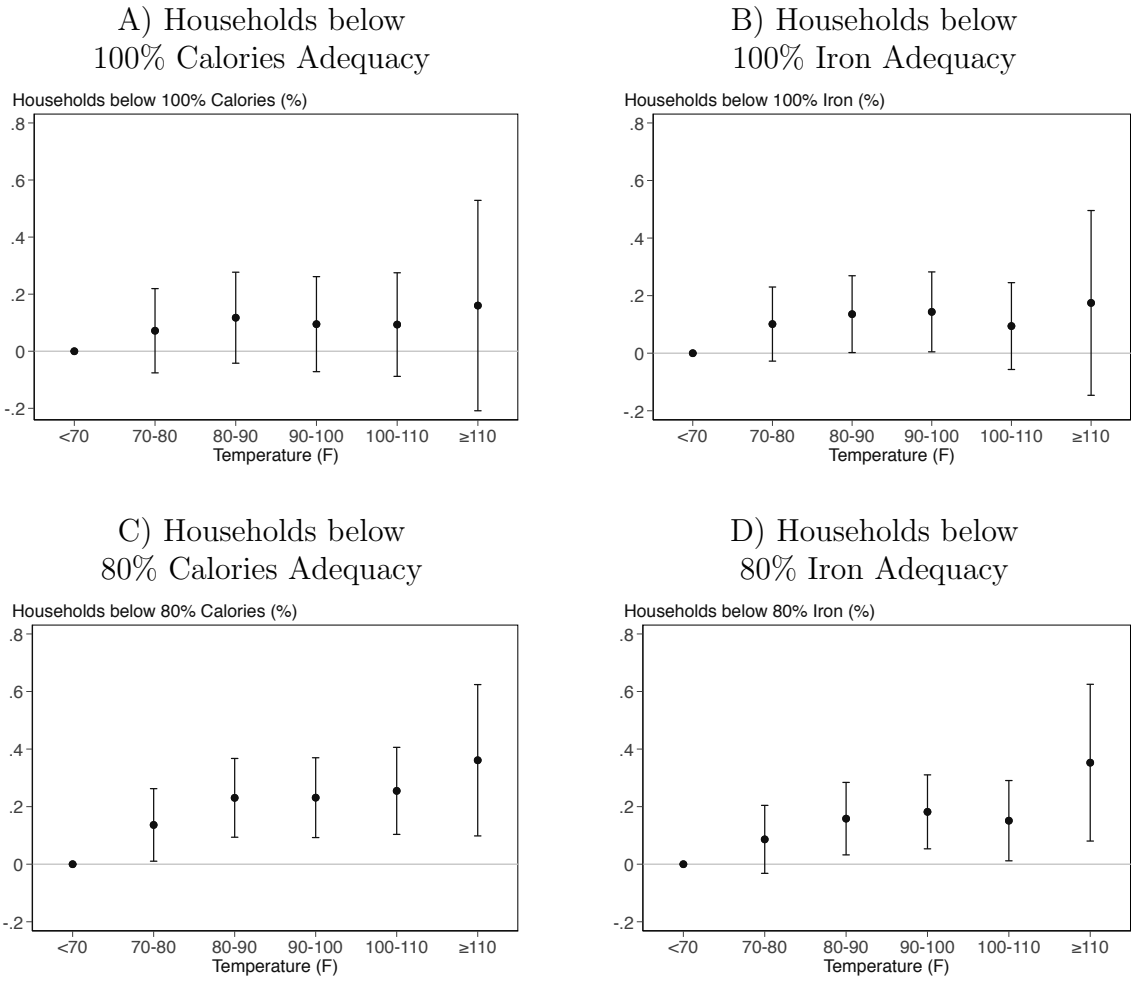


Figure 3.3: Calorie and Iron Undernourishment

Panel A shows the change in the percentage of households consuming below 100 percent of the recommended level of calories, based on household composition, for each additional day at a given temperature during last year's growing season relative to a day below 70°F. Panel B shows the same relationship for the 100 percent iron threshold. Panels C and D show the same as A and B, respectively, but for the 80 percent threshold. We match household-year-month nutritional outcomes with weather data at the district-year-season level. All models include 10°F temperature bins from 70-80°F. to above 110°F for this year and last year's growing season (June to December) and non-growing season (March to May). They omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. The models include district, month, and year fixed effects, state time trends, rainfall shocks, demographic (religion, social group, and head of household education) and household size controls. Standard errors (error bars represent the 95% confidence interval) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. All models include 314,425 observations. These coefficients correspond to those in Table 3.2, columns 3-6.

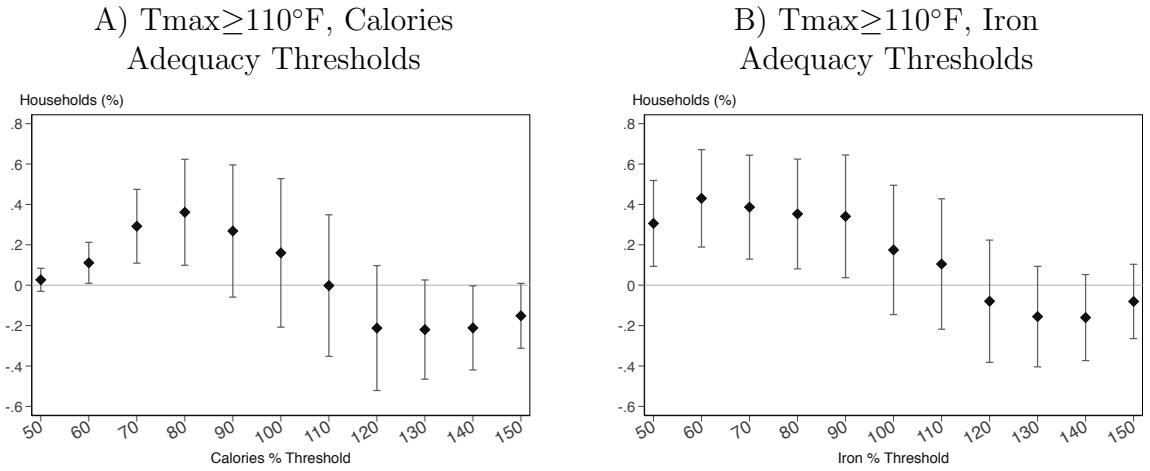


Figure 3.4: Calorie and Iron Adequacy Thresholds

Panel A shows the change in the percentage of households below different calorie consumption thresholds, between 50 percent and 150 percent, for each additional day above 110°F during the previous growing season relative to a day below 70°F. Panel B shows the same as Panel A, but for iron consumption thresholds. We match household-year-month nutritional outcomes with weather data at the district-year-season level. All models include 10°F temperature bins from 70-80°F to above 110°F for this year and last year's growing season (June - December) and non-growing season (March - May). They omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. The models include district, month, and year fixed effects, state time trends, rainfall shocks, and demographic controls (house size, religion, social group, and education). Standard errors (error bars represent the 95% confidence interval) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. All models include 314,425 observations.

Potential Mechanisms

First, we examine a potential adaptation mechanism. If households experience a shock to their home-grown sources of food, such as a local weather shock, they may respond by purchasing more food. We test this hypothesis by estimating equation 3.1 for home-grown and purchased foods separately. We estimate these regressions as levels rather than logs since 35 percent of households consume no home-grown foods.

In Figure 3.5, we provide suggestive evidence for this adaptation mechanism. Panel A plots the coefficient estimates from the previous year's growing season from equation 1, for the dependent variable daily calorie consumption from home-grown foods per capita. Panel C plots them for daily calorie consumption from purchases per capita (See Table 3.2 columns 7 and 8 for the coefficient numbers). The coefficients from Panel A suggest that households consume fewer home-grown calories as a result of hotter temperatures. They consume 2.5kCal per person per day fewer calories due to a single day between 100 and 110°F relative to one below 70°F, and 4kCal (slightly more than 1g of uncooked rice, where one serving is 45g) per person per day fewer for a day above 110°F. Panel C shows that households purchase more calories after a hot growing season. A day between 100 and 110°F increases purchased calories by 1.9kCal per person per day and a day above 110°F increases them by 5.3kCal per person per day. Figure 3.5 points to an important yet understudied agricultural household adaptation to heat: households compensate calorically for these home-grown losses by purchasing more food.

However, these additional purchased calories are not iron-dense. On average, calories that households in our sample purchase contain 24.1 percent less iron than home-grown calories. Figure 3.5 Panel B plots coefficient estimates for daily iron from home per capita, and Panel D plots them for daily purchased iron per capita.[10] Figure 3.5B shows that temperatures above 70°F decrease home-grown iron consumption, with estimates ranging from -0.020 to -0.027 mg/person-day. Similar to Figure 3.2B, the effect does not increase with temperature.

[10]See Table 3.2 columns 9 and 10 for the coefficient numbers.

Panel D shows that, despite increasing calories consumption, households do not experience an increase in iron consumption. The point estimates are all statistically insignificant and very close to 0. This finding suggests that faced with shocks to consumption, households tend to adapt by focusing on replacing their lost calories with foods that are not rich in iron.

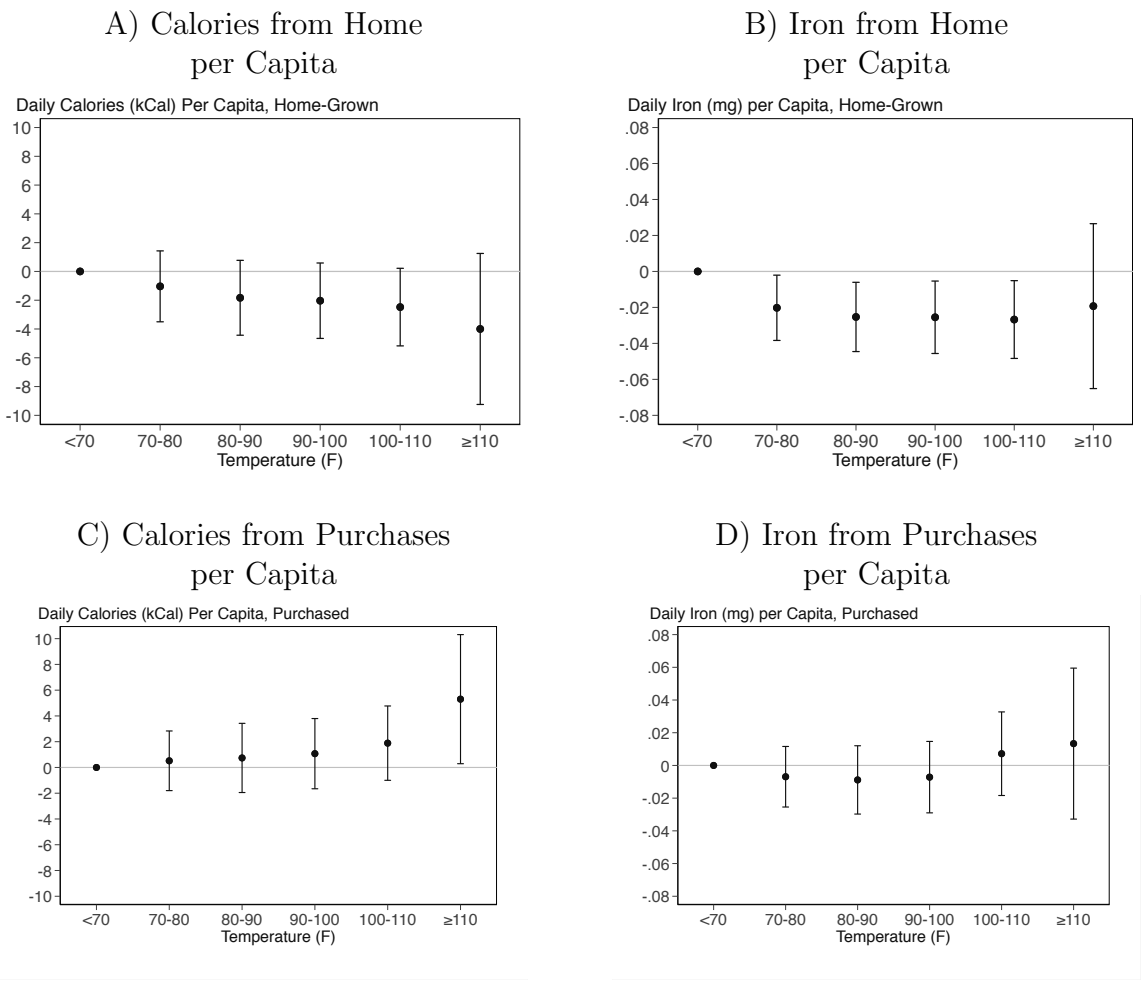


Figure 3.5: Sources of Calories and Iron

Panel A shows the change in per capita daily calorie consumption from home-grown foods for each additional day at a given temperature during last year's growing season relative to a day below 70°F. Panel C shows the same relationship for per capita calorie consumption from food purchases. Panels B and D show the same relationships as panels A and C, respectively, but for iron. We match household-year-month nutritional outcomes with weather data at the district-year-season level. All models include 10°F temperature bins from 70-80°F to above 110°F for this year and last year's growing season (June to December) and non-growing season (March to May). They omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. The models include district, month, and year fixed effects, state time trends, rainfall shocks, demographic controls (religion, social group, and head of household education), and controls for household composition. Standard errors (error bars represent the 95% confidence interval) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. The model in panel A includes 314,425 observations, panel B includes 314,425, panel C includes 314,425, and panel D includes 314,425. These coefficients are the same as in Table 3.2, columns 7-10.

Extreme heat could affect household diets through several different channels. One possibility is that extreme heat's negative impact on yields (Taraz, 2018) decreases available food for households to consume. Other possible mechanisms include heat's effects on physiological well-being (Carleton et al., 2022), fertility (Barreca and Schaller, 2020), and labor force participation and productivity (Somanathan et al., 2021). By analyzing the difference in the effects of extreme heat based on its timing (i.e., during the growing season vs. the non-growing season), we test which mechanisms are most likely.

First, we test for the possibility that extreme heat pushes households into extreme undernourishment primarily due to its effect on crop yields. If crop yield reductions are a key mechanism, then extreme heat during the non-growing season, which we confirm does not affect yields[11], should not affect our nutritional outcomes. In Appendix Table B11, we show the coefficient estimates for last year's non-growing season weather for all of our nutritional outcome variables. None of the coefficients are statistically significant at conventional levels apart from 2 exceptions: 1) days above 90°F decrease the percentage of households below the 80 percent calories threshold, and 2) days above 110°F decreases consumption of homegrown iron. These coefficients are also generally closer to 0 than the coefficient estimates for last year's growing season. This suggests that temperatures during the previous growing season are more important determinants of household food consumption this year, consistent with the hypothesis that heat's impact on yields is driving the results.

Next, we consider other channels through which heat might impact food consumption. Heat in the prior period could change a household's capacity to gain income through labor, for example through decreased health (Carleton et al., 2022) or a change in a woman's current likelihood of being pregnant or raising an infant (Barreca and Schaller, 2020). Extreme heat can also decrease labor supply and labor productivity, which would have direct impacts on income (Somanathan et al., 2021). These reductions in labor participation could then translate into decreased food consumption in the period after a heat shock. We test the

[11]See Appendix Table B13 for details.

fertility channel by estimating regressions with the number of children and number of household members as outcomes. These regressions are identical to our main specification but do not include household composition controls. Appendix Figure B4A plots the coefficient estimates for the impact of last year's growing season temperature on the number of children present in a household. Panel B plots the coefficients for the dependent variable house size. They show no effect of heat on either outcome, which helps rule out the fertility channel.

In addition, if non-agricultural mechanisms are important, we would expect to see impacts of heat on nutrition during the entire prior year, not just during the previous growing season. While we cannot fully rule out other effects of heat in the previous growing season on labor income or health, the limited effects of non-growing season weather suggest that the agricultural mechanism is the strongest.

Table 3.2: Effects of Last Year's Growing Season Weather on Calorie and Iron Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.020	-0.196	0.160	0.361***	0.175	0.353**	-3.997	5.303**	-0.019	0.013
	(0.107)	(0.141)	(0.188)	(0.134)	(0.163)	(0.139)	(2.671)	(2.551)	(0.023)	(0.024)
Days 100-110°F	-0.037	-0.160**	0.093	0.255***	0.094	0.151**	-2.478*	1.888	-0.027**	0.007
	(0.053)	(0.072)	(0.092)	(0.077)	(0.077)	(0.071)	(1.373)	(1.469)	(0.011)	(0.013)
Days 90-100°F	-0.048	-0.208***	0.095	0.231***	0.144**	0.182***	-2.034	1.072	-0.025**	-0.007
	(0.049)	(0.067)	(0.085)	(0.071)	(0.071)	(0.065)	(1.334)	(1.388)	(0.010)	(0.011)
Days 80-90°F	-0.050	-0.213***	0.118	0.231***	0.136**	0.158**	-1.831	0.739	-0.025**	-0.009
	(0.048)	(0.065)	(0.081)	(0.070)	(0.068)	(0.064)	(1.324)	(1.367)	(0.010)	(0.011)
Days 70-80°F	-0.023	-0.158***	0.072	0.136**	0.101	0.086	-1.036	0.517	-0.020**	-0.007
	(0.046)	(0.061)	(0.075)	(0.064)	(0.066)	(0.060)	(1.254)	(1.179)	(0.009)	(0.009)
Rainfall Shock	-0.052	0.434	-0.167	-0.123	-0.163	-0.304	-0.445	1.271	-0.008	0.082
	(0.227)	(0.302)	(0.391)	(0.309)	(0.307)	(0.325)	(4.945)	(5.281)	(0.041)	(0.050)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year, and month fixed effects, demographic controls, rainfall shocks, and state time trends. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Robustness tests

We test the validity of our model through a series of robustness checks that we present in the Appendix. First, in Table B3 we show that our results do not change much when including district-month and year-month fixed effects. These fixed effects account for the possibility that each district, and each year, has a unique seasonality in nutritional outcomes. For example, wealthier districts may have a better ability to smooth consumption across the year, whereas poorer districts may see food consumption decrease as time from the most recent harvest increases. The coefficient for a day above 110°F increases for most outcomes, but the other temperature coefficients are largely the same. This consistency is unsurprising, because our identifying variation is driven by district-year, not monthly, variations in weather.

Second, we show in Table B4 that replacing state-year with district-year trends results in qualitatively similar coefficients. Third, we test the effect of removing this year's weather as a control. Table B5 shows that we find similar results qualitatively to our main specification, though many of the coefficients are smaller. One explanation for this difference is the correlation between temperatures across years, which could lead to bias in the coefficients without controlling for this year's weather. Fourth, we show that our results are robust to the inclusion of different rainfall variables (Tables B6, B7, and B8). Fifth, we ensure that they are robust to calculating per capita consumption by weighting children as half or a third of an adult, as in past work (Deaton, 1997). Table B9 shows the results for children weighted as half-adults, and Table B10 for children weighted as third-adults. They show that the results are similar to our main specification, though the outcomes in levels (columns 3-6) are slightly bigger, as are the mean outcomes.

Last, we also present the coefficients for the current year's growing season weather from our main specification in Table B12. In columns 1-6, we see no effects of hot weather this year on overall calories or iron consumption. However, in columns 7-10, we find that hot weather in the current year's growing season decreases home-grown calorie and iron consumption, and it increases purchases of calories and iron by a similar magnitude. This finding may

possibly be due to contemporaneous changes in health or labor productivity. For example, home-grown crops likely require more effort to prepare than purchased foods. Therefore, in periods of extreme heat, households may choose to purchase more foods to reduce effort.

We are unable to analyze two important connections between household food consumption and individual health outcomes due to the nature of our data. First, we cannot take waste (during cooking or after), different preparation styles, or combinations of foods into account. It is possible that levels of waste differ after a heat shock, as households may waste less when less food is available. The survey does not indicate what foods were eaten when, which matters as iron absorption depends partly on the presence of vitamin C in the same meal (Lane and Richardson, 2014). Second, we are unable to say anything about intra-household allocation of resources, which are notoriously biased against female children in India (Kaul, 2018; Azam and Kingdon, 2013). It is also possible that this allocation is actually sensitive to weather. For example, families may decide to shift food away from the parents and to the children during difficult periods.

3.5 Discussion

Understanding household responses to rising temperatures is a key aspect of designing climate adaptation policy. Our findings suggest that, on average, rural Indian households have been able to mitigate the loss of calorie consumption. They do so by purchasing more food than normal in order to compensate for losses in home-grown foods. Importantly, these households may face significant longer-term consequences that we cannot observe. We find that these increased food purchases, as measured by calories, do not come with a concurrent increase in iron from purchased foods, a key mineral in the rural Indian context. Households may be purchasing cheaper, less nutritious foods to make up for caloric losses, leading to losses in diet quality. In addition, prior research documents that poor rural households often lack access to credit or other means of smoothing consumption across periods in response to income shocks (Castells-Quintana et al., 2018). Along with our findings, this suggests house-

holds may be sacrificing non-food purchases, or reducing long-term investments in human capital.

In addition, many households, such as those without savings, may be unable to afford the costs of such a response. This constraint is especially binding for the poorest households, who are likely to already have low levels of food consumption. We find evidence that households near the lowest quintile of calorie consumption (around the 80% adequacy threshold) do actually consume fewer calories overall subsequent to a hot growing season.

Studying the costs and barriers to adaptation, alongside the first-order effects of extreme weather events, is important for two reasons. First, these costs should be included in assessments of damages from climate change, such as when calculating the social cost of carbon. Second, this knowledge can help policymakers respond to climate change by developing or extending social programs to lower barriers to adaptation. For example, programs that help buffer crop losses, such as already-existing workfare (e.g., India's National Rural Employment Guarantee Scheme) or microfinance programs, may become increasingly important as the climate becomes warmer. Policymakers could also choose to directly assist those who cannot adapt, for example by distributing food in areas which have recently experienced extreme heat.

3.6 Extension: Climate Projections

I calculate projections for increases in the percentage of households experiencing extreme iron undernourishment for 2030 to 2049 relative to 2000 to 2019 under the Shared Socioeconomic Pathway 3-7.0 scenario (SSP3-7.0). I show that the within-model, or "internal," variability accounts for roughly 20 percent of the total uncertainty within SSP3-7.0 projections. Not including this source of uncertainty leads to a 95^{th} percentile of damages that is 12 percent lower in magnitude than when including all model members, and a worst-case scenario that is 27 percent lower. This result shows the importance of including multiple members per model in projection ensembles, which economists rarely do. Doing so is low in additional

computational and time costs relative to previous methods which include many fewer climate model members.

Motivation

In 2015, a team of economists and climate scientists published an article describing economists' use of climate projections (Burke et al., 2015). In it, they stress the importance of accounting for model uncertainty (i.e., the uncertainty that exists between different models) in climate projections, something that economists at the time rarely did. For example, in the 188 economics studies they survey, the median number of climate models used is 2. They then show that failing to use more than one model often understates the probability of worst case (and best case) scenarios (Burke et al., 2015). Though a full review of recent literature is outside the scope of this analysis, my sense is that economists today have improved in this respect. Many papers include multi-model ensembles in their climate projections.

However, there is an additional source of uncertainty that economists have largely ignored. This source, referred to as internal variability, comes from the natural, non-linear processes that drive the climate (Deser et al., 2012). A common tactic to assess its magnitude is to use multi-member ensembles of the same climate model that all have slightly different initial conditions (McKinnon and Deser, 2018). Schwarzwald and Lenssen (2022) show that internal variability's contribution to total uncertainty is especially large for early and mid-century climate damage projections and for impacts driven by nonlinear relationships. Both of these conditions are met here.

Data and Methods

To illustrate the typical process in climate economics, I briefly recap the methods from Chapter 2 of this dissertation. First, researchers choose a climate scenario (e.g., SSP3.70) or two, and create an ensemble of around 20 model projections, usually picked for ease of

access.[12] This ensemble rarely, if ever, includes more than one member per scenario-model. Researchers then calculate the difference in the temperature distribution between a reference period (e.g., 2000 to 2019) and a period in the future (e.g., 2080 to 2099). By multiplying this difference by the damage function estimated in the paper, researchers then present estimates of the climate-change induced changes in the outcome of interest.

Here, I download daily temperature projections for India's growing season (June-December) from 204 model members across 30 models from Coupled Model Intercomparison Project Phase 6 (CMIP6). I chose these models, listed in Appendix Table B14, based on their availability on the Pangeo server, an online open-access collection of models maintained by Google. I only include projections from the SSP3-7.0 scenario, thereby ignoring scenario uncertainty. I do so to focus on the relative contributions of internal variability and model uncertainty to total uncertainty. For a more comprehensive account of all three sources of uncertainty in economic projections of climate damages, see Schwarzwald and Lenssen (2022). While I focus on projections for 2030 to 2049, I also compare them to end of century projections. Therefore, I drop ensemble members that end before 2100, resulting in a final ensemble of 190 members across 26 models.

In addition to incorporating many more model members, I focus on different outcomes than in Chapter 2. In Chapter 2, we report the mean impact of climate change on disconnections for 2080 to 2099 (though we graph them for the whole century). Instead, here I focus on the upper tail of potential damages for the middle of the century. From a policy perspective, understanding the upper tail of potential damages is important in order to be prepared for the potential worst case scenario. As I show, incorporating as many members as possible is important to characterizing these extreme, low-probability scenarios.

I focus on the middle of the century because adaptation policy is likely more meaningful for time periods closer in the future.[13] In part, this is because climate damage projections

[12] In the case of Chapter 2, we chose a set of projections that my undergraduate thesis advisor's graduate student at the time, Duo Chan, had been using.

[13] Climate mitigation policy, on the other hand, is more likely to benefit from longer-term thinking. Because

become much more uncertain further into the future. For example, Schwarzwald and Lenssen (2022) finds that the uncertainty in heat-induced increases in mortality in the US for 2070 to 2099 is over 6 times larger than for 2040 to 2069. The authors' calculation is based solely on the uncertainty in the warming levels due to uncertainty in future emissions pathways and the challenges of modeling a system as complex as our climate. This warming uncertainty is compounded by fact that the relationship between temperature and the outcome of interest could be significantly different by the end of the century, even absent policy interventions (Carleton et al., 2022). Focusing on a time period in the near future increases the importance of internal variability. In the Schwarzwald and Lenssen (2022) mortality projections, internal variability is 56 percent of the total uncertainty for 2010 to 2039, 12 percent in 2040 to 2069, and only 3 percent in 2070 to 2099.

Working with climate scientist Will Krantz made incorporating so many climate models possible. When calculating climate projections for the Los Angeles area in Chapter 2, I had to first open the entire model, which requires significant space and memory (the .nc file for a single daily temperature CMIP5 projection is over 3GB). Even using UCLA's Hoffman2 cluster, the process was cumbersome and time-consuming enough that including 190 climate model members would have been prohibitively costly.

Here, as recommended by Will Krantz, I use the Python function *xarray*, which allows me to load in the portion of the dataset I am working with (India, 2000 to 2100, June-December) in a user-friendly way. In Chapter 2, I had to 1) identify the correct latitudes and longitudes for Southern Califonia, and 2) measure time by the number of days since January 1st 1800. The latter task was complicated by the fact that models alternated in their treatment of leap years. *xarray*, combined with another function called *regionmask*, allowed me to subset the models by year and country, thereby reducing the potential for human error. In addition, because I did not need to load the entire dataset, I was able to work with the data entirely on my laptop.

---

of the slow-moving response of the climate to greenhouse gas emissions, curbing emissions today can have a meaningful effect on end-of-century warming.

I present all damage projections below relative to a reference period of 2000 to 2019. Importantly, I calculate future changes in the temperature distribution for each ensemble member relative to its own projections for 2000 to 2019. Doing so corrects for any within-member average biases in temperature, ensuring that I am comparing the spread in the projected *change* in temperatures across members, rather than the spread in temperatures themselves.

I calculate projections of climate change-induced increases in extreme undernourishment as follows. First, for each model member I calculate the pixel-year number of days in each 10°F temperature bin from < 70°F to ≥ 110°F (and ≥ 100°F for expositional purposes) for the growing season. I calculate the district-year temperature distributions using functions from Schwarzwald and Lenssen (2022) that assign each pixel to districts based on their spatial overlap. Using 2020 Indian population data (Tiwari, 2020), I calculate India-wide district population-weighted average temperature distributions for each ensemble member and year.[14] For each ensemble member, I calculate the difference in its temperature distribution in each year from 2020 to 2100 and the average distribution for my baseline period of 2000 to 2019. Finally, I multiply these temperature distribution changes by my damage function estimates from equation 3.1 and Table 3.2 Column 6.

To simulate results using methods from Chapter 2, I also present damage projections using an ensemble with a single member per model. I use the "r1i1p1f1[15]" member, which we use in Chapter 2. Six models from the full ensemble do not include r1i1p1f1 members that end in 2100, so this r1i1p1f1 ensemble only includes 20 models.

[14]Note that, in cases where temperature bins are not used to estimate the damage function (e.g., a cubic spline), one would need to first assign damage projections to each district-year before taking the average. Because in this case the damage function is just a linear combination of temperature bins, calculating the damage projections before or after taking the India-wide average produces identical results.

[15]This naming convention is common to all models, and this name refers to the **r**ealisation (ensemble member), **i**nitialisation method, **p**hysics, and **f**orcing.

Results

First, I present yearly changes in the temperature distribution. In Figure 3.6, I plot the projected change in the number of growing season (June-December) days above 100°F ("hot days") per year relative to the mean from 2000 to 2019. In Panel A, I plot all 190 ensemble members from 2020 to 2100 in grey. I plot the mean change in hot days, weighted by the inverse of the number of members per model, in red. For the year 2100, the weighted mean change in hot days is 13.8, with a range from below 0 to above 45. In Panel B, I plot the 20 r1i1p1f1 members in grey, with the mean across these members in red. The mean change in hot days across the r1i1p1f1 models in 2100 is 15.1, with a range from below 0 to 32.2.

Comparing the same member's position in the overall distribution between Panels A and B illustrates the importance of using multiple members per model. In both panels, I plot the r1i1p1f1 member of CanESM5, which projects 32.2 more hot days in 2100 than in the baseline period, in blue. Panel B shows that this member projects the highest increase in hot days for 2100 among all r1i1p1f1 models. In contrast, Panel A includes many members that project higher increases in 2100 than does CanESM5 r1i1p1f1. Many of these hotter members are from CanESM5, but 5 are from 4 other models. The r1i1p1f1 ensemble assigns a probability of zero to the event of over 32.2 more hot days in 2100 than in the baseline period. In contrast, the full ensemble assigns this event a non-zero probability. Therefore, only considering a single member per model may understate the chance of low-probability extreme events.

Next, I consider heat-induced changes in the percentage of households below 80 percent iron adequacy ("households experiencing extreme iron undernourishment"). I multiply changes in the temperature distribution for each year, relative to 2000 to 2019, by the damage function estimates from equation 3.1 and Table 3.2 Column 6. Figure 3.6 Panels C and D show these changes for all members and for r1i1p1f1 members, respectively. In 2100, the full ensemble projects a weighted mean (plotted in red) increase in households experiencing extreme iron undernourishment of 7.2 percentage points, with a range across members from

3.6 to 11.1. The mean projected increase in 2100 for the r1i1p1f1 ensemble is 6.9, with a range from 3.6 to 10.1. Once again, the r1i1p1f1 ensemble understates the most extreme scenarios.

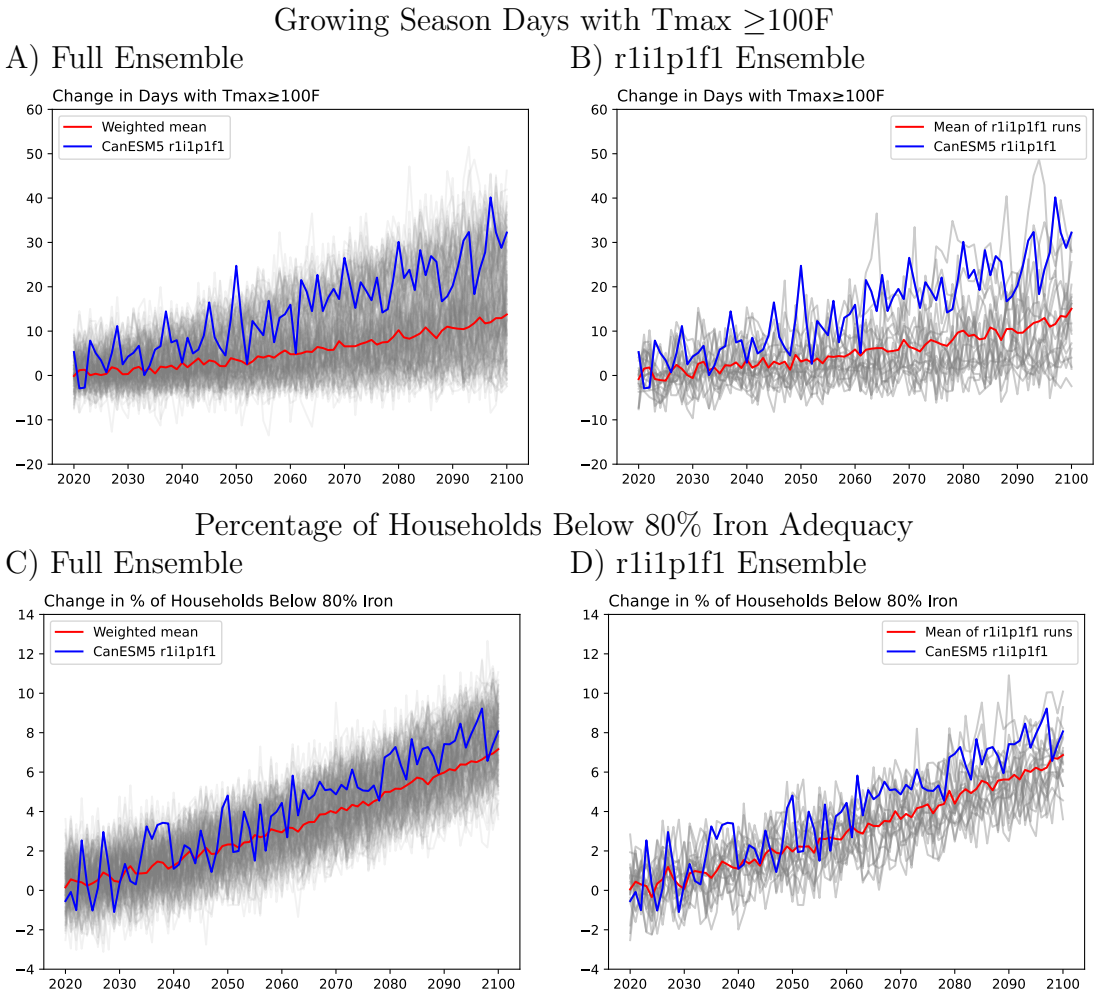


Figure 3.6: Projected Change Relative to 2000-2019

This figure shows the projected change relative to 2000 to 2019 in two variables: the number of growing season (June-December) days above 100°F (Panels A and B), and the heat-induced percentage of households below 80% iron adequacy (Panels C and D). The numbers in C and D are determined by calculating, for each ensemble member, the change in each temperature bin for each year's growing season, and then multiplying this change by the damage function estimates from and Table 3.2 Column 6. Panels A and C include the full ensemble of 190 members across 26 models in grey, and the weighted (by $\frac{1}{n}$, where n is the number of members per model) mean across members in red. Panels B and D plot the 20 members of the ri1ip1f1 ensemble, and the mean in red. In all panels, the blue line represents the r1i1p1f1 member of CanESM5.

Economists generally calculate average changes over longer periods in the future rather than for a single year. In part, this preference is due to the yearly variability of projections, as exemplified by CanESM5's r1i1p1f1 member in Figure 3.6. In some years this ensemble member is near the very top of the distribution, and in others it is closer to the middle. This variability makes understanding the bigger picture by looking at a single year difficult. Therefore, I calculate average temperature distributions, and subsequent increases in extreme iron undernourishment, for two periods: middle (2030 to 2049) and late (2080 to 2099).

Projections nearer in the future require more careful consideration of within-model variation. Figure 3.7 shows the sources of uncertainty in the projected increase in households experiencing extreme iron undernourishment. I limit this analysis to the 5 models with at least 10 members with daily temperature projections until 2100. This choice balances having enough members per model to calculate internal variability with having enough total models to calculate model uncertainty. I show in Appendix Figure B5 that including the 10 models with at least 5 members results in qualitatively similar results. Following Schwarzwald and Lenssen (2022), for each time period, I calculate internal variability by 1) calculating the variance in damages within each model, then 2) averaging across these variances. I calculate model uncertainty by 1) calculating the mean damages for each model, then 2) taking the variance across these model means. Panel A plots the total uncertainty for each time period, broken up by model and internal sources. Total uncertainty for this given scenario is much larger for the late period (2.5) than for the middle period (0.6). Panel B, which plots percentage contributions of each source of uncertainty, shows that this growth is due to an increase in model uncertainty. While internal variability is 20% of the total uncertainty in the middle period, it is only 3.3% for in the late period. These findings are consistent with Schwarzwald and Lenssen (2022), which shows that internal uncertainty is more significant in earlier projections across a variety of settings.

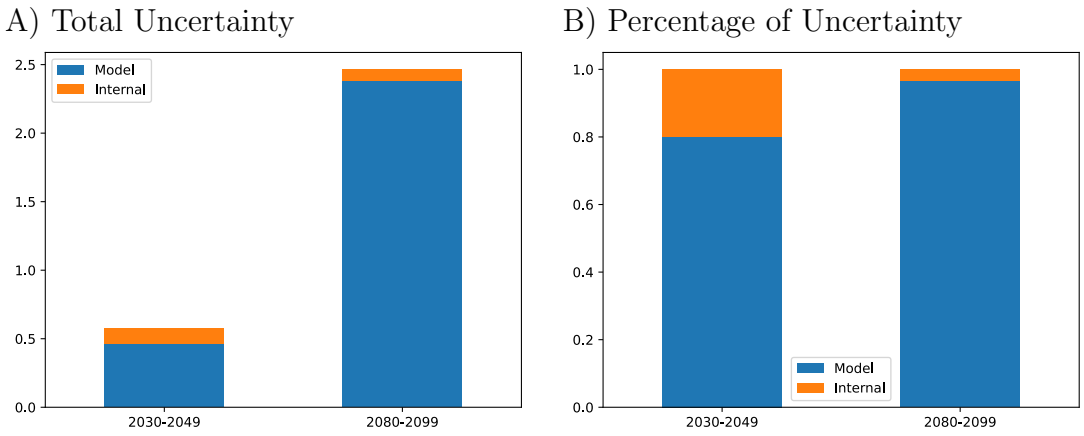


Figure 3.7: Climate Projection Uncertainty

This figure graphs the contributions to the uncertainty in the warming-induced changes, relative to 2000 to 2019, in the percentage of households below the 80% adequacy thresholds for iron and calories. Panel A graphs uncertainty in levels, with uncertainty for estimates for the 2030 to 2049 period in the bar on the left, and for the 2080 to 2099 period on the right. Panel B graphs the percentage of uncertainty contributed by each source. These calculations include the 5 models with at least 10 members for daily temperature in the SSP3-7.0 projection scenario: CanESM5, EC-Earth3, IPSL-CM6A-LR, MPI-EMS1-2-HR, and MPI-ESM1-2-LR. The model uncertainty is calculated as the variance across the model means, and the internal variability is calculated as the mean of the within-model (across members) variance.

Next, I demonstrate the importance of considering variation across model members. In Figure 3.8, I plot the spread in the projected increase in the percentage of households experiencing extreme iron undernourishment across ensemble members for 2030 to 2049. The geren box-and-whisker plot includes the full ensemble, and the purple one includes the r1i1p1f1 ensemble. The plots include the median (middle line in white), the 25^{th} and 75^{th} percentiles (the box), the 5^{th} and 95^{th} (the whiskers), and any individual values outside these percentiles (circles). For the full set of models, the median projected increase in the percentage of households experiencing extreme iron undernourishment is 1.43, which corresponds to 13 million people, using India's 2022 total rural population of 909 million (World Bank, 2024). The median for the r1i1p1f1 projections is similar, at 1.36 percentage points.

However, the extremes differ substantially. The r1i1p1f1 ensemble's 95^{th} percentile (2.12) is 12 percent smaller than the full ensemble's (2.42), and its maximum (2.22) is 27 percent

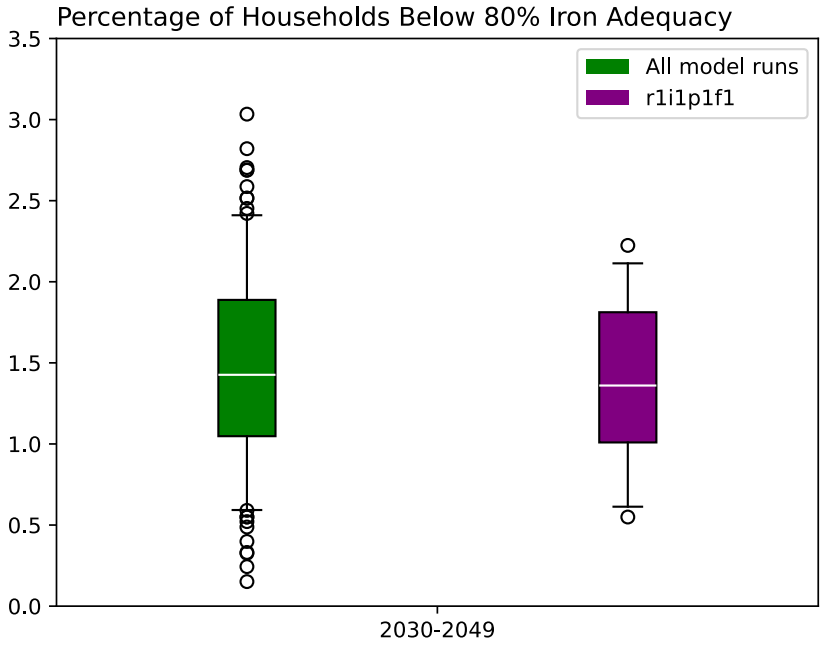


Figure 3.8: Climate Change: Households Below 80% Iron Adequacy

This figure shows the spread in projections of the warming-induced changes, relative to 2000 to 2019, in the percentage of households below the 80% adequacy thresholds for iron for 2030 to 2049. The projections are calculated for each ensemble member by multiplying its projected change in the temperature distribution by the damage function estimates from and Table 3.2 Column 6. The green box-and-whiskers plot includes all 190 ensemble members, and the purple one includes the 20 r1i1p1f1 ensemble members. The plots includes the median (middle line in white), the 25^{th} and 75^{th} percentiles (the box), the 5^{th} and 95^{th} (the whiskers), and any individual values outside these percentiles (circles).

smaller than the full ensemble's (3.03). Therefore, focusing only on the r1i1p1f1 members would largely underestimate the probability of the most extreme scenarios, which include up to 27.5 million people being pushed into extreme iron undernourishment. In Appendix Figure B6A, I show that this underestimate is less extreme for end of century projections: for the late period, the r1i1p1f1 ensemble's maximum is only 5 percent smaller than the full ensemble's. In Figure B6B, I show the damage projections for extreme calorie undernourishment.

Discussion

In this section, I show the importance of incorporating multiple members from the same climate model in economic damage projections. Doing so is especially important when

calculating projections closer into the future, which is more likely to be useful for designing adaptation policy.

There are some limitations to this approach. The first is that it is difficult to anticipate future changes in the relationship between temperature and the outcome of interest. As the case of extreme iron undernourishment exemplifies, it is sometimes unclear whether this relationship will ameliorate or worsen. On the one hand, households could become more resilient to extreme heat events, for example through increased income or the adoption of more heat resilient crops. On the other hand, it is possible that there is a non-linear and increasing relationship between crop yield damages and household diet quality. For example, a household may be able to withstand the crop losses from a growing season with 5 days above 100°F, but not those from one with 10 such days. Heat may also threaten household capacity to respond to yield losses, for example through direct impacts on health. Therefore, it is important to approach these damage projections with caution and understand that the true numbers may lie above or below the projected range across all models.

In addition, these projections do not take into account the uncertainty in the damage function I estimate for 2003-2012. Instead, they just take the point estimates of the damage function. I do so in order to focus on the importance of incorporating multi-member ensembles in damage projections. In future work, however, economists (including myself) should also include the uncertainty of their damage function estimates in their projections of future damages.

The second concern is that different models perform better in different parts of the globe. For example, some models might be worse at incorporating how a shift in global mean temperature will influence the Indian monsoon, and therefore are worse at projecting temperature changes in the Indian subcontinent. If this is the case, then it might be better to ignore these models for this exercise. While true in theory, doing so is difficult in practice and involves a deeper understanding of the physics behind our climate system than an average climate economist possesses (Goldenson et al., 2023). Even among climate scientists, there

is hardly a consensus about which models are "good" at certain tasks. Therefore, unless such a consensus emerges in the future, climate economists should continue to include all possible models.

There are other approaches to communicating climate projections that are becoming popular among climate scientists. The first is to measure local changes in terms of "warming levels" rather than in terms of time (Rahimi et al., 2024). For example, one could count the projected number of growing season days in India above 100°F at a 5.4°F (3°C) higher global mean temperature than pre-industrial levels. The appeal to this approach is that it allows one to avoid details about when exactly this warming level will be reached. It also abstracts away from emissions pathways, which are difficult to predict. However, this approach poses some difficulties for economists for whom thinking in terms of time is central to inter-temporal cost-benefit analysis.

Another option is the "storyline" approach, which involves a detailed description of one or more single physically consistent climate projections. For example, one could pick the scenario which represents the 95^{th} percentile of potential damages to household diet quality. This approach allows for easy exposition of the joint risks of multiple changes in regional climate. For example, one could focus on the co-occurrence of both extreme heat and drought or the instances of multiple hot days in a row (rather than just the total number of hot days in a period). Proponents of this approach note that it helps to easily communicate the risks associated with climate change in a salient way to non-experts, thereby strengthening their ability to make informed policy decisions (Shepherd et al., 2018). As climate economists continue to incorporate damage projections into their work, it is worth considering these alternative approaches. However, for now, the easiest way to improve our process is to incorporate all possible projections into our analyses.

4

Occupational Mobility and Climate Adaptation: Evidence from France

4.1 Introduction

The extent to which climate change will exacerbate economic inequality is a growing concern. A future with more frequent extreme weather events theoretically threatens the welfare of environmentally exposed workers, who already have below average wages (Xiao, 2023b). Past work in this area has largely focused on the on-the-job effects of extreme weather on exposed workers and short-term adaptations. However, little is known about long-term adaptations for workers, such as changing occupations. In addition, the potential spillover effects of adaptations to non-exposed workers, who might themselves have low socioeconomic status, have received little attention. In part, this lack of attention is due to a dearth of reliable data with sufficient granularity to analyze the relationships between occupational mobility and environmental exposure. To help assess climate change's effects on the labor market, I use individual-level panel data from France to study occupational mobility's potential as an adaptation strategy.

As the climate gets warmer, workers and firms will seek to adapt. Workers in climate-exposed occupations[1] will suffer losses in productivity (Adhvaryu et al., 2020), safety (Dillender, 2021), and comfort (Graff Zivin and Neidell, 2014). The extent of these losses, and how much they increase inequality between high and low exposure occupations, depend on workers' ability to adapt to changing climatic conditions. These adaptations can fall into three main categories: 1) workers could alter their schedules or working conditions[2], but remain in the same occupation, 2) households of high exposure workers could engage in intra-household substitution[3], or 3) workers could change occupations. Past research on

[1]The effects of climate change on workers will transcend industrial boundaries. In France, I find that no industry is completely devoid of workers in high exposure occupations, and 11 industries out of 36 have at least 5 percent of their workers in high exposure occupations. How climate change will affect entire industries, for example by increasing demand (e.g., electricity for air conditioning), is outside the scope of this paper. Here, I focus on the consequences of climate change for on-the-job utility, worker adaptations, and the spillover effects of these adaptations.

[2]This category includes workers either pushing their firms to make safety investments such as purchasing shade structures, or buying personal protective equipment, such as umbrellas, on their own.

[3]For example, households could reduce the amount of time members in exposed occupations work (includ-

worker adaptation to climate change has focused on option 1. Workers in exposed industries reduce the amount of time they work on unusually warm days (Graff Zivin and Neidell, 2014; LoPalo, 2020). Analyses of workers' ability to mitigate heat exposure on the job show mixed results, with some papers finding limited capacity for adaptation (Dillender, 2021), and others finding evidence of it (Park et al., 2021). My study is the first quantitative analysis, to my knowledge, of occupational mobility as a form of climate adaptation. My analysis focuses only on worker-driven adaptations and their spillover costs to other workers, though workers' responses will likely influence firm responses to climate change as well.[4]

Occupational mobility based adaptation has important implications for climate change's effects on the entire labor market, not just on exposed workers. Theory suggests that as the disutility of climate-exposed work increases, worker preferences will shift towards work in climate controlled environments. This shift would increase labor supply in low exposure occupations, possibly suppressing their wages. However, strong occupational mobility frictions raise the possibility that workers may not be able to switch out of climate-exposed work (Cortes and Gallipoli, 2017). Determining inter-exposure mobility rates is therefore necessary to fully account for the spillover costs of climate change to low exposure workers, which should be included in aggregate climate cost assessments. It is also important for understanding the distributional costs of climate change, not just between high exposure and low exposure workers, but also between more and less wealthy low exposure workers. Due to the prevalence of low-skilled workers in high exposure work (Maestas et al., 2017), low exposure occupations with more low-skilled workers (and lower wages) are more likely to experience labor supply increases and wage suppression.

I study occupational mobility using 9 years of individual panel data from the *Déclaration Annuelle de Données Sociales (DADS) Panel*, which is compiled from firms' responses to

ing stopping work altogether) and increase it for other members. However, work considering this question in China finds no evidence of intra-household substitution (Garg et al., 2020a).

[4]For example, if many workers are leaving high exposure occupations, firms could respond by increasing compensation to retain workers. I leave an inquiry into these responses for future work.

a mandatory survey. I merge these data with an index of environmental exposure at the occupation level, which I construct from a survey of French workers. I show that "high exposure" occupations, which I define in several specific ways, span a variety of industries and skill requirements. I quantify occupational mobility between high and low exposure occupations in two complementary ways. First, I calculate the probability of high exposure workers remaining in high exposure work year after year. High persistence[5] in high exposure work could reflect one of two phenomena, or both. On the one hand, it could mean that workers rarely leave high exposure occupations, indicating that they are relatively attractive (Cortes and Gallipoli, 2017). On the other hand, it could mean that, when they do change occupations, high exposure workers rarely transition into low exposure occupations. To test which explanation is more likely, I directly analyze occupational transitions. Specifically, for each occupation, I calculate the percentage of its departing workers whose new job is low exposure. In addition, I test for differences in inter-exposure mobility between workers in the same occupation.[6]

This dataset has two main advantages that allow me to conduct these analyses. First, the administrative nature of the data enables me to capture a more representative view of the labor market than past work on the flow of workers between occupations. For example, datasets that rely on scraped resume data are likely to underrepresent certain groups of workers, such as those in manual occupations or without a college education (Schubert et al., 2022). Because many high exposure occupations are manual in nature, having reliable data on manual workers is crucial for this analysis. In addition, a worker's education may affect their ability to move from high to low exposure occupations. Therefore, having a more representative sample of workers bolsters the external validity of these findings on inter-exposure mobility, even just within France. Second, the size of the data is such that it allows me to calculate reliable flows of workers between occupations on a relatively granular

[5] i.e., workers remaining in high exposure work for long periods of time.

[6] Specifically, I test whether relatively high-paid workers are more likely to transfer to low exposure occupations than lower paid workers in the same occupation.

level. For the 368 occupations I study, I observe a median of 2,100 departing workers.[7] The size of this dataset is especially important for within-occupation analyses of inter-exposure mobility.

This project has four important findings. First, I find that employment in high exposure occupations exhibits strong persistence. For example, French workers in a high exposure occupation have around a 77 percent chance of still being in a high exposure occupation 8 years later. I find that these rates increase slightly with age, from 75 percent for workers in their thirties to 79 percent for workers in their fifties. Analyzing occupational transitions reveals that this is not simply a question of workers staying in the same occupation. When leaving a high exposure occupation, workers end up in a different high exposure occupation 49 percent of the time. This rate is surprisingly high since high exposure work, by my preferred definition, only comprises around 9 percent of the overall labor market. These findings suggest strong segmentation between the labor markets for exposed and non-exposed workers, and that changing occupations is unlikely to be a viable adaptation strategy.

Second, I find suggestive evidence that job tasks are a partial, but insufficient, explanation for this segmentation. I use data from the Occupational Information Network (O*NET) to measure the relevance of 17 tasks, separated into 5 groups[8], to each occupation. I show novel evidence that an occupation's exposure level is strongly and positively correlated with its manual task intensity (both routine and non-routine). This finding suggests that human capital, particularly in manual tasks, plays a role in persistence in high exposure work. However, analyzing occupational transitions shows that the labor markets between high and low exposure work are more segmented than those between high and low manual intensity work. Persistence in high manual intensity work is lower than in high exposure work, and

[7]The number of workers I observe departing each occupation ranges from 23 (occupation code 441b "Clergy") to 28,431 (occupation code 542a "Secretaries")

[8]Following past work from France, the groups I choose are non-routine analytical, non-routine interpersonal, non-routine manual, routine manual, and routine cognitive (Le Barbanchon and Rizzotti, 2020).

only 38 percent of workers leaving a high routine manual occupation[9] end up in a different high routine manual occupation.[10]

In addition, I show that the low exposure occupations that high exposure workers often transition into[11] are disproportionately high in routine task intensity.[12] I find that 29 percent of jobs in these occupations are in the highest quartile of routine task intensity, compared to 17 percent of jobs in all other low exposure occupations. Routine task-intensive jobs have declined across Western Europe and the United States over the past few decades (Goos et al., 2014; Autor and Dorn, 2013). This finding raises the possibility that low exposure opportunities for high exposure workers have decreased over time, contributing to the present-day segmentation.

Third, I show that any spillover effects of a climate shock to low exposure workers will likely be concentrated on those in low wage occupations. I use inter-exposure mobility rates to calculate how a climate shock that only directly affects high exposure workers would propagate throughout the labor market. Past work shows that a worker's outside occupation options, defined as their job opportunities outside of their current occupation, affect their current wages (Schubert et al., 2022). The same paper also shows that transitions between occupations are a relevant measure of these outside occupation options. For each occupation, I use the percentage of its high exposure outside occupation options (i.e., the probability that a worker leaving a job in that occupation moves to a high exposure occupation) as a proxy for its vulnerability to the spillover effects of a climate shock. For example, despite both

[9]Consistent with my definition of high environmental exposure, I define a "high" task intensity as 1.5 standard deviations above the mean across occupations, and close to 9 percent of jobs are in this category for most tasks.

[10]As noted in the previous paragraph, 49 percent of high exposure workers leaving their occupation remain in high exposure work.

[11]An occupation counts for the "often transitions into" category if either 1) at least 9 percent of its incoming transfers are from high exposure occupations, or 2) it has at least 100 incoming workers who left a high exposure occupation. Occupations in this category account for 53 percent of low exposure jobs and 85 percent of switches from high exposure occupations to low exposure ones.

[12]Routine task intensity is a measure of an occupation's 1) intensity in routine tasks and 2) its lack of intensity in non-routine tasks.

being low exposure occupations, 22 percent of electrical equipment assemblers' outside occupation options are high exposure, compared to only 1.7 percent of cooks' outside occupation options. These numbers mean that electrical equipment assemblers are more vulnerable to climate shock spillover effects, through their outside occupation options, than are cooks. I show that, across low exposure occupations, workers in below median wage occupations are nearly 5 times more vulnerable to these indirect effects than workers in above median wage occupations.

Fourth, I find evidence for within-occupation differences in inter-exposure mobility. When leaving their occupation, workers with relatively high wages for their occupation are more likely to transfer to a low exposure occupation. For high exposure workers, a top decile earner is 23 percent more likely to move to a low exposure occupation than is a bottom decile earner. This observation suggests that better paid workers are more likely to be able to adapt to climate change by changing occupations. For low exposure workers, a bottom decile earner is 65 percent more likely than a top decile earner to transfer to a high exposure occupation. Therefore, low earning low exposure workers are more vulnerable to the spillover effects of a climate shock to high exposure occupations. These findings motivate future work to understand workers' vulnerability to climate change at levels even more granular than the occupation.

This project contributes to two main literatures. First, it contributes to the occupational mobility literature by documenting evidence of labor market segmentation, a policy-relevant topic which has been of interest since at least the 1970s (Reich et al., 1973; Eichhorst et al., 2017). I show that this segmentation is particularly strong between labor markets for exposed and non-exposed workers. In addition, I add to a general understanding of outside occupation options. I provide evidence that workers' outside occupation options, as measured by worker transitions, differ across workers within the same occupation in ways that matter for their vulnerability to climate change. This project is also the first paper, to my knowledge, to test the outside occupation options elasticity of wages in France, a country with a relatively rigid occupational structure and strong worker protections. I find that a 1 percent increase

in an occupation's average outside options' wages increases its own wages by 0.052 percent, compared to 0.1 percent in the US (Schubert et al., 2022). I then show one way for future research to use this elasticity, together with the composition of each occupation's outside options, to calculate how a climate shock's effects would propagate throughout the labor market.

Second, my study adds to the climate adaptation literature, which to date has focused primarily on on-the-job adaptations and their costs to exposed workers and firms (Park et al., 2021; Rode et al., 2023). Recent work considers long-term adaptation strategies for firms (Xiao, 2023b), but I add to a sparse and largely qualitative literature on long-term worker adaptations (Farbotko et al., 2022). In particular, I show that changing occupations is unlikely to be a successful adaptation strategy in France. In addition, I provide suggestive evidence that the automation of routine tasks could further dampen workers' adaptation prospects. The substitution of capital for labor for routine tasks has been used to explain important phenomena such as labor market polarization and the rising college wage premium, both in the US (Autor and Dorn, 2013) and in Europe (Goos et al., 2014). However, less is understood about how this automation might affect workers' and firms' climate adaptation decisions. I show that low exposure job options for high exposure workers are particularly high in routine tasks. The computerization of routine work might further decrease inter-exposure mobility's potential as a successful adaptation strategy, exacerbating inequality concerns between high and low exposure workers.

4.2 Data

DADS All Employees Panel: Worker Job Spells

I create an individual-level panel of job spells using the *Déclaration Annuelle de Données Sociales (DADS)*, yearly linked employer-employee data compiled by the French National Institute of Statistics and Economic Studies (INSEE). It is based on a mandatory survey of

all employers, excluding the public sector. The *Panel Tous Salariés* (All Employees Panel) is a subset of the *DADS* that only includes employees born in October, so it is roughly a 1/12 sample. Entries are at the employer-employee-year level and include the employee's age, gender, municipality of residence, department of birth, start and end dates (within the calendar year), four digit occupation (PCS4) code, contract type, gross earnings, and hours worked.

Following Marinescu, Ouss, and Pape (2021) and Le Barbanchon, Rathelot, and Roulet (2021), I impose several restrictions on the data. I limit the sample to the years 2011 through 2019.[13]. I remove any job spells that are classified as secondary or are shorter than 6 months long.[14] I also remove any jobs spells in public sector jobs or occupations that are primarily in the public sector.[15] I restrict my sample to a single employment spell active on January 1st, per worker, per year. If a worker has more than one job spell listed for a given year, I keep the one with the most hours in the year, then the highest monthly salary, and then the latest end date within that year.[16]

To focus on occupational transitions that are representative of a typical career trajectory, I keep job spells for workers between the ages of 25 and 60. Younger workers are more likely to be engaged in work that is not representative of their career or their typical inter-exposure

[13]2010 is the first year that does not suffer from response-rate concerns for job codes. Commuting zone definitions changed in 2010, and are consistent from 2011 onwards (Marinescu et al., 2021). Though 2020 data are available, I stop at 2019 to avoid changes in occupational mobility during the COVID-19 pandemic.

[14]This removal raises potential sample selection issues. For example, if a high exposure worker enters temporary low exposure employment to reduce exposure, I will not capture that behavior, potentially undercounting inter-exposure mobility. To include as many transitions between occupations as possible, I keep workers in the sample even if they are missing years of full-time long-term employment. In the Appendix, I show that my occupational transition results are robust to only including workers with observable full-time long-term employment for all 9 years of my sample period.

[15]This includes PCS4 codes that start with 33 "public service executives", 45 "administrative public service occupations", and 52 "public service agents". Despite the data excluding the public sector, some job spells still list these occupations. They are unlikely to be a representative sample of these occupations, leading me to exclude these job spells.

[16]I drop the remaining 159 worker-years with more than 1 observation.

mobility.[17] I exclude workers from the end of their careers to avoid potential biases from retirement. For example, a high exposure worker could decide to leave high exposure work by retiring instead of by changing to a low exposure occupation. This worker would drop out of my sample, potentially leading to an underestimate of inter-exposure mobility. In the Appendix, I show that my results are robust to including workers between 18 and 67 as well.

The *DADS* does not record exhaustive data for self-employed workers (Harrigan et al., 2021), so I exclude them from my analysis. I do so using the first digit of the 4-digit PCS-ESE occupation codes which the *DADS* uses. The first digit of the PCS4 corresponds to the following broad categories: 1) Agricultural and farming directors, 2) Independent artisans, small business owners, and CEOs, 3) Managers and intellectuals, 4) Intermediate professions, 5) Non-manual workers, and 6) Manual laborers. I exclude Agricultural and farming directors, artisans, small business owners, and CEOs (categories 1 and 2) from my analysis.[18] Finally, as I am studying occupational transitions, I exclude workers I observe for only one year. My analysis dataset includes 9.2 million observations for 1.6 million individuals across 368 occupations from 2011 to 2019.

DADS Postes: Labor Market Metrics

I create aggregate labor market metrics such as employment, wages, and labor market concentration using the *DADS Postes* dataset. This dataset is derived from the same employer survey as the *DADS Panel*, but it includes all employment spells and changes worker identifiers every year. Consistent with past work, I keep workers between 18 and 67 years old (Marinescu et al., 2021). As with the *DADS Panel*, I limit my sample to 2011-2019, and I keep a single primary employment spell, active on January 1st, per year for each individual. I calculate employment for each PCS4- commuting zone (CZ)-year as the number of individ-

[17]Take, for example, young workers who are lifeguards for a year after college before beginning work in a bank. Including them would inflate my estimates of inter-exposure mobility.

[18]While not all workers in these two PCS1 categories are self-employed, many of them are. These occupations are underrepresented in my sample, so I exclude them from my analysis.

uals whose primary employment is in that PCS4, in that CZ, and in that year. I calculate average hourly wages, which I convert to 2019 real terms, for each PCS4-CZ-year.[19] I also calculate average yearly employment and hourly wages across France for each PCS4.

CT-Individus: Occupational Climate Exposure

I create environmental exposure scores for each occupation using the *Conditions de Travail - Volet Individus* (*CT-Individus*), a survey of French workers produced by INSEE and the *Direction de l'animation de la recherche, des études et des statistiques* under the French Ministry of Labor. I use the 2019 vintage, the only version available that includes questions on exposure to heat or humidity on the job. The survey asks workers across most occupations in France about their working conditions, including the following two questions: 1) Are you exposed to heat on the job? (yes/no), and 2) Are you exposed to humidity on the job? (yes/no).

I assign each occupation a heat and humidity score equal to the percentage of yes respondents for each of these questions. There is an average of 42 respondents per PCS4. If a PCS4 has fewer than 10 respondents[20], I assign it heat and humidity scores by using the average exposure score across respondents in the same 3-digit occupation category, including that occupation itself.[21] Figure C2 in the Appendix plots the correlation between exposure scores at the PCS4 level and the PCS3 level, including those with between 1-10 observations at the PCS4 level. As is clear, there is a strong relationship between the scores at the PCS3 and PCS4 levels.

I define the environmental exposure score for each occupation, θ, as the minimum of its

[19]Following Marinescu, Ouss, and Pape (2021), I drop the bottom 5 percent and top 1 percent of hourly wages in each year. Doing this ensures that no hourly wages are below the minimum wage and drops any outliers. For the *DADS Panel*, I set wages outside of these thresholds to missing, but I retain the observations for any analyses that do not involve wages.

[20]115 out of 368 occupations have fewer than 10 respondents, including 45 with 0 respondents.

[21]Even after this method, there remains 1 occupation (531b, municipal police officers) that is in the *DADS* with no exposure score. I exclude it from this analysis.

humidity and heat scores. I do this because, for some occupations, the humidity or heat score is a function of the work tasks rather than of the environment within which the work takes place. Table 4.1 shows example occupations, their humidity scores, their heat scores, and their θ. It illustrates the importance of taking the minimum of the heat and humidity scores for each occupation. For example, flight attendants have a heat score of 100, but a humidity score of 10. This suggests that the heat the flight attendant respondents refer to is likely a function of their job (e.g., heat from the hot meals that they serve), rather than of any environmental conditions. I assign flight attendants a θ of 10. In contrast, roofers, who generally work outdoors, have a humidity score of 94 and a heat score of 100, so their θ is 94. In the Appendix, I show that my results are robust to defining exposure as either just the humidity score or just the heat score.

In my primary definition, I designate occupations with $\theta \geq 50$ as "high exposure," and all others as "low exposure."[22] Table C1 in the Appendix lists all 47 occupations (out of 368) that I classify as high exposure. Around 9 percent of workers are employed in high exposure occupations. Low exposure workers are generally paid better than high exposure ones. The median high exposure worker earns €15.5/hour (2019 value), compared to €19.2/hour for the median low exposure worker. Ninety-five percent of high exposure workers in my sample are men.

[22]I choose 50 as my primary cutoff because this means that workers are more likely exposed than not. In the Appendix, I show that my results are robust to using a cutoff of 40.

Table 4.1: Example Exposure Scores (θ)

Occupation	Humidity	Heat	θ
Roofers	94	100	94
Agricultural Managers	69	85	69
Plumbers	82	62	62
Crane and Tower Operators	58	58	58
Cashiers	15	35	15
Flight Attendants	10	100	10
Butchers	52	9.7	9.7
Pharmacists	3.7	9.3	3.7

This table lists example occupations, their humidity score, their heat score, and their exposure score (θ). The humidity and heat scores are from the CT-Individus survey of workers. θ is the minimum of these two scores.

***O*NET*: Occupational Tasks**

I study the relationship between occupations' environmental exposure and task composition using occupation-level task scores from the Occupational Information Network (O*NET). O*NET, the United States' primary source of occupational information, administers nationally representative surveys of workers and occupation experts. The surveys include information on the task requirements and work environments of occupations, classified by their Standard Occupation (SOC) code. I use 21 O*NET rounds from 2011 to 2019 to calculate the various tasks involved in different French occupations.[23] I assign each O*NET occupation at the 8-digit SOC-code level a score, based on its average across surveys, for 17 tasks that fit into the following five task groups[24]: 1) Non-routine analytical, 2) non-routine interpersonal, 3) non-routine manual, 4) routine manual, and 5) routine cognitive.

I create a crosswalk between PCS4 and SOC-codes by matching job titles and descriptions. In cases where there are multiple possible SOC matches for a given PCS4, I pick one most likely occupation. In the Appendix, I validate this crosswalk by comparing scores from

[23]I use O*NET, rather than the *CT-Individus*, which also includes some job characteristics, because the *CT-Individus* does not include all necessary questions to make it comparable to past job task papers. Past work, including in France, has used O*NET (Le Barbanchon and Rizzotti, 2020)

[24]I present the specific makeup of each group in Tables C2 and C3.

O*NET for tasks that are also included in the *CT-Individus* survey. I also verify that my main results are robust to an alternative crosswalk that averages task scores across all potential PCS4-O*NET matches. I assign each PCS4 with task scores using my PCS4-SOC-code crosswalk. Then, following Le Barbanchon and Rizzotti (2020), I normalize the scores, using average yearly employment by PCS4 as weights, so that they have mean 0 and standard deviation 1. I add up these normalized scores for each PCS4 within each of the five task groups, then once again normalize the scores for each group.

4.3 Inter-Exposure Mobility

4.3.1 Methods

Persistence in High Exposure Work

To show differences in occupational exposure over the life course, I plot the distribution of workers in high exposure work by age. I merge the All-Employees Panel with occupation-level θ scores, then I calculate the percentage of workers in high exposure occupations ($\theta \geq 50$) for each age between 25 and 60. Next, I test persistence in high exposure work as follows. First, I limit my sample to workers who are engaged in high exposure work at any point and who are present for all 9 years of my data.[25] Second, I calculate the percentage of these workers who are still in high exposure work n years later, where n runs from 1 to 8. In order to detect any differences in high exposure persistence over the life course, I repeat this calculation separately for each age in year t, from 25-60.

Occupational Transitions

The persistence test combines two dimensions of occupational mobility: 1) worker occupation departure frequency and 2) the probability of finding low exposure work when changing

[25] I do this to exclude workers who leave the labor market, potentially in response to climate exposure.

occupations. If persistence is high, that could mean that at least one of these two dimensions is low. However, the implications of each dimension being low are somewhat contradictory. On the one hand, high exposure workers could leave their occupation at relatively low rates. This could indicate that high exposure occupations are relatively attractive (Cortes and Gallipoli, 2017). On the other hand, high exposure workers could leave their occupation at a normal (or high) rate, but often switch into different high exposure work. This could indicate one of two characteristics of high exposure work. First, it could mean that high exposure workers generally have better job match quality in high exposure work, and so they choose to stay in high exposure occupations. Second, it could mean that there are high barriers to entry for low exposure work, and high exposure workers remain in high exposure work due to a lack of options rather than by choice.

To test the reasons behind the persistence levels I observe, I focus directly on occupational transitions. I define a transition between origin occupation o and destination occupation $p \neq o$ if an individual is in occupation o on January 1st, year t, and in occupation p on January 1st, year $t+1$. First, I test whether high exposure occupations do have lower rates of worker departure. Second, I analyze the relationship between the exposure scores of a transitioning worker's origin and destination occupations. I calculate two statistics for each occupation from the sample of workers who leave it: 1) the average exposure level of all the destination occupations ($\bar{\theta_p}$) and 2) the percent of transitions to low exposure occupations ($\mathrm{P}(\theta_p < 50)$).

4.3.2 Results

Persistence in High Exposure Work

Figure 4.1 shows that workers of all age groups are engaged in high exposure work, including a substantial fraction of middle-aged workers. As environmentally exposed work has relatively low wages and requires physical fitness, one would expect the probability of engaging in high exposure work to decline with age. I show that, initially, around 15 percent of 25 year olds

are engaged in high exposure work. As expected, this percentage steadily declines to between 6.5 and 7 percent by around age 35. However, the percentage remains surprisingly constant between ages 35 and 50, at which point it restarts a slow decline. This finding suggests that either 1) a significant fraction of high exposure workers remain in high exposure work for long periods of time or 2) different workers are frequently moving in and out of high exposure work. My hypothesis is that the former explanation is more likely.

I test this hypothesis by analyzing how long workers remain in high exposure occupations. Figure 4.2A graphs the percentage of high exposure workers from year t still employed in high exposure work in years $t + n$, for n from 1 to 8. It shows that, on average, 77 percent of high exposure workers are still employed in high exposure work 8 years later. Because the probability of high exposure employment differs across the life cycle, it is possible that persistence in high exposure work does too. I test this hypothesis in Figure 4.2B, where I plot persistence against age on the x-axis. The y-axis is the percent of workers, out of those in high exposure work in year t, still in a high exposure occupation in year $t + n$. Each line represents a different value of n between 1 and 8. It shows that older workers are slightly more likely to remain in high exposure work: 93 percent of workers aged 50 in year t remain in high exposure work one year later $(t + 1)$, compared to 92 percent for workers aged 25. This gap gets larger with n. 82 percent of workers aged 50 in year t are still in high exposure work eight years later $(t + 8)$, compared to only 76 percent of workers aged 25. These percentages are quite high, and although my data do not allow me to explicitly test the following hypothesis, they suggest that many workers likely remain in high exposure work for multiple decades.

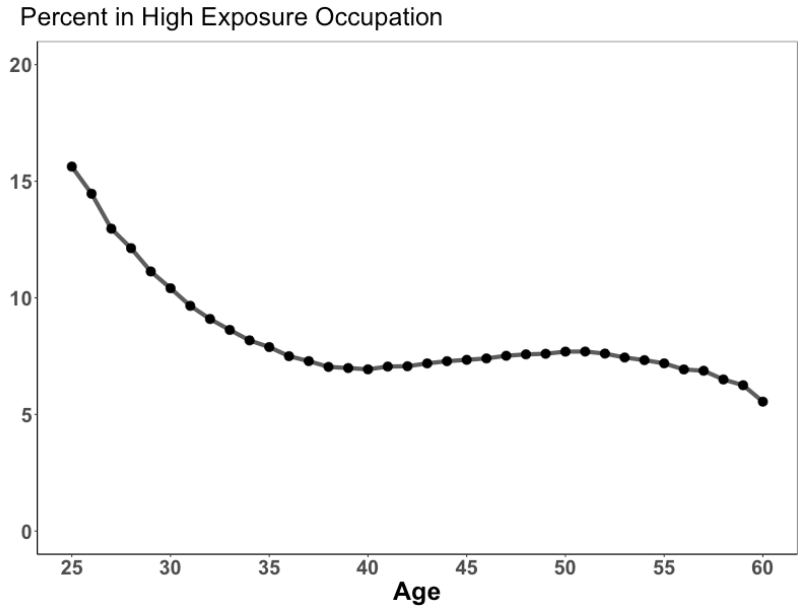


Figure 4.1: High Exposure Work and Age

This figure shows the percentage of workers in high exposure occupations ($\theta \geq 50$) by age. The sample runs from 2011-2019, and I limit it to the 498,472 workers I observe for all 9 years.

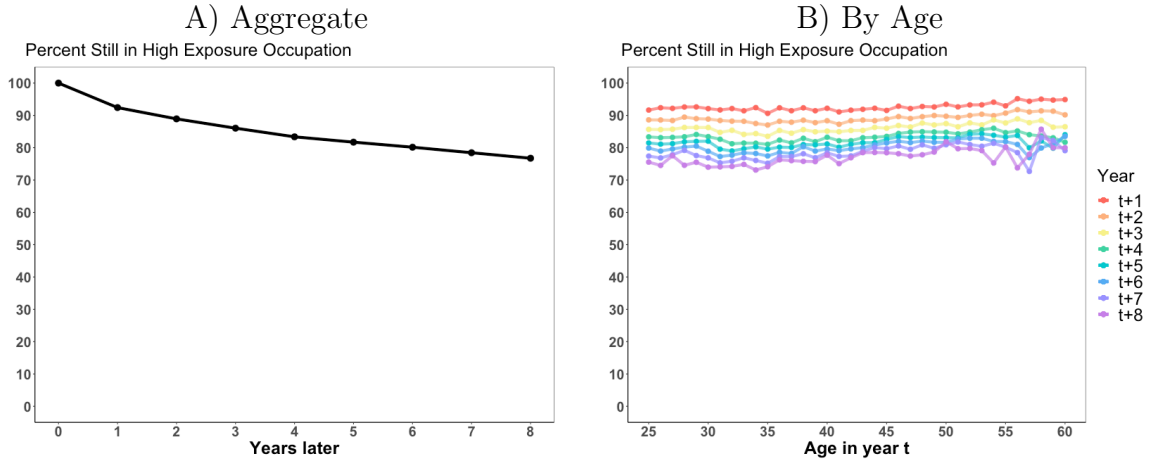


Figure 4.2: Persistence in High Exposure Work

Panel A shows the percentage of workers still in high exposure occupations ($\theta \geq 50$) 1-8 years after being in a high exposure occupation. Panel B graphs this percentage separately by age in year t. The sample in both panels is limited to the 55,850 workers I observe for all 9 years (2011-2019) and who worked in a high exposure occupation at least once between 2011 and 2018.

Occupational Transitions

I test the reasons behind persistence in high exposure work by analyzing occupational transitions. First, I rule out that high exposure workers leave their occupations less frequently than low exposure workers. I find that, on average, high exposure occupations have a departure rate of 17.2 percent, compared to 16.4 percent for low exposure occupations. This finding suggests that persistence in high exposure work is high because workers leaving a high exposure occupation frequently move to a different high exposure occupation.

The data shown in Figure 4.3 confirm this hypothesis. In Panel A, I plot the exposure level of each origin occupation, θ_o, on the x-axis. I plot the average exposure level of all destination occupations, $\bar{\theta_p}$, for workers leaving a given origin occupation on the y-axis. Each dot represents a single PCS4 occupation, and its size represents average yearly national employment from 2011-2019. The figure shows that as θ_o increases, $\bar{\theta_p}$ increases alongside it, meaning that workers leaving more exposed occupations end up in occupations which are, on average, more exposed. For example, the average worker leaving an occupation with θ_o between 10 and 20 ends up in an occupation with θ_p of 17. For workers leaving an occupation with θ_o between 60 and 70, $\bar{\theta_p}$ is 52. To put these values of $\bar{\theta_p}$ into perspective, I compare them to the average exposure levels of occupations with similar wages to the origin occupations.[26] I find that workers leaving an occupation with θ_o between 10 and 20 end up in occupations with lower exposure than those with similar wages, and workers leaving an occupation with θ_o between 60 and 70 end up in occupations with higher exposure than those with similar wages.

Panel B shows the relationship between θ_o and the probability that θ_p is below 50 (i.e., a low exposure occupation). It shows that the vast majority of workers leaving low-exposure occupations move to a different low-exposure occupation. It also indicates that while many

[26]First, I find the distribution of wages for occupations with θ between 10 and 20. I find the 25^{th} percentile of wages for these occupations is €13.4/hour, and the 75^{th} percentile is €20.7/hour. I then take the mean exposure level of occupations with wages between these two values, which is 27.1. I repeat the same exercise for occupations with θ between 60 and 70. In this case, the 25^{th}-75^{th} wage range is from 13.9 to 16.2 €/hour, and the average θ for occupations with wages in this range is 32.5

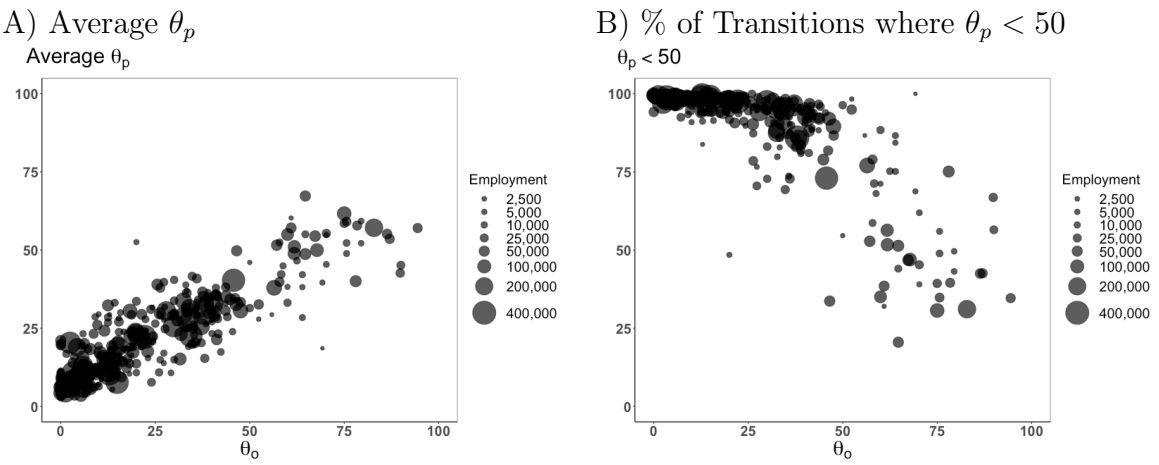


Figure 4.3: Occupational Transitions and Climate Exposure

Panel A graphs the relationship between an occupation's θ_o and the average θ_p of destination occupations for workers leaving that occupation. Each dot represents a single occupation and its size is based on average yearly employment from 2011-2019. Panel B graphs the relationship between an occupation's θ_o and the percentage of transitions where $\theta_p < 50$. the data include 1,156,194 transitions from 2011-2019 across 368 occupations.

workers leaving a high exposure occupation find low exposure work in their next job, the majority of workers leaving an occupation with $\theta > 75$ do not.

The data shown in Figures 4.2 and 4.3 indicate strong segmentation between high exposure and low exposure labor markets. Workers in high exposure jobs tend to remain in them for long periods of time. In addition, workers leaving high exposure occupations often find new work in a different high exposure occupation. These observations suggest that changing occupations is a costly strategy with respect to climate change adaptation. These costs could be due to the difference in job match quality between high and low exposure work for high exposure workers, or they could be due to other barriers to entry into low exposure occupations.

4.3.3 Exposure and Task Intensities

One potential reason for the high costs to inter-exposure mobility is that job tasks differ between high and low exposure occupations. Task intensity (i.e., the extent to which a

given task is relevant for a given occupation) can be related to worker occupational choice in two main ways. First, workers might initially be attracted to jobs involving particular tasks, either due to preferences or qualifications. In particular, jobs with certain tasks may offer relatively high wages for low-skilled workers. For example, more manual jobs have a higher risk of injury, which employers may compensate for with higher wages. Second, as workers accrue job experience, they are more likely to stay in work that requires similar tasks (Gathmann and Schönberg, 2010).

Manual Tasks and Exposure

In Figure 4.4, I plot novel evidence on the correlation between each occupation's exposure and its five O*NET task group intensities.[27] One main finding stands out. High exposure work is strongly manual. Panel A plots the relationship between an occupation's routine manual task intensity and its exposure. Each dot represents a percentile of routine manual intensity and the corresponding average θ, controlling for the other four tasks and weighted by average yearly employment. It shows that with a 1 standard deviation increase in routine manual intensity, exposure increases by 6.8. Non-routine manual intensity (Panel B) shows an even stronger coefficient of 9.3. These strong relationships are largely due to the lack of outdoor occupations that are non-manual. For example, 97 percent of high exposure jobs have an above-average routine manual intensity, and 100 percent of them have an above-average non-routine manual intensity. In addition, 40 percent of jobs that are one standard deviation above the mean in non-routine manual intensity are high exposure, compared to only 2 percent of all other jobs.[28] This finding suggests that workers with a preference for more manual work[29] are more likely to work in high exposure occupations. That being said,

[27]Routine manual, routine cognitive, non-routine manual, non-routine analytical, and non-routine interpersonal.

[28]These numbers are 28 percent and 4 percent, respectively, for routine manual jobs.

[29]Either because they prefer it outright to non-manual work or because it offers them higher compensation than their other options.

there is a significant number of jobs which are both highly manual and low exposure.

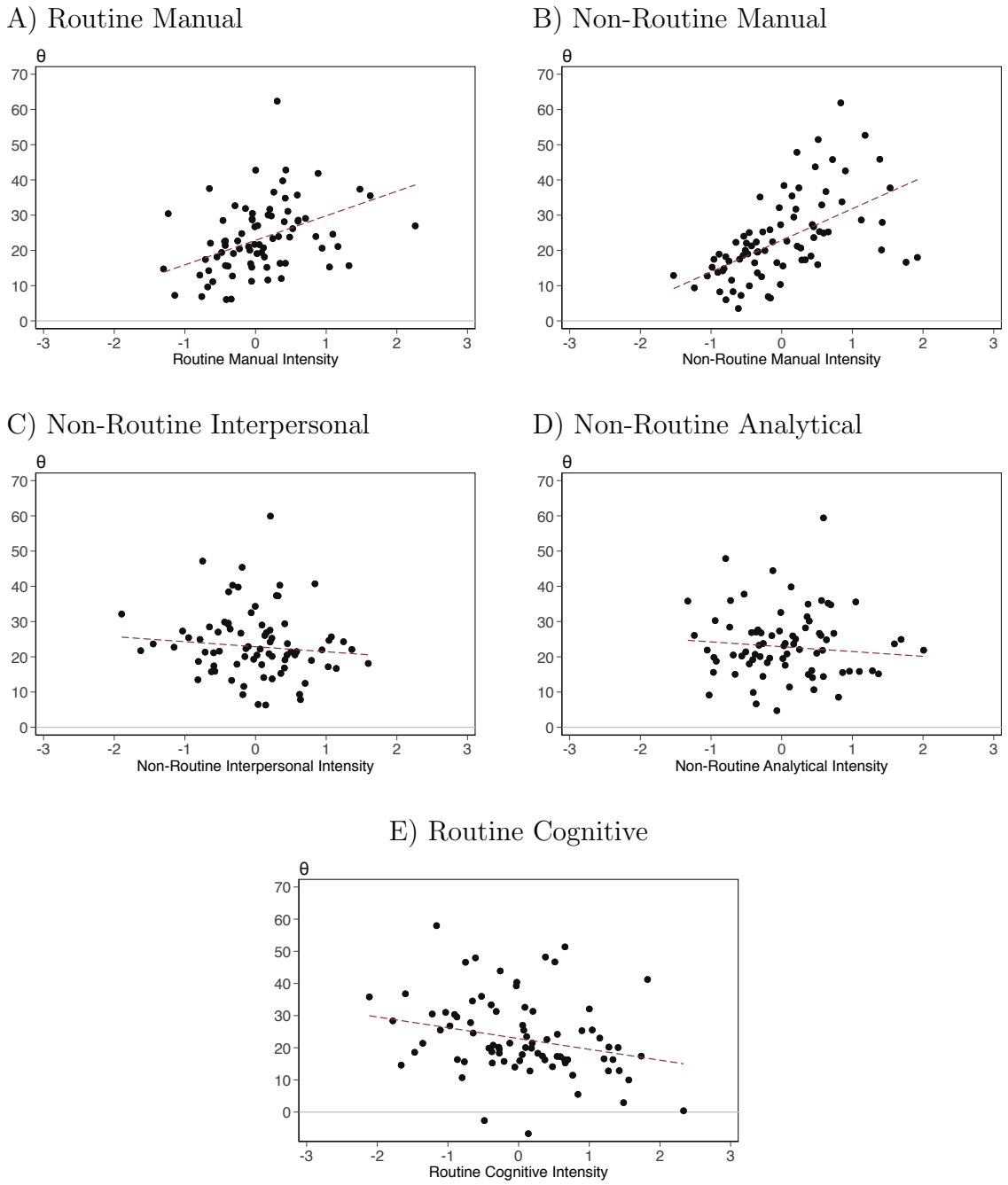


Figure 4.4: Task Intensity and Exposure

This figure shows the relationship between an occupation's different task intensities and its environmental exposure (θ). Panel A is for routine manual, Panel B for non-routine manual, Panel C for non-routine interpersonal, Panel D for non-routine analytical, and Panel E for routine cognitive. All graphs control for all other task intensities and are weighted by yearly national employment between 2011 and 2019. The data include 368 occupations.

The strong relationship between manual tasks and exposure raises the concern that the results in Figures 4.2 and 4.3 are just capturing workers staying in highly manual work for long periods of time. To address this concern, I compare persistence in high exposure work to persistence in work with high task intensities that proxy for manual labor. I define an occupation as having "high" intensity in a manual task if its score is at least 1.5.[30] Eight point three percent of jobs are in high routine manual work, and 10 percent are in high non-routine manual work. For each manual task group, I calculate the percentage of workers still in a high intensity occupation n year later. I then plot these percentages against n to compare them against persistence in high exposure work.

Figure 4.5A shows that workers stay in high exposure occupations for longer than they stay in occupations with high levels of any other task intensity. 77 percent of high exposure workers are still in high exposure work 8 years later, compared to 73 percent for non-routine manual intensity and 72 percent for routine manual intensity. This gap is not very large. However, it widens when restricting the sample to workers who change occupations sometime during my sample period (2011-2019) (Figure 4.5B). For the "occupation changers" sample, 56 percent of high exposure workers are still in high exposure work 8 years later, compared to 36 percent for non-routine manual intensity and 45 percent for routine manual intensity. This finding suggests that when changing occupations, high exposure workers move to a different high exposure occupation at a uniquely high rate. Table C4 in the Appendix confirms this hypothesis. Fourty-nine percent of workers leaving a high exposure occupation find work in a different high exposure occupation, compared to just 38 percent for the closest task group (routine manual).

[30]By construction, this is 1.5 standard deviations above the mean. $\theta = 50$ is also 1.5 standard deviations above the mean θ of 20.

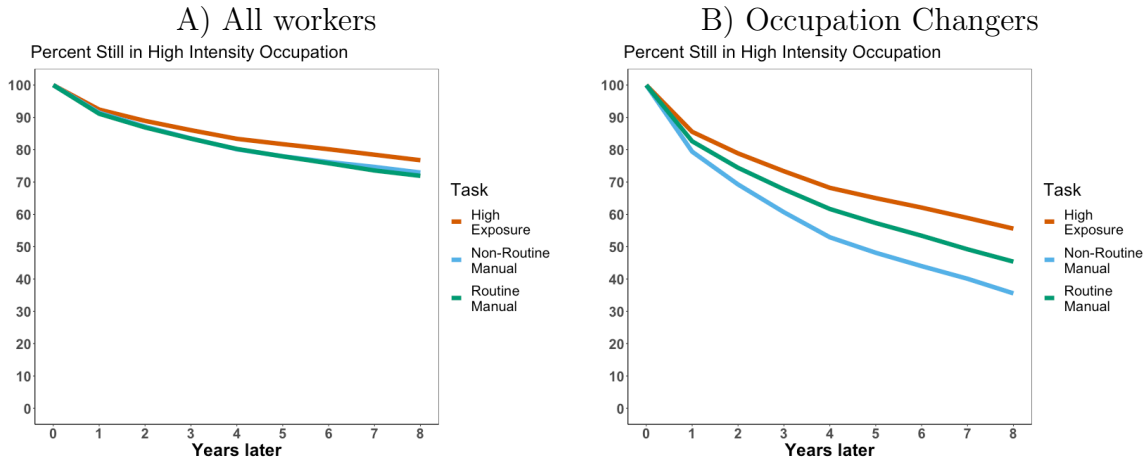


Figure 4.5: Persistence: Manual Intensity vs. Exposure

Panel A graphs worker persistence in occupations with high manual intensities (both routine and non-routine) alongside environmental exposure. The red line represents high exposure, the blue line represents high non-routine manual, and the green line represents high routine manual. For each task, the sample includes all workers present for all 9 years of my data who ever work in an occupation that I label as high in that task. An occupation has a high task intensity if its score is 1.5 standard deviations above the mean. For θ, this corresponds to 50, and for the task groups, this corresponds to 1.5 (by construction). Panel B shows the same comparison, but limits the sample to workers who have changed occupations at least once during my sample period.

In sum, while it may offer some explanation, the distribution of tasks between high and low exposure work is unlikely to fully account for the high degree of labor market segmentation between them. Other factors such as professional networks, language proficiency, or preferences for outdoor work may also play a role in explaining this persistence. In the Appendix, I show that workers without a bachelor's degree, as well as foreign-born workers, are disproportionately likely to work in high exposure jobs. Both groups also exhibit particularly strong persistence in high exposure work. However, more work is needed to fully understand the explanations for the segmentation between high and low exposure labor markets.

4.3.4 Low Exposure Options for High Exposure Workers

In this section, I characterize low exposure occupations that are popular destinations for high exposure workers. While inter-exposure mobility rates are low, Figure 4.3 shows that many workers leaving high exposure occupations do manage to find low exposure work. However, it is unlikely that low exposure occupations receive high exposure workers equally. Understanding which low exposure occupations receive more high exposure workers is important for two reasons. First, it can provide suggestive evidence on recent trends in inter-exposure mobility. For example, in light of the growing automation of routine tasks and consequent labor market polarization of the past few decades (Autor and Dorn, 2013), it is possible that these low exposure options for high exposure workers have become more or less prevalent. Second, it has important implications for how climate change might affect inequality within low exposure occupations. For example, lower wage low exposure occupations may be more likely to receive high exposure workers. If this is the case, then climate-induced occupational mobility to low exposure work would suppress wages in these lower wage occupations more than in the higher wage ones, exacerbating existing inequalities.

For the purpose of this exercise, I calculate each occupation's routine task intensity (RTI). This measure captures both an occupation's intensity in routine tasks and its lack of intensity in non-routine tasks. It is a common proxy for an occupation's vulnerability to automation

(Autor and Dorn, 2013). It is therefore useful to help understand how low exposure options for high exposure workers have changed due to automation and how they may continue to evolve in the future. I define RTI with the following equation:

$$RTI = RC + RM - NRA - NRI - NRM, \tag{4.1}$$

where RC is an occupation's routine cognitive score, RM is its routine manual score, NRA is its non-routine analytical score, NRI is its non-routine interpersonal score, and NRM is its non-routine manual score. Across all jobs in the *DADS Postes* sample, RTI has an employment-weighted mean of 0 and a standard deviation of 2.7.

I separate low exposure jobs into two categories: those which high exposure workers often transition into ("options"), and those which they do not often transition into ("non-options"). I include two criteria in this definition, and an occupation counts for the "options" category if it satisfies either one. First, I check if at least 9 percent (the percentage of high exposure jobs overall) of an occupation's incoming transfers are from high exposure occupations. However, this threshold excludes occupations that, despite receiving a large number of high exposure workers, also receive too many low exposure workers to meet the 9 percent cutoff. For example, in my sample, 789 workers left high exposure occupations to become "hand packers and packagers", an occupation with an exposure score of 28. Despite only making up 7.6 percent of transfers into this occupation, 789 is higher than the number of incoming workers for 95 percent of low exposure occupations, including 34 that meet the 9 percent cutoff. To address this type of issue, I also include any occupations with at least 100 incoming transfers from high exposure occupations. Occupations satisfying at least one of these criteria account for 52 percent of low exposure jobs and 85 percent of switches from high exposure occupations to low exposure ones.

Table 4.2 presents the characteristics of occupations in three categories: 1) low exposure "non-options", 2) low exposure "options", and 3) high exposure occupations. The 111 low exposure occupations that are options for high exposure workers make up 47.6 percent of

the labor market, compared to 43.4 percent for the low exposure occupations that are not.

The difference in wages between the "options" and 'non-options" groups suggests that inter-exposure mobility could increase inequality between low exposure occupations. The mean wage for "options" jobs is €17/hour, compared to €22/hour for "non-options" jobs. To the extent that it is possible, mobility from high to low exposure occupations will primarily be concentrated in lower wage destination occupations. This increased labor supply could suppress wages in the "options" occupations, widening inequality between them and the "non-options" occupations.

One would expect high exposure occupations to be more similar to the "options" group, as workers tend to find new work in similar occupations (Gathmann and Schönberg, 2010). What is less clear is which exact characteristics they have in common, and-perhaps more importantly-where they differ. I find that they are more similar in almost every characteristic I test. They both have higher exposure, lower wages, higher manual task intensity, and lower analytical and interpersonal task intensities than the "non-options" group.

However, one difference stands out. While high exposure occupations have low routine task intensity, "options" occupations have relatively high routine task intensity. While only 14.6 percent of high exposure jobs, and 16.8 percent of "non-options" jobs, are in the highest quartile of routine task intensity, 29.4 percent of "options" jobs are. This finding makes sense in light of the fact that both high exposure work and routine work are more easily available to low-skilled workers (Maestas et al., 2017; Autor and Dorn, 2013). In addition, occupational exposure and routine cognitive task intensity are negatively correlated (Figure 4.4E). This observation suggests that many of the job opportunities for low-skilled workers fall into one of two buckets: 1) high exposure work with high manual intensity, and 2) low exposure work with high routine cognitive intensity, and therefore high routine task intensity.

Table 4.2: Low Exposure Options for High Exposure Workers

	(1) Low θ Non-Options	(2) Low θ Options	(3) High θ
Number of Occupations	210	111	47
Percent of Jobs	43.4	47.6	9.1
Percent of High to Low Transfers	14.9	85.1	-
Mean θ	10.2	25.6	69.4
Mean Wages (€/hour)	21.9	17.0	15.5
Routine Tasks (RTI)	-1.1	0.89	0.67
Routine Manual (RM)	-0.51	0.25	1.1
Routine Cognitive (RC)	-0.20	0.20	-0.12
Non-Routine Analytical (NRA)	0.46	-0.33	-0.48
Non-Routine Interpersonal (NRI)	0.44	-0.30	-0.53
Non-Routine Manual (NRM)	-0.50	0.19	1.4
RTI %ile > 75	16.8	29.4	14.6
RC %ile > 75	21.7	26.2	13.5
RM %ile > 75	4.3	28.2	97.6
NRA %ile > 75	44.0	12.3	0.3
NRI %ile > 75	41.7	12.9	7.0
NRM %ile > 75	6.1	31.9	65.6

This table characterizes occupations according to three groups. Column 1 includes the low exposure ($\theta < 50$) occupations that high exposure workers do not frequently transfer into ("non-options"). Column 2 includes low exposure ($\theta < 50$) occupations high exposure workers do frequently transfer into ("options"). An occupation counts for the "options" category if either 1) at least 9% of its incoming transfers are from high exposure occupations, or 2) it has at least 100 incoming workers who left a high exposure occupation. Column 3 includes high exposure (θ) occupations. All task scores and exposure capture the mean level for the corresponding group. By construction, the task scores, other than routine tasks (RT) have a mean of 0 and a standard deviation of 1. RTI is equal to RC + RM - NRA - NRI - NRM, and has a mean of 0 and standard deviation of 2.7.

As past work has extensively documented, routine work is on the decline in the United States and many Western European countries. In France, for example, middle-wage jobs, many of which have high RTI, saw an 8.6 percentage point drop in employment share from 1993-2010. The same paper finds that across Europe, a one standard deviation higher RTI was associated with a 0.9 percentage point slower yearly growth rate (Goos et al., 2014). I confirm that routine intensive jobs have been declining in France in the period from 2011 to 2019. Figure 4.6A plots the employment share of occupations in the top quartile of routine task intensity across France over time. It shows an overall decrease of around 1 percent in this period. Panel B, which plots individual task intensities, shows that both a decrease in the prevalence of routine cognitive tasks and an increase the prevalence of non-routine tasks help explain this decline. Routine cognitive task intensity has consistently decreased from 2011-2019, while non-routine analytical and non-routine interpersonal intensities have grown over this period. Routine manual and non-routine manual tasks have evolved almost identically, offsetting each others' effects on changes in RTI over this period. These graphs show that, even in a relatively short time frame, routine task intensive jobs have diminished in prevalence in France.

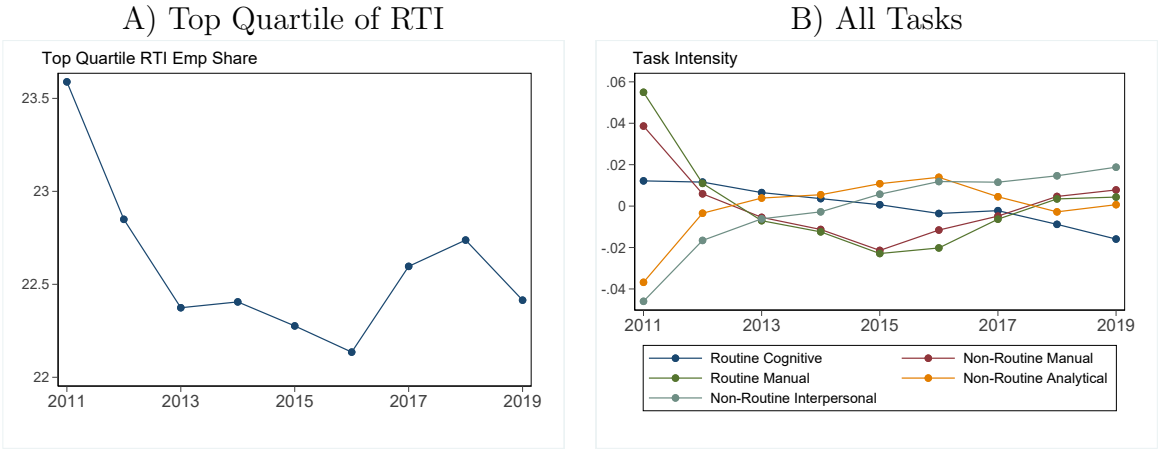


Figure 4.6: Task Intensity by Year

Panel A shows the employment share of occupations in the top quartile of routine task intensity (RTI) from 2011 to 2019. Panel B shows different task intensities per year from 2011-2019. By construction, task intensities have a mean of 0 and a standard deviation of 1 across the entire sample.

These observations have implications for both the past and the future of inter-exposure mobility. First, they suggest that low exposure options for low-skilled workers have decreased alongside the decades-long decline in routine task-intensive jobs. This observation may offer some explanation for today's low rates of inter-exposure mobility. Second, this decline in routine task-intensive jobs is projected to continue into the future (Hensvik and Nordström, 2023). Together with the increasing disutility of environmentally exposed work in the future, this projection paints a worrying picture for low-skilled workers. Because my data only cover 9 years, I cannot test for a causal relationship between the decades-long decline in routine task employment and inter-exposure mobility. However, the importance of this link to the future welfare of low-skilled workers motivates such an analysis in future research.

4.4 Outside Occupation Options and Climate Shock Spillovers to Low Exposure Workers

Estimates of climate damages to worker welfare generally focus on the direct effects to high exposure workers. However, despite the strong segmentation between high and low exposure

labor markets, there are likely to be considerable spillover effects to low exposure workers. Understanding the nature of these spillovers and how they are distributed is important to designing comprehensive climate adaptation policy. In Section 3, I discussed the potential spillover effects of high exposure workers moving to low exposure occupations. Specifically, I showed that the low exposure occupations that might see an increase in labor supply from workers leaving high exposure jobs are predominantly low wage occupations. This phenomenon could increase inequality between workers in low exposure occupations. In this section, I illustrate a second channel through which climate shocks could affect inequality between different low exposure occupations: workers' outside occupation options.

I conduct a back of the envelope calculation to simulate the indirect effects of a climate shock on low exposure occupations' wages through their outside occupation options. Past work shows that a worker's outside occupation options, defined as their job opportunities outside their occupation, influence their current wages. It also shows that observing the flows of workers departing a given occupation provides a valid measure of the outside options for workers in that occupation (Schubert et al., 2022). Figure 4.3 shows that certain low exposure occupations have a larger percentage of high exposure outside occupation options than others. This finding suggests that the indirect effects of climate change on low exposure workers, through their outside occupation options, will not be distributed equally. If these effects are concentrated on low wage low exposure destination occupations, then a climate shock could indirectly widen inequality between different low exposure occupations.

This calculation requires an estimate of the influence of outside occupation options on wages in France. France has a more rigid labor market with stronger worker protections than the US, where past work estimates that a 1 percent increase in the wages of an occupation's outside options increases its own wages by 0.1 percent (Schubert et al., 2022). I would expect this coefficient to be smaller in France due to its labor market conditions. Despite the necessity of this calculation, it is not the focus of this project. In the Appendix, I describe the methodology for this calculation and its results in detail. I find that in France the outside occupation options elasticity of wages is 0.052 percent. This elasticity is a little more than

half that found in the US. In the rest of this section, I use this elasticity to calculate the indirect effects of a climate shock on the wages of low exposure occupations, through their outside occupation options.

Methods

I calculate the indirect effects of climate changes on the wages of low exposure occupations, through their outside occupation options, using the following equation:

$$\Delta_{\%} w_{o,ind} = \beta_1 \cdot \Delta_{\%} OOI = \beta_1 \cdot \Delta_{\%} \sum_{p \neq o}^{N_{occs}} \pi_{o \to p} \cdot \frac{s_{p,m,t}}{s_{p,t}} \cdot w_{p,m,t}, \tag{4.2}$$

where $\Delta_{\%} w_{o,ind}$ is the indirect wage change (in percentage terms) for each occupation, $\Delta_{\%} OOI$ is the change in its outside occupation options index[31], and β_1 is the elasticity of wages to outside occupation options quality from equation C.4 in the Appendix. The second equality holds from the definition of the outside occupation options metric from in equation C.2 in the Appendix. Since I am interested in an aggregate assessment of indirect wage impacts, I abstract away from local relative employment shares of each occupation in each CZ ($\frac{s_{p,m,t}}{s_{p,t}}$) by setting them equal to one. Next, I assume that national-level occupational mobility flows ($\pi_{o \to p}$[32]) remain constant. So, workers' outside occupation options indices only change as a weighted average of the wage changes of those outside occupation options, weighted by $\pi_{o \to p}$. Equation 4.2 reduces to

$$\Delta_{\%} w_{o,ind} = \beta_1 \cdot \Delta_{\%} \sum_{p \neq o} w_p \cdot \pi_{o \to p}, \tag{4.3}$$

[31]This is equal to the average wages of a worker's outside occupation options, weighted by a worker's probability of moving to that occupation given they are changing occupations. For more details, see equation C.2 in the Appendix.

[32]$\pi_{o \to p}$ is equal to the probability that a worker leaving a job in occupation o moves to occupation p. In practice, this is equal to the product of two quantities: 1) the "leave share" of occupation o (the percentage of workers who leave occupation o when they leave their job) and 2) the percentage of workers who move to occupation p when leaving occupation o.

The purpose of this exercise is not to predict actual wage losses to low exposure workers. Rather, it is to calculate the losses to low exposure workers *relative* to the damages to high exposure worker wages. Therefore, for this exercise, I assume a uniform 1 percent decrease in high exposure occupations' wages, and no decrease for low exposure wages. Using national-level average wages for each occupation from 2011-2019, I calculate the resultant change in OOI as follows. First, I calculate each occupation's baseline "national-level outside occupation options index" ($OOI_o^{base} = \sum_{p \neq o} w_p \cdot \pi_{o \to p}$). Then, after applying the 1 percent wage reduction for each high exposure occupation, I recalculate a new OOI_o^{shock}. The percentage change between OOI_o^{base} and OOI_o^{shock} is $\Delta_{\%} OOI$. By multiplying this by β_1, I get each occupation's indirect percentage wage change due to a 1 percent decrease for all high exposure occupations.

I compare the indirect effects on wages between occupations with average wages above and below the median for low exposure workers.

Results

I find that the indirect impacts of climate change are substantially higher for low wage low exposure occupations than they are for high wages ones. Figure 4.7A shows the distribution of indirect wage shocks separately for low exposure workers above and below the median of wages. For a 1 percent decrease in high exposure occupation wages, the median low wage low exposure worker experiences a 0.0016 percent decrease in wages, compared to 0.00034 percent for the median high wage low exposure worker. The lower wage group experiences higher losses in real hourly wages as well, despite having lower wages (Panel B). I calculate aggregate hourly wage losses by multiplying the real hourly wage losses for each occupation by its average employment from 2011-2019. I find that, despite earning less in general, below median wage occupation aggregate losses are over twice as large (4,174 people-€/hour vs. 1,739 people-€/hour) as above median wage occupation aggregate losses. As I show in the Appendix, the qualitative result of this exercise is robust to altering the assumptions used

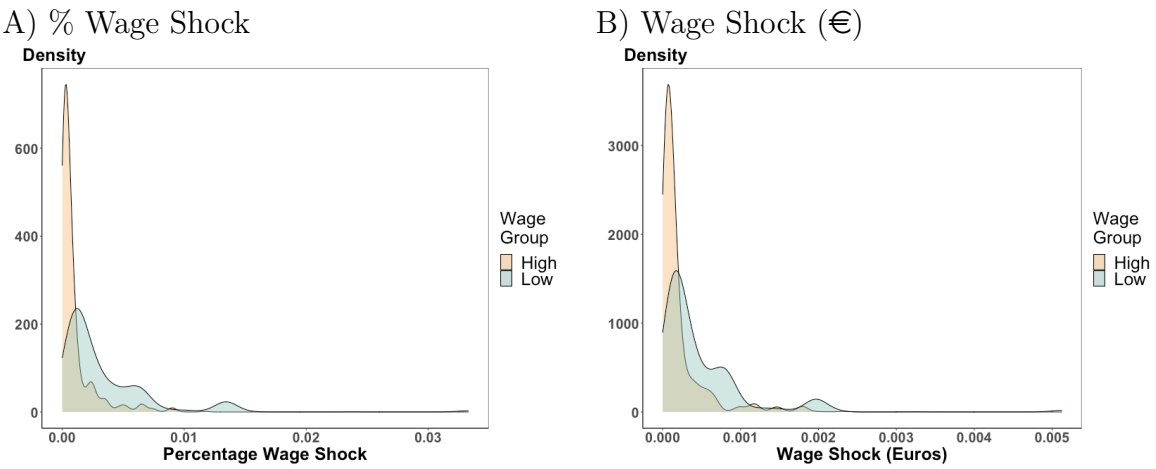


Figure 4.7: Indirect Climate Shock Effects to Low Exposure Occupations

Panel A graphs the employment-weighted distribution of indirect wage percentage shocks after a 1% decrease in all high exposure occupations' wages. I show the distribution separately for low exposure occupations with above and below median wages. Panel B graphs the wage shock in 2019 €.

for the calculation.

Through the outside occupation options channel at least, climate change will likely act as an inequality widener between low exposure workers. The magnitude of these losses, even to the below median wage group, are not especially high. However, it is important to note that outside occupation options are only one channel for spillovers of climate shocks to affect low exposure workers. For example, this channel is separate from the movement of high exposure workers into low-wage work, which is also likely to increase inequality between high and low wage low exposure occupations. These findings motivate more research on climate spillovers across the labor market, which are important to consider when designing climate adaption policy for our labor markets.

4.5 Within-Occupation Differences

The analysis so far has focused on outcomes at the occupation level. However, workers *within* the same occupation might have different probabilities of transitioning between high and low exposure occupations. These differences may have important implications for climate

change's effects on inequality. For example, imagine two groups of workers in the same high exposure occupation. Group 1 earns more than Group 2. Group 1 workers also have certain qualities that make them more likely to be able to transfer to low exposure work. They will be less vulnerable to the effects of climate change, which might widen inequality between the two groups.

There are many characteristics which may influence both a worker's wages and their propensity to move from high to low exposure work. For example, workers who speak French more proficiently[33] might be more likely to earn higher wages (relative to the rest of their occupation) and to be promoted to a low exposure occupation.[34] Identifying heterogeneity in inter-exposure mobility based on each of these numerous possible characteristics, or attempting to create an "index of inter-exposure mobility potential" using observable demographic characteristics, is unlikely to be a fruitful approach.

Instead, I calculate the relationship between a worker's propensity for inter-exposure mobility and their position in their within-occupation wage distribution. This approach has two advantages. The first is that it lends itself to a straightforward analysis of climate change's effects on economic inequality. Specifically, using this framework, I can easily compare inter-exposure mobility differences between workers in the same occupation with different levels of earnings. The second advantage is theoretical. As past work has argued, a worker's within-occupation relative earnings, after controlling for observable characteristics, are a relevant proxy for a worker's unobservable labor market abilities (Groes et al., 2015).

Groes, Kircher, and Manovskii (2015) (henceforth GKM) shows that these unobservable

[33]In the Appendix, I show that high exposure workers who are college educated or born in France are more likely to transition to low exposure work than those who are not. The *DADS Panel* does not include the education or nationality of workers, so I use a subsample of workers who applied to unemployment benefits from the *FH-DADS* for the education and nationality analysis. The *FH-DADS* is not suitable for this within-occupation analysis, however, as it 1) does not offer a representative view of the labor market and 2) does not include enough observations per occupation.

[34]Importantly, many of these characteristics are unobservable in administrative datasets large enough to conduct within-occupation analyses of occupational mobility. Additional examples include charisma, rapport with one's managers, dedication, and adaptability.

labor market abilities, as revealed by a worker's relative wage, are related to occupational mobility outcomes. In Denmark, the setting of their study, the likelihood of changing occupations is highest for high earners and low earners within an occupation. High earners within an occupation tend to move to occupations with higher average wages. In contrast, low earners within an occupation tend to move to lower-paid occupations. This finding has potential implications for inter-exposure mobility as well. High exposure workers moving to a higher-paid occupation likely have more low exposure jobs available to them than those moving to a lower-paid occupation. Figure 4.8 shows that there 91 low exposure occupations, employing 20 percent of low exposure workers, with higher wages than the highest paid high exposure occupation.[35] However, 230 low exposure occupations are below this threshold, and many are similarly paid to low wage high exposure occupations. Seventy-five low exposure occupations and 40 percent of all low exposure employment are in occupations with wages below the median for high exposure occupations of €15.6/hour.

[35]This is occupation 480b, "Merchant marine and fishing captains", with an average wage of €23.3/hour.

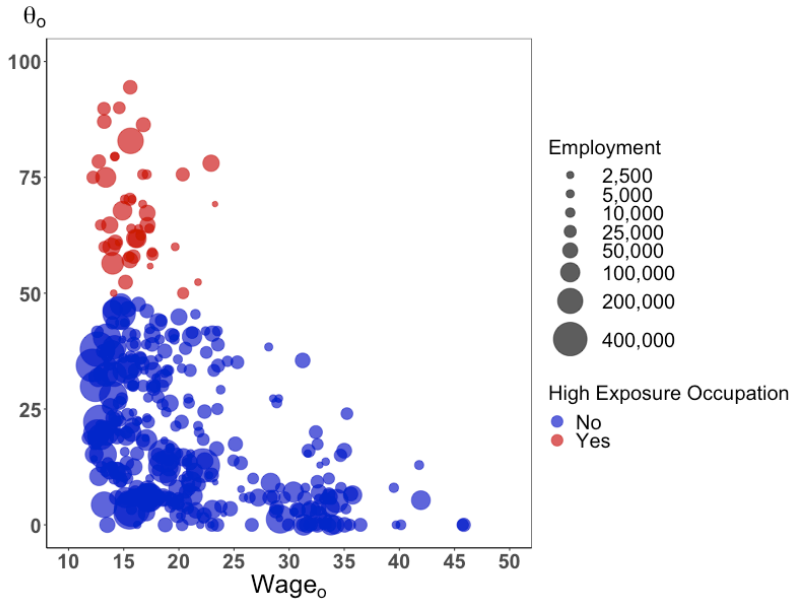


Figure 4.8: Exposure and Wages

This figure shows the relationship between an occupation's exposure level and its average yearly wages from 2011-2019 (in 2019 €). Each dot represents a single occupation and its size is based on average yearly employment from 2011-2019. High exposure occupations are in red, and low exposure occupations are in blue. The data include 368 occupations.

The distribution of wages across high and low exposure occupations and the findings from GKM lead me to posit the following hypotheses:

Hypothesis 1: When changing occupations, high earners within their high exposure occupation will be more likely to move into low exposure work than low earners. This hypothesis follows from the idea that high earners tend to move into higher paid work, which is more likely to be low exposure. If true, this finding would suggest that occupational mobility based climate adaptation prospects are better for workers with higher wages relative to others in their occupation.

Hypothesis 2: When changing occupations, low earners within their low exposure occupation will be more likely to move into high exposure work than high earners. This hypothesis follows from the same observations as Hypothesis 1. If true, this would mean that low earning low exposure workers are more vulnerable to the spillover effects of a climate shock than are high earning low exposure workers. If confirmed, this finding would suggest that these spillovers would widen inequality not just between low exposure occupations, as I show in Section 4, but within these occupations as well.

Methods

For this analysis I use the sample of individual worker panel data from the *DADS Panel*. Following GKM, I restrict the sample to men in order to avoid capturing the impact of fertility decisions on occupational mobility. The bulk of my analysis focuses only on workers in, leaving, or moving into high exposure occupations, and 95 percent of workers in these occupations are men. In the Appendix, I show that my results are robust to including women in the sample.

I impose additional restrictions on the dataset to better match the *Small Sample* described in GKM. I drop workers after their first absence from the *DADS Panel*, which indicates that they were not employed on January 1^{st} that year by a private company in a

spell that lasted for at least 6 months.[36] I also drop workers after their first year of missing information on wages, occupation, firm, experience, or firm tenure. I then drop 9 occupations with fewer than 10 remaining observations in each year.[37] After these restrictions, the sample includes 3,727,121 observations for 777,715 men.

The main difference between my main sample and the *Small Sample* from GKM is that their sample only includes workers for 15 years after they finish their education. Doing so allows them to construct occupation and industry tenure variables and to ensure a consistent distribution of labor market experience across the years in their sample. I do not do so, for two reasons. First, it would preclude me from including older workers, whose occupational mobility I am interested in capturing as well. Second, I do not have worker education information, so I cannot reliably determine the age at which they finished their schooling.

I calculate each worker's position within the distribution of wages for their occupation and year as follows. First, I estimate the following regression:

$$\ln w_{it} = X_{it}\beta + \epsilon_{it}, \tag{4.4}$$

where w_{it} is worker i's wage in year t.[38] X_{it} is a series of controls including year fixed effects, number of years of labor market experience (raw, squared, and cubed), age, and firm tenure (raw and squared). I winsorize labor market experience and firm tenure at the 1 percent and 99 percent level. Relative to GKM, this regression lacks variables capturing education and occupation or industry tenure. The *DADS Panel* does not include any of these variables. However, for many workers, age and labor market experience together proxy for years of education. GKM construct tenure variables by observing workers' entire labor market histories since finishing their education. Without education information, or reliable

[36]Following GKM, even if a worker reappears in the sample after an absence, I do not include these subsequent years.

[37]These occupations have PCS4 codes 343a, 344c, 351a, 352b, 421a, 431c, 431e, 441b, and 562a.

[38]In constructing my original dataset, I set the bottom 5 percent and top 1 percent of wages within each year to missing.

data going far back enough, I cannot do so. Therefore, I can only include firm tenure, which the *DADS Panel* provides, in equation 4.4.

I estimate this equation separately for each occupation and obtain residual wages for each worker-year. These residuals more plausibly represent a worker's non-observable labor market abilities relative to others in the same occupation than do their raw wages. By conditioning on observables such as firm tenure, age, and labor market experience, I ensure that I am not just comparing workers at different stages of their careers. For example, a high exposure worker with more experience might have higher wages and be more likely to be promoted to a low exposure occupation. Consistent with GKM, however, I show that my main results are similar when ranking workers within each occupation-year using raw wages. This robustness check alleviates concerns that the results of this study are overly dependent on the X_{it} variables.

I then use these within-occupation wage percentiles to test how occupational mobility patterns and outcomes differ for workers within the same occupation. First, I replicate analyses from GKM by plotting two outcomes as a function of a worker's position in the distribution of wages in his occupation-year. The first outcome is the probability of a worker leaving his occupation. The second outcome is the probability of a departing worker moving to a higher "ranked" occupation. I rank occupations based on their average wages by estimating the regression in equation 4.4 without a constant term. I take the coefficients on the year dummies to indicate an occupation's average wage for a given year, adjusted for the characteristics of its workforce. I then rank the occupations within each year based on these coefficients. I plot both outcomes separately for all occupations and for high exposure occupations alone.

Next, I test the relationship between inter-exposure mobility and within-occupation relative earnings. To test Hypothesis 1, I only include workers leaving a high exposure occupation. I plot these workers' probability of moving to a low exposure occupation against their position in their within-occupation wage distribution. I test Hypothesis 2 with workers

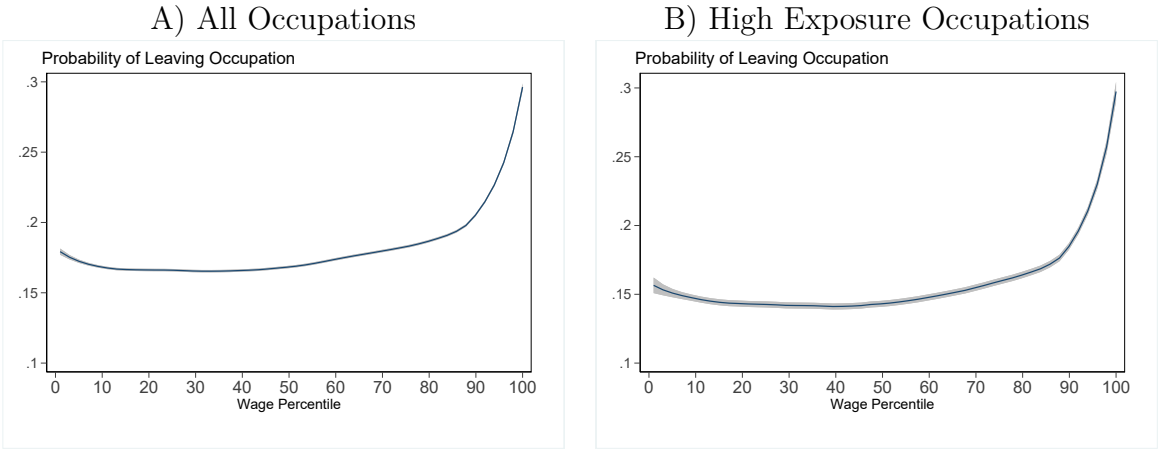


Figure 4.9: Occupational Mobility vs. Relative Earnings

Panel A shows the probability that a worker leaves their occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on the residuals from equation 4.4. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95 percent confidence intervals are shaded in gray. Panel B is the same as Panel A, but limits the sample to high exposure occupations. The sample in Panel A includes 3,727,121 observations for 777,715 male workers, and Panel B includes 519,278 observations for 141,225 male workers.

leaving low exposure occupations. I plot their probability of moving to a high exposure occupation against their relative earnings.

Results

Replication of GKM

First, I show how occupational mobility differs across the within-occupation wage distribution. Figure 4.9 Panel A shows that French workers at the high end of their within-occupation wage distribution are the most likely to leave their occupation. The highest earners have a 30 percent chance of leaving their occupation, whereas all other workers have between a 15 percent and a 20 percent chance. As compared to Denmark, French workers at the low end of the within-occupation distribution are relatively unlikely to change occupations (Groes et al., 2015). In the Appendix, I discuss potential reasons for this difference. Panel B shows that the same general shape holds for workers in high exposure occupations, with a slightly sharper increase at the higher end than is the case for all occupations.

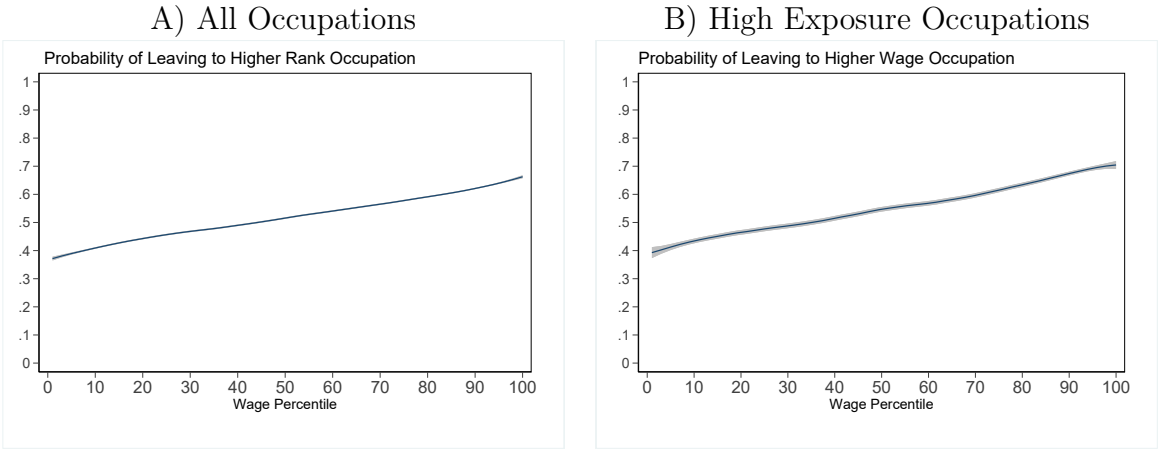


Figure 4.10: Moving to a Higher Ranked Occupation vs. Relative Earnings

Panel A shows the probability of a departing worker changing to an occupation with a higher rank (in the year of the switch). The wage percentile is calculated within each occupation and year based on the residuals from equation 4.4. Occupation ranks are determined by comparing year dummy coefficients from equation 4.4. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95 percent confidence intervals are shaded in gray. The sample is limited to workers leaving their occupations. Panel B is the same as Panel A, but limits the sample to high exposure occupations. The sample in Panel A includes 668,479 observations for 349,982 male workers, and Panel B includes 81,333 observations for 60,363 male workers.

Despite the differences in low earners' propensity to change occupations between France and Denmark, the wage related outcomes for workers who do switch are remarkably similar. Figure 4.10 Panel A shows that workers at the low end of their within-occupation wage distribution tend to move to a lower ranked occupation, while high earners tend to move to a higher ranked occupation. The observed range of probabilities of moving to a higher ranked occupation runs from below 40 percent for the lowest earners to above 65 percent for the highest earners. This range is slightly wider than that observed in Denmark (Groes et al., 2015). Panel B shows that this pattern is the same for high exposure workers.

High to Low Exposure Mobility

Next, I show that high earners leaving a high exposure occupation are more likely to move to a low exposure occupation. Figure 4.11A plots the probability that a leaving worker's new job is low exposure relative to his within-occupation wage percentile. It shows that this probability is relatively stable between the 1^{st} and 80^{th} percentiles, after which it begins a

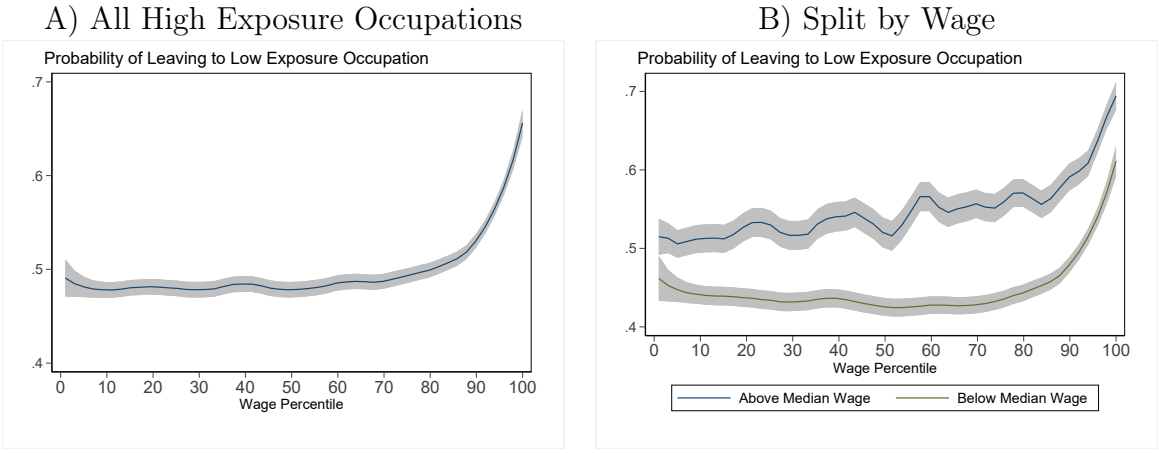


Figure 4.11: Moving to a Low Exposure Occupation vs. Relative Earnings

Panel A shows the probability that a worker leaves to a low exposure occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on the residuals from equation 4.4. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95 percent confidence intervals are shaded in gray. The sample is limited to workers leaving a high exposure occupation. Panel B shows the same graphs, with occupations separated into occupations above and below the median of average wages for high exposure occupations. The sample in Panel A includes 81,333 observations for 60,363 male workers. The blue line in Panel B (above the median) includes 39,210 observations for 32,176 male workers. The green line in Panel B (below the median) includes 42,123 observations for 34,669 male workers.

rapid increase. Whereas a leaver anywhere below the 80$^{\text{th}}$ percentile of the wage distribution has below a 50 percent chance of moving to a low exposure occupation, a leaver above the 90$^{\text{th}}$ percentile has a 59 percent chance. This non-linearity stands in stark contrast to the relatively linear increase in a worker's change in occupational rank with his place in the within-occupation wage distribution (Figure 4.9B).

To ensure that a small fraction of occupations with extreme differences in inter-exposure mobility are not driving this result, I calculate how many occupations display a similar pattern. For the purpose of this exercise, I define "median earners" as those between the 40$^{\text{th}}$ and 60$^{\text{th}}$ percentiles, and "high earners" as those above the 90$^{\text{th}}$ percentile. For each high exposure occupation, I test whether high earners are at least 10 percent more likely to transition to a low exposure occupation than median earners when changing occupations. I find that this is the case for 31 of 47 high exposure occupations making up 73 percent of high exposure employment.

This prevalent pattern can be explained by one of two possibilities, or both. The first is the fact that, above a certain wage threshold, workers can only increase their wages by switching to a low exposure occupation (See Figure 4.8). Once above this threshold, the probability of moving to a low exposure occupations rapidly increases. The second possibility is that these highest earners (above the 90^{th} percentile) have unobservable characteristics that their managers especially value. Workers with a better rapport with their managers could be more likely to earn larger raises and therefore higher wages than observably similar workers. They may also be more likely to be promoted into management roles, which are more likely to be low exposure.

To help explore these potential explanations, I split high exposure workers into two groups. Figure 4.11B replicates Panel A separately for workers in occupations with average wages above (blue line) and below (green line) the median for high exposure workers.[39] By definition, workers in the below median group have many high exposure options with higher average wages than their current occupation. Therefore, if a lack of higher wage high exposure options fully accounted for the heightened inter-exposure mobility among high earners, I would expect the pattern found in Panel A to be less true among this lower wage group. However, high earners, even in low wage high exposure occupations, transition to low exposure occupations at significantly higher rates than their lower wage counterparts. This finding suggests that high earners' unobservable characteristics provide at least a partial explanation for their high inter-exposure mobility.

Regardless of the explanation, the strong within-occupation differences in inter-exposure mobility have important implications. This evidence, which confirms Hypothesis 1, suggests that the highest earners may have a stronger ability to adapt to climate change by changing occupations. This raises concerns about how workers' ability to adapt to climate change might exacerbate existing inequality even within occupations. Many inquiries into workers' ability to adapt to climate change are conducted at the occupation level. The findings

[39]I calculate these categories using the average national-level wages for men in high exposure occupations from 2011-2019.

presented in this section motivate future work at a more granular level. In addition, more work is needed to fully understand the reasons behind the differences in inter-exposure mobility across the earnings distribution. Designing optimal policies to aid those least able to adapt to climate change on their own requires understand exactly which factors, at an individual level, influence a worker's ability to adapt.

Low to High Exposure Mobility

Figure 4.12A shows that low earning low exposure workers are more likely to move to a high exposure occupation when changing occupations. While a worker in the bottom decile of within-occupation earnings has a 8.1 percent chance of doing so, a worker in the top decile only has a 4.9 percent chance. Panel B splits plots the graph in Panel A separately for two groups, those in high average wage (above the median) and those in low average wage (below the median) low exposure occupations. It shows, as discussed in Section 4.3, that workers in low wage low exposure occupations are more likely to transition into high exposure occupations. It also shows that even within these low wage occupations, a worker's percentage of high exposure outside occupation options differs across the wage distribution. This pattern is common across most low exposure occupations. For 237 occupations comprising 74 percent of low exposure employment, a bottom decile earner is at least 1.5 times as likely to transfer to a high exposure occupation than a top decile earner.

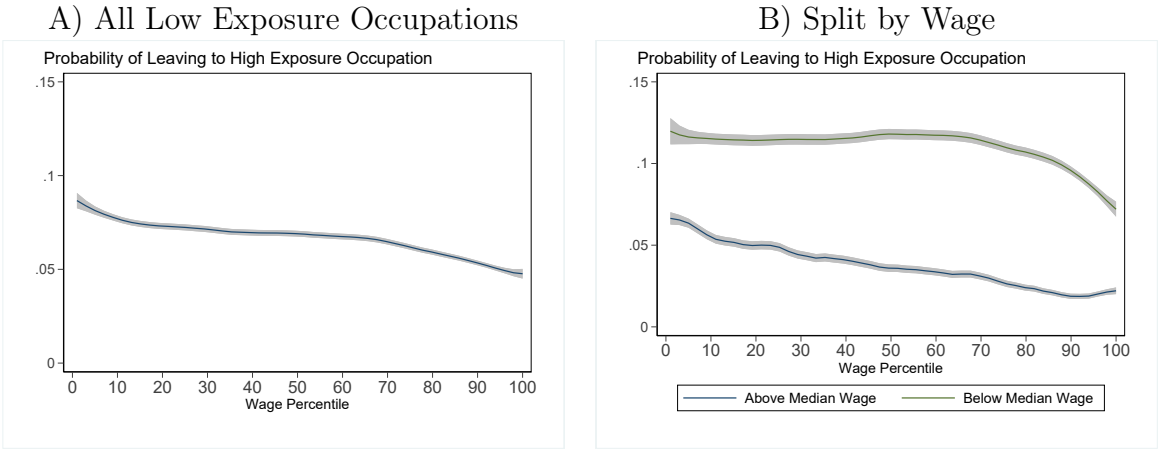


Figure 4.12: Moving to a High Exposure Occupation vs. Relative Earnings

Panel A shows the probability that a worker leaves to a high exposure occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on the residuals from equation 4.4. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95 percent confidence intervals are shaded in gray. The sample is limited to workers leaving a low exposure occupation. Panel B shows the same graphs, with occupations separated into occupations above and below the median of average wages for low exposure occupations. The sample in Panel A includes 587,146 observations for 316,446 male workers. The blue line in Panel B (above the median) includes 352,054 observations for 200,669 male workers. The green line in Panel B (below the median) includes 235,092 observations for 159,807 male workers.

This finding suggests that low earners will be more vulnerable to the spillover effects of a climate shock than high earners within the same low exposure occupation. Table 4.3 shows the percentage of high exposure destination occupations[40] for 4 different groups of workers. I split occupations into two groups: those above and below the median of average wages for low exposure workers. Within these two groups of occupations, I show the percentage of high exposure outside occupation options for those in the top and bottom deciles of within-occupation earnings. Only 2 percent of the outside occupation options for top decile earners in above median occupation are high exposure, compared to 6.2 percent for the bottom decile of earners in these same occupations. For below median occupations, these numbers are 8.2 percent and 11.6 percent. Therefore, as climate change affects the quality of high exposure work, the spillover effects will likely increase inequality not just across low exposure

[40]This measure is slightly different than outside occupation options, as it does not include the leave share of each occupation. However, it is still an important measure of an occupation's relationship to high exposure occupations.

Table 4.3: Percentage of High Exposure Destination Occupations

	Above Median Occupation	Below Median Occupation
Top Decile Worker	2.0	8.2
Bottom Decile Worker	6.2	11.6

This table shows the percentage of high exposure destination occupations for 4 sets of workers: those in occupations with average wages above and below the median for low exposure occupations, each split into top and bottom decile earners within their own occupation.

occupations, but also within them.

4.6 Conclusion

In this paper, I show that high exposure work exhibits strong persistence: workers in high exposure work remain in high exposure work for long periods of time. This finding indicates strong segmentation between the labor markets for high exposure and low exposure work and that the costs to inter-exposure mobility are higher than expected. Therefore, changing occupations is unlikely to be a successful adaptation strategy, exacerbating concerns that climate change will widen existing inequalities between high and low exposure workers. In addition, my findings suggest limited spillover effects of climate change to non-exposed workers. However, to the extent that these spillovers occur, they will be more concentrated in low wage occupations, and especially low wage workers in those occupations, increasing economic inequality between low exposure workers.

This paper is just one step in assessing the multiple ways that climate change's direct effects on exposed workers can spillover to other parts of the economy. While I focus on outside occupation options, there are other channels through which climate change could have indirect effects on worker welfare. For example, decreased productivity of high exposure workers could push firms to reallocate work, when possible, to more temperate climates, or to substitute labor with capital (Xiao, 2023a). These firm-level adaptation decisions could have important consequences for the local labor markets that see a resulting decrease in high

exposure work. In addition, any climate-related costs that firms do incur might be passed on to consumers. As high exposure work is involved in the production of many vital goods, such as food and housing, this possibility could raise the financial burden faced by low-income households.

More research on this topic is needed in other countries. The setting of my research is France, a country that has experienced a unique economic trajectory over that last half century. In most OECD countries, wage inequality has risen sharply. France is the major exception: wage inequality has fallen by around 25 percent during that period (Bozio et al., 2020). In 2019, earners in the top wage decile in France earned about 3.5 times more than earners in the bottom decile, a below median ratio for the EU, and much smaller than the US' 6.3. Research credits this low and shrinking wage inequality to high coverage of workers under binding collective agreements, a high and dynamic minimum wage, and progressive payroll tax regimes (Bosch and Weinkopf, 2017; Bozio et al., 2020). French workers also have relatively low occupational mobility: about two thirds the average level for European workers (Bachmann et al., 2020). On average, they have high levels of labor market power, likely dampening workers' desire and ability to move between occupations at different levels of environmental exposure. Therefore, changing occupations might be a more successful adaptation strategy in other countries. However, the outside occupation options elasticity of wages in France, which I show to be half the size compared to that for the US, is likely near the bottom of the range for similar countries. This, in combination with the possibility that low exposure occupations in other countries have a larger percentage of high exposure outside options, suggests that spillover effects of climate change to low exposure workers might be higher in other countries as well.

5

Conclusion

In this dissertation, I explore three settings where climate change threatens human well-being. In the first, my coauthors and I show that extreme heat increases financial distress for low-income households in California. In the second, my coauthors and I show evidence of heterogeneous dietary impacts of hot weather during the most recent crop growing period in rural India. We find that those negatively affected are likely already worse off than those who are able to adapt. In the third, I show that there exists a high degree of segmentation between the labor markets for high and low exposure occupations in France. This finding suggests that adapting to climate change by changing occupations will likely be difficult.

Together, these studies show substantial within-geography heterogeneity in climate impacts. They suggest that the costs of adaptations to extreme weather events are, for some, prohibitively high. Understanding the nature of these costs and how to mitigate them is an important step in designing appropriate policy responses. In some cases, like the labor market, crafting the right policies will require an understanding of the power dynamics between different groups.

Future Work: Power Dynamics and Climate Adaptation

In March 2024, California Governor Gavin Newsom decided to yet again delay adopting heat illness prevention regulations for indoor workers. California has had similar legislation for outdoor workers since 2006, and there is evidence that heat can also affect indoor workers (Park et al., 2021). The Newsom administration withdrew approval for the upcoming

vote because it learned that the rule would be too expensive for state prisons. Workers in California's warehouses, manufacturing plants and restaurants will have to wait for legal protections against heat exposure in order to accommodate the budgetary needs of prisons. The California Division of Occupational Safety and Health (Cal/OSHA) has since decided to revise the legislation with an exemption for prisons, a process that could take several weeks (Kuang, 2024a). This law would be so expensive for prisons because it would require significant infrastructure investments. California's prisons, including many in the desert, lack air conditioning, a growing problem under a warming climate (Kuang, 2024b). In order to keep temperatures during heat waves at levels optimal for human health, prisons would have to pay for prisoners' comfort.

This dynamic represents an extreme version of the climate adaptation decision settings that I am most interested in studying in the future: ones where the costs and benefits of adaptation investments belong to different groups. These situations arise within the labor market, but also in other settings such as landlord-tenant relationships and, indeed, in correctional facilities. Power imbalances between these groups make it less likely for mutually optimal levels of adaptation investments to materialize. For example, landlords or firms with high market share are less likely to need to install air conditioners to retain workers or tenants.

My interest in how power dynamics affect adaptation was spurred by the experiences of food systems workers during the COVID-19 pandemic. Due to meatpacking firms' reluctance to slow down production for worker safety, over 50,000 meatpacking workers in the US contracted COVID-19, with 250 deaths (Marks, 2022). During California's wildfire season of 2020, which coincided with the height of the pandemic, there were widespread accounts of farmworkers laboring in smoky conditions without a mask (Mahoney, 2020). As reports of these types of incidents multiplied, I sought to better understand why workers in such an important industry seemed to have so little protection during this crisis, despite the costs of some adaptations, such as providing masks to farmworkers, seeming relatively low.

Through historical and ethnographic accounts of agricultural and meatpacking workers, I learned that labor power dynamics might offer an explanation for this lack of protections. The nature of these dynamics in the meatpacking industry, for example, has undergone a cyclical evolution. In the early 1900s, the industry was dominated by a small number of large firms that hired from an ever-growing pool of Eastern European immigrants. As described in Upton Sinclair's 1906 novel *The Jungle*, worker safety and pay were abysmal. By the 1970s, however, the industry was comprised of many small firms and a heavily unionized workforce. Workers reaped the benefits of this new regime: wages for meatpacking workers were 20 percent higher than those of the average factory worker (Schlosser, 2006). Today, however, the industry is even more consolidated than in the era of *The Jungle*. Currently, the four largest companies control over 80% of the product market (Bolotova, 2022). In addition, meatpacking plants have broken worker unions through a combination of tactics including relocating from urban areas to rural regions with low employment prospects, and hiring Latin American immigrants, many of whom are undocumented. Meatpacking workers today are paid 24 percent *lower* than the typical factory worker (Schlosser, 2006).

Climate adaptation is a form of worker compensation that differs from wages in some important ways that motivate studying it separately. First, in theory, firms have more flexibility to dictate levels of adaptation investments. In contrast to minimum wage laws, heat illness prevention laws are relatively uncommon (OSHA, 2024). To the extent that these laws do exist, they are also more difficult to enforce: while minimum wage violations can be documented with pay stubs and rectified retroactively, these options are both more complicated for heat protection. In California's scattered farmlands and construction sites, an understaffed Cal/OSHA has a limited ability to enforce its heat illness protections (Riley et al., 2012). In addition, showing that an injury was caused by lack of heat protections is difficult, thereby complicating any grievance process. Second, as I note in my introduction, firms themselves may underestimate the benefits of climate control, whereas they are more likely to understand the tradeoffs involved with setting wages. Enhanced power, for example via stronger unions, might help workers enforce existing laws or bargain for adequate heat

protections. Impending climate change only makes understanding these dynamics more urgent.

As I learned, the intense power imbalances in the American food industry have deep roots. Large-scale agriculture in the United States began with the labor of enslaved Africans and Indigenous People (Reséndez, 2016). The 1935 National Labor Relations Act, which grants the right to form unions, excludes agricultural workers from its protections.[1] Today, an estimated 50% of farm laborers in the United States are working without proper legal authorization (Hernandez and Gabbard, 2019). In addition, through the H-2A visa program, which offers no pathway to citizenship, over 200,000 foreign workers obtain visas for seasonal work under a specific agricultural employer (Luckstead and Devadoss, 2019). These workers' legal status in the US is entirely dependent on working for the assigned employer, setting the stage for the well-documented abuse of workers in the program (Stockdale, 2012).

The power dynamics between food workers and their employers are driven by structural forces which extend far beyond the workplace itself. Agricultural producers face market pressures such as overseas competition or food distributors with the ability to dictate prices. These pressures require cost cutting, and hiring workers with few other options for employment helps keep labor costs low (Holmes, 2013). This situation makes addressing the power imbalances at the workplace level difficult. Unions are one option that workers can typically use to counteract firm power and demand better compensation. However, there is little scope for farms to raise their wages, due to their thin profit margins. A farm with a unionized and more expensive workforce, unable to raise its prices, becomes more likely to go out of business (Kumar, 2020). Ultimately, workers who are most vulnerable to the effects of climate change are suffering the combined weight of multiple complementary forces. My goal is to understand where policy can help to alleviate some of these pressures.

As I see it, quantitative researchers can play in important role in this area. First, by helping to understand how power dynamics causally affect adaptation decisions, we could

[1]By many accounts, this exclusion was a pre-condition for its passage by Southern Democrats who sought to carve out industries with a high number of Black workers (Coffman, 2022).

further motivate policies to address power disparities. Second, we can help to identify groups of workers whose power relative to their employer is especially limited and are therefore vulnerable to climate change. Conducting applied microeconomics research in settings with large power disparities between workers and their employers is difficult. Data on workers outside of traditional labor markets, such as undocumented or incarcerated workers, are rarely suitable for causal inference. That being said, there are still several ways that applied microeconomics can contribute to designing policies to reduce these workers' vulnerability to climate change.

One option is for applied microeconomists to study labor market laws and institutions that influence worker-firm power relationships. For example, situations like the high level of consolidation in the meatpacking industry have led labor economists to wonder whether we should consider antitrust remedies, which have traditionally focused on product market power, in the labor market (Naidu et al., 2018; Marinescu and Hovenkamp, 2019). Other scholars have pushed back, noting that traditional labor market remedies, such as minimum wages, may be a preferable approach to addressing these issues (Dimick, 2023). However, this idea misses the role of non-wage compensation, an important determinant of worker welfare (Maestas et al., 2023). An open question in this area is the extent to which firms may respond to minimum wage legislation by decreasing non-wage compensation, which could include climate adaptation investments (Clemens, 2021). If they can and do, in the absence of complementary legislation to address the root causes of suboptimal compensation, minimum wage legislation is less likely to substantially ameliorate overall worker welfare.

Much of this work will need to occur in traditional labor market settings where data are more readily available. For example, in ongoing work, my coauthors and I test how outside occupation options affect the temperature-injury relationship. We show that a standard deviation increase in the wages of a worker's outside occupation options decreases their injury sensitivity to heat by 14%. The data we use likely do not include workers at the furthest margins of our labor markets: undocumented migrant workers are unlikely to file workers' compensation claims despite being legally allowed to (Rebecca Smith, 2012), and the

resume data used to calculated outside occupation options underrepresent manual workers and those without a BA (Schubert et al., 2022). However, by showing how power dynamics generally affect adaptation decisions, I hope that we can motivate future research into areas with more extreme power imbalances. For example, we still have much to learn about the tradeoff between policies that mandate adaptations (and could therefore impose higher levels of adaptation investment than workers themselves want) and ones that attempt to level the bargaining playing field. In theory, the latter may be preferable from the point of view of market efficiency. In practice, however, there are likely many situations where the gulf in power is so large that a mandate is much more effective.

Finally, applied microeconomists can conduct research in data-poor settings by collecting their own. For example, Maestas et al. (2023) conduct a survey of 3,000 workers to help determine how working conditions such as schedule flexibility are distributed across workers in the US. Doing so allows them to present statistics on disparities in the level of workplace amenities enjoyed by different demographic groups. These findings would be impossible with already existing administrative data. While certainly no straightforward task, engaging in similar data collection efforts could increase applied microeconomics' usefulness in addressing questions to help those most marginalized adapt to climate change.

APPENDIX A

Appendix to Chapter 2

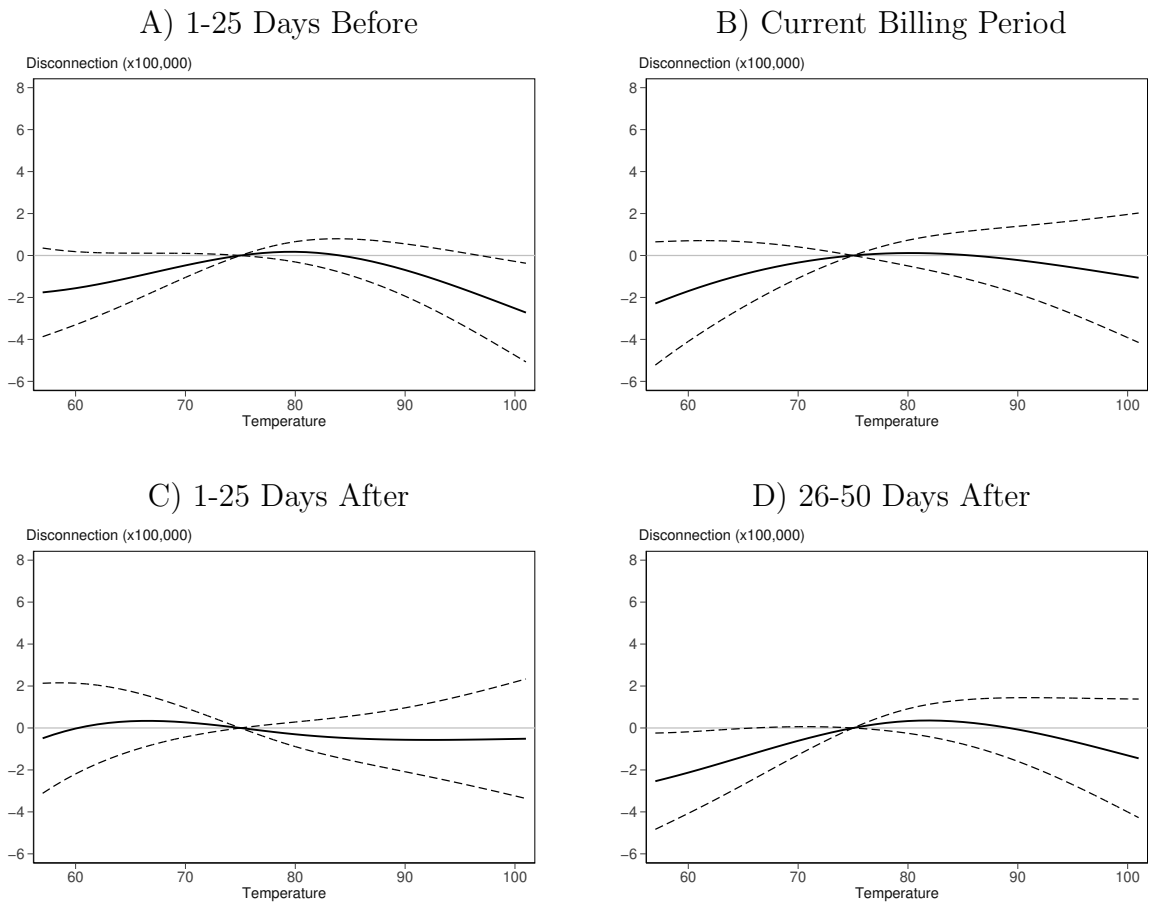


Figure A1: Effect of Temperature on Subsequent Disconnection Risk (1)

Panel A shows the change in disconnections (x100,000) for an additional day with a given maximum temperature relative to 75°F with a spline temperature model, for 1-25 days before the billing period. Panel B shows the same for the current period, Panel C for 1-25 days after and Panel D for 26-50 days after. All models include billing start date, zip by calendar week, age of account, and individual account fixed effects, as well as precipitation controls. Standard errors are clustered at the zipcode level. The dashed lines are 95% confidence intervals centered on the estimated coefficient. The models (in order from a-d) include 13,329,877, 13,329,871, 13,237,428, and 12,993,175 observations. The sample sizes are somewhat smaller with disconnection probabilities farther into the future since we are missing data after August 2017.

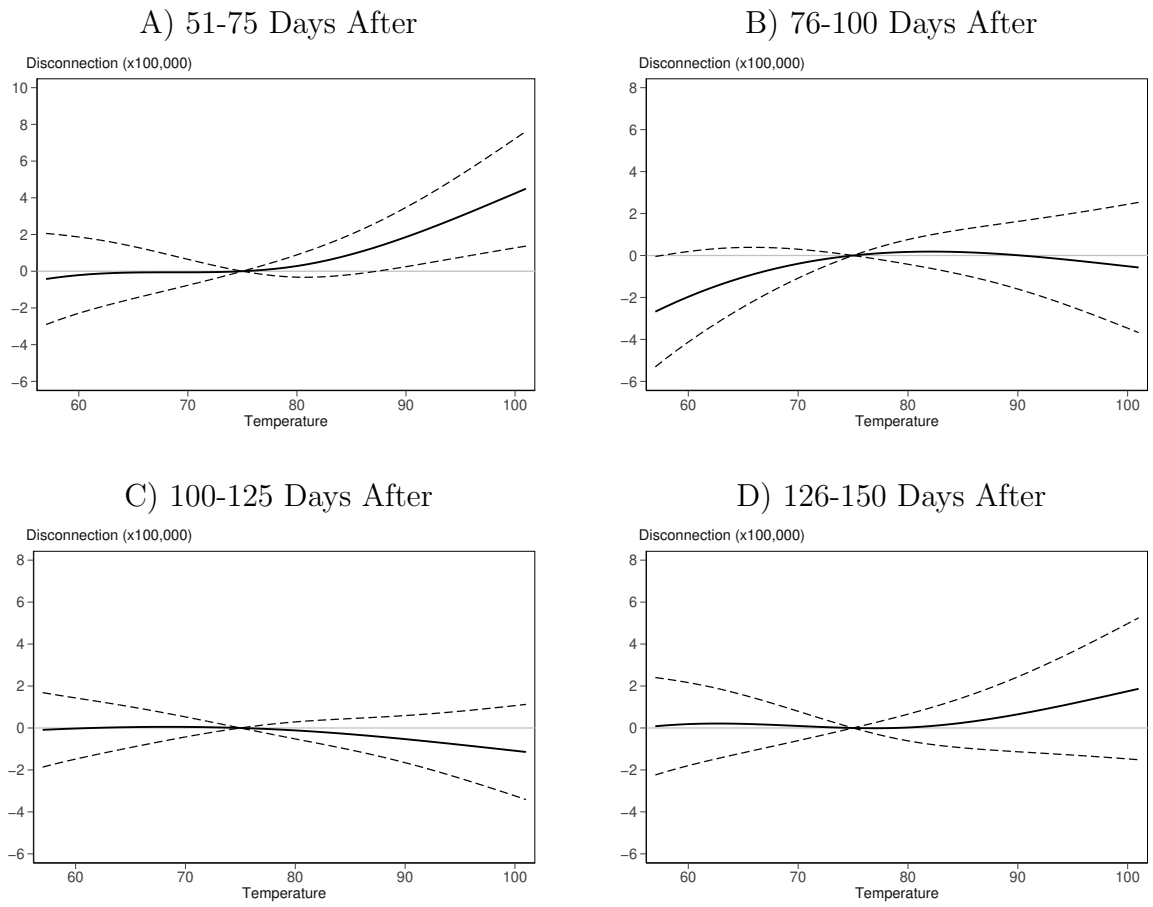


Figure A2: Effect of Temperature on Subsequent Disconnection Risk (2)

Panel A shows the change in disconnections (x100,000) for an additional day with a given maximum temperature relative to 75°F with a spline temperature model, for 51-75 days after the billing period. Panel B shows the same for 76-100 days after, Panel C for 101-125 days after and Panel D for 126-150 days after. All models include billing start date, zip by calendar week, age of account, and individual account fixed effects, as well as precipitation controls. Standard errors are clustered at the zipcode level. The dashed lines are 95% confidence intervals centered on the estimated coefficient. The models (in order from a-d) include 12,758,335, 12,522,895, 12,393,882, and 12,170,748 observations. The sample sizes are somewhat smaller with disconnection probabilities farther into the future since we are missing data after August 2017.

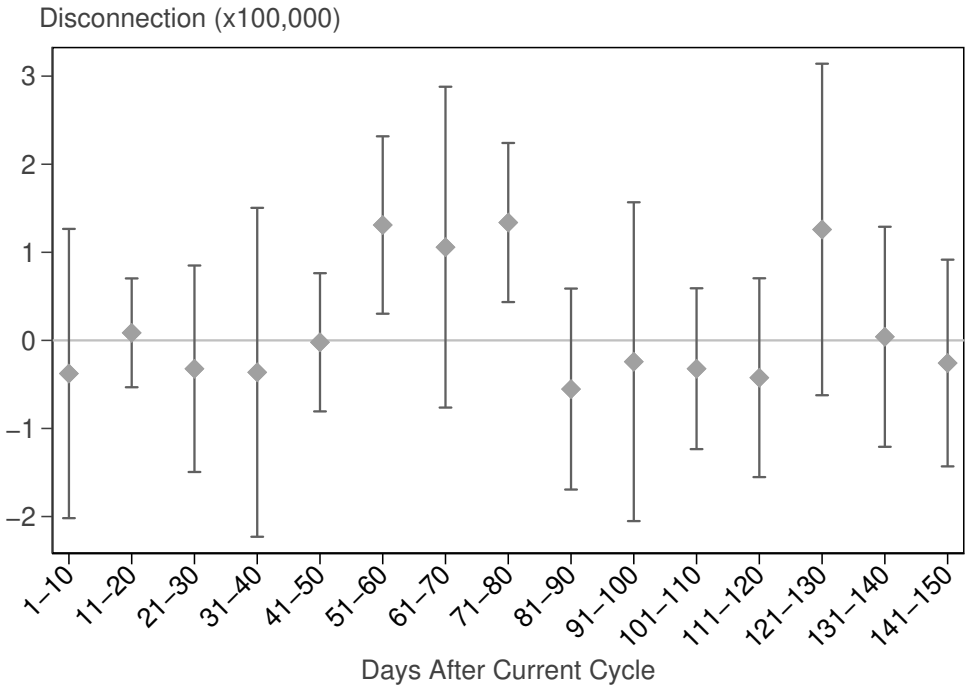


Figure A3: Effect of Day at 95°F on Subsequent Disconnection Risk

The change in disconnections for an additional day at 95°F relative to 75°F. All models include billing start date, zip by calendar week, age of account, and individual account fixed effects, as well as precipitation controls. Standard errors are clustered at the zipcode level. The brackets are 95% confidence intervals centered on the estimated coefficient. The largest sample size is 13,329,871 billing-period observations. The sample sizes are somewhat smaller with disconnection probabilities farther into the future since where we are missing data after August 2017.

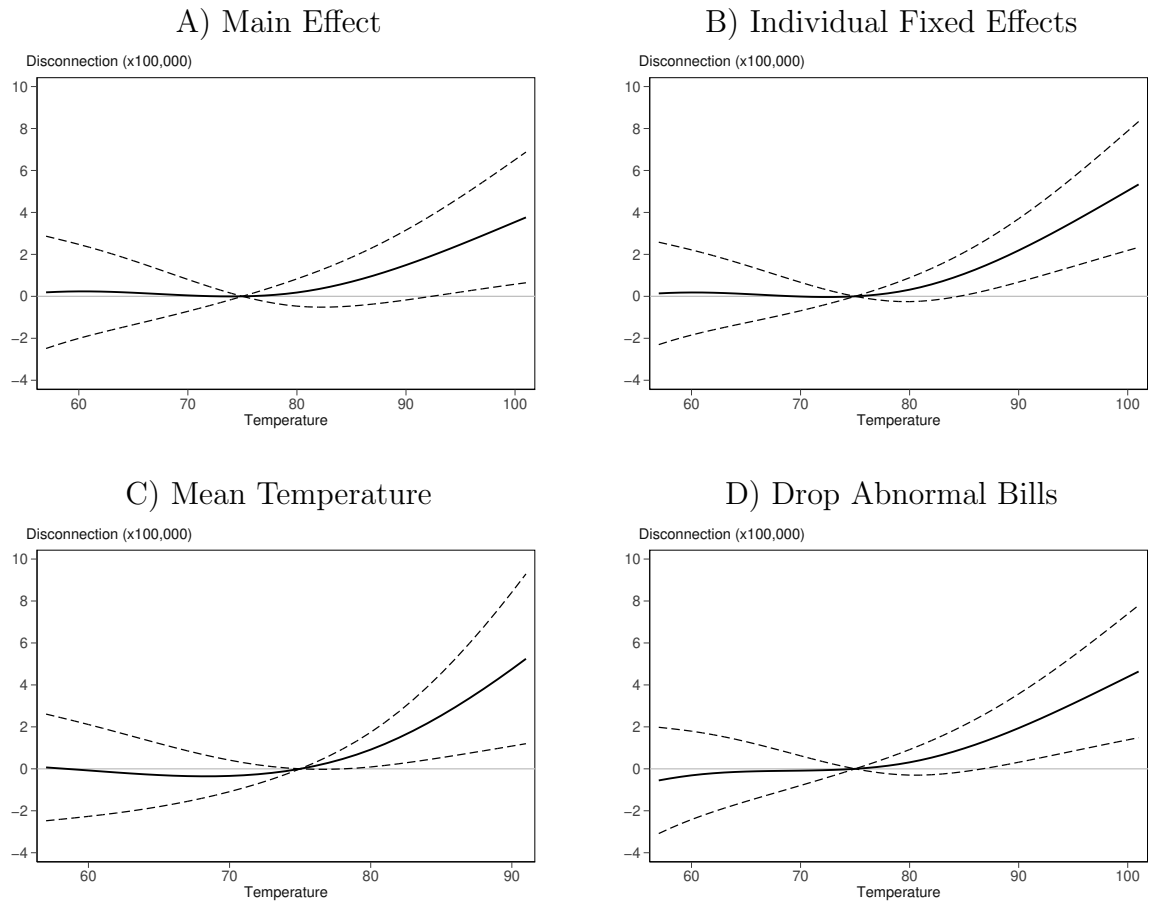


Figure A4: Robustness Checks on Temperature-Disconnection Relationship 51-75 Days Later

Panel A shows the change in disconnections 51-75 days after the close of the billing cycle for an additional day with a given maximum temperature relative to 75°F. The model includes billing start date, zip by calendar week, and age of account fixed effects, as well as precipitation controls. Panel B shows the same relationship as Panel A with the addition of individual account fixed effects and controlling for weather 30 days prior to and after the billing period. Panel C shows the same relationship as Panel A for mean temperature rather than maximum temperature, and with the addition of individual account fixed effects. Note that the x axis runs until 91°F. Panel D shows the same relationship as Panel A but dropping bills with abnormal lengths, and with the addition of individual account fixed effects. Standard errors are clustered at the zipcode level. The dashed lines are 95% confidence intervals centered on the estimated coefficient. Panel A includes 12,758,335 observations, Panel B includes 12,758,335 observations, Panel C includes 12,758,335 observations, and Panel D includes 12,616,659 observations.

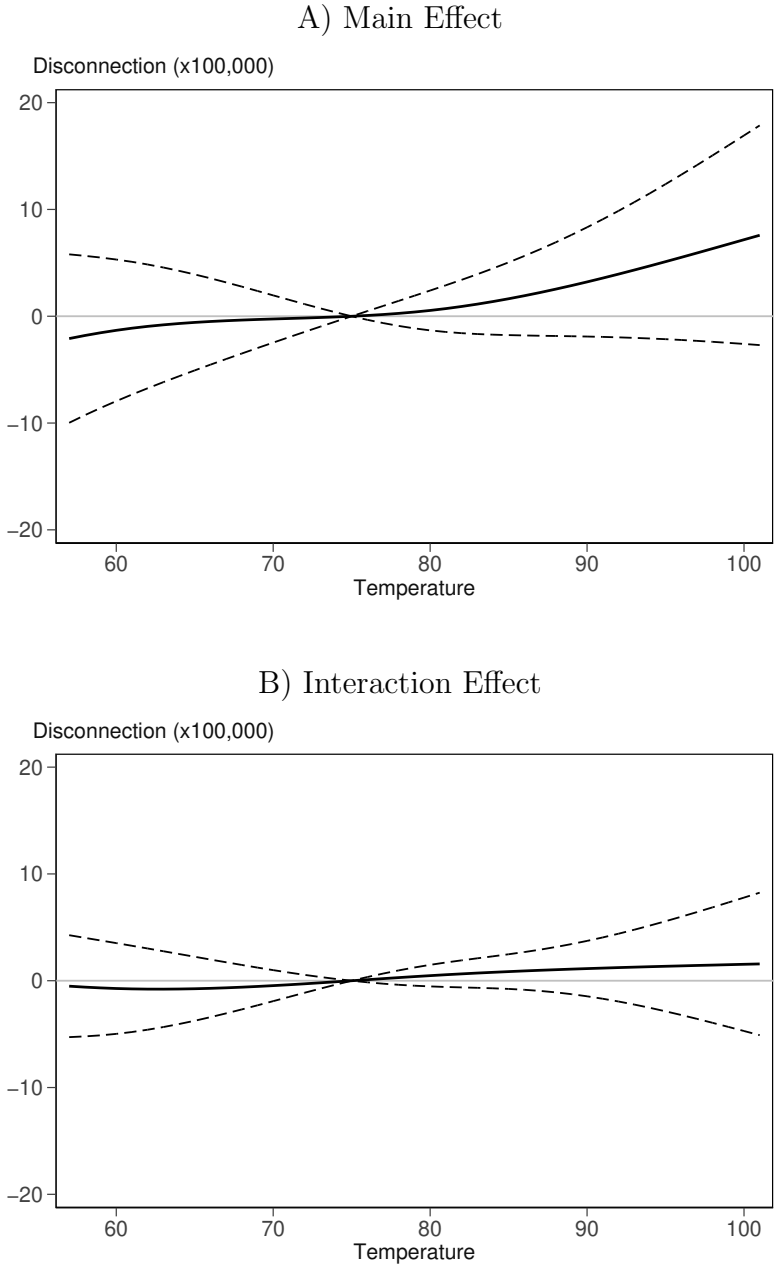


Figure A5: Modifying Effect of CARE on Temperature-Disconnection relationship 51-75 days later

Panel A shows the main effect of one day's temperature relative to one 75°F day on disconnection risk 51-75 days later. Panel B shows the interaction effect of temperature after the 26th period relative to one 75°F day. The model includes temperatures in the current billing period only. The model includes billing start date, zip by calendar week, and individual account fixed effects, as well as precipitation controls. And, we interact the temperature variables with a linear trend for time on CARE. Standard errors are clustered at the zipcode level. The dashed lines are 95% confidence intervals centered on the estimated coefficient. The sample size is 1,948,564 billing-period observations.

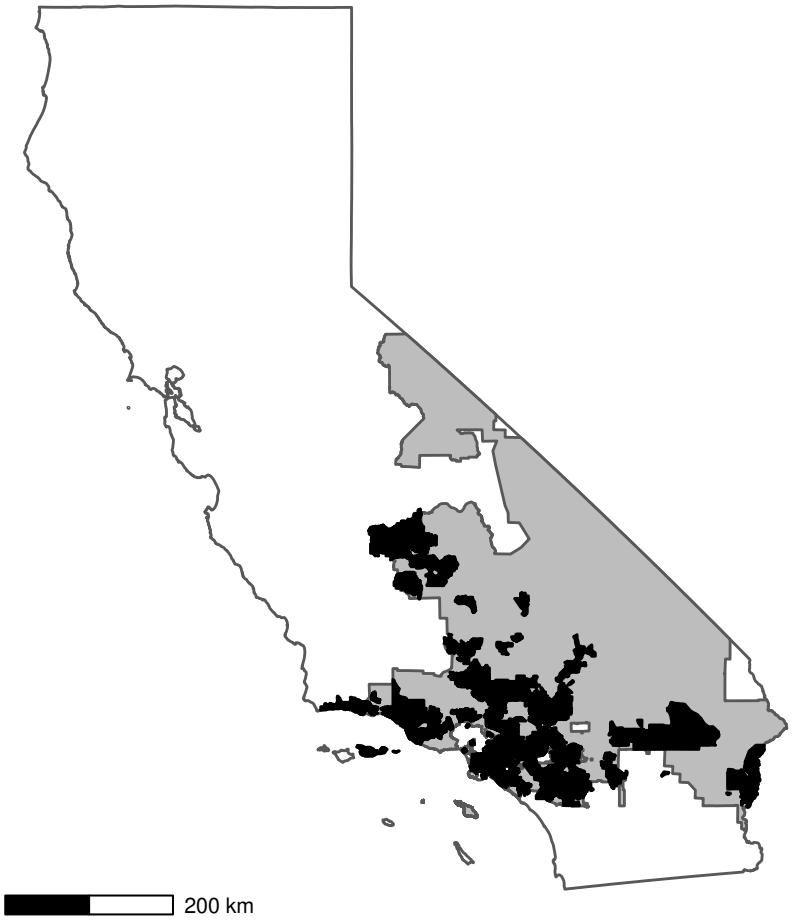


Figure A6: Map of SCE Service Area

The black area denotes the zip codes in SCE's service area that are in our sample. The gray area encompasses SCE's service area that is not in our sample. Shapefiles for this map come from the California Energy Commission's Electric Load Serving Entities (IOU and POU) map (https://cecgis-caenergy.opendata.arcgis.com/datasets/electric-load-serving-entities-iou-pou) and the 2018 Census Gazetteer (https://www.census.gov/geographies/mapping-files/time-series/geo/kml-cartographic-boundary-files.html).

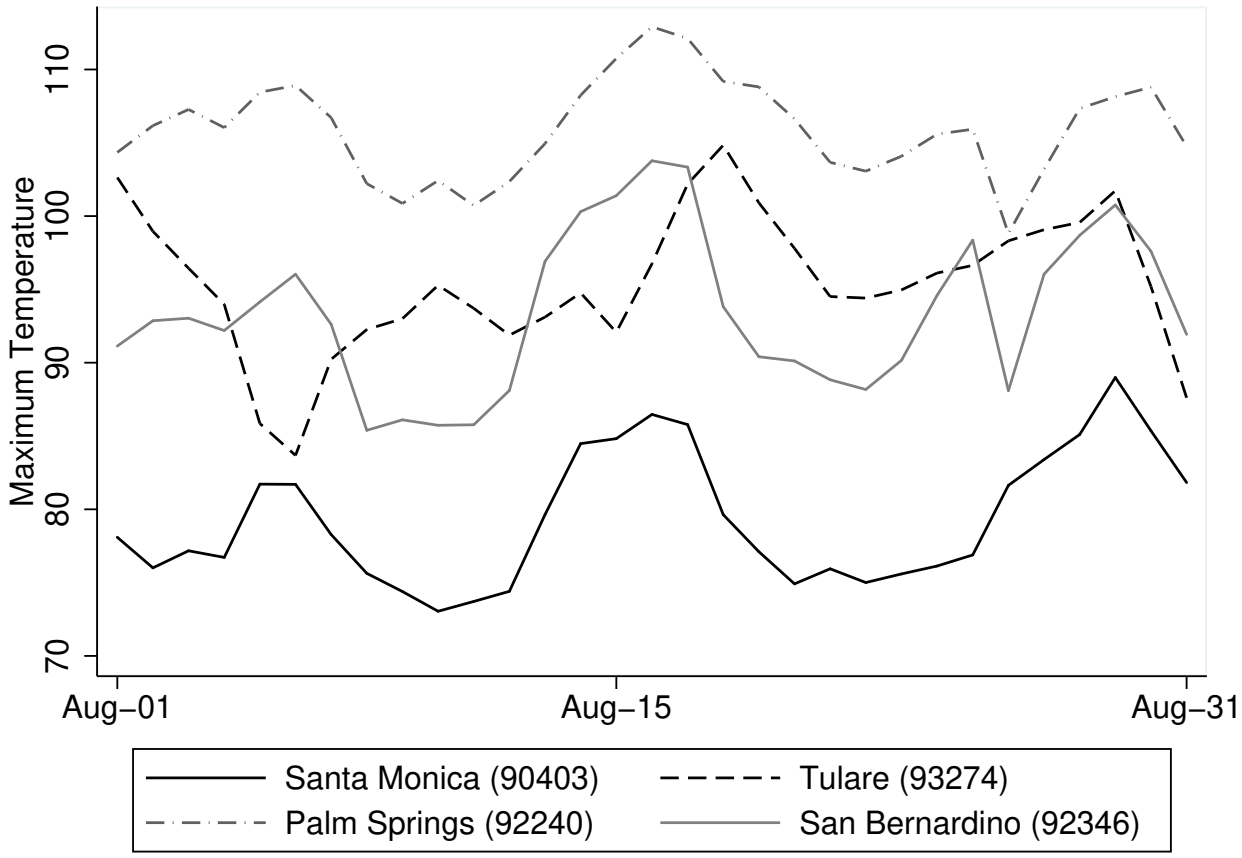


Figure A7: Temperatures for 4 Zip Codes, August 2015

This figure illustrates daily maximum temperatures using PRISM weather data for 4 zip codes in August 2015. Our full sample includes weather data for 300 zip codes from 2012 through 2017.

APPENDIX B

Appendix to Chapter 3

Additional Data Details

The raw NSS data include over 317,000 rural households. We drop around 1,700 households in India's far off islands (Lakshwadeep, Andoman and Nicobar), as well as any households in districts with very few observations (12 or fewer) or in districts that only show up in a single round. We also drop around 100 households with no survey date, 300 who do not list consumption of any food items, 180 missing core demographic information, and 800 with unusually high numbers of meals eaten away from home or of meals served to non-household members. We use district codes that are geographically consistent over time, originally constructed by Supreet Kaur. These codes take into account any changes in districts (e.g., a district split into two, or two districts combined into one), allowing one to create a panel of districts.

The NSS data also include the number of meals served to non-household members and the number of outside meals consumed. Following Carpena (2019), we assume that on average, meals served to others and acquired outside the house have similar diet composition to in-house consumption, and we adjust nutrient consumption accordingly. We use the adjustment factor $\frac{(N_h+N_a)}{(N_h+N_o)}$, where N_h is the number of meals household members have eaten at home in the last 30 days, and N_a and N_o are the same for meals away from home and meals given to non-household members.[1]

We determine each household member's caloric and iron needs by their age, gender, and occupation, and then add up the total requirements for the household. We assign the adults in each household to a particular "effort category," which affects caloric needs, based on the head of household's occupation (individual occupations are not included in these data).

We define our main rainfall control as the "rainfall shock." Following Shah and Millett Steinberg (2017), this takes on a value of 1 if rainfall in a given district-season-year exceeds the 80th percentile of its historical distribution (defined as 1979-2012) of rainfall for that

[1]Following Carpena (2019), we drop any households with adjustment factors above 2 or below 0.5. This results in dropping 817 households, or 0.26 percent of the sample.

district-season, and a value of -1 if rainfall is below the 20th percentile of the historical distribution (0 otherwise). We create three other measures of precipitation to use as controls across different model specifications. The first involves creating separate wet and dry shocks, which each take on a value of 1 if the rainfall is above the 80th or below the 20th percentile, respectively. The second measures the deviation above or below the long-run mean, capturing each separately. The third is total precipitation.

Weather and Agricultural Yields

We test whether hot weather impacts crop yields, the first step in the chain from heat to nutritional outcomes. We use agricultural data from the Village Dynamics in South Asia Meso dataset from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT 2015). The dataset includes price-weighted yield for the six major crops (rice, wheat, sugarcane, groundnut, sorghum, and maize), and the five major monsoon crops (excludes wheat), at the district-year level (Garg et al., 2020b). The data include yields from 1981 to 2011, the last year of the NSS data.

We match district-year level yields with district-year-season level weather variables for 1981-2011 to estimate the following model:

$$log(Y_{dy}) = \sum_{g=0}^{1}\sum_{j=1}^{J} \beta_g^j T_{d,y,g}^j + \sum_{g=0}^{1} \alpha_g P_{y,g} + \gamma_d + \tau_y + \epsilon_{dy}, \tag{B.1}$$

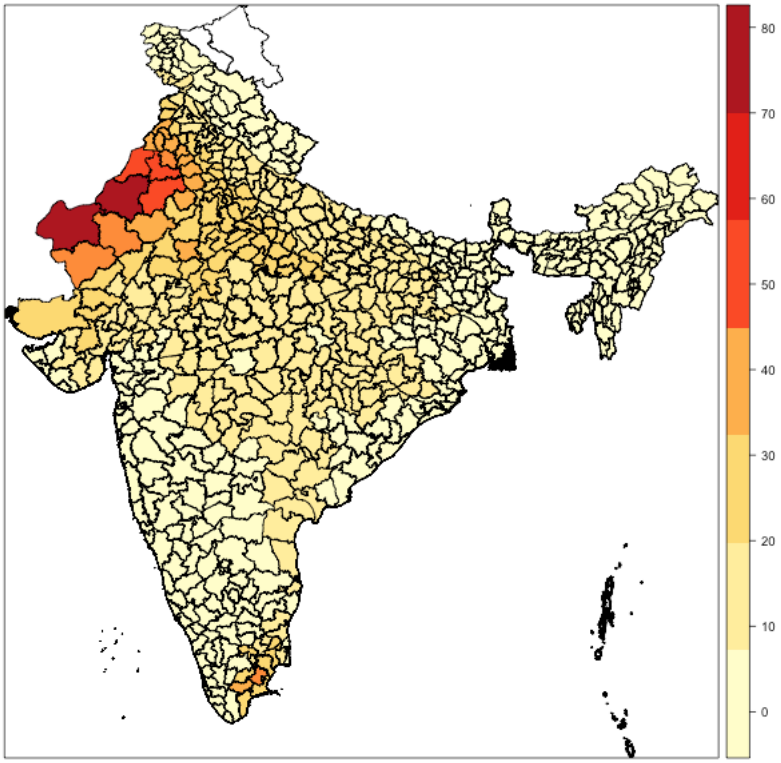
where Y is the yield for district d in year y.$T_{d,y,g}^j$ and $T_{d,y,ng}^j$ are the number of days in temperature bin j in district d during year y's growing ($g = 1$) and non-growing ($g = 0$) seasons. We include 10 degree temperature bins running from 70-80°F to above 100°F, and exclude the bin below 70°F. Therefore, all coefficients can be interpreted relative to a day below 70°F. We also include precipitation controls (P), district fixed effects (γ_d), and year fixed effects (τ_y). We cluster standard errors at the district level to account for any within-district correlation of the error terms. We estimate this regression for the yields of two

different groups of crops: 1) the top 6 crops (rice, wheat, sugarcane, groundnut, sorghum, and maize), and 2) the top 5 monsoon crops (which excludes wheat from 1)).

Figure B3 shows the effects of hot temperatures during the growing season on yields. Panel A shows results for the top six crops in India, and Panel B only includes monsoon crops. We multiply the log yield outcome variables by 100 so the coefficients can be interpreted as percentage changes. We find statistically significant decreases in overall yields for days above 70°F, with larger damages at higher temperatures. Panel A shows that a day above 110°F decreases yields for the top six crops by roughly 1 percent relative to a day below 70°F. This is 2.7 times larger than the effects of a day between 80 and 90°F. These results for the impact of growing season temperature are consistent with past literature (Garg et al., 2020b). Table B13 shows the coefficients for the non-growing season. Heat above 70°F negatively impacts crop yields, but these impacts do not increase with higher temperatures as sharply as during the growing season.

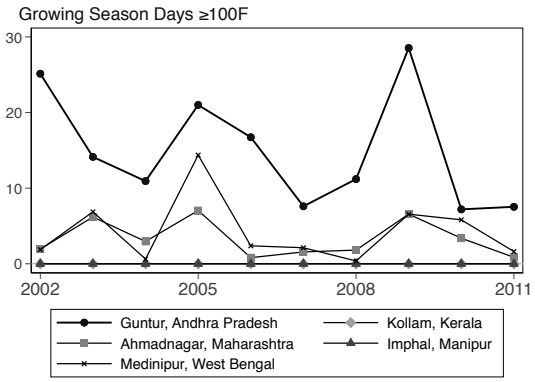
Appendix Figures

A) Yearly Growing Season Days with Tmax ≥ 100°F , 2002-2011



Days with Tmax ≥ 100°F by Year, Select Districts

B) Growing Season



C) Full Year

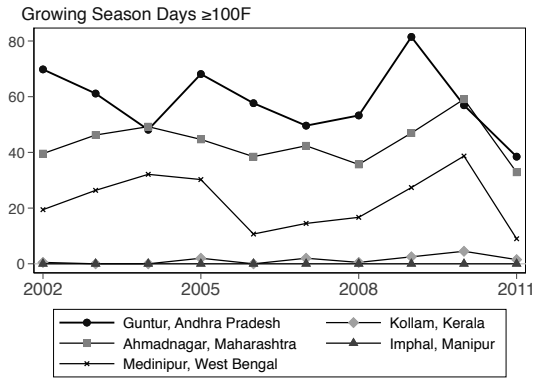


Figure B1: Days with Tmax ≥100°F by District (2002-2011)

Panel A shows the average number of days per growing season (June-December) with a maximum temperature of at least 100°F by district across India. Panel B shows the number of growing season Days ≥100°F by year from 2002 to 2011 for select districts with the highest number of surveyed households. Panel C shows the same but for the whole year. Temperature data are from ERA5.

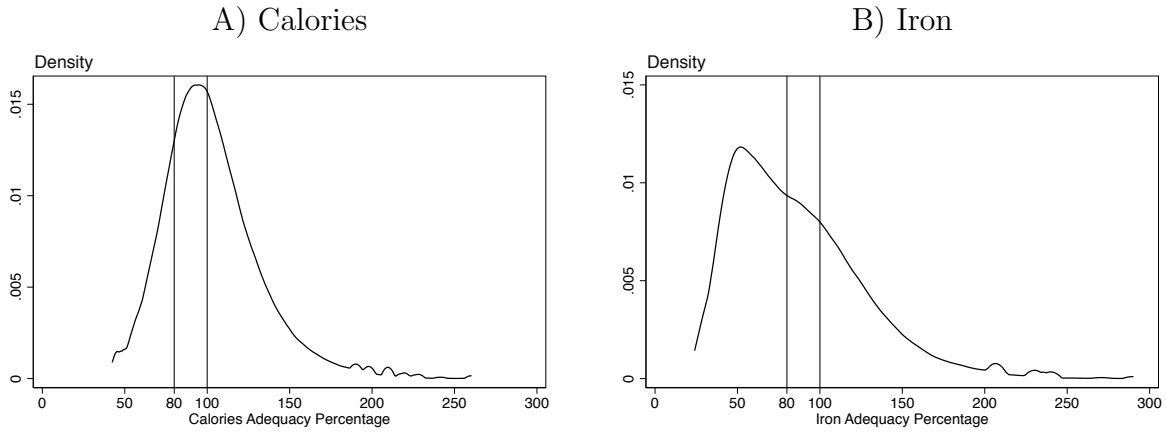


Figure B2: Calorie and Iron Adequacy Distribution

Panel A shows the distribution of calorie adequacy across households in the sample. Panel B shows the distribution of iron adequacy across households in the sample. The cutoffs at 80 percent and 100 percent adequacy correspond to the cutoffs for our two main outcome variables from Figure 3.3.

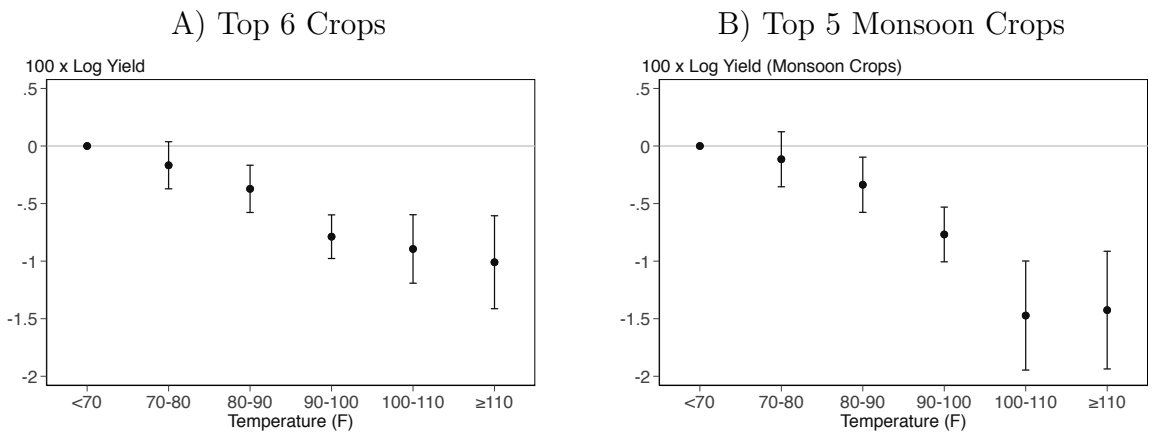


Figure B3: Yields

These graphs show the change in 100 times log yields for each additional day at a given temperature during this year's growing season relative to a day below 70°F. Panel A is for the top 6 crops in India (rice, wheat, sugarcane, groundnut, sorghum, and maize). Panel B is for the top 5 monsoon-dependent crops (excludes wheat). The coefficients can be interpreted as percentage changes. We match district-year yields with weather data at the district-year-season level. The model includes 10°F temperature bins from 70-80°F to above 110°F for this year's growing season (June - December) and non-growing season (March - May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. The models include district and year fixed effects and rainfall shocks. Standard errors (errors bars represent the 95 percent confidence interval) are clustered at the district level to account for any inter-district correlation of the error terms. The models include data from 1981-2011. The model in Panel A includes 9,141 observations, and Panel B includes 9,138. The mean outcome variable in panel A is 507.88, and the mean outcome variable in panel B is 496.74.

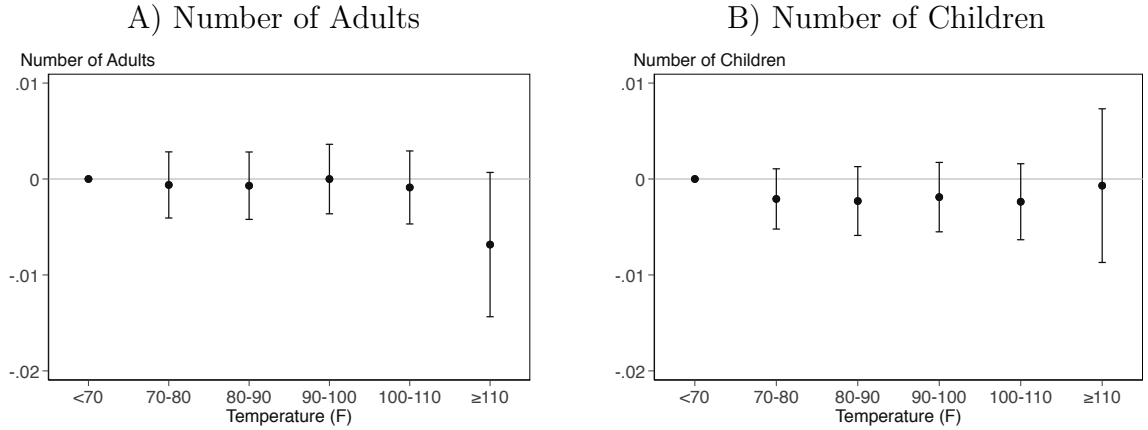


Figure B4: Household Composition

Panel A shows the change in 100 times the number of adults for each additional day at a given temperature during last year's growing season relative to a day below 70°F. Panel B shows the same relationship for the number of children. We match household-year-month data with weather data at the district-year-season level. Both models include 10°F temperature bins from 70-80°F to above 110°F for this year and last year's growing season (June - December) and non-growing season (March - May). They omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. The models include district, month, and year fixed effects, rainfall shocks, and demographic controls (religion, social group, and head of household education). Standard errors (error bars represent the 95% confidence interval) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. The model in panel A includes 314,425 observations, and panel B includes 314,425.

Table B1: Additional Rainfall Controls and Non-Growing Season Weather

	Mean	SD
Dry growing season (dummy)	0.16	0.37
Wet growing season (dummy)	0.27	0.44
Growing season dry deviation (m)	0.07	0.12
Growing season wet deviation (m)	0.13	0.18
Non-growing season days $\geq$ 110°F	1.15	2.94
Non-growing season days 100-110°F	23.40	20.05
Non-growing season days 90-100°F	37.25	19.71
Non-growing season days 80-90°F	19.57	18.08
Non-growing season days 70-80°F	5.91	12.02
Non-growing season days <70°F	4.72	16.96
Non-growing season rainfall (m)	0.22	0.30
Non-growing season rainfall shock	0.02	0.62
Dry non-growing season (dummy)	0.18	0.39
Wet non-growing season (dummy)	0.21	0.40
Non-growing season dry deviation (m)	0.02	0.05
Non-growing season wet deviation (m)	0.04	0.09

This table shows means and standard deviations of alternative precipitation controls and non-growing season (March to May) weather variables from 2002 to 2011. Weather variables are from ERA5 and are separated by growing season (June to December) and non-growing season (March to May). The dry season dummy represents a season with less than the 20th historical (from 1979-2012) percentile of rainfall for a particular district. The wet season dummy indicates rainfall above the 80th percentile. The rainfall shock takes a value of -1 for a dry season, 1 for a wet season, and 0 otherwise. The dry deviation represents the amount of rainfall (in m) below the historical median (and is 0 in a wet season). The wet deviation is analogous for above median rainfall.

Table B2: Additional Calorie and Iron Adequacy Thresholds

	Mean	SD
CAP < 150% (dummy)	93.12	25.32
CAP < 140% (dummy)	89.73	30.36
CAP < 130% (dummy)	84.49	36.20
CAP < 120% (dummy)	76.81	42.21
CAP < 110% (dummy)	65.82	47.43
CAP < 100% (dummy)	51.51	49.98
CAP < 90% (dummy)	35.49	47.85
CAP < 80% (dummy)	20.64	40.48
CAP < 70% (dummy)	10.05	30.07
CAP < 60% (dummy)	3.98	19.54
CAP < 50% (dummy)	1.41	11.78
IAP < 150% (dummy)	92.74	25.95
IAP < 140% (dummy)	90.06	29.92
IAP < 130% (dummy)	86.39	34.29
IAP < 120% (dummy)	81.49	38.84
IAP < 110% (dummy)	75.35	43.10
IAP < 100% (dummy)	67.90	46.69
IAP < 90% (dummy)	59.48	49.09
IAP < 80% (dummy)	50.36	50.00
IAP < 70% (dummy)	40.64	49.12
IAP < 60% (dummy)	29.86	45.76
IAP < 50% (dummy)	18.15	38.54

This table shows means and standard deviations of our calorie and iron threshold outcome variables from 2003 to 2012. CAP stands for calorie adequacy percentage, and IAP stands for iron adequacy percentage.

Table B3: District-Month and Year-Month Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	-0.056	-0.277*	0.208	0.456***	0.245	0.427***	-5.513*	5.381**	-0.030	0.009
	(0.114)	(0.145)	(0.203)	(0.141)	(0.179)	(0.140)	(2.837)	(2.646)	(0.025)	(0.025)
Days 100-110°F	-0.043	-0.159**	0.089	0.268***	0.110	0.153**	-2.848**	2.106	-0.028**	0.007
	(0.056)	(0.075)	(0.097)	(0.080)	(0.081)	(0.075)	(1.417)	(1.522)	(0.012)	(0.014)
Days 90-100°F	-0.050	-0.202***	0.097	0.240***	0.150**	0.178***	-2.400*	1.371	-0.026**	-0.007
	(0.052)	(0.069)	(0.088)	(0.073)	(0.074)	(0.068)	(1.383)	(1.424)	(0.011)	(0.011)
Days 80-90°F	-0.042	-0.196***	0.102	0.230***	0.135*	0.143**	-2.180	1.281	-0.025**	-0.007
	(0.050)	(0.068)	(0.084)	(0.072)	(0.071)	(0.067)	(1.384)	(1.418)	(0.011)	(0.011)
Days 70-80°F	-0.011	-0.138**	0.050	0.122*	0.106	0.066	-1.327	1.036	-0.019*	-0.006
	(0.048)	(0.063)	(0.079)	(0.066)	(0.069)	(0.062)	(1.302)	(1.231)	(0.010)	(0.010)
Rainfall Shock	-0.035	0.408	-0.201	-0.039	-0.181	-0.218	-1.263	2.713	-0.010	0.088*
	(0.245)	(0.314)	(0.418)	(0.332)	(0.323)	(0.342)	(5.321)	(5.525)	(0.044)	(0.053)
Observations	314,415	313,781	314,416	314,416	314,416	314,416	314,416	314,416	314,416	314,416
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District-Month FE	X	X	X	X	X	X	X	X	X	X
Year-Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district-month and year-month fixed effects (as opposed to Table 3.2, which includes district, month, and year fixed effects), demographic controls, rainfall shocks, and state time trends. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B4: District-Specific Year Trends

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.010	-0.285**	0.224	0.381***	0.179	0.324**	-2.798	3.873	-0.010	-0.016
	(0.113)	(0.138)	(0.196)	(0.138)	(0.178)	(0.136)	(2.877)	(2.522)	(0.025)	(0.024)
Days 100-110°F	-0.029	-0.219***	0.068	0.221***	0.082	0.141**	-2.501*	2.067	-0.031***	0.002
	(0.050)	(0.067)	(0.088)	(0.072)	(0.074)	(0.067)	(1.342)	(1.352)	(0.011)	(0.013)
Days 90-100°F	-0.027	-0.215***	0.053	0.194***	0.107	0.158***	-2.465*	1.962	-0.029***	-0.002
	(0.046)	(0.062)	(0.081)	(0.065)	(0.067)	(0.061)	(1.311)	(1.302)	(0.010)	(0.011)
Days 80-90°F	-0.029	-0.208***	0.084	0.194***	0.093	0.132**	-2.166*	1.562	-0.027***	-0.004
	(0.044)	(0.059)	(0.077)	(0.063)	(0.064)	(0.058)	(1.302)	(1.274)	(0.010)	(0.010)
Days 70-80°F	-0.019	-0.146**	0.048	0.124**	0.056	0.064	-1.459	1.062	-0.020**	-0.002
	(0.044)	(0.057)	(0.073)	(0.059)	(0.063)	(0.056)	(1.260)	(1.134)	(0.009)	(0.009)
Rainfall Shock	-0.013	0.135	-0.378	-0.345	-0.148	-0.271	-3.852	5.145	-0.063	0.086*
	(0.231)	(0.312)	(0.392)	(0.314)	(0.305)	(0.338)	(5.265)	(5.374)	(0.043)	(0.052)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
District time trends	X	X	X	X	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year, and month fixed effects, demographic controls, and rainfall shocks. In contrast to our main estimates (Table 3.2), these models include district-specific year trends, as opposed to state-specific year trends. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B5: Remove This Year's Weather

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	-0.009	-0.213*	0.101	0.276**	0.164	0.235**	-2.820	2.903	-0.012	-0.009
	(0.087)	(0.113)	(0.153)	(0.112)	(0.136)	(0.112)	(2.148)	(2.194)	(0.019)	(0.020)
Days 100-110°F	-0.035	-0.135**	0.107	0.210***	0.048	0.092	-2.240*	1.774	-0.023**	0.008
	(0.045)	(0.063)	(0.079)	(0.067)	(0.067)	(0.060)	(1.243)	(1.319)	(0.011)	(0.012)
Days 90-100°F	-0.045	-0.164***	0.099	0.182***	0.086	0.112**	-2.154*	1.349	-0.024***	0.000
	(0.042)	(0.060)	(0.072)	(0.062)	(0.062)	(0.056)	(1.211)	(1.263)	(0.009)	(0.010)
Days 80-90°F	-0.043	-0.162***	0.115*	0.180***	0.073	0.087	-1.878	1.060	-0.022**	-0.002
	(0.041)	(0.058)	(0.069)	(0.060)	(0.060)	(0.055)	(1.205)	(1.237)	(0.009)	(0.009)
Days 70-80°F	-0.017	-0.120**	0.069	0.099*	0.045	0.021	-1.115	0.833	-0.018**	-0.001
	(0.041)	(0.056)	(0.068)	(0.056)	(0.060)	(0.053)	(1.192)	(1.104)	(0.009)	(0.008)
Rainfall Shock	-0.138	0.299	-0.068	-0.137	-0.107	-0.244	-0.730	-0.416	-0.019	0.074
	(0.211)	(0.281)	(0.361)	(0.285)	(0.285)	(0.304)	(4.625)	(4.977)	(0.039)	(0.048)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. In contrast to our main estimates (Table 3.2), these models do not include this year's weather. Each model includes district, year, and month fixed effects, demographic controls, rainfall shocks, and state time trends. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B6: Dry and Wet Shocks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.020	-0.187	0.159	0.356***	0.181	0.337**	-4.388	5.669**	-0.022	0.017
	(0.107)	(0.141)	(0.187)	(0.134)	(0.163)	(0.138)	(2.677)	(2.563)	(0.023)	(0.024)
Days 100-110°F	-0.033	-0.150**	0.082	0.252***	0.091	0.143**	-2.501*	2.008	-0.027**	0.009
	(0.053)	(0.072)	(0.092)	(0.078)	(0.077)	(0.071)	(1.371)	(1.479)	(0.011)	(0.013)
Days 90-100°F	-0.054	-0.214***	0.103	0.243***	0.151**	0.185***	-2.080	0.989	-0.026**	-0.007
	(0.049)	(0.067)	(0.085)	(0.071)	(0.071)	(0.066)	(1.333)	(1.398)	(0.010)	(0.011)
Days 80-90°F	-0.057	-0.221***	0.126	0.242***	0.141**	0.165**	-1.822	0.607	-0.026***	-0.009
	(0.048)	(0.065)	(0.082)	(0.070)	(0.068)	(0.065)	(1.320)	(1.375)	(0.010)	(0.011)
Days 70-80°F	-0.031	-0.168***	0.083	0.149**	0.107	0.095	-1.012	0.350	-0.020**	-0.008
	(0.046)	(0.061)	(0.075)	(0.065)	(0.066)	(0.060)	(1.250)	(1.181)	(0.009)	(0.009)
Dry Shock	0.684**	0.531	-1.054*	-0.831*	-0.185	-0.645	-0.227	10.619	0.027	0.002
	(0.332)	(0.437)	(0.575)	(0.433)	(0.461)	(0.482)	(7.034)	(7.974)	(0.060)	(0.075)
Wet Shock	0.449	1.180***	-1.121**	-0.895**	-0.441	-1.038**	-0.606	10.276	0.012	0.142**
	(0.314)	(0.412)	(0.529)	(0.437)	(0.408)	(0.434)	(6.610)	(7.241)	(0.054)	(0.070)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Dry and Wet shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year and month fixed effects, demographic controls, and state time trends. As compared to our main specification (Table 3.2), these models include separate dry and wet shocks rather than a single rainfall shock. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B7: Dry and Wet Deviations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.028	-0.217	0.156	0.340***	0.175	0.369***	-4.050	5.483**	-0.019	0.011
	(0.106)	(0.139)	(0.185)	(0.131)	(0.165)	(0.136)	(2.651)	(2.469)	(0.023)	(0.023)
Days 100-110°F	-0.042	-0.174**	0.119	0.264***	0.097	0.167**	-2.348*	1.642	-0.025**	0.004
	(0.052)	(0.069)	(0.088)	(0.075)	(0.074)	(0.068)	(1.370)	(1.447)	(0.011)	(0.013)
Days 90-100°F	-0.063	-0.233***	0.136*	0.252***	0.157**	0.211***	-1.858	0.573	-0.024**	-0.012
	(0.048)	(0.065)	(0.082)	(0.069)	(0.069)	(0.064)	(1.348)	(1.379)	(0.010)	(0.011)
Days 80-90°F	-0.064	-0.237***	0.150*	0.250***	0.149**	0.185***	-1.661	0.279	-0.024**	-0.014
	(0.048)	(0.065)	(0.080)	(0.069)	(0.067)	(0.064)	(1.343)	(1.365)	(0.010)	(0.011)
Days 70-80°F	-0.038	-0.185***	0.103	0.160**	0.116*	0.110*	-0.886	0.049	-0.019**	-0.012
	(0.046)	(0.060)	(0.075)	(0.064)	(0.065)	(0.060)	(1.270)	(1.184)	(0.009)	(0.009)
Dry Deviation	3.437***	4.290***	-6.698***	-2.602*	-1.972*	-3.845***	-20.039	83.994***	-0.008	0.458**
	(1.030)	(1.428)	(1.881)	(1.423)	(1.194)	(1.413)	(23.411)	(26.187)	(0.158)	(0.205)
Wet Deviation	1.643**	4.005***	-3.113**	-2.315**	-2.113**	-2.949***	-10.306	47.156**	0.004	0.575***
	(0.817)	(1.094)	(1.415)	(1.169)	(0.983)	(1.090)	(17.922)	(18.853)	(0.152)	(0.173)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Dry, Wet deviations	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year and month fixed effects, demographic controls, and state time trends. As compared to our main specification (Table 3.2), these models include separate dry and wet deviations rather than a single rainfall shock. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B8: Total Rainfall

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.021	-0.223	0.168	0.354***	0.178	0.373***	-4.061	5.370**	-0.020	0.010
	(0.106)	(0.139)	(0.185)	(0.131)	(0.165)	(0.136)	(2.650)	(2.485)	(0.023)	(0.023)
Days 100-110°F	-0.043	-0.180***	0.123	0.262***	0.099	0.171**	-2.349*	1.605	-0.025**	0.003
	(0.051)	(0.069)	(0.088)	(0.075)	(0.074)	(0.067)	(1.363)	(1.432)	(0.011)	(0.013)
Days 90-100°F	-0.056	-0.226***	0.125	0.242***	0.152**	0.203***	-1.934	0.773	-0.024**	-0.011
	(0.048)	(0.065)	(0.082)	(0.070)	(0.069)	(0.064)	(1.331)	(1.365)	(0.010)	(0.011)
Days 80-90°F	-0.056	-0.225***	0.135*	0.240***	0.143**	0.173***	-1.772	0.558	-0.024**	-0.012
	(0.048)	(0.065)	(0.080)	(0.069)	(0.067)	(0.063)	(1.325)	(1.353)	(0.010)	(0.011)
Days 70-80°F	-0.028	-0.171***	0.085	0.148**	0.109*	0.098	-0.951	0.303	-0.019**	-0.010
	(0.046)	(0.061)	(0.075)	(0.064)	(0.065)	(0.060)	(1.254)	(1.173)	(0.009)	(0.009)
Total Rainfall	-0.250	0.938	0.531	-0.466	-0.596	-0.425	1.109	-1.750	0.004	0.191
	(0.568)	(0.776)	(0.946)	(0.796)	(0.702)	(0.768)	(13.641)	(13.981)	(0.108)	(0.121)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Total Rainfall	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year and month fixed effects, demographic controls, and state time trends. As compared to our main specification (Table 3.2), these models include total rainfall rather than a rainfall shock. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B9: Children Weighted as $\frac{1}{2}$ Adults

	(1)	(2)	(3)	(4)	(5)	(6)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.028	-0.188	-4.784	6.783**	-0.022	0.021
	(0.107)	(0.141)	(3.366)	(3.185)	(0.029)	(0.029)
Days 100-110°F	-0.040	-0.161**	-2.981*	2.249	-0.033**	0.010
	(0.053)	(0.071)	(1.690)	(1.784)	(0.014)	(0.016)
Days 90-100°F	-0.053	-0.212***	-2.460	1.174	-0.032**	-0.009
	(0.048)	(0.066)	(1.644)	(1.684)	(0.013)	(0.014)
Days 80-90°F	-0.056	-0.217***	-2.221	0.760	-0.032***	-0.011
	(0.048)	(0.065)	(1.632)	(1.659)	(0.012)	(0.013)
Days 70-80°F	-0.030	-0.166***	-1.297	0.452	-0.026**	-0.010
	(0.045)	(0.060)	(1.558)	(1.451)	(0.011)	(0.012)
Rainfall Shock	-0.087	0.415	0.324	-0.432	-0.003	0.094
	(0.224)	(0.300)	(5.998)	(6.412)	(0.050)	(0.060)
Observations	314,424	313,790	314,425	314,425	314,425	314,425
Mean Outcome	783.15	287.25	733.97	1,860.94	5.37	13.98
District FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Month FE	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
State time trends	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 3-6 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year, and month fixed effects, demographic controls, rainfall shocks, and state time trends. When calculating per-capita outcomes, we weight children as one half of an adult, as opposed to a full adult in our main specification (Table 3.2). Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B10: Children Weighted as $\frac{1}{3}$ Adults

	(1)	(2)	(3)	(4)	(5)	(6)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Days ≥110°F	0.031	-0.183	-5.157	7.470**	-0.024	0.026
	(0.108)	(0.141)	(3.737)	(3.546)	(0.033)	(0.033)
Days 100-110°F	-0.042	-0.164**	-3.261*	2.427	-0.037**	0.010
	(0.053)	(0.071)	(1.856)	(1.960)	(0.015)	(0.018)
Days 90-100°F	-0.056	-0.215***	-2.701	1.222	-0.036***	-0.011
	(0.048)	(0.066)	(1.805)	(1.850)	(0.014)	(0.015)
Days 80-90°F	-0.059	-0.220***	-2.443	0.762	-0.036***	-0.013
	(0.048)	(0.065)	(1.793)	(1.823)	(0.013)	(0.014)
Days 70-80°F	-0.034	-0.171***	-1.444	0.407	-0.030**	-0.012
	(0.045)	(0.060)	(1.717)	(1.607)	(0.013)	(0.013)
Rainfall Shock	-0.093	0.410	0.921	-1.284	0.001	0.101
	(0.225)	(0.300)	(6.550)	(7.017)	(0.054)	(0.066)
Observations	314,424	313,790	314,425	314,425	314,425	314,425
Mean Outcome	791.11	295.21	800.19	2,026.74	5.86	15.26
District FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
Month FE	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X
State time trends	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's growing season relative to a day below 70°F (Equation 1). The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 3-6 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year, and month fixed effects, demographic controls, rainfall shocks, and state time trends. When calculating per-capita outcomes, we weight children as one third of an adult, as opposed to a full adult in our main specification (Table 3.2). Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B11: Main Specification: Last Year's Non-Growing Season Coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
Non-GS Days ≥110°F	-0.004	-0.124	0.034	-0.233**	0.249*	0.041	-3.723	3.224	-0.041**	0.018
	(0.089)	(0.130)	(0.161)	(0.118)	(0.129)	(0.127)	(2.302)	(2.210)	(0.019)	(0.021)
Non-GS Days 100-110°F	0.017	0.042	0.041	-0.172*	0.062	-0.047	-1.150	1.291	-0.002	0.003
	(0.069)	(0.099)	(0.129)	(0.092)	(0.095)	(0.102)	(1.948)	(1.851)	(0.014)	(0.017)
Non-GS Days 90-100°F	0.013	0.049	0.023	-0.168**	0.056	-0.096	-0.624	0.653	-0.000	0.004
	(0.063)	(0.089)	(0.121)	(0.085)	(0.087)	(0.092)	(1.804)	(1.724)	(0.013)	(0.016)
Non-GS Days 80-90°F	-0.034	0.016	0.103	-0.120	0.093	-0.045	-1.318	0.434	-0.005	0.002
	(0.059)	(0.083)	(0.117)	(0.083)	(0.079)	(0.087)	(1.710)	(1.592)	(0.013)	(0.015)
Non-GS Days 70-80°F	0.006	0.013	0.116	-0.048	0.044	-0.061	0.931	-0.748	0.025*	-0.023
	(0.065)	(0.094)	(0.126)	(0.085)	(0.080)	(0.087)	(2.132)	(1.960)	(0.014)	(0.016)
Non-GS Rainfall Shock	-0.159	-0.133	0.436	0.432	0.265	0.206	2.153	-5.947	0.044	-0.062
	(0.235)	(0.324)	(0.378)	(0.331)	(0.292)	(0.321)	(5.146)	(5.758)	(0.042)	(0.052)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during last year's non-growing season relative to a day below 70°F (Equation 1). It is the same specification as Table 3.2, which shows last year's growing season weather. The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year, and month fixed effects, demographic controls, rainfall shocks, and state time trends. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B12: Main Specification: This Year's Growing Season Coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	100x Log of Daily per Capita Calories	100x Log of Daily per Capita Iron	Percent of Households Below 100% Calories Adequacy	Percent of Households Below 80% Calories Adequacy	Percent of Households Below 100% Iron Adequacy	Percent of Households Below 80% Iron Adequacy	Daily Calories per Capita, Home-Grown	Daily Calories per Capita, Purchased	Daily Iron per per Capita, Home-Grown	Daily Iron per Capita, Purchased
TY Days ≥110°F	-0.027	0.104	0.202	0.159	-0.005	0.059	-5.135**	5.160**	-0.054***	0.078***
	(0.090)	(0.122)	(0.153)	(0.119)	(0.150)	(0.115)	(2.349)	(2.383)	(0.020)	(0.020)
TY Days 100-110°F	-0.039	0.073	0.038	0.068	-0.007	-0.085	-3.544**	2.469	-0.042***	0.044***
	(0.063)	(0.083)	(0.107)	(0.082)	(0.074)	(0.072)	(1.385)	(1.621)	(0.011)	(0.014)
TY Days 90-100°F	-0.074	-0.027	0.109	0.110*	0.039	-0.024	-2.720**	0.760	-0.037***	0.023*
	(0.052)	(0.071)	(0.093)	(0.066)	(0.066)	(0.060)	(1.226)	(1.426)	(0.009)	(0.012)
TY Days 80-90°F	-0.067	-0.033	0.095	0.095	0.019	-0.029	-2.309*	0.492	-0.035***	0.021*
	(0.050)	(0.069)	(0.089)	(0.063)	(0.064)	(0.059)	(1.209)	(1.388)	(0.009)	(0.012)
TY Days 70-80°F	-0.100**	-0.062	0.149*	0.112*	0.037	0.015	-2.153*	-0.373	-0.038***	0.016
	(0.047)	(0.066)	(0.082)	(0.059)	(0.063)	(0.058)	(1.163)	(1.313)	(0.009)	(0.011)
TY Rainfall Shock	0.473**	0.899***	-0.481	-0.012	-0.628**	-0.606**	1.484	9.556*	0.056	0.087*
	(0.202)	(0.270)	(0.349)	(0.293)	(0.273)	(0.266)	(4.997)	(5.667)	(0.041)	(0.047)
Observations	314,424	313,790	314,425	314,425	314,425	314,425	314,425	314,425	314,425	314,425
Mean Outcome	763.87	267.98	51.51	20.64	67.90	50.36	602.70	1,538.69	4.39	11.53
District FE	X	X	X	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X	X	X	X
Month FE	X	X	X	X	X	X	X	X	X	X
Demographic controls	X	X	X	X	X	X	X	X	X	X
State time trends	X	X	X	X	X	X	X	X	X	X
Rainfall shocks	X	X	X	X	X	X	X	X	X	X

This table shows the change in our main outcome variables for an additional day at a given temperature during this year's (TY) growing season relative to a day below 70°F (Equation 1). It is the same specification as Table 3.2, which shows last year's growing season weather. The outcomes in columns 1 and 2 are multiplied by 100 so the coefficients can be interpreted as percentage changes. The outcomes in columns 7-10 are in levels due to the high number of households who do not consume home-grown foods. The models all include 10°F temperature bins from 70-80°F to above 110°F for this and last year's growing season (June - December) and non-growing season (March-May). We omit the <70°F bins so the coefficients can be interpreted relative to this temperature range. Each model includes district, year, and month fixed effects, demographic controls, rainfall shocks, and state time trends. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Table B13: This Year Growing vs. Non-Growing Seasons: Yields

	(1)	(2)
	100 x Log of yield (top 6 crops)	100 x Log of yield (top 5 monsoon crops)
GS Days ≥110°F	-1.010***	-1.425***
	(0.205)	(0.260)
GS days 100-110°F	-0.894***	-1.473***
	(0.151)	(0.241)
GS days 90-100°F	-0.788***	-0.768***
	(0.096)	(0.121)
GS days 80-90°F	-0.372***	-0.337***
	(0.104)	(0.122)
GS days 70-80°F	-0.167	-0.115
	(0.104)	(0.121)
GS Rainfall shock	1.762***	0.912
	(0.500)	(0.606)
Non GS Days ≥110°F	-0.690***	-0.748***
	(0.255)	(0.239)
Non GS days 100-110°F	-0.815***	-0.627***
	(0.224)	(0.180)
Non GS days 90-100°F	-0.839***	-0.704***
	(0.223)	(0.182)
Non GS days 80-90°F	-0.858***	-0.578***
	(0.220)	(0.159)
Non GS days 70-80°F	-0.711***	-0.434**
	(0.238)	(0.176)
Non GS Rainfall Shock	-1.520***	-2.891***
	(0.540)	(0.879)
Observations	9,141	9,138
Mean Outcome	507.88	496.74
District, Year FE	X	X
Demographic controls	X	X
State time trends	X	X
Rainfall shocks	X	X

This table shows the change in (100x) the log of yield for the top 6 crops in India (column 1 - rice, wheat, sugarcane, groundnut, sorghum, and maize), and the top 5 monsoon crops (column 2 - excludes wheat) for an additional day at a given temperature during this year's growing and non growing seasons relative to a day below 70°F on the log of yield. The models all include 10°F temperature bins from 70 to 80°F to above 110°F for last year's growing seasons (June - December). Each model includes district and year fixed effects, and rainfall shocks. Standard errors (shown in parentheses) are clustered at the district level to account for any inter-district correlation of the error terms. All regressions include household-level outcomes from 2003-2012. *p<0.1 **p<0.05 ***p<0.01

Climate Projections

Table B14: Summary Statistics: CMIP6

(1) Model Name	(2) Number of Members	(3) Latest End Year	(4) Includes r1i1p1f1
ACCESS-CM2	3	2100	Yes
AWI-CM-1-1-MR	5	2100	Yes
CESM2	3	2100	No
CESM2-WACCM	2	2100	Yes
CMCC-CM2-SR5	1	2100	Yes
CMCC-ESM2	1	2100	Yes
CNRM-CM6-1	6	2100	No
CNRM-CM6-1-HR	1	2100	No
CNRM-ESM2-1	5	2100	No
CanESM5	50	2100	Yes
EC-Earth3	57	2100	Yes
EC-Earth3-AerChem	2	2100	Yes
EC-Earth3-Veg	3	2100	No
EC-Earth3-Veg-LR	3	2100	Yes
FGOALS-g3	3	2100	Yes
GFDL-ESM4	1	2100	Yes
INM-CM4-8	1	2100	Yes
INM-CM5-0	5	2100	Yes
IPSL-CM6A-LR	11	2100	Yes
MIROC-ES2L	1	2100	No
MIROC6	3	2100	Yes
MPI-ESM1-2-HR	10	2100	Yes
MPI-ESM1-2-LR	10	2100	Yes
MRI-ESM2-0	5	2100	Yes
NorESM2-LM	3	2100	Yes
NorESM2-MM	1	2100	Yes
BCC-ESM1	3	2055	Yes
CAMS-CSM1-0	1	2099	No
MPI-ESM-1-2-HAM	3	2055	Yes
IITM-ESM	1	2098	Yes

This table includes summary statistics for each model used in this analysis. All models are at the daily level and begin at least as early as 2000. Column 1 lists the model name and column 2 includes the number of members on Pangeo at the time of download (February 2024). Column 3 lists the latest final year of data across the members within that model. Note that not all members of the same model end in the same year. For example, for NorESM2-LM, r2i1p1f1 ends in 2054, but r1i1p1f1 ends in 2100. Column 4 lists whether that model includes an r1i1p1f1 member.

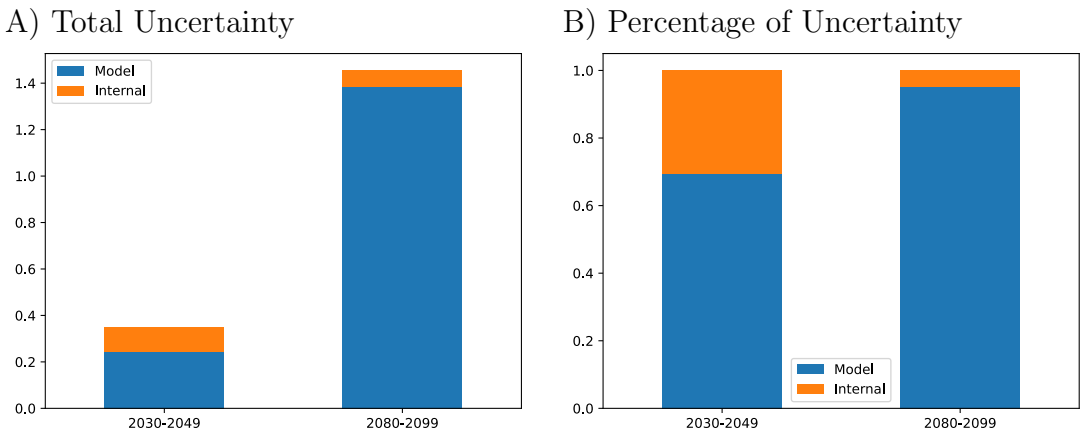
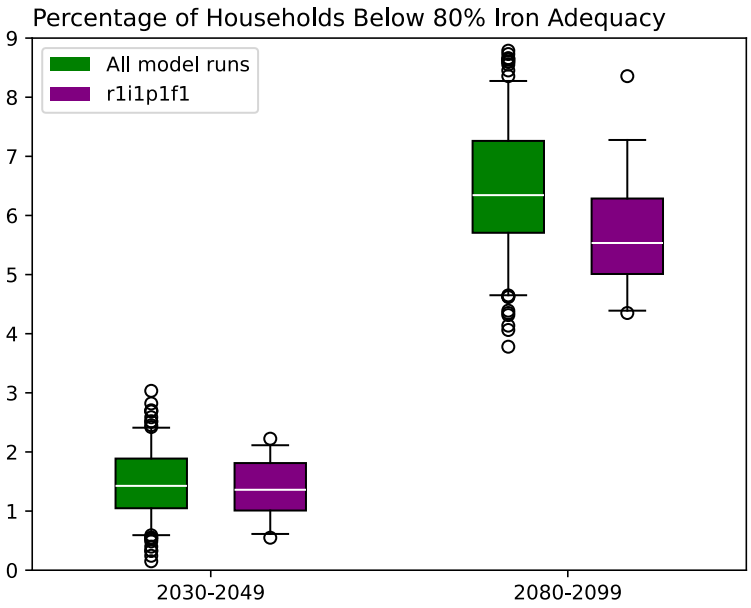


Figure B5: Climate Projection Uncertainty: Models with Over 5 Runs

This figure graphs the contributions to the uncertainty in the warming-induced changes, relative to 2000 to 2019, in the percentage of households below the 80% adequacy thresholds for iron and calories. Panel A graphs uncertainty in levels, with uncertainty for estimates for the 2030 to 2049 period in the bar on the left, and for the 2080 to 2099 period on the right. Panel B graphs the percentage of uncertainty contributed by each source. These calculations include the 9 models with at least 5 members for daily temperature in the SSP3-7.0 projection scenario: AWI-CM-1-1-MR, CanESM5, EC-Earth3, INM-CM5-0, IPSL-CM6A-LR, MPI-EMS1-2-HR, MPI-ESM1-2-LR, and MRI-ESM2-0. The model uncertainty is calculated as the variance across the model means, and the internal variability is calculated as the mean of the within-model (across members) variance.

A) Iron



B) Calories

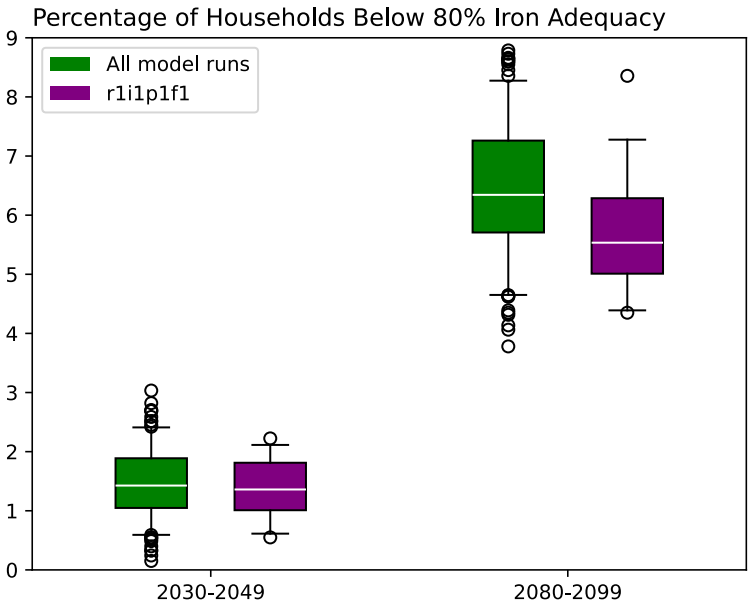


Figure B6: Climate Projections: Mid and Late Century

This figure shows the spread by model in the warming-induced changes, relative to 2000 to 2019, in the percentage of households below the 80% adequacy thresholds for iron (Panel A) and calories (Panel B). The two box-and-whisker plots to the left are for the period 2030 to 2049, and the two to the right are for the period 2080 to 2099. The projections are calculated for each ensemble member by multiplying its projected change in the temperature distribution by the damage function estimates from and Table 3.2 Columns 4 and 6. In each time period, the green box-and-whiskers plot includes all 190 ensemble members, and the purple one includes the 20 r1i1f1p1 ensemble members. The plots includes the median (middle line in white), the 25^{th} and 75^{th} percentiles (the box), the 5^{th} and 95^{th} (the whiskers), and any individual values outside these percentiles (circles).

APPENDIX C

Appendix to Chapter 4

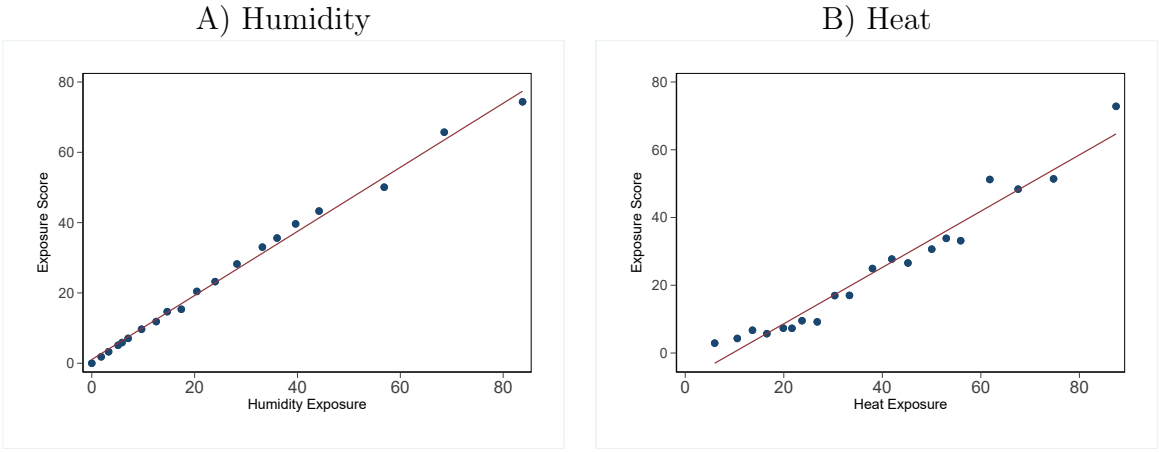


Figure C1: Humidity, Heat and Exposure

Panel A shows the relationship between an occupation's humidity score and its exposure score (θ, the minimum of heat and humidity). Each dot represents a quantile of the humidity score and the corresponding average exposure score. Panel B shows the corresponding relationship for the heat score and exposure score.

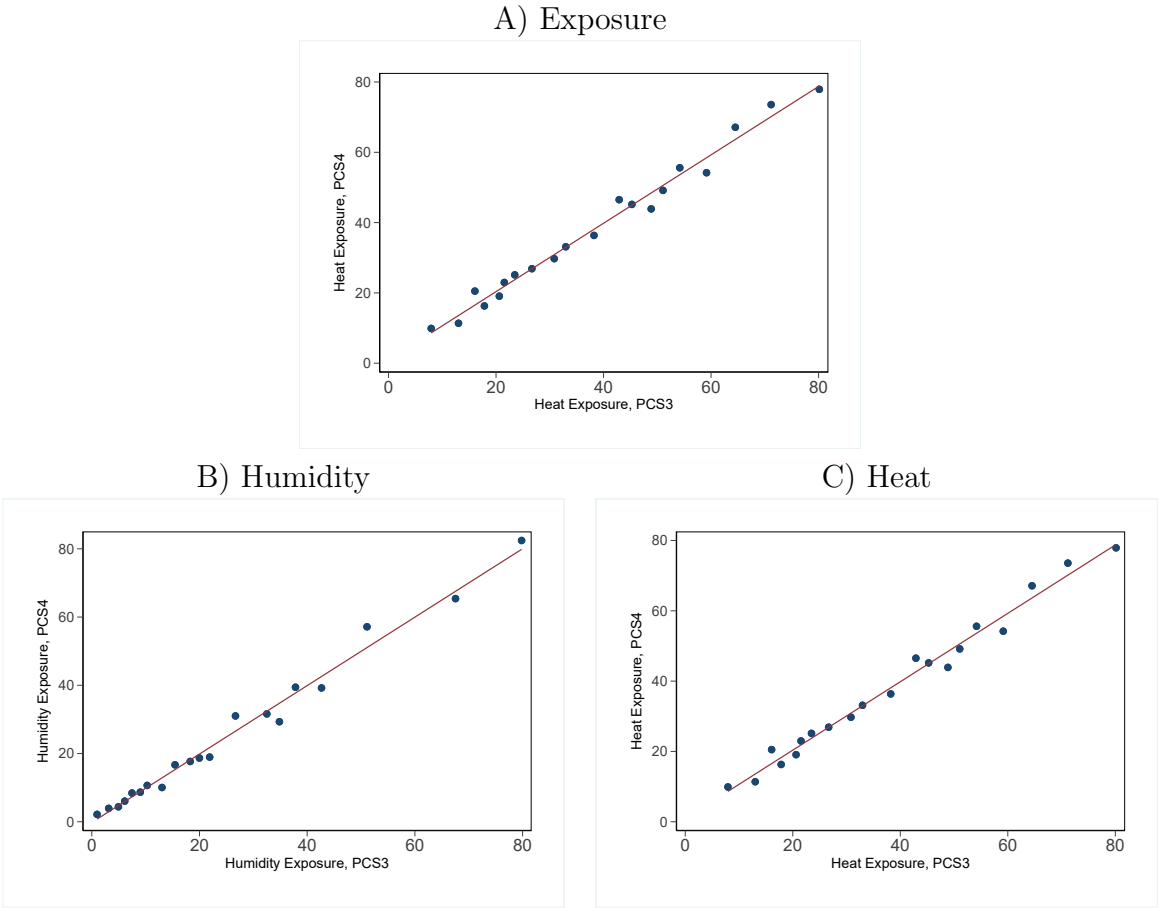


Figure C2: Exposure Metrics: PCS3 vs. PCS4

Panel A shows the relationship between an occupation's exposure score at the 4-digit level and its exposure score at the 3-digit level. Each dot represents a quantile of the 3-digit occupation exposure score and the corresponding average exposure score at the 4-digit level. Panel B shows the same relationship for the humidity score, and Panel C shows the same relationship for the heat score.

Table C1: High Exposure Occupations (θ)

Occupation Code	Title	θ	Employment (K)
632e	Roofers	94	37.9
684b	Refuse and Recycling Collectors (unskilled)	90	23.5
631a	Gardeners	90	29.1
691d	Tree Farmworkers	87	38.5
621e	Public Works Laborers (skilled)	86	42.7
632a	Brickmasons	83	209.4
691a	Agricultural Equipment Operators	79	7.1
691f	Forestry Workers	79	6.1
691e	General Farmworkers	78	36
481b	Construction Supervisors	78	64.5
621a	Public Works Supervisors	76	33.5
621d	Telecomms Line Installers and Repairers	76	13.6
621g	Skilled Mine Workeres	76	8.7
681a	Construction Helpers	75	111.4
691c	Gardening and Horticulture Workers	75	27.7
632b	Stonemasons	70	7.5
632c	Construction Carpenters	70	26.1
632h	Floor Layers	70	4.3
480a	Agricultural Supervisors	69	4.4
480b	Merchant Navy Boat Masters	69	0.2
632g	Painters	68	95
621c	Construction Equipment Operators	67	58.6
621b	Cement Masons	65	51.6
621f	Construction Laborers	65	64.4
691b	Animal Farmworkers	65	15.5
633c	Vehicle Electricians	64	4.8
633b	Electronic Home Entertainment Equipment Repairers	64	8.7
633d	Non-Industrial Electricians	64	9.3
655a	Electricians	63	12.6
633a	Skilled Electricians	62	87.7
632f	Plumbers	62	85.5
671c	Cement Workers	61	48.3
671d	Mining Helpers	61	1.3
471b	Agriculture Control Technicians	60	4.3
675b	Woodworking Workers	60	18
681b	Unskilled Secondary Building Workers	60	77.6
651b	Railyard Wagon Drivers	59	9.6
651a	Crane and Tower Operators	58	21
628e	Refuse and Recycling Collectors (skilled)	58	14.8
634b	Metal Workers	58	37
632j	Drywall and Insulation Installers	57	56.3
632k	Building Maintenance Workers	56	139.6
533b	Park Rangers	56	0.5
533a	Firefighters	52	1.8
634a	Autobody Repairers	52	40.8
625a	Production Installation Workers	50	19.5
671a	Public Works Laborers (unskilled)	50	1.5

This table lists every high exposure $\theta \geq 50$ occupation, its θ, and its average yearly employment in France. I calculate the employment based on the number of workers in the *Dads Postes* sample from 2011-2019.

Individual Characteristics: Education and Place of Birth

FH-DADS: Subsample with Education

The *DADS Panel* does not include workers' education or nationality, so to conduct the following analyses I use the *FH-DADS*, which does. This dataset is the combination of two data sources: 1) the regular *DADS Panel*, and 2) the "Fichier Historique," which is derived from a survey that workers must complete when applying for unemployment benefits. Therefore, any results that include education rely on this sample of workers who at some point applied for unemployment insurance. Results from this section, while informative, may not be generalizable to the entire labor market. The data in this sample span the years 2010 to 2012, the last year of *FH-DADS* data.

I examine persistence in high exposure occupations separately by two characteristics: education and nationality. I find that 2.6 percent of workers with a BA work in high exposure occupations, compared to 11.3 percent of workers without a BA. Figure C3A plots the relationship between θ_o and $\bar{\theta_p}$, separately for workers with and without a BA. It shows that workers with a BA 1) work in less exposed occupations in general and 2) have a less steep relationship between their origin occupation exposure and that of their destination occupation. Panel B plots the relationship between θ_o and the probability that $\theta_p < 50$, separately by BA-attainment. Twenty one percent of workers with a BA transitioning out of a high exposure occupation remain in high exposure work, compared to 51 percent of non-BA workers. This finding suggests that education is likely a contributing factor to the segmentation between high and low exposure labor markets.

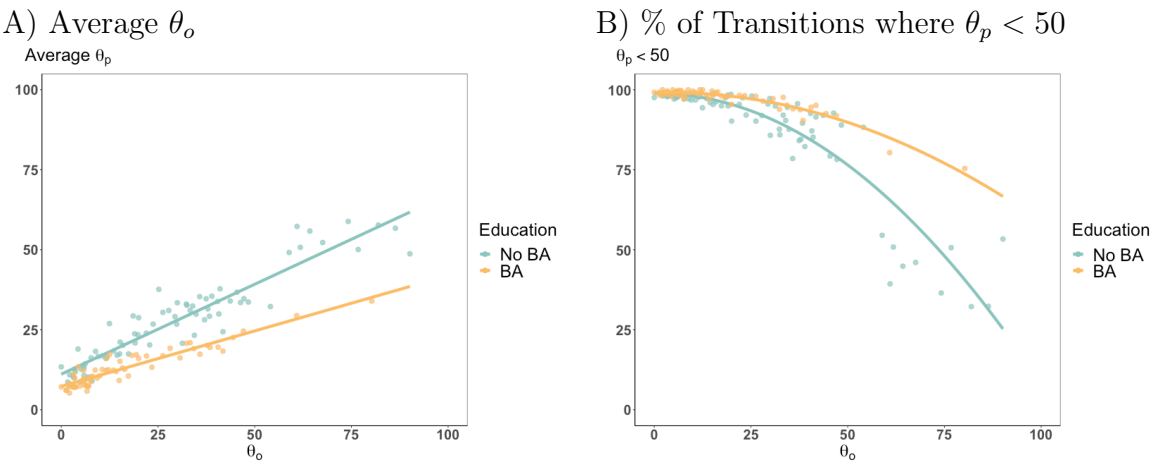


Figure C3: Origin Exposure and Average Destination Exposure: By Education

Panel A shows the average destination exposure score θ_p for each percentile of the origin exposure score θ_o, separately by workers with and without a BA. The lines are linear best fits. Panel B differs only in the y-axis, which represents the percentage of destination occupations that have $\theta < 50$. The data are from 2010 to 2012, and include 136,678 transitions.

Next, I test the relationship between worker nationality and inter-exposure mobility. Nationality may affect a worker's ability to find low exposure work in multiple ways. For example, non-French workers may be less likely to speak French proficiently or to have well-developed occupational networks. Both of these factors could influence a worker's ability to find low exposure work. Of the workers in the *FH-DADS*, 8.8 percent of French workers and 18.8 percent of non-French workers are engaged in high exposure work. Figure C4 shows that the relationship between origin occupation exposure and destination occupation exposure is slightly stronger for non-French workers than for French workers. However, this difference is not as stark as is the case for BA vs. non-BA workers. I find that a 46 percent of French workers leaving a high exposure occupation move to a different high exposure occupation, compared to 64 percent of non-French workers.

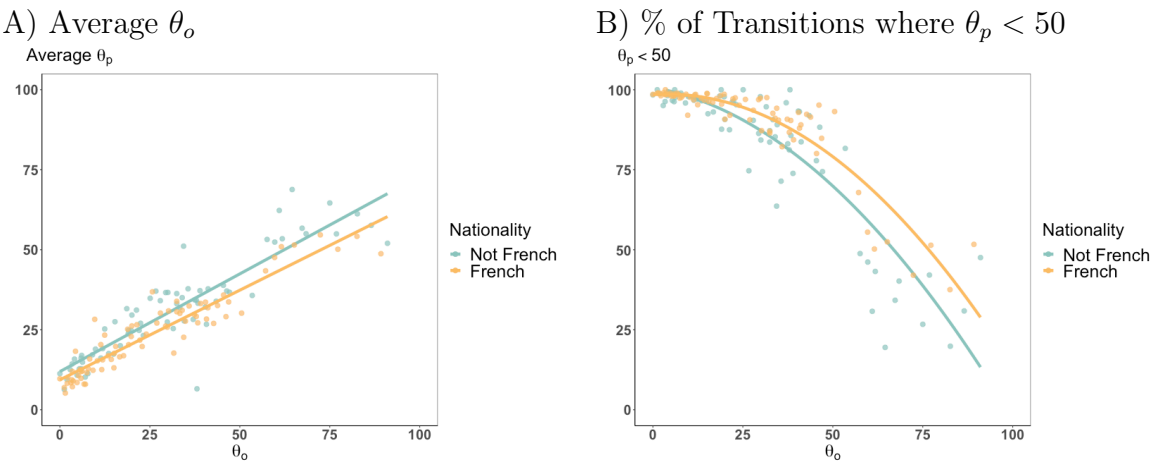


Figure C4: Origin Exposure and Average Destination Exposure: By Nationality

Panel A shows the average destination exposure score θ_p for each percentile of the origin exposure score θ_o, separately by nationality (French or not). The lines are linear best fits. Panel B differs only in the y-axis, which represents the percentage of destination occupations that have $\theta < 50$. The data are from 2010-2012, and include 131,310 observations.

Occupational Transitions and Task Intensities

Table C2: Tasks Used from O*NET (1)

Non-routine analytical	
4.A.2.a.4 Analyzing data/information	Identifying the underlying principles reasons, or facts of information by breaking down information or data into separate parts.
4.A.2.b.2 Thinking creatively	Developing, designing, or creating new applications, ideas, relationships, systems, or products, including artistic contributions.
4.A.4.a.1 Interpreting information for others	Translating or explaining what information means and how it can be used.
Non-routine interpersonal	
4.A.4.a.4 Establishing and maintaining personal relationships	Developing constructive and cooperative working relationships with others, and maintaining them over time.
4.A.4.b.4 Guiding, directing and motivating subordinates	Providing guidance and direction to subordinates, including setting performance standards and monitoring performance.
4.A.4.b.5 Coaching/developing others	Identifying the developmental needs of others and coaching, mentoring, or otherwise helping others to improve their knowledge or skills.

This table lists the five broad task groups and the specific O*NET skills that comprise each category.

Table C3: Tasks Used from O*NET (2)

Non-routine manual	
4.A.3.a.4 Operating vehicles, mechanized devices, or equipment	Running, maneuvering, navigating, or driving vehicles or mechanized equipment, such as forklifts, passenger vehicles, aircraft, or water craft.
4.C.2.d.1.g Spend time using hands to handle, control or feel objects, tools or controls	How much does this job require using your hands to handle, control, or feel objects, tools or controls?
1.A.2.a.2 Manual dexterity	The ability to quickly move your hand, your hand together with your arm, or your two hands to grasp, manipulate, or assemble objects.
1.A.1.f.1 Spatial orientation	The ability to know your location in relation to the environment or to know where other objects are in relation to you.
2.B.1.a Social Perceptiveness	Being aware of others reactions and understanding why they react as they do.
Routine manual	
4.C.3.d.3 Pace determined by speed of equipment	How important is it to this job that the pace is determined by the speed of equipment or machinery? (This does not refer to keeping busy at all times on this job.)
4.A.3.a.3 Controlling machines and processes	Using either control mechanisms or direct physical activity to operate machines or processes (not including computers or vehicles).
4.C.2.d.1.i Spend time making repetitive motions	How much does this job require making repetitive motions?
Routine cognitive	
4.C.3.b.7 Importance of repeating the same tasks	How important is repeating the same physical activities (e.g., key entry) or mental activities (e.g., checking entries in a ledger) over and over, without stopping, to performing this job?
4.C.3.b.4 Importance of being exact or accurate	How important is being very exact or highly accurate in performing this job?
4.C.3.b.8 Structured v. Unstructured work	To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals?

This table lists the five broad task groups and the specific O*NET skills that comprise each category.

Past work finds that workers tend to transition between occupations with similar task requirements (Gathmann and Schönberg, 2010). I confirm this is true in my sample as well with Figure C5. In Panel A, I plot the average routine manual task intensity of the destination occupations for workers leaving a given occupation against the occupation's routine

manual task intensity itself. Panels B-E do the same for non-routine manual, non-routine interpersonal, non-routine analytical, and routine cognitive task intensity, respectively. They confirm that workers leaving an occupation with a higher intensity of a particular task group end up in occupations with a higher intensity of that particular task.

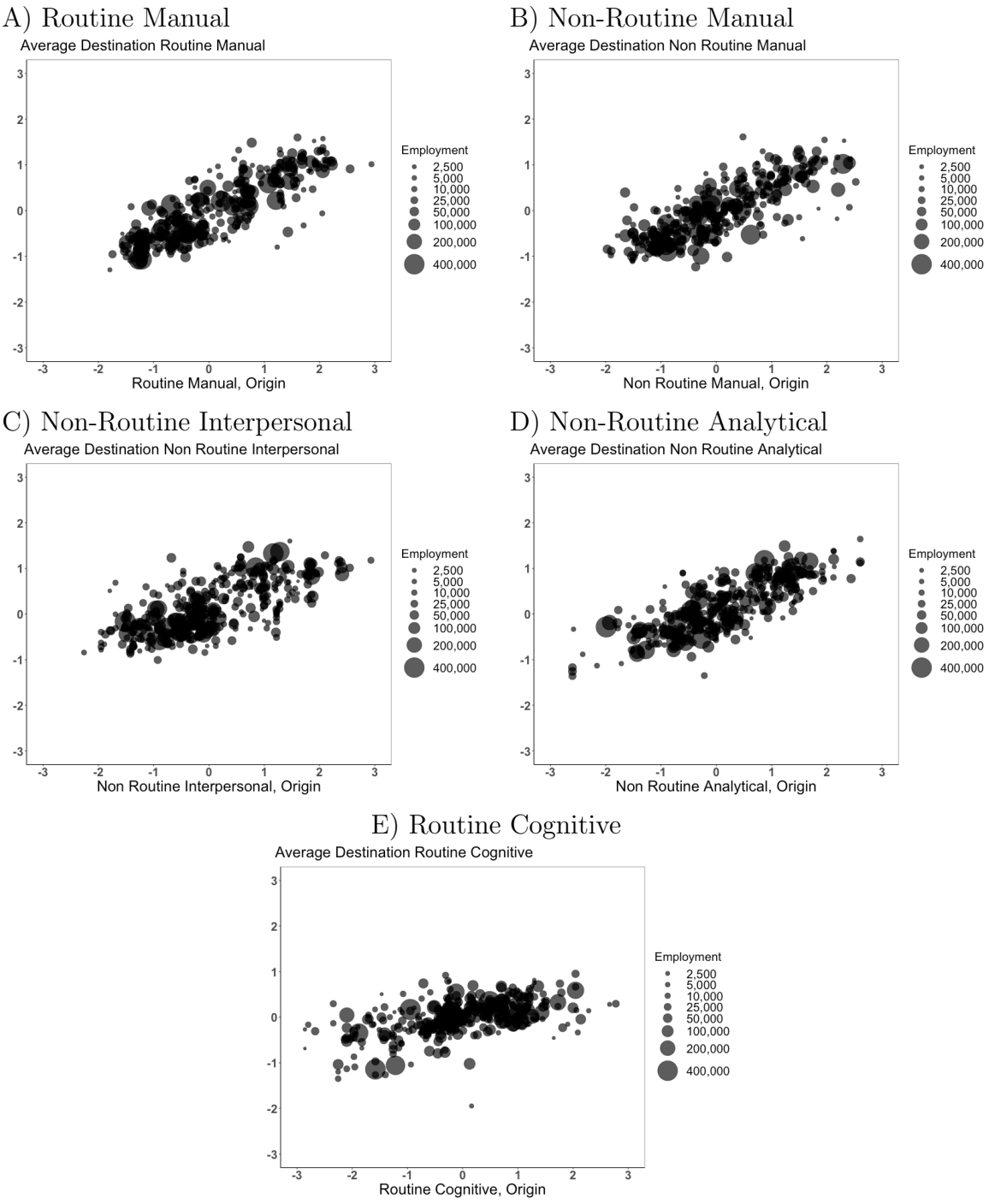


Figure C5: Occupational Transitions and Task Intensities

This figure shows the relationship between the average task intensity of destination occupations for workers leaving a particular occupation and the origin occupation's intensity for the same task. Panel A shows routine manual, Panel B shows non-routine manual, Panel C shows non-routine interpersonal, Panel D shows non-routine analytical, and Panel E shows routine cognitive. Each dot represents a single occupation and its size represents average employment in that occupation from 2011 to 2019. The data include 1,156,194 transitions from 2011-2019 across 368 occupations.

Given this finding, and the relationship between high exposure work and manual task intensity, I have to ensure that my findings on occupational transitions and exposure are not just due to workers staying in manual tasks. I present evidence suggesting that this is not the case in Section 4.3. Here, I extend that analysis in four ways.

First, I look at persistence in individual tasks. Figure C6 expands on Figure 4.5 with Panels C and D, which include all 17 individual tasks rather than just the 5 task groups. For the sake of clarity, I only highlight high exposure in red, and leave all other tasks in gray. I find that workers stay in high exposure occupations longer than they do in occupations with high intensities of any of these tasks. For these 17 tasks, persistence 8 years later ranges from 56 percent (creative tasks) to 76 percent (spatial awareness). Again, the difference is even starker with workers who leave their occupation, where the persistence for high task intensities ranges from 25 percent (repetitive motions) to 53 percent (data analysis).

Second, I confirm that these findings are true across all age groups. Figure C7 shows that workers remain in high exposure work longer than in any other task group across all age groups. The gap between persistence in high exposure work and high task intensity work shrinks as workers get older. This is consistent with the idea that as workers get more experience in the labor market, they get more entrenched in jobs with specific tasks. Relative to occupations with high task intensity, persistence in high exposure work does not increase as much with age. This is consistent with the fact that task composition of occupations alone does not explain why workers remain in high exposure occupations for so long.

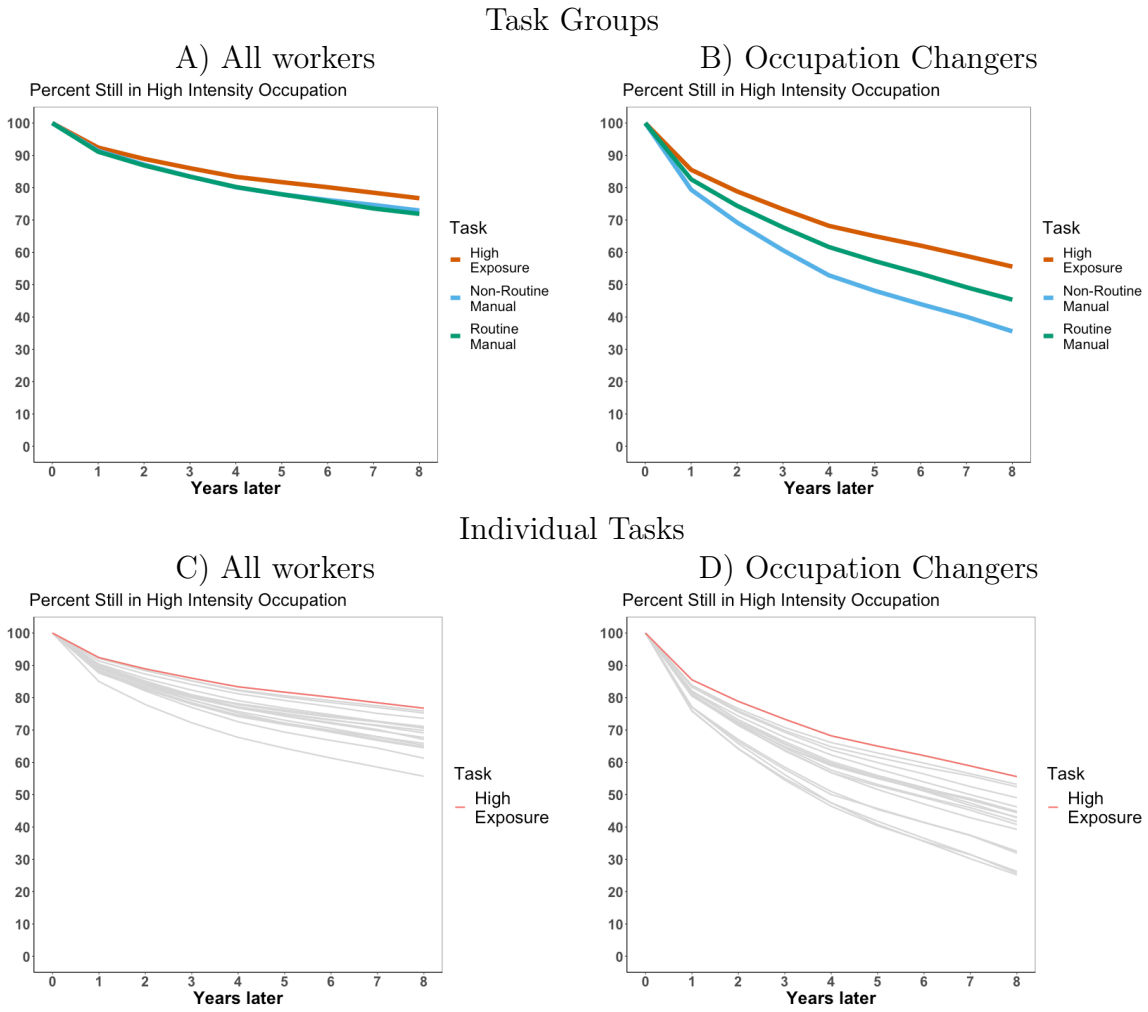


Figure C6: Persistent High Task Comparison

Panel A graphs worker persistence in occupations with high intensities in manual tasks alongside environmental exposure. For each task, the sample includes all workers present for all 9 years of my data who ever work in an occupation that I label as high in that task. An occupation has a high task intensity if its score is 1.5 standard deviations above the mean. For θ, this corresponds to 50, and for the task groups, this corresponds to 1.5 (by construction). Panel B shows the same comparison, but limits the sample to workers who have changed occupations at least once during my sample period. Panels C and D are the same as A and B, respectively, but instead of the 5 task groups, they show the 17 individual tasks outlined in Tables C2 and C3.

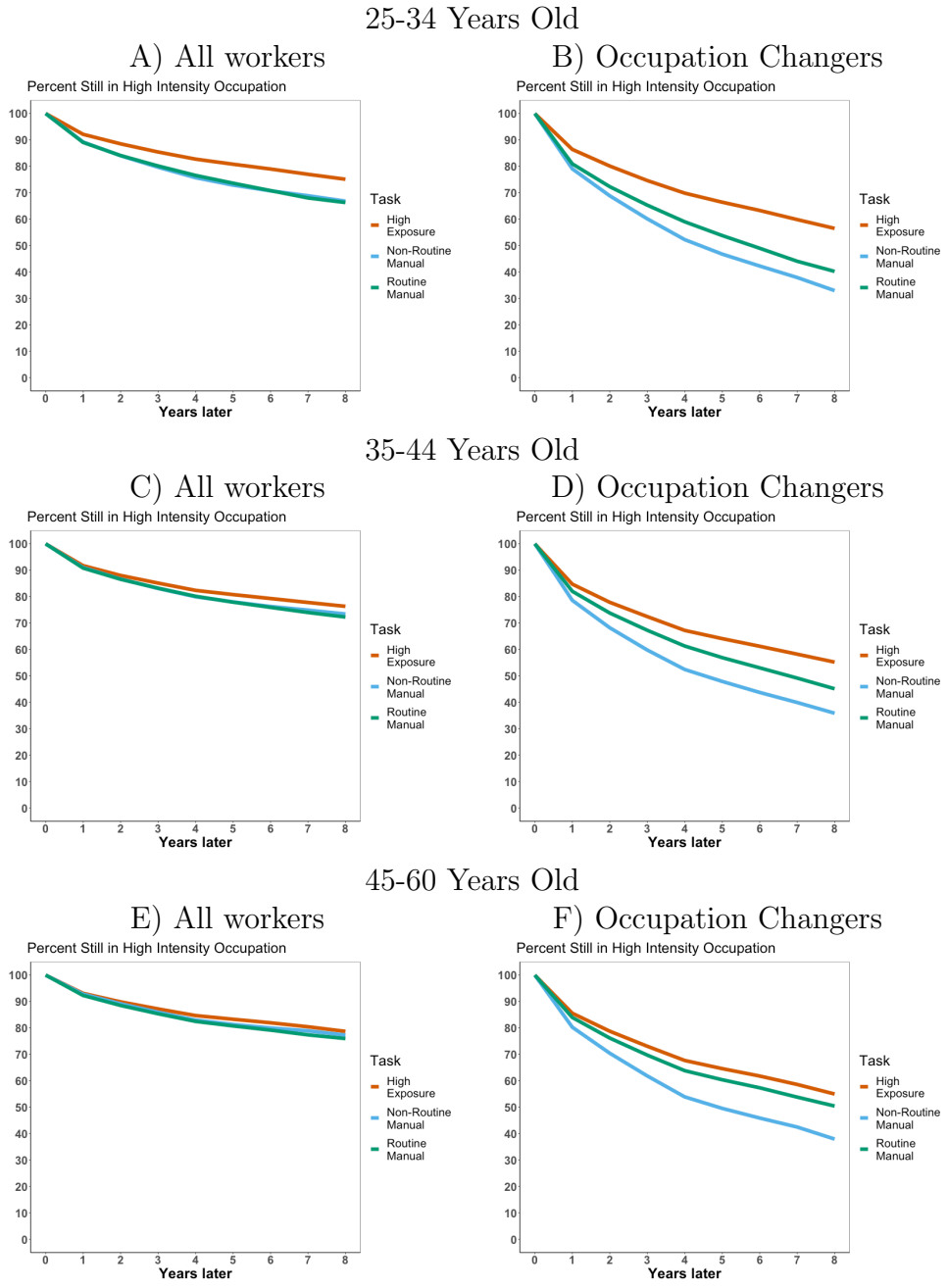


Figure C7: Persistent High Task Comparison by Age Group

Panel A graphs worker persistence in occupations with high intensities in manual tasks alongside environmental exposure. For each task, the sample includes all workers aged 25-34 present for all 9 years of my data who ever work in an occupation that I label as high in that task. An occupation has a high task intensity if its score is 1.5 standard deviations above the mean. For θ, this corresponds to 50, and for the task groups, this corresponds to 1.5 (by construction). Panel B shows the same comparison, but limits the sample to workers who have changed occupations at least once during my sample period. Panels C and D are the same as A and B, respectively, but for workers between 35 and 44. Panels E and F are the same as A and B, respectively, but for workers between 45 and 60.

Third, I analyze transitions in and out of high exposure work relative to those in and out of high task intensity occupations. Table C4 shows the percentage of employment comprised by high intensity jobs for a given task. This percentage ranges from 3 to 15 percent. It also shows the percentage of "high-to-high" transitions by task (e.g., the number of workers moving between two occupations with a high intensity in that task divided by the total outflow from occupations with a high intensity in that task). It shows that high exposure is, by far, unique in terms of high-to-high transitions. The task group with the largest high-to-high percentage is routine manual, with 38 percent, The individual task with the largest high-to-high percentage is data analysis, with 39 percent, which is still 11 percentage points lower than high exposure.

Fourth, I analyze the relationship between the manual task intensity of an origin occupation and the θ of its destination occupations. I would expect there to be a positive relationship between these two variables. However, if manual task intensity does not alone explain low rates of inter-exposure mobility, I would also expect that, for a given level of origin manual task intensity, the more exposed the origin occupation is, the more exposed its destination occupations would be too. I test this hypothesis in Figure C8. Panel A plots the relationship between the routine task intensity of an origin occupation and the average θ_p of its destination occupations. The dots are colored by the θ of the origin occupation. Two phenomena stand out. First, workers leaving occupations with higher levels of routine manual task intensity end up in more exposed occupations. Second, there is a clear positive relationship between θ_o and $\bar{\theta_p}$ for any given level of routine manual task intensity. This finding suggests that manual task intensity is not enough to fully explain workers remaining in high exposure tasks. Panel B is the same as Panel A, but replaces routine manual intensity with non-routine manual intensity.

Table C4: High Task Intensities

Task	(1) % of Employment	(2) $P(High_p \mid High_o)$
Climate Exposure (θ)	9.1	48.6
Non-Routine Analytical	6.2	27.8
Non-Routine Interpersonal	9.3	30.2
Non-Routine Manual	9.8	25.0
Routine Manual	8.3	37.5
Routine Cognitive	5.0	8.2
Repetitive Motions	6.4	12.5
Accuracy	8.5	29.4
Equipment Sets Pace	8.9	39.1
Manual Dexterity	6.3	20.0
Spatial Orientation	8.8	29.9
Repetitive Tasks	8.5	31.2
Use of Hands	3.4	19.1
Data Analysis	8.4	38.5
Creative Thinking	8.7	15.3
Interpretation	10.0	27.6
Personal Relationships	6.6	30.0
Motivation	12.0	37.4
Coaching	14.9	31.4
Machines	9.5	34.7
Vehicles	11.0	33.8
Social Perceptiveness	8.7	24.6
Structured	10.0	18.6

This table lists each task group and individual task, along with the percentage of employment that its high intensity occupations comprise, and the probability of a worker's new occupation (p) being in that high category, given that they left an occupation (o) in that high category. The first column includes all workers 18-67, and the second column includes transitions from workers between 25 and 60.

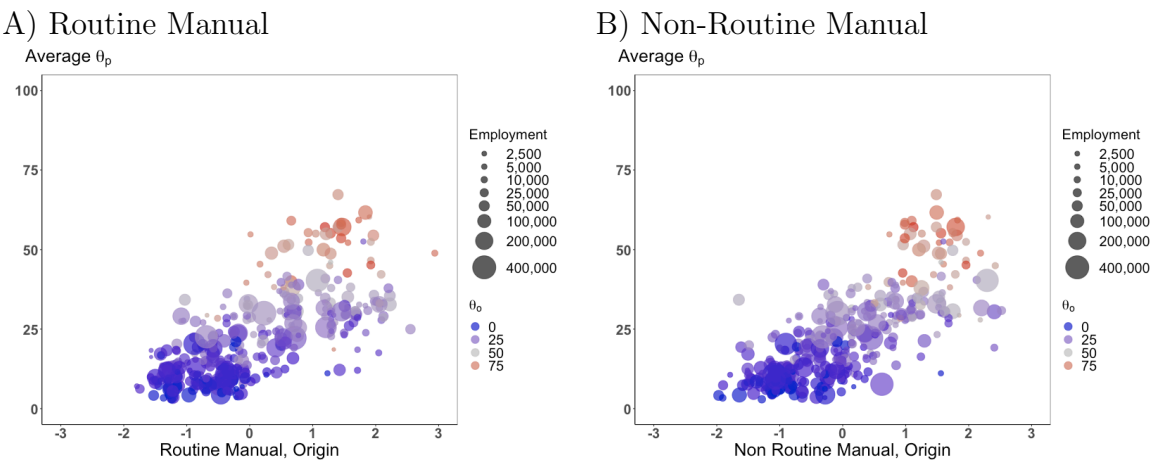


Figure C8: Origin Manual Task Intensity and Destination Exposure

Panel A shows the relationship between the average exposure level of destination occupations for workers leaving a particular occupation and the routine manual task intensity of the origin occupation. Each dot represents a single 4-digit occupation and its size represents average employment for the occupation from 2011-2019. The dots are colored by the exposure score of the origin occupation. Panel B is the same as Panel A, but for non-routine manual task intensity. The data include 1,156,194 transitions from 2011-2019 across 368 occupations.

Persistence and Transitions: Robustness

In this section, I present robustness checks for my main high exposure persistence and inter-exposure mobility results. My first set of robustness checks are based on different definitions of exposure and of the high exposure threshold. I check that my results are robust to defining exposure based solely on an occupation's humidity score or solely on its heat score. I also check that they are robust to defining high exposure occupations as those with $\theta \geq 40$. My second sets of robustness checks is based on three additional samples of workers: 1) workers 18-67, 2) workers I observe for all 9 years[1], and 3) workers I observe for all 9 years and who were in a high exposure occupation in 2011.[2]

Exposure and Age

Figure C9 shows that the percentage of workers in high exposure occupations by age is robust to a number of different specifications. Panel A defines high exposure as $\theta \geq 40$, Panel C defines it based on humidity, and Panel D defines it based on heat. In all cases the shape remains the same, though overall percentages are higher in the case of $\theta \geq 40$ and heat. These data further demonstrate that humidity is often the binding element in an occupation's definition of exposure. Panel B shows the relationship between high exposure and age for the sample from 18 to 67 years old. It shows that workers under 25 have even higher exposure than those older than them.

[1]This sample definition is the default sample for the persistence results, so this test only applies to the transition results.

[2]This test is only for the persistence results.

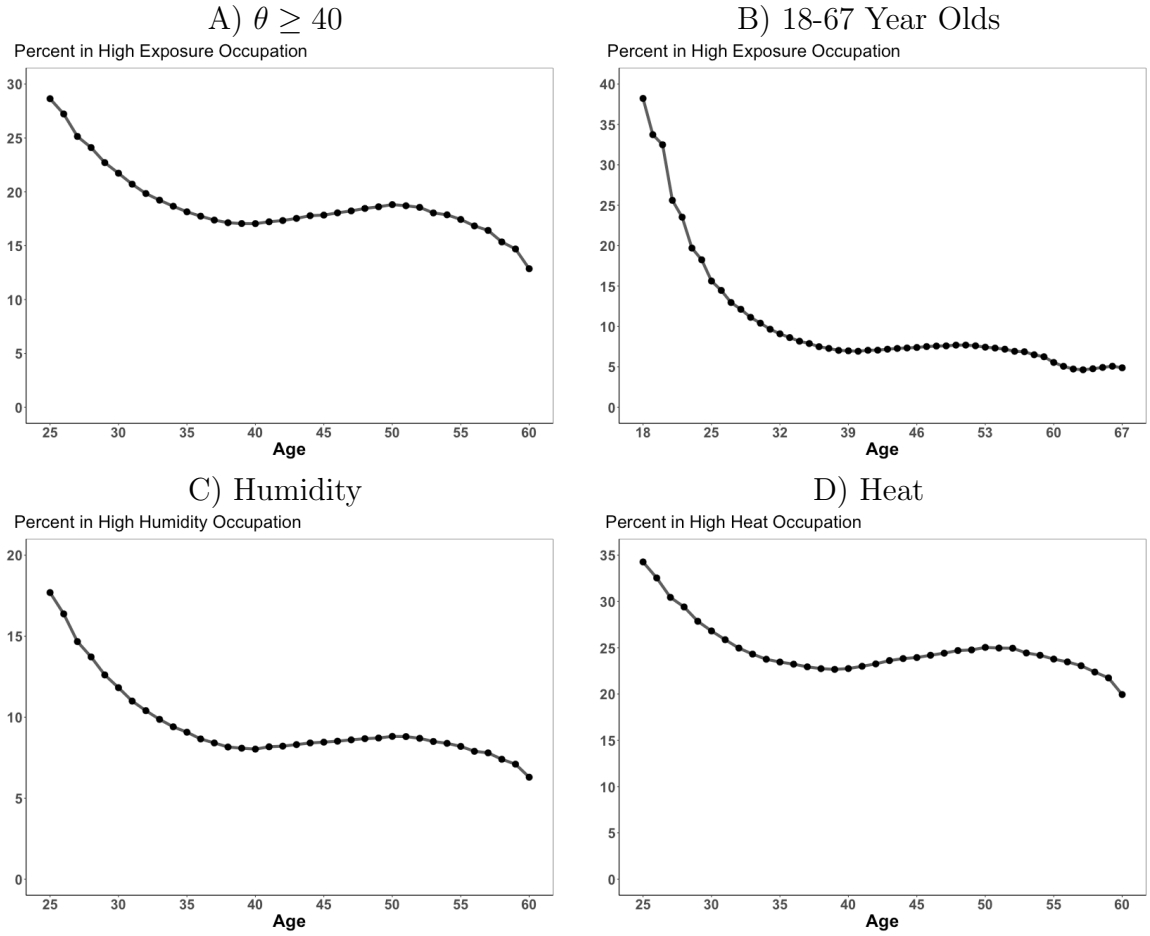


Figure C9: Exposure and Age: Robustness

This figure shows the percent of high exposure workers by age, with different samples of workers and different definitions of high exposure. Panel A redefines high exposure as $\theta \geq 40$ rather than the preferred threshold of 50. Panel B includes workers between 18 and 67 years old. Panel C defines exposure based solely on an occupation's humidity score (high exposure is humidity ≥ 50), and Panel D defines exposure based solely on an occupation's heat score (high exposure is heat ≥ 50). Panels A, C, and D include 498,472 workers, and panel B includes 500,165 workers.

High Exposure Persistence

Figure C10 shows that worker persistence in high exposure work is robust to a number of different specifications. Panel A defines high exposure as $\theta \geq 40$, Panel C defines it based on humidity, and Panel D defines it based on heat. Panel B expands the sample to workers between 18 and 67, and Panel E restricts it to workers who are in high exposure work in 2011, limiting the data to one starting year per worker. In all cases, the shape remains the same.

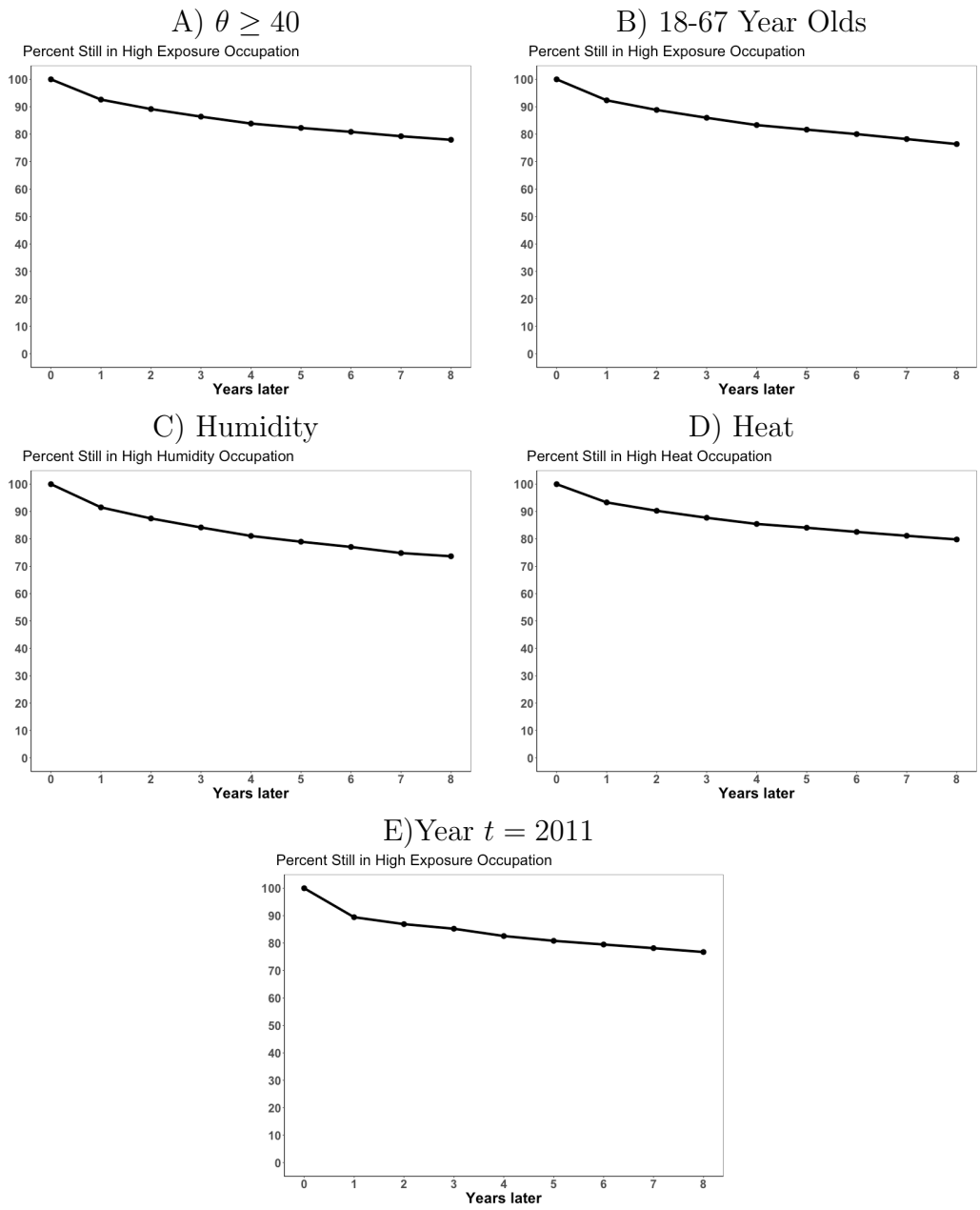


Figure C10: Persistence in High Exposure Work: Robustness

This figure shows the percent of high exposure workers by age, with different samples and different definitions of high exposure. Panel A redefines high exposure as $\theta \geq 40$ rather than the preferred threshold of 50. Panel B includes workers between 18 and 67 years old. Panel C defines exposure based solely on an occupation's humidity score (high exposure is humidity ≥ 50), and Panel D defines exposure based solely on an occupation's heat score (high exposure is heat ≥ 50). In all cases, the sample is limited to workers I observe for 9 years and who work in a high exposure occupation at least once. The sample in Panel E is further limited to workers who are in high exposure work in 2011, and it only uses 2011 as a starting year t. Panels A includes 125,651 workers, Panel B includes 65,674 workers, Panel C includes 57,004 workers, Panel D includes 158,140 workers, and Panel E includes 36,439 workers.

Transitions and Exposure

Figure C11 shows that the relationship between an occupation's exposure and the average exposure of the destination occupations for departing workers is robust to different exposure definitions and sample selections. Panel A defines exposure as humidity and Panel B defines it as heat. Panel C includes workers from 18-67, and Panel D restricts the sample to only includes workers I observe for all 9 years. In all cases, the average destination exposure rises with the origin exposure.

Figure C12 shows that the relationship between the probability of moving to a low exposure occupation and one's origin occupation exposure is robust to a variety of definitions of high exposure and sample selections. It shows the same robustness checks as Figure C11, with the addition of Panel E, which defines high exposure as $\theta \geq 40$.

In Panel E, the overall findings remain the same. However, one major difference stands out relative to Figure 4.3B. First, workers in occupations with lower levels of θ are much more likely to transition to occupations with $\theta \geq 40$ than ≥ 50. For example, 6 percent of workers leaving an occupation with $\theta < 25$ transition to occupations with $\theta \geq 40$, compared to only 1.7 percent with $\theta \geq 50$.

This observation suggests a meaningful difference between the 40 and 50 thresholds. In fact, even workers leaving occupations with θ between 40 and 50 are much more likely to transition to occupations with θ between 40 and 50 than above 50. Twenty-five percent of transitions out of these occupations end up in occupations with θ between 40 and 50, compared to only 13 percent above 50. This is despite only 13 percent of jobs being in the former group, compared to 9 percent in the latter.

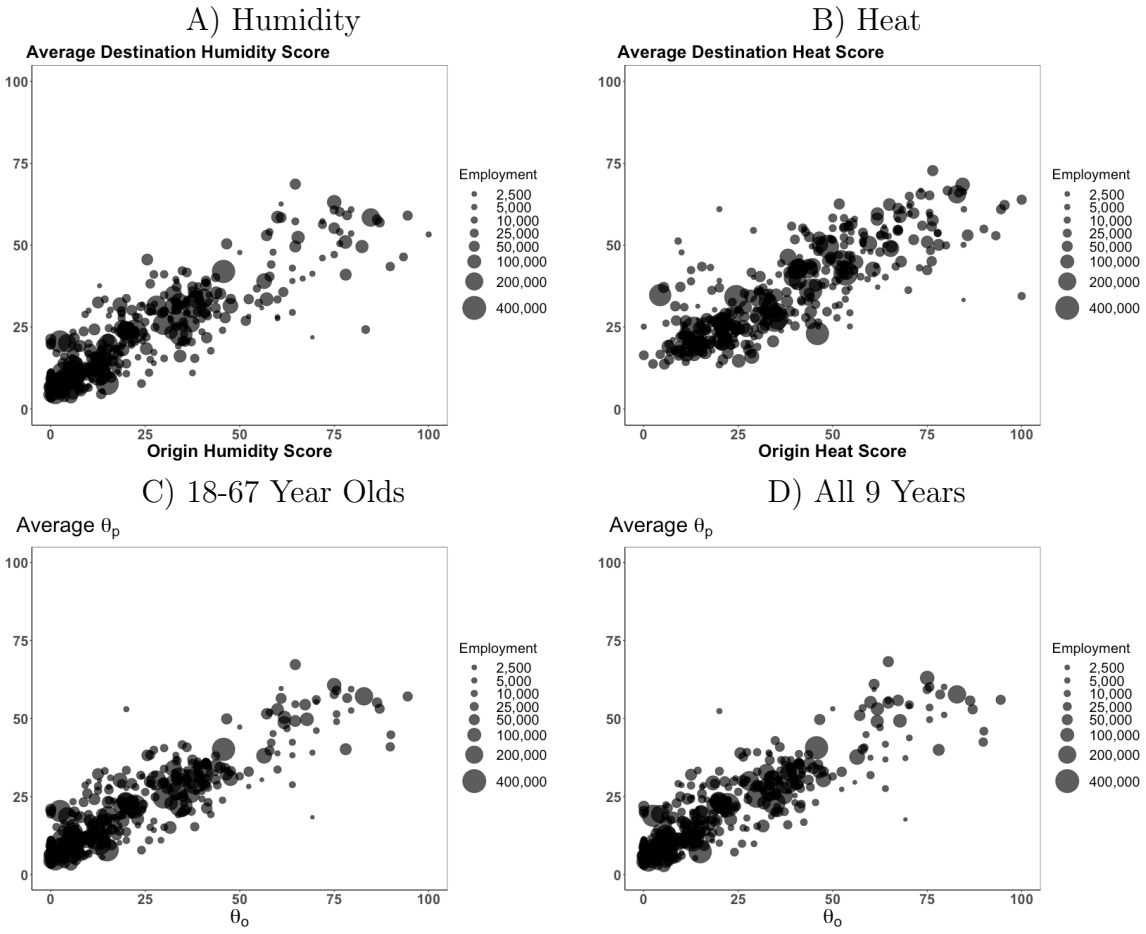


Figure C11: Transitions and Exposure: Robustness (1)

All panels in this figure graph the relationship between an occupation's exposure score and the average exposure score of the destination occupations for workers leaving that occupation. Panel A redefines exposure as solely based on an occupation's humidity score, and Panel B defines it as solely based on its heat score. Panel C includes workers from 18-67 years old, and Panel D only includes workers I observe for all 9 years. Each dot represents a single occupation and its size is based on average yearly employment from 2011-2019. All panels include 368 occupations from 2011-2019. Panels A and B include 1,156,194 transitions, Panel C includes 1,315,535 transitions, and Panel D includes 576,399 transitions.

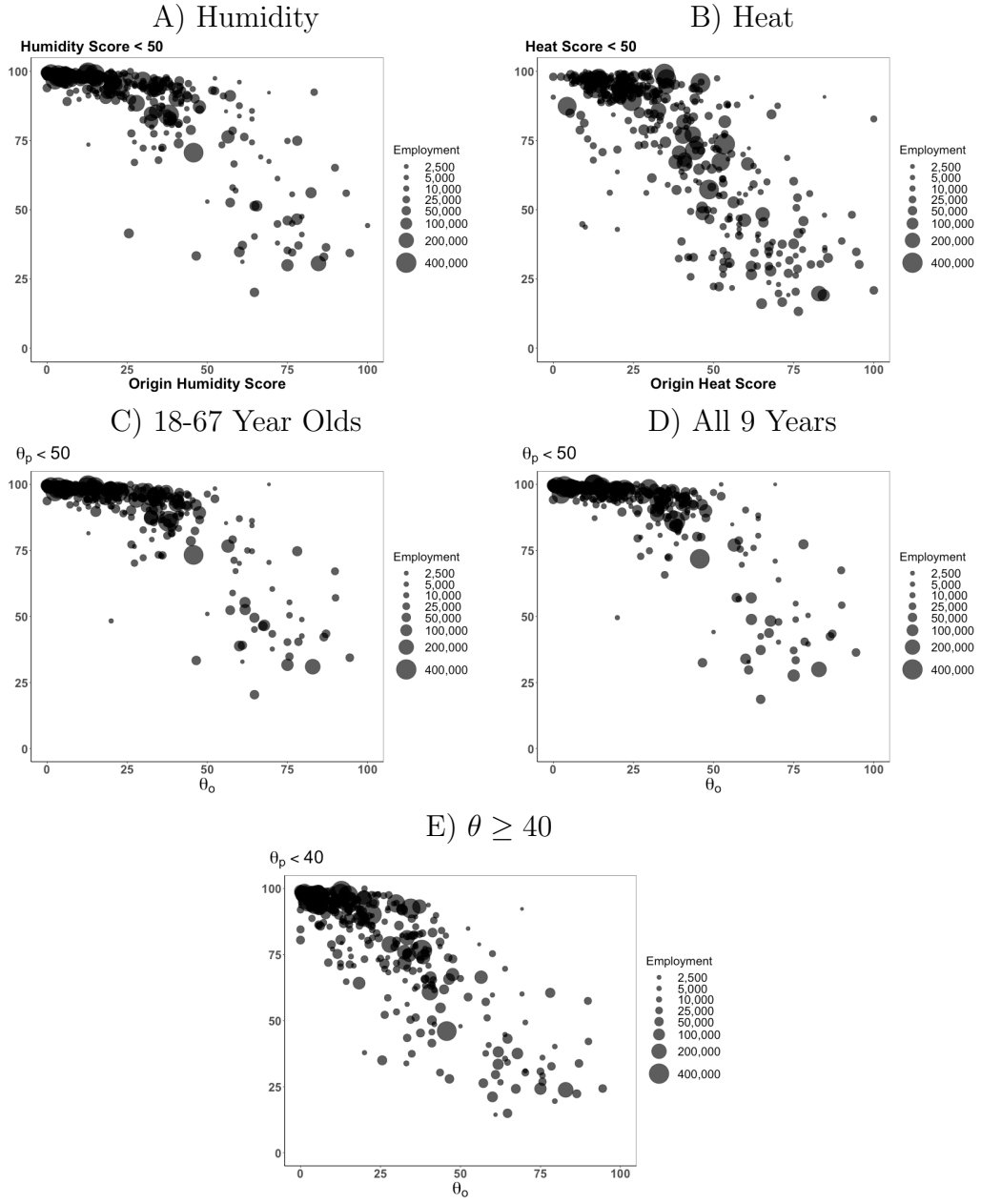


Figure C12: Transitions and Exposure: Robustness (2)

All panels in this figure graph the relationship between an occupation's exposure score and the percentage of transitions to a low exposure occupation for workers leaving that occupation. Panel A redefines exposure as solely based on an occupation's humidity score, and Panel B defines it as solely based on its heat score. Panel C includes workers from 18-67 years old, and Panel D only includes workers I observe for all 9 years. Each dot represents a single occupation and its size is based on average yearly employment from 2011-2019. Panel E defines high exposure as $\theta \geq 40$ instead of 50. All panels include 368 occupations from 2011-2019. Panels A, B, and E include 1,156,194 transitions, Panel C includes 1,315,535 transitions, and Panel D includes 576,399 transitions.

PCS4-SOC-code Crosswalk Validation

Tasks Scores: Average of Multiple PCS4-O*NET Matches

In this section, I show that the results of my task-based analyses are robust to setting a PCS4's task scores equal to the average of its possible SOC-code matches, rather than just the preferred match. Figure C13 shows that the comparison between persistence in high exposure work and work that is high in other tasks remains the same. Figure C14 shows that the correlations between exposure (θ) and other task intensities remain the same.

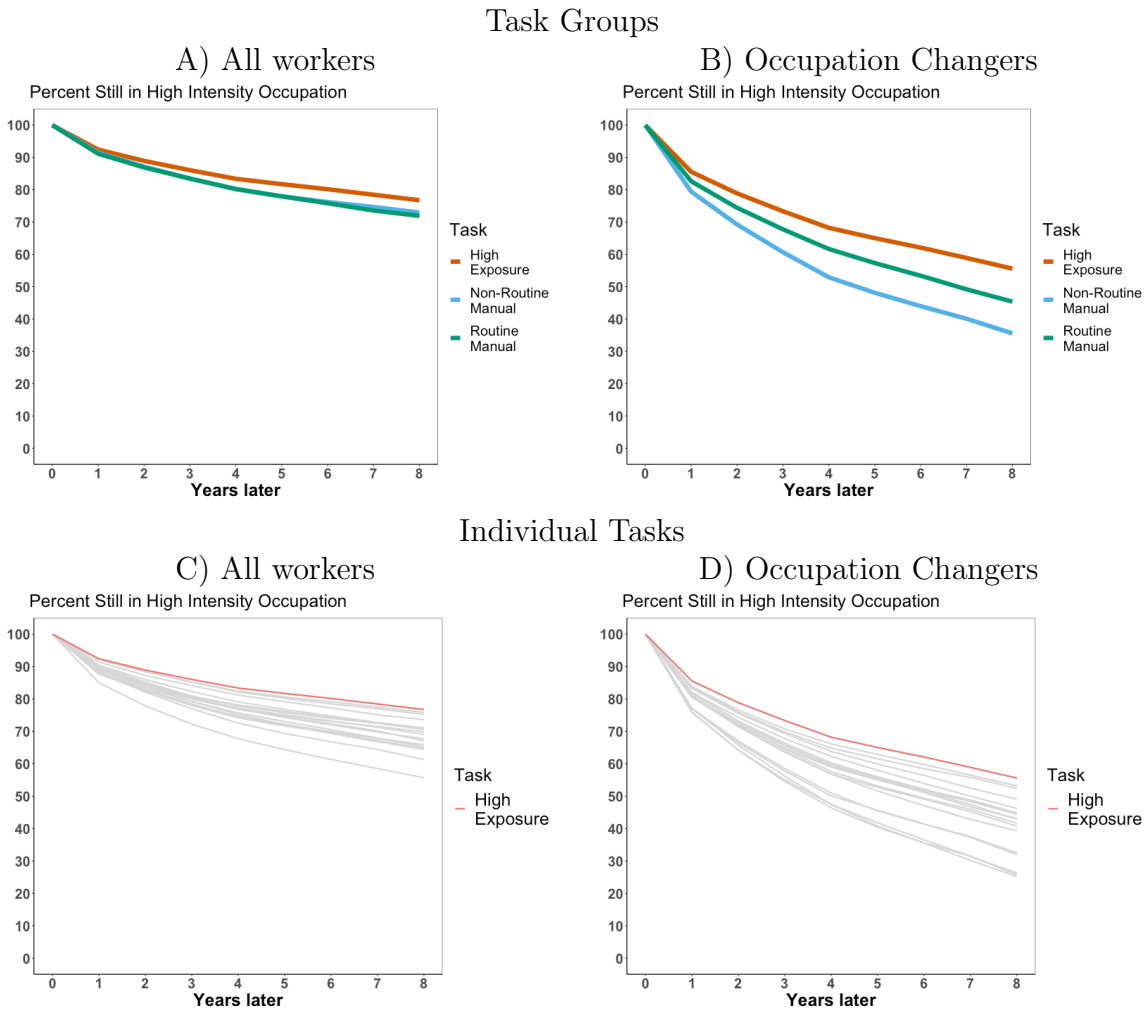


Figure C13: Persistent High Task Comparison: Average Score of SOC-code Matches

Panel A graphs worker persistence in occupations with high intensities in various tasks alongside environmental exposure. The line in red represents worker persistence in high environmental exposure, and the others represent the 5 task groups: non-routine analytical, non-routine interpersonal, non-routine manual, routine manual, and routine cognitive. For each task, the sample includes all workers present for all 9 years of my data who ever work in an occupation that I label as high in that task. An occupation has a high task intensity if its score is 1.5 standard deviations above the mean. For θ, this corresponds to 50, and for the task groups, this corresponds to 1.5 (by construction). Panel B shows the same comparison, but limits the sample to workers who have changed occupations at least once during my sample period. Panels C and D are the same as A and B, respectively, but instead of the 5 task groups, they show the 17 individual tasks outlined in Tables C2 and C3. The sample in all panels is limited to workers I observe for all 9 years. This figure differs from the main results by how I assign task scores to PCS4 occupations. In the main estimation, I match 1 SOC-code to each PCS4, whereas here I assign each PCS4 a score based on the average of all possible matches.

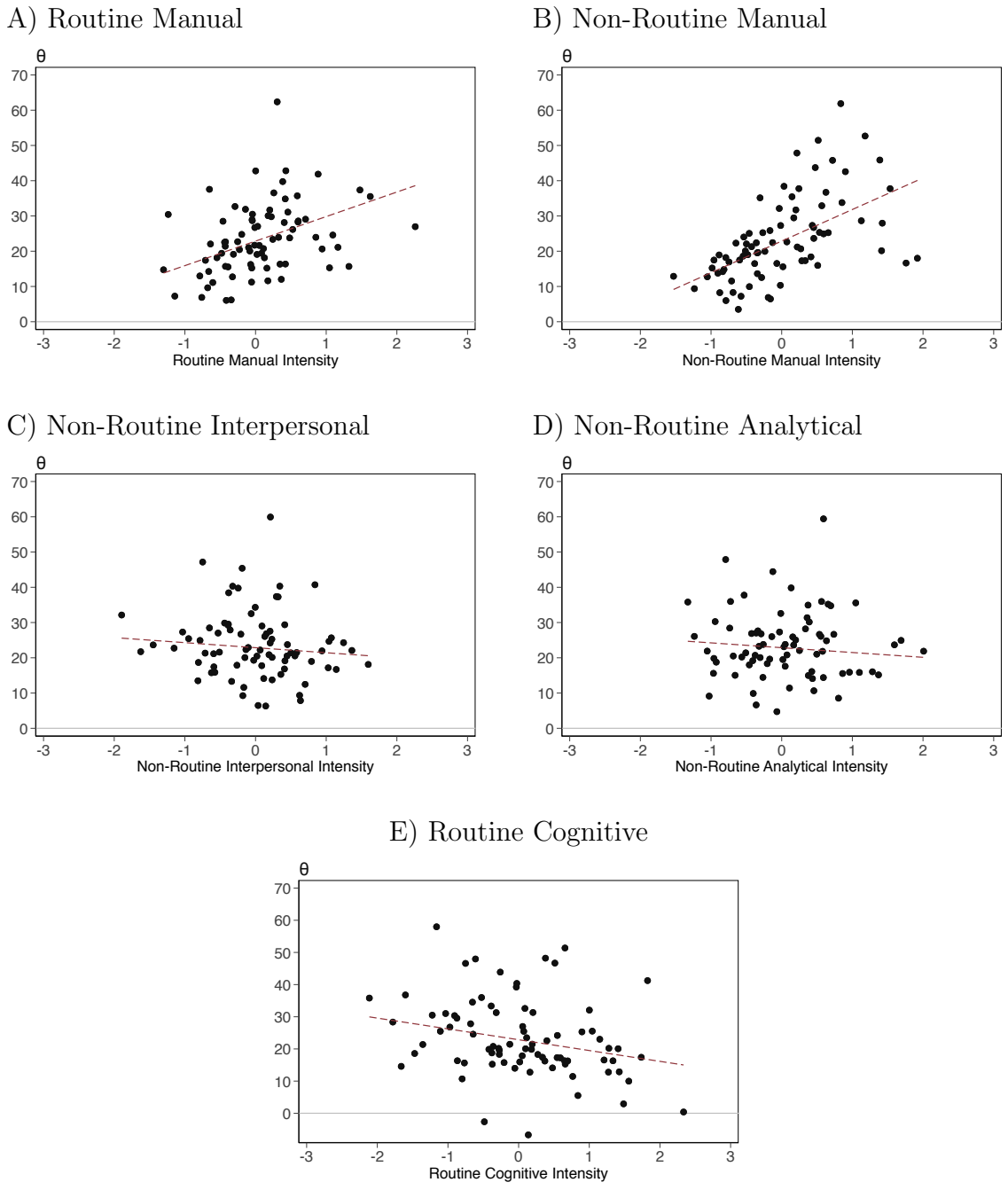


Figure C14: Task Intensity and Transitions: Average Score of SOC-code Matches

This figure shows the relationship between different an occupation's task intensities and its environmental exposure. Panel A is for routine manual, Panel B is for non-routine manual, Panel C is for non-routine interpersonal, Panel D is for non-routine analytical, and Panel E is for routine cognitive. All graphs control for all other task intensities and are weighted by employment. The data include 368 occupations. This figure differs from the main results by how I assign task scores to PCS4 occupations. In the main estimation, I match 1 SOC-code to each PCS4, whereas here I assign each PCS4 a score based on the average of all possible matches.

Similar Task Comparison

In this section, I compare scores for tasks that are common to both the O*NET survey and the *CT-Individus* survey. This exercise tests two aspects of my research. First, it tests the crosswalk I have created to translate between occupation SOC-codes (in O*NET) and occupation PCS4 codes (in *CT-Individus*). Second, it tests the assumption that a survey from the US is a good proxy for occupation characteristics from France.

Figures C15 and C16 show that, generally speaking, the scores line up quite well. In most cases, the task intensities for similar tasks have strong positive correlations.

In the case of exposure (Panel A), the questions are slightly different. For American occupations, I use responses to an O*NET question that asks workers the percentage of days they spend time in an air-conditioned environment. I assign each occupation an exposure score equal to 100 minus this score. This means that while my PCS4 exposure score is measures a percent of workers, the O*NET exposure score measures a percentage of days. In addition, while the O*NET score captures *potential* exposure to extreme weather, the PCS4 score captures *actual* exposure. Therefore, the exposure scores from O*NET are likely larger than those from the *CT-Individus*. Panel A confirms this hypothesis, especially for occupations at high levels of exposure.

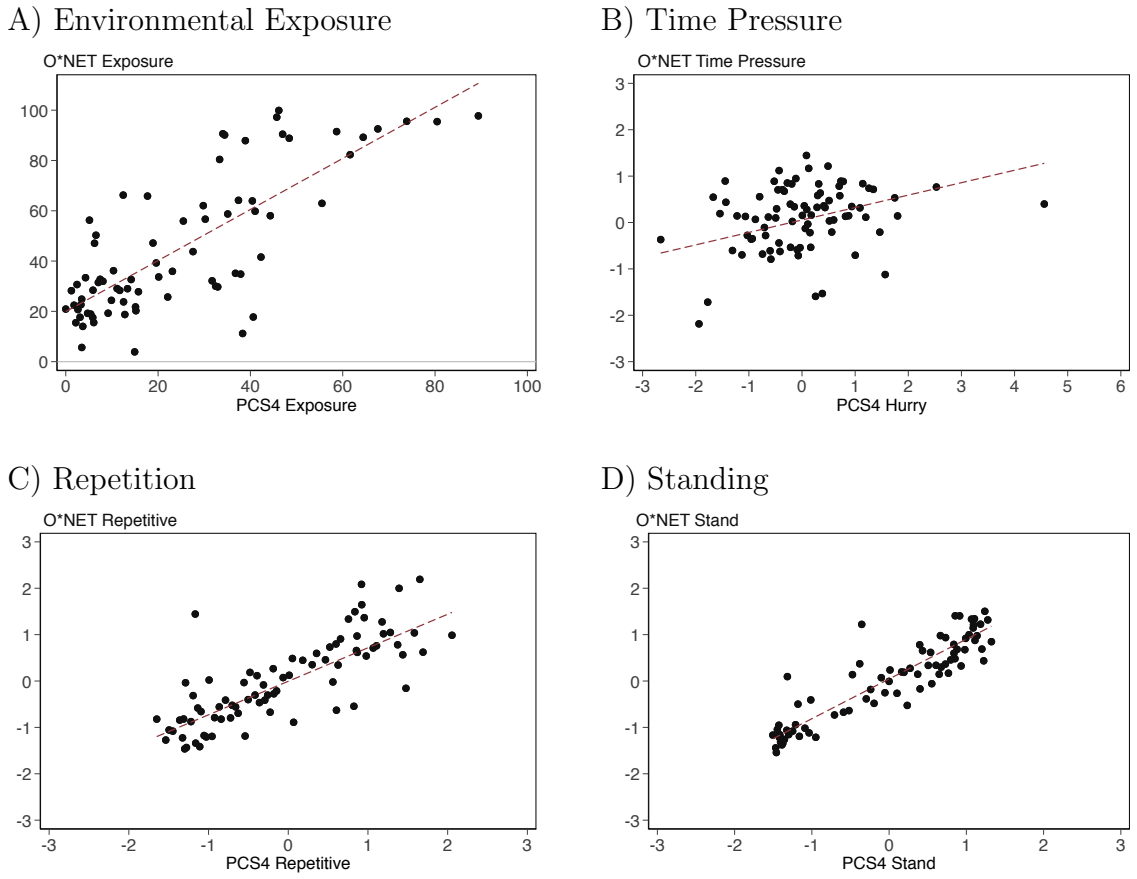


Figure C15: Task Score: O*NET vs. PCS4 (1)

These figures compare scores for tasks that are common to both the O*NET survey and the *CT-Individus* survey. Other than the exposure scores, which are out of 100, the scores are all normalized to have mean 0 and standard deviation 1. Panel A compares exposure from the *CT-Individus* (the minimum of the percentage of workers who report experiencing heat and the percentage of workers who report experiencing humidity) to exposure from O*NET (the average percent of days a worker spends not in a climate controlled environment). Panel B is for time pressure, Panel C is for repetitive, and Panel D is for time spent standing. All panels use employment weights.

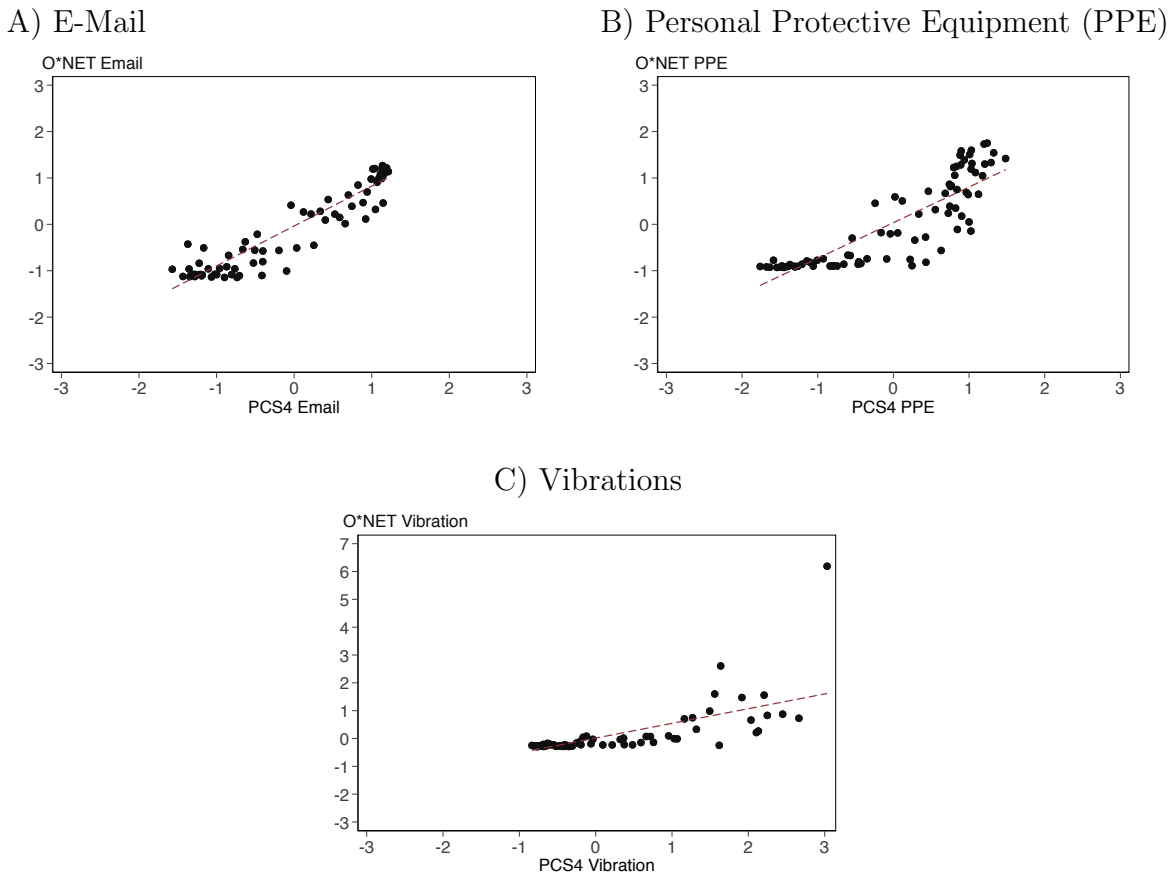


Figure C16: Task Score: O*NET vs. PCS4 (2)

These figures compare scores for tasks that are common to both the O*NET survey and the *CT-Individus* survey. The scores are all normalized to have mean 0 and standard deviation 1. Panel A compares the use of E-mail from the *CT-Individus* to that in O*NET, Panel B is for the use of personal protective equipment (PPE), and Panel C is for exposure to vibrations. All panels use employment weights.

Outside Options, Market Concentration, and Wages

Here, I describe the analysis I conduct to calculate the effect of an occupation's outside occupation options' wages on its own wages. For my preferred specificaition, I follow the methodology of Schubert, Stansbury, and Taska (2022). In robustness checks, I expand on their methodology in two main ways. First, I estimate regressions at the individual level, with the inclusion of individual fixed effects. Second, I use an alternative instrument for the Herfindahl-Hirschman Index (HHI) from Marinescu, Ouss, and Pape (2021).

Methods

Using the *DADS Postes*, I calculate two labor market quality metrics: labor market concentration and outside occupation options quality. The former is a measure of the local quality of one's current occupation, whereas the latter metric captures job opportunities outside of a one's current occupation. I calculate both metrics at the occupation (o) by CZ (m) year (t) level.

I construct the outside occupation options index (OOI) as formulated in Schubert, Stansbury, and Taska (2022). The OOI is the average local wage in occupations outside a worker's current one, weighted by their probability of transitioning to each occupation:

$$\text{OOI}_{o,m,t} = \sum_{p \neq o}^{N_{occs}} P(o \to p) \cdot w_{p,m,t}, \tag{C.1}$$

where

$$P(o \to p) = \pi_{o \to p} \cdot \frac{s_{p,m,t}}{s_{p,t}},$$

and $\pi_{o \to p}$ is the national probability that a worker moves to occupation p given that they leave their job in occupation o. Following the definition of Schubert, Stansbury, and Taska (2022), this is equal to the product of two quantities:

1. The leave share of occupation o. This is equal to the percentage of workers who leave

occupation o when changing jobs.[3]

2. The percentage of workers who, leaving occupation o, move to occupation p.

For each of these two measures, I use the occupational mobility data from the *DADS Panel* in Section 3. $\frac{s_{p,m,t}}{s_{p,t}}$ is the ratio of the employment share of occupation p in CZ m to the national employment share, in year t, which I calculate from the *DADS Postes*. This helps account for any differences in the geographical distribution of a given occupation. The full equation for the outside-option index is as follows:

$$\text{OOI}_{o,m,t} = \sum_{p \neq o}^{N_{occs}} \pi_{o \to p} \cdot \frac{s_{p,m,t}}{s_{p,t}} \cdot w_{p,m,t}, \tag{C.2}$$

I show summary statistics for the OOI in 2019 in Table C5. The OOI to wage ratio varies from 0.28 to 2.5, which is larger than that found in past work in the US, largely due to the different definition of the leave share used in my paper (Schubert et al., 2022).

Next, following Marinescu, Ouss, and Pape (2021), I measure labor market concentration with the Herfindahl-Hirschman Index (HHI) for hires. This measures how few firms make hires for a given occupation in the local area, and it is equal to the sum of the squares of the shares of new hires for each employer $j \in J$:

$$HHI_{o,m,t} = \sum_{j \in J} \sigma^2_{j,o,m,t}, \tag{C.3}$$

where

$$\sigma_{j,o,m,t} = \frac{N_{j,o,m,t}}{\sum_{k \in J} N_{k,o,m,t}}$$

[3]There is one major difference between my leave share measure and that used in Schubert, Stansbury, and Taska (2022). Unlike them, I am unable to see within-occupation job title changes. So, I categorize a worker leaving a job if they change occupations (PCS4) or firm establishment (SIRET). This results in significantly lower numbers of workers observed as leaving their job, and therefore significantly higher leave shares (and OOIs). However, this difference does not affect the comparability of the results between our papers. By construction, the leave share is constant across time and space for each occupation o, and is therefore removed by occupation fixed effects (because I use the log of OOI). My results are identical regardless of the distribution of leave shares across occupations.

is the number of hires ($N_{j,o,m,t}$) for employer j in the given labor market. I define a hire as an employee in occupation o for employer j in CZ m and year t, but not in year $t-1$. The index runs between 0 (exclusive) and 1 (inclusive). I show summary statistics for HHI in 2019 in Table C5. The HHI values span nearly the entire possible range. The median occupation-CZ has an HHI of 0.056 and a mean of 0.154, similar to past research in France (Marinescu et al., 2021). Per the American Department of Justice and Federal Trade Commission, and HHI between 0.15 and 0.25 designates a moderately concentrated labor market, and above 0.25 a highly concentrated one (Marinescu et al., 2021). This definition suggests that roughly 25 percent of occupation-CZ units were at least moderately concentrated in 2019.

Table C5: Summary Statistics: OOI and Wages, Full Sample, 2019

Percentile	1	5	10	25	50	75	90	95	99
Hourly Wage (€)	12.0	12.4	12.8	14.1	16.9	22.4	32.3	35.5	40.9
Employment	1	4	8	20	56	170	467	885	2,839
HHI	0.002	0.004	0.006	0.019	0.056	0.16	0.46	0.84	1
OOI (€)	5.2	6.9	8.2	11.0	14.3	20.4	32.9	45.7	62.0
$\frac{OOI}{wage}$	0.28	0.41	0.48	0.66	0.84	1.1	1.4	1.7	2.5
Observations	84,636								

This table lists the distribution of employment, and employment-weighted wages, HHI, OOI, and the OOI to wage ratio at the CZ-Occ level for the year 2019.

Empirical Specification

I test the impact of local labor market concentration and outside options quality on an occupation's wages. I merge occupation-CZ-year level average wages from the *DADS Postes* with the corresponding OOI and HHI. My empirical strategy is as follows:

$$\ln w_{n,o,m,t} = \alpha_{o,t} + \alpha_{m,t} + \beta_1 \ln OOI_{o,m,t} + \beta_2 \ln HHI_{o,m,t} + \epsilon_{n,o,m,t}, \tag{C.4}$$

where $w_{n,o,m,t}$ are wages in 2019 €. $\alpha_{o,t}$ is an occupation by year fixed effect, which accounts for any common shocks to an entire occupation and its outside occupation options in a given year. For example, if 2016 has particularly low demand for construction nation-wide,

lowering wages for both construction workers and construction managers (one of their outside occupation options), this could lead to a spurious positive correlation between OOI and wages. $\alpha_{m,t}$ is a CZ by year fixed effect, which accounts for any shocks to a local area in a given year. Together, these fixed effects ensure that the regression compares OOI and wage deviations relative to their occupation-year and CZ-year means. ϵ is an error term. The standard errors are clustered at the CZ level to account for any correlation between the error terms of occupations in the same location. This specification is the same as the preferred specification in Schubert, Stansubry, and Taska (2022).

I test the robustness of the model to a range of specifications in the Appendix. Using individual wage data from the *DADS Panel*, I also estimate this equation at the individual level with the inclusion of individual fixed effects. This specification helps address the possibility that the worker composition of an occupation-CZ may change alongside changes in HHI or OOI, affecting average productivity and wages.

OOI and HHI Instruments

The coefficient estimates for the impact of both the OOI and the HHI on wages suffer from potential endogeneity bias. I therefore instrument for local labor market concentration and local outside occupation options' quality, using the 'Bartik' shift-share instruments detailed in Schubert, Stansbury, and Taska (2022). β_1, the coefficient for the OOI, could be biased by the fact that the wages for a worker's current occupation and one of their outside options could both be increased by an increase in local demand for a common product (e.g., local construction demand). Along with endogeneity bias, β_1 also suffers from reverse causality as two occupations could both be outside options for one another. The OOI instrument consists of replacing two of the three OOI terms. I instrument for local wages ($w_{p,m,t}$) with the national mean wage in occupation p and year t, excluding CZ m ($w_{p,\not{m},t}$). I also instrument for the local relative occupation share with the value from 2011 ($\frac{s^{2011}_{p,m}}{s^{2011}_{p}}$), the first year included in my sample. The transition probabilities remain the same. Thus, my instrument for the

OOI is as follows:

$$\text{OOI}^{inst}_{o,m,t} = \sum_{p \neq o}^{N_{occs}} \pi_{o \to p} \cdot \frac{s^{2011}_{p,m}}{s^{2011}_{p}} \cdot w_{p,\not{m},t}. \tag{C.5}$$

For the instrument to be valid, the national leave-one-out mean wages for all outside option occupations p must be positively correlated with local wages in p, but not affect wages in o other than through its outside occupation options, conditional on the fixed effects. The instrument excludes local wages as well as local changes in relative employment, lending credibility to these assumptions.

The sign of the bias for β_2, the affect of HHI on wages, is ambiguous. On the one hand, high labor market concentration could be due to the presence of highly productive firms (who could pay higher wages). On the other hand, lagging local productivity would both reduce wages and make the area less attractive to new firms, increasing concentration. I instrument for the HHI by calculating a predicted change in HHI for each local area based on local firm-level hiring growth. I predict this local firm-level occupational growth with the national hiring growth rate of each firm (excluding the local area). The instrument for HHI, HHI^{inst}, is as follows:

$$HHI^{inst} = ln\Bigg(\sum_{j} \sigma^2_{j,o,m,t-1} \Big(\frac{(1 + \tilde{g}_{j,o,\not{m},t})^2}{(1 + \tilde{g}_{o,\not{m},t})^2} - 1 \Big) \Bigg), \tag{C.6}$$

where $\tilde{g}_{o,\not{m},t} = \sum_j \sigma_{j,o,m,t-1} \tilde{g}_{j,o,\not{m},t}$ is the predicted local hires growth rate. $\tilde{g}_{j,o,\not{m},t}$ is the national growth in hires in a given occupation for a firm, excluding the local area. $\sigma_{j,o,m,t-1}$ is that firm's local occupation hiring share in the prior year. Following Schubert, Stansbury, and Taska (2022), I only include positive values of $\tilde{g}_{j,o,\not{m},t}$.

By construction, this instrument requires, for any given occupation-CZ-year, at least one firm that hires new workers in this CZ and at least one other CZ (for national leave-one-out data), both last year and this year (for growth). This imposes restrictions on my data, particularly in smaller CZs and smaller PCS4 occupations. In addition, since it requires data from year $t-1$, it limits this analysis to 2012-2019. This reduces my sample size

from 744,511 occupation-CZ-year cells to 105,922. However, the cells I do retain include 58 percent of employment from 2012-2019. In the Appendix, I show that I find similar results using an HHI instrument from Marinescu, Ouss, and Pape (2021), which allows me to use the full sample.

The key assumptions for this instrument are that the firm's national leave-one-out hiring growth is correlated with local hiring growth but uncorrelated with occupation-specific local wages through any other channel, conditional on occupation-year and CZ-year fixed effects. Following Schubert, Stansbury, and Taska (2022), I also control for three quantities to address potential concerns with these assumptions. First, I control for the growth rate of local hires in an occupation-CZ. Second, I control for the predicted growth rate of local hires based on national-level firm hiring growth rates. I add a third control variable that addresses a concern with non-linear shift share instruments identified in past work (Borusyak and Hull, 2021). The control involves constructing the expected value of the HHI instrument by creating 1000 random counterfactual instruments. For each counterfactual, I randomly re-assign national hiring growth rates across firms within an occupation-year, then construct the instrument. I average the instruments and take the log. For a more comprehensive discussion of endogeneity, these instruments, their construction, and the necessary assumptions, see Schubert, Stansbury, and Taska (2022).

I show summary statistics for the sample that includes non-missing values for the log of the HHI instrument in Table C6. This includes 13,440 observations for 2019, compared to 84,332 for the full sample. As expected, 2019 employment is larger in this sample than in the full one, ranging from 25 in the 1^{st} percentile to 9,000 in the 99^{th} percentile, as opposed to 1 and 2,800 for the full 2019 sample. In addition, HHI is lower, running from 0.001 to 0.37, with a median of 0.028 instead of 0.054. The OOI distribution remains similar to the full sample.

Table C6: Summary Statistics: OOI and Wages, SST Sample, 2019

Percentile	1	5	10	25	50	75	90	95	99
Hourly Wage (€)	12.1	12.5	12.8	13.8	16.2	21.8	33.5	36.2	39.5
Employment	25	54	83	166	362	829	1,786	2,914	8,982
HHI	0.001	0.003	0.004	0.010	0.028	0.065	0.13	0.19	0.37
OOI (€)	5.3	7.0	8.1	10.9	14.0	21.2	39.8	51.4	63.7
$\frac{OOI}{wage}$	0.34	0.45	0.52	0.70	0.86	1.1	1.4	1.6	2.2
Observations	13,440								

This table lists the distribution of employment, and employment-weighted wages, HHI, OOI, and the OOI to wage ratio at the CZ-Occ level for the year 2019. It only includes observations with valid HHI instruments as per Schubert, Stansbury, and Taska (2022).

Results

Table C7 shows that a higher outside occupation options index leads to higher wages. Panel A presents the coefficients from an OLS estimation, but I focus on Panel B, which presents the results from an IV estimation. Column 1, my preferred specification, mirrors that in Schubert, Stansbury and Taska (2022). It includes CZ-year and occupation-year fixed effects, as well as HHI controls. The estimate for β_1 is 0.052, and it is statistically significant at the 0.01 level. It means that a 1 percent increase in the local outside occupation options index increases wages by 0.052 percent. This coefficient is just over half the size of the coefficient found for the US, which matches the hypothesis that it would be smaller in France. The estimate for β_2 is -0.013 (though it is not statistically significant at conventional levels), about half the size of that found in the US (-0.025). Column 2 shows that these findings are robust to changing the fixed effects to CZ, occupation, and year. Column 3 shows that they are robust to removing the employment weighting. Column 4 shows that they are also robust to removing the HHI instrument controls.

I present and discuss results from three alternative specifications. First, I present results using individual level results in Table C8. Next, I present results from constructing the HHI instrument following Marinescu, Ouss, and Pape (2021), which also uses *DADS Postes Data*. I present results from a specification at the CZ-occupation-year level (Table C9) and at the individual-year level (Table C10).

Table C7: Wages, HHI, and OOI, SST HHI

Dependent variable: Log Wage				
	(1)	(2)	(3)	(4)
Panel A: OLS regression				
Log(OOI)	0.054***	0.054***	0.039***	0.054***
	(0.004)	(0.003)	(0.003)	(0.004)
Log(HHI)	-0.002*	-0.003***	-0.003***	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B: 2SLS IV regression				
Log(OOI)	0.052***	0.056***	0.042***	0.054***
	(0.004)	(0.004)	(0.006)	(0.003)
Log(HHI)	-0.013	-0.004	0.008	-0.008
	(0.015)	(0.012)	(0.009)	(0.011)
Predicted Vacancy Growth	0.003	-0.001	0.007	
	(0.008)	(0.006)	(0.005)	
Vacancy Growth	0.026	-0.018	-0.067***	
	(0.061)	(0.041)	(0.023)	
Expected HHI	0.004	-0.002	-0.001	
	(0.007)	(0.004)	(0.003)	
Fixed effects	CZ-Year	CZ, Occ	CZ-Year	CZ-Year
	Occ-Year	Year	Occ-Year	Occ-Year
Employment Weighting	Yes	Yes	No	Yes
HHI Inst Controls	Yes	Yes	Yes	No
Mean				
Observations	105,922	105,922	105,922	105,922
F-Stat	22.8	37.1	72.3	12.3

This table shows the change in the log wage for a change in the log outside occupation options index (OOI) and the log Herfindahl-Hirschman Index (HHI). Panl A is an OLS estimation, and Panel B is a 2SLS IV estimation. The HHI and OOI instruments are made following Schubert, Stansbury, and Taska (2022) (SST). The outcome, explanatory variables, and instruments are all at the commuting zone (CZ) by occupation by year level. Standard errors clustered at the CZ level in parentheses. The reported F-stat for the 2SLS IV regressions is the Kleibergen-Paap Wald F statistic. * (p<0.10), ** (p<0.05), *** (p<0.01).

Robustness

Alternative Specifications

I repeat the analysis of outside occupation options and labor market concentration on wages in three ways:

1. At the individual level, using the Schubert, Stansbury and Taska (2022) HHI instrument (Table C8). This allows me to include individual fixed effects, controlling for the potential movement of workers (with different levels of productivity) in and out of local labor markets in ways that are correlated with local HHI and OOI. In my primary specification, I find a smaller effect of OOI on wages, of 0.017. It is smaller across all models except in the specification with no individual fixed effects, where I find a very similar coefficient (0.052 percent).

2. At the CZ-occupation-year using the Marinescu, Ouss, and Pape (2021) HHI instrument (Table C9). This allows me to include any observations that are dropped due to insufficient data or a negative Schubert HHI instrument. This specification includes over 700,000 observations, rather than 100,000. However, it does not allow me to control for year-occupation fixed effects, as the instrument is negatively correlated with HHI within a year-occupation, by construction. I find similar coefficients across the board for the effect of OOI on wages. I find a coefficient of -0.005 for the effect of HHI on wages. This coefficient is within the range of coefficients found in Marinescu, Ouss, and Pape (2021).

3. At the individual level using the Marinescu, Ouss, and Pape (2021) HHI instrument (Table C10). Again, I find a statistically significant relationship between outside options and wages, though it is less than a fifth the size of the main specification. Without individual fixed effects, I find a similar coefficient to the main specification.

Table C8: Individual Wages, HHI, and OOI, SST HHI, Individual Sample

Dependent variable: Log Wage				
	(1)	(2)	(3)	(4)
Panel A: OLS regression				
Log(OOI)	0.018***	0.022***	0.053***	0.018***
	(0.002)	(0.003)	(0.003)	(0.002)
Log(HHI)	0.000	-0.002*	0.001	0.000
	(0.000)	(0.001)	(0.001)	(0.000)
Panel B: 2SLS IV regression				
Log(OOI)	0.017***	0.034***	0.052***	0.019***
	(0.005)	(0.003)	(0.004)	(0.004)
Log(HHI)	0.001	0.025***	-0.004	0.005
	(0.010)	(0.007)	(0.013)	(0.007)
Predicted Vacancy Growth	0.003	-0.016***	0.009	
	(0.004)	(0.004)	(0.008)	
Vacancy Growth	0.011	-0.099***	0.059	
	(0.046)	(0.036)	(0.062)	
Expected HHI	0.002	-0.003	0.005	
	(0.003)	(0.002)	(0.007)	
Fixed effects	CZ-Year	CZ, Occ	CZ-Year	CZ-Year
	Occ-Year	Year	Occ-Year	Occ-Year
	Person	Person		Person
HHI Inst Controls	Yes	Yes	Yes	No
Mean				
Observations	4,574,692	4,574,748	4,936,627	4,574,692
F-Stat	14.5	25.8	30.5	6.4

This table shows the change in the log wage for a change in the log outside occupation options index (OOI) and the log Herfindahl-Hirschman Index (HHI). Panl A is an OLS estimation, and Panel B is a 2SLS IV estimation. The HHI and OOI instruments are made following Schubert, Stansbury, and Taska (2022) (SST). Wages are at the individual by year level. The explanatory variables and instruments are all at the commuting zone (CZ) by occupation by year level. Standard errors clustered at the CZ level in parentheses. The reported F-stat for the 2SLS IV regressions is the Kleibergen-Paap Wald F statistic. * ($p<0.10$), ** ($p<0.05$), *** ($p<0.01$).

Table C9: Wages, HHI, and OOI, MOP HHI

Dependent variable: Log Wage			
	(1)	(2)	(3)
Panel A: OLS regression			
Log(OOI)	0.045***	0.045***	0.055***
	(0.003)	(0.003)	(0.002)
Log(HHI)	-0.005***	-0.005***	0.005***
	(0.002)	(0.002)	(0.001)
Panel B: 2SLS IV regression			
Log(OOI)	0.045***	0.045***	0.048***
	(0.002)	(0.002)	(0.003)
Log(HHI)	-0.005***	-0.005***	-0.004***
	(0.001)	(0.001)	(0.001)
Fixed effects	CZ,Occ	CZ-Year	CZ,Occ
	Year	Occ	Year
Employment Weighting	Yes	Yes	No
Mean			
Observations	744,511	744,511	744,511
F-Stat	1,440.5	1,294.9	6,877.4

This table shows the change in the log wage for a change in the log outside occupation options index (OOI) and the log Herfindahl-Hirschman Index (HHI). Panl A is an OLS estimation, and Panel B is a 2SLS IV estimation. The OOI instrument is made following Schubert, Stansbury, and Taska (2022). The HHI instrument is made following Marinescu, Ouss, and Pape (2021) (MOP). The outcome, explanatory variables, and instruments are all at the commuting zone (CZ) by occupation by year level. Standard errors clustered at the CZ level in parentheses. The reported F-stat for the 2SLS IV regressions is the Kleibergen-Paap Wald F statistic. * ($p<0.10$), ** ($p<0.05$), *** ($p<0.01$).

Table C10: Individual Wages, HHI, and OOI, MOP HHI, Individual Sample

Dependent variable: Log Wage			
	(1)	(2)	(3)
Panel A: OLS regression			
Log(OOI)	0.016***	0.015***	0.047***
	(0.003)	(0.002)	(0.002)
Log(HHI)	-0.002*	-0.001*	0.001**
	(0.001)	(0.001)	(0.001)
Panel B: 2SLS IV regression			
Log(OOI)	0.009***	0.007**	0.044***
	(0.003)	(0.003)	(0.002)
Log(HHI)	-0.011***	-0.011***	-0.000
	(0.004)	(0.004)	(0.002)
Fixed effects	CZ, Occ Year Person	Region-Occ CZ-Year Person	CZ, Occ Year
Mean			
Observations	9,027,295	9,027,280	9,311,392
F-Stat	2,108.2	1,274.0	2,913.7

This table shows the change in the log wage for a change in the log outside occupation options index (OOI) and the log Herfindahl-Hirschman Index (HHI). Panl A is an OLS estimation, and Panel B is a 2SLS IV estimation. The OOI instrument is made following Schubert, Stansbury, and Taska (2022). The HHI instrument is made following Marinescu, Ouss, and Pape (2021) (MOP). Wages are at the individual by year level. The explanatory variables and instruments are all at the commuting zone (CZ) by occupation by year level. Standard errors clustered at the CZ level in parentheses. The reported F-stat for the 2SLS IV regressions is the Kleibergen-Paap Wald F statistic. * ($p<0.10$), ** ($p<0.05$), *** ($p<0.01$).

First Stages

I present first-stage regressions for my instruments in the tables below. Tables C11 and C12 are first-stage regressions using the HHI instrument from Schubert, Stansbury and Taska (2022). Tables C13 and C14 use the HHI instrument from Marinescu, Ouss, and Pape (2021). Across all specifications, the relationship between the OOI Instrument and the OOI is strong and similar in magnitude to that in Schubert, Stansbury and Taska (2022). The HHI instrument also has a statistically significant relationship with HHI, though it is smaller than that in Schubert, Stansbury and Taska (2022). It is stronger in the case of the Marinescu, Ouss, and Pape (2021) instrument, though the empirical specification is quite different.

Table C11: First Stage: Outside Occupation Options Index, SST HHI

Dependent variable: Log Outside Occupation Options Index				
	(1)	(2)	(3)	(4)
Panel A: OLS regression				
Log(HHI)	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Log(OOI)	0.903***	0.903***	0.941***	0.917***
	(0.022)	(0.022)	(0.015)	(0.018)
Predicted Vacancy Growth	-0.022*	-0.033***	-0.008	-0.026*
	(0.013)	(0.013)	(0.006)	(0.014)
Vacancy Growth	0.125**	0.122**	0.035*	0.142**
	(0.056)	(0.053)	(0.021)	(0.059)
Expected HHI	0.049	0.042	0.005	0.043
	(0.038)	(0.032)	(0.006)	(0.032)
Sample	CZ-Year-Occ	CZ-Year-Occ	Individual	Individual
Fixed effects	CZ-Year	CZ, Occ	CZ-Year	CZ-Year
	Occ-Year	Year	Occ-Year	Occ-Year
			Person	
Employment Weighting	Yes	Yes	NA	NA
HHI Inst Controls	Yes	Yes	Yes	Yes
Observations	105,922	105,922	4,574,692	4,936,627

The outcome variable in this table is the Log Outside Occupation Options Index (OOI). This table shows the first stage estimations for the 2SLS IV regressions of log wages on log OOI and log Herfindahl-Hirschman Index (HHI). The HHI and OOI instruments are made following Schubert, Stansbury, and Taska (2022) (SST). Standard errors clustered at the CZ level in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01).

Table C12: First Stage: HHI, SST HHI

Dependent variable: Log HHI	(1)	(2)	(3)	(4)
Panel A: OLS regression				
Log(HHI)	0.025***	0.029***	0.012***	0.026***
	(0.004)	(0.003)	(0.002)	(0.003)
Log(OOI)	-0.391***	-0.392***	-0.464***	-0.412***
	(0.031)	(0.032)	(0.022)	(0.023)
Predicted Vacancy Growth	-0.039	0.041	0.066	-0.024
	(0.067)	(0.133)	(0.057)	(0.062)
Vacancy Growth	4.046***	3.680***	4.461***	4.320***
	(0.366)	(0.310)	(0.386)	(0.391)
Expected HHI	0.392***	0.229**	0.180**	0.414***
	(0.110)	(0.095)	(0.074)	(0.107)
Sample	CZ-Year-Occ	CZ-Year-Occ	Individual	Individual
Fixed effects	CZ-Year	CZ, Occ	CZ-Year	CZ-Year
	Occ-Year	Year	Occ-Year	Occ-Year
			Person	
Employment Weighting	Yes	Yes	NA	NA
HHI Inst Controls	Yes	Yes	Yes	Yes
Observations	105,922	105,922	4,574,692	4,936,627

The outcome variable in this table is the Log HHI. This table shows the first stage estimations for the 2SLS IV regressions of log wages on log OOI and log Herfindahl-Hirschman Index (HHI). The HHI and OOI instruments are made following Schubert, Stansbury, and Taska (2022) (SST). Standard errors clustered at the CZ level in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01).

Table C13: First Stage: Outside Occupation Options Index, MOP HHI

Dependent variable: Log Outside Occupation Options Index				
	(1)	(2)	(3)	(4)
Panel A: OLS regression				
HHI Instrument	0.004***	0.005***	0.006	0.009***
	(0.001)	(0.001)	(0.004)	(0.003)
Log(OOI Instrument)	0.935***	0.935***	0.948***	0.945***
	(0.014)	(0.014)	(0.008)	(0.010)
Sample	CZ-Year-Occ	CZ-Year-Occ	Individual	Individual
Fixed effects	CZ-Year	CZ, Occ	CZ, Occ	CZ, Occ
	Occ	Year	Year	Year
			Person	
Employment Weighting	Yes	Yes	NA	NA
Observations	744,511	744,511	9,027,295	9,311,392

The outcome variable in this table is the Log Outside Occupation Options Index (OOI). This table shows the first stage estimations for the 2SLS IV regressions of log wages on log OOI and log Herfindahl-Hirschman Index (HHI). The OOI instrument is made following Schubert, Stansbury, and Taska (2022). The HHI instrument is made following Marinescu, Ouss, and Pape (2021) (MOP). Standard errors clustered at the CZ level in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01). * (p<0.10), ** (p<0.05), *** (p<0.01).

Table C14: First Stage: HHI, MOP HHI

Dependent variable: Log HHI				
	(1)	(2)	(3)	(4)
Panel A: OLS regression				
HHI Instrument	1.345***	1.346***	0.842***	0.861***
	(0.027)	(0.025)	(0.014)	(0.012)
Log(OOI Instrument)	-0.605***	-0.604***	-0.622***	-0.608***
	(0.023)	(0.023)	(0.041)	(0.026)
Sample	CZ-Year-Occ	CZ-Year-Occ	Individual	Individual
Fixed effects	CZ-Year	CZ, Occ	CZ, Occ	CZ, Occ
	Occ	Year	Year	Year
			Person	
Employment Weighting	Yes	Yes	NA	NA
Observations	744,511	744,511	9,027,295	9,311,392

The outcome variable in this table is the Log HHI. This table shows the first stage estimations for the 2SLS IV regressions of log wages on log OOI and log Herfindahl-Hirschman Index (HHI). The OOI instrument is made following Schubert, Stansbury, and Taska (2022). The HHI instrument is made following Marinescu, Ouss, and Pape (2021) (MOP). Standard errors clustered at the CZ level in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01). * (p<0.10), ** (p<0.05), *** (p<0.01).

Back of the Envelope Calculations

Assumptions

In my back of the envelope calculation of indirect wage decreases due to climate change, I make a number of assumptions:

1. There is an initial 1 percent decrease in the wages of all high exposure occupations and no change to low exposure occupation wages.
2. The geographical distribution, and therefore the local relative employment shares ($\frac{s_{p,m,t}}{s_{p,t}}$), of occupations do not change.
3. National-level flows of workers between occupation pairs do not change.

I discuss each of these assumptions in turn.

1) A climate shock is unlikely to uniformly affect all high exposure occupations and not affect low exposure occupations at all. In particular, two occupations on either side of the $\theta = 50$ threshold are likely to be affected more similarly than an occupation with $\theta = 50$ and $\theta = 90$. Here, I show how the calculation would differ if the initial climate shock was proportional to each occupation's θ:

$$\Delta_{\%} w_{p \neq o} = \eta \cdot \theta_p,$$

where η is a scaling factor that determines the size of the shock relative to θ. For the purpose of this exercise, I set $\eta = \frac{1}{0.75}$, so that an occupation at $\theta = 75$ (the midpoint of the high exposure range) experiences a direct wage decrease of 1 percent. I calculate the percentage change in each occupation's OOI after this shock, then multiply it by 0.052 to get its percentage wage change.

As I show in Figure C17, I find that the damages are still higher for the lower wage groups. However, this difference is lower than in the case of a uniform 1 percent shock to

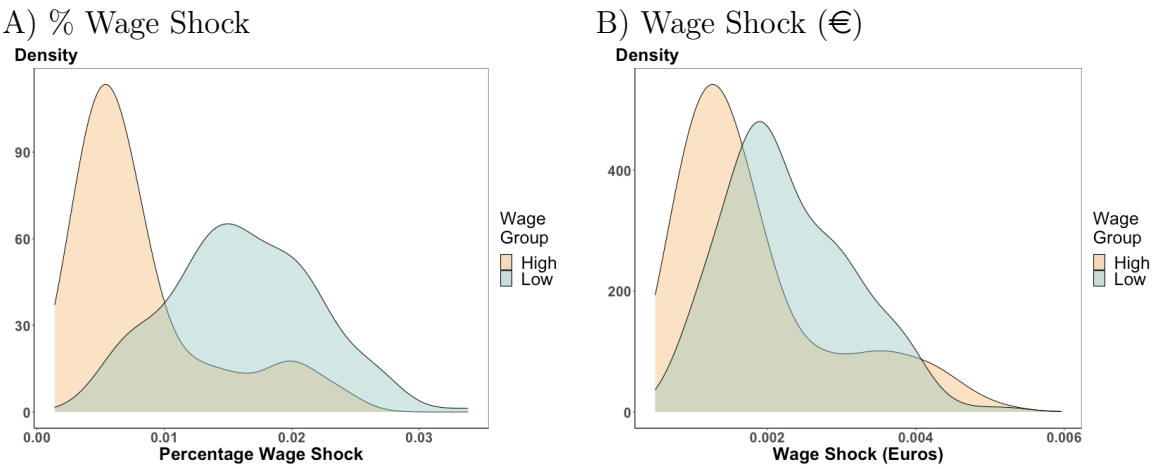


Figure C17: Indirect Climate Shock Effects to Low Exposure Occupations: Alternative Shock

Panel A graphs the employment-weighted distribution of indirect wage percentage shocks after a 1% decrease in all high exposure occupations' wages. I show the distribution separately for low exposure occupations with above and below median wages. Panel B graphs the wage shock in 2019 €.

high exposure wages. For example, the median worker in the low wage group experiences an indirect wage loss which is 2.6 times as large as the median high wage worker. In the original scenario, this difference was a factor of 4.7.

2) Relative employment of high exposure occupations will likely change as the climate warms. For example, if a high exposure occupation is involved in the production of a tradable good, firms may reallocate production to a more temperate area. This reallocation would result in an increase in the relative employment share of high exposure occupations in temperature areas, and a decrease in warmer areas.

This reallocation interacts with the assumption about the nature of the climate shock as well. For example, a high exposure occupation in a relatively temperature area may not experience much of a shock at all. This point suggests that reallocation of workers could mitigate spillovers to low exposure workers. Specifically, the influence of high exposure outside options would be lower in areas with higher climate shocks. In this case, my calculations could be considered an upper bound of the actual spillover effects. However, the distribution of these spillover effects within low exposure occupations is not clearly affected by this

reallocation. In all likelihood, low wage low exposure occupations would still experience the bulk of the spillovers. In addition, many industries (such as construction) with high levels of occupational exposure cannot easily reallocate the bulk of their production across space.

3) In the future, national-level inter-exposure mobility is likely to be different than it is today, though the sign of the change is difficult to predict. For example, a warmer climate might decrease the flow of workers from low to high exposure occupations. However, a decrease in the prevalence of routine cognitive occupations could increase this flow. For the purposes of this exercise, I abstract away from any future changes in inter-exposure mobility.

Within-Occupation Differences: Robustness

In Section 3.4, I present evidence on the occupational mobility behavior of workers across their within-occupation earnings distribution. The levels of occupational mobility are not exactly comparable between my study and Groes, Kircher, and Manovskii (2015) (GKM). For one, the occupation codes I use are slightly more disaggregated than those in GKM. While I have 368 occupations in my sample, they have 271. This helps explain the fact that, while France is known to have low levels of occupational mobility (Bachmann et al., 2020) and Denmark is known to have high levels (Groes et al., 2015), I find a slightly higher rate of workers leaving overall.

However, the patterns of occupational mobility are more easily comparable. I show that, compared to Denmark, low earning workers in France are relatively unlikely (relative to other workers) to change occupations. There are three main reasons that this might be the case. First, France has slightly higher rates of coverage by collective agreements and lower rates of low wage workers than does Denmark (Bosch and Weinkopf, 2017). Second, France has a high and dynamic minimum wage (Bozio et al., 2020). These factors mean that low earners may be unlikely to have an incentive to change occupations, as the earnings floor is relatively high. Third, younger workers may have different occupational mobility patterns than older workers. My sample is considerably older than the one in GKM, which limits workers to a

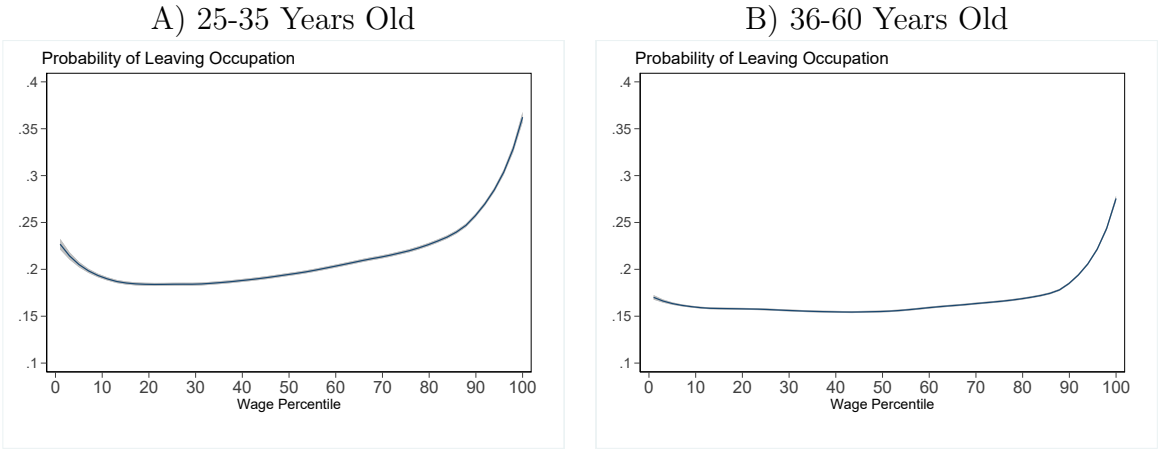


Figure C18: Occupational Mobility vs. Relative Earnings (1)

Panel A shows the probability that a worker leaves their occupation compared to their position in their within-occupation wage distribution. It is limited to workers between 25 and 35 years old. The wage percentile is calculated within each occupation and year based on raw wages. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95% confidence intervals are shaded in gray. Panel B is the same as Panel A, but for workers between 36 and 60 years old.

maximum of 15 years after entering the labor market.

I directly test the possibility that younger workers have different mobility patterns by separating workers into those 35 and below and those above 35. I choose this threshold so that the average age in my younger sample matches that in GKM (30 - I only have workers at least 25 years old in my sample). Figure C18 shows that relatively low earners in the younger group do leave their occupation more often than middle-wage earners. However, the highest earners still leave their occupations at significantly higher rates than do the lowest earners. Due to differences in rates of occupational departure between the two groups, I confirm that inter-exposure mobility, given an occupational departure, is not different between these two groups. Figure C19 shows that patterns of within-occupation inter-exposure mobility are consistent across age groups.

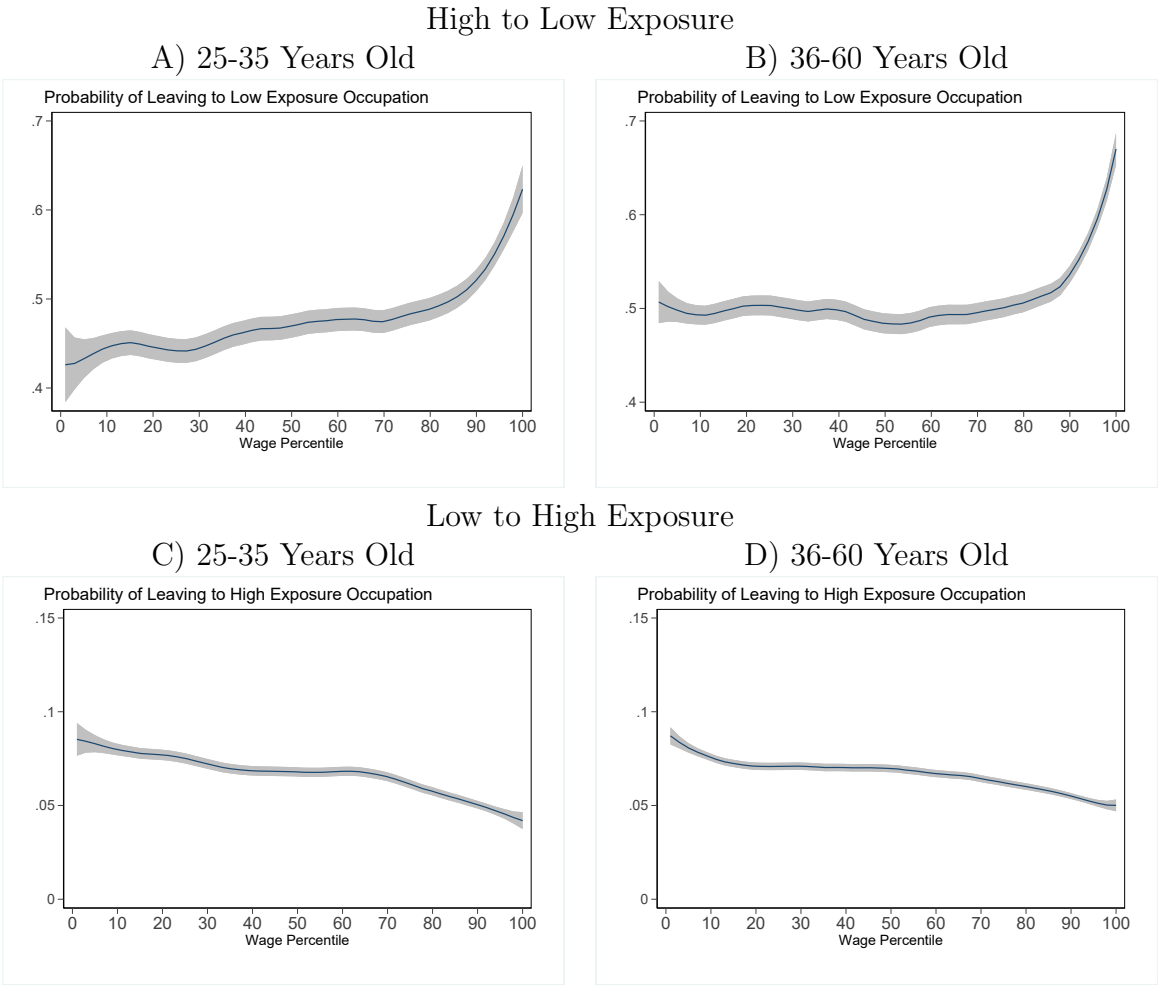


Figure C19: Occupational Mobility vs. Relative Earnings (2)

Panel A shows the probability that a worker leaves to a low exposure occupation compared to their position in their within-occupation wage distribution. It is limited to workers between 25 and 35 years old. The wage percentile is calculated within each occupation and year based on raw wages. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95% confidence intervals are shaded in gray. Panel B is the same as Panel A, but for workers between 36 and 60 years old. Panels C and D are the same as Panels A and B, respectively, but for workers leaving from a low exposure to a high exposure occupation.

Raw Wages and All Workers

In this section, I present robustness checks to the within-occupation differences in occupational mobility. In the main specification, I calculate each worker's position in their within-occupation wage distribution using the residuals from equation 4.4 for each year and occupation. Here, I calculate their position using the raw wages within each year and occupation. Overall, the results (in Figures C20 and C21) are very similar to those from the main specification.

Next, I present the results including both men and women. This is only expected to matter, if at all, for the aggregate results, because 95 percent of high exposure jobs in my sample are conducted by men. In constructing the residual wages, I control for the gender of the worker. I therefore drop any occupations if there are fewer than 10 workers per year of the same gender.[4]

Overall, the results for the full sample of workers is very similar. The one notable difference is in Figure C23C, which graphs a worker's probability of moving to a high exposure occupation when leaving a low exposure one vs. their relative earnings. As expected, in general, the probabilities of a worker transferring to a high exposure occupation are lower across the board relative to the all-male sample. However, the overall pattern remains the same: low earners are more likely than high earners to transfer to a high exposure occupation.

[4]Because there are many more observations in general, I only drop 2 occupations for not having 10 observations per year: 352b and 441b. I drop 671d for not having at least 10 women per year.

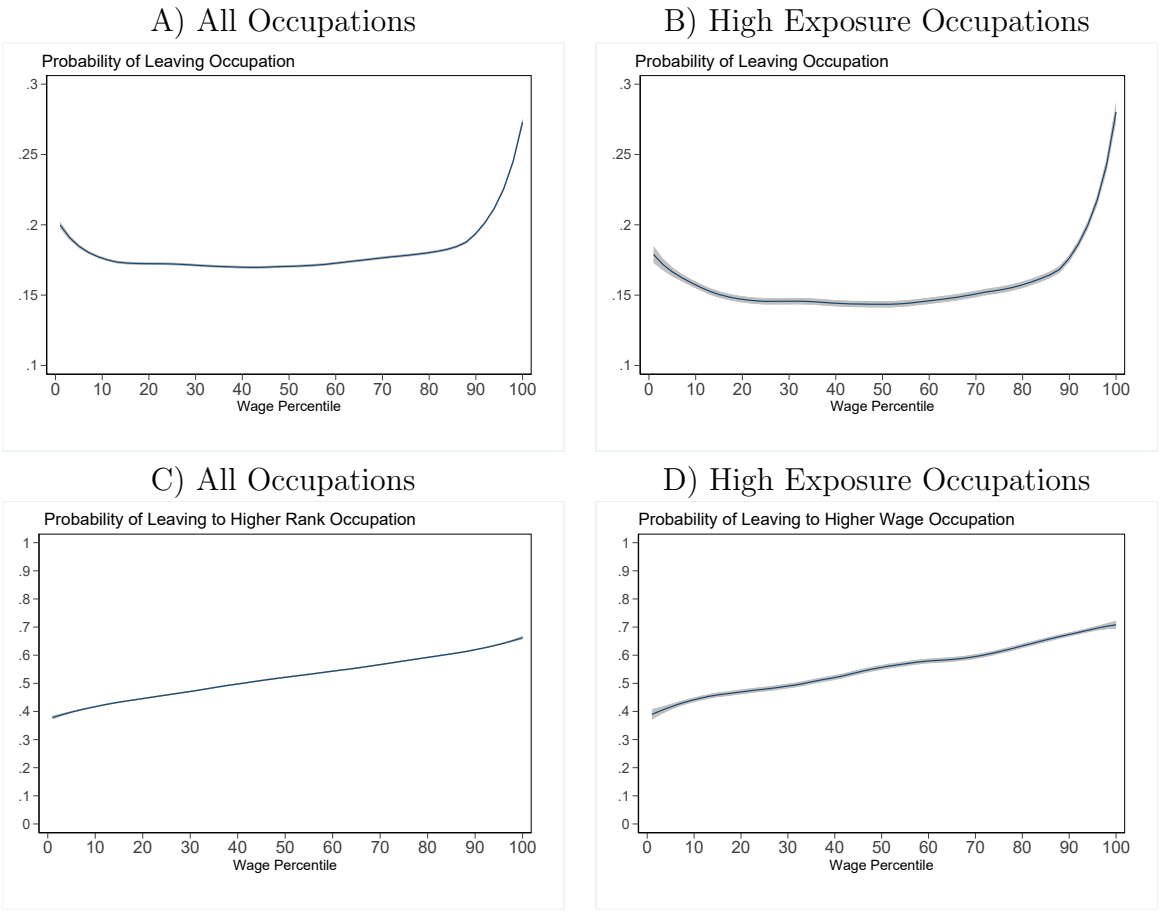


Figure C20: Robustness: Raw Wages (1)

Panel A shows the probability that a worker leaves their occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on raw wages. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95% confidence intervals are shaded in gray. Panel B is the same as Panel A, but limits the sample to high exposure occupations. Panel C shows the probability of a departing worker changing to an occupation with a higher rank (in the year of the switch). Occupation ranks are determined by comparing year dummy coefficients from equation 4.4. The sample is limited to workers leaving their occupations. Panel D is the same as Panel A, but limits the sample to high exposure occupations. The sample in Panel A includes 3,721,121 observations for 777,715 male workers, Panel B includes 519,278 observations for 141,225 male workers, Panel C includes 668,479 observations for 349,982 male workers, and Panel D includes 81,333 observations for 60,363 male workers.

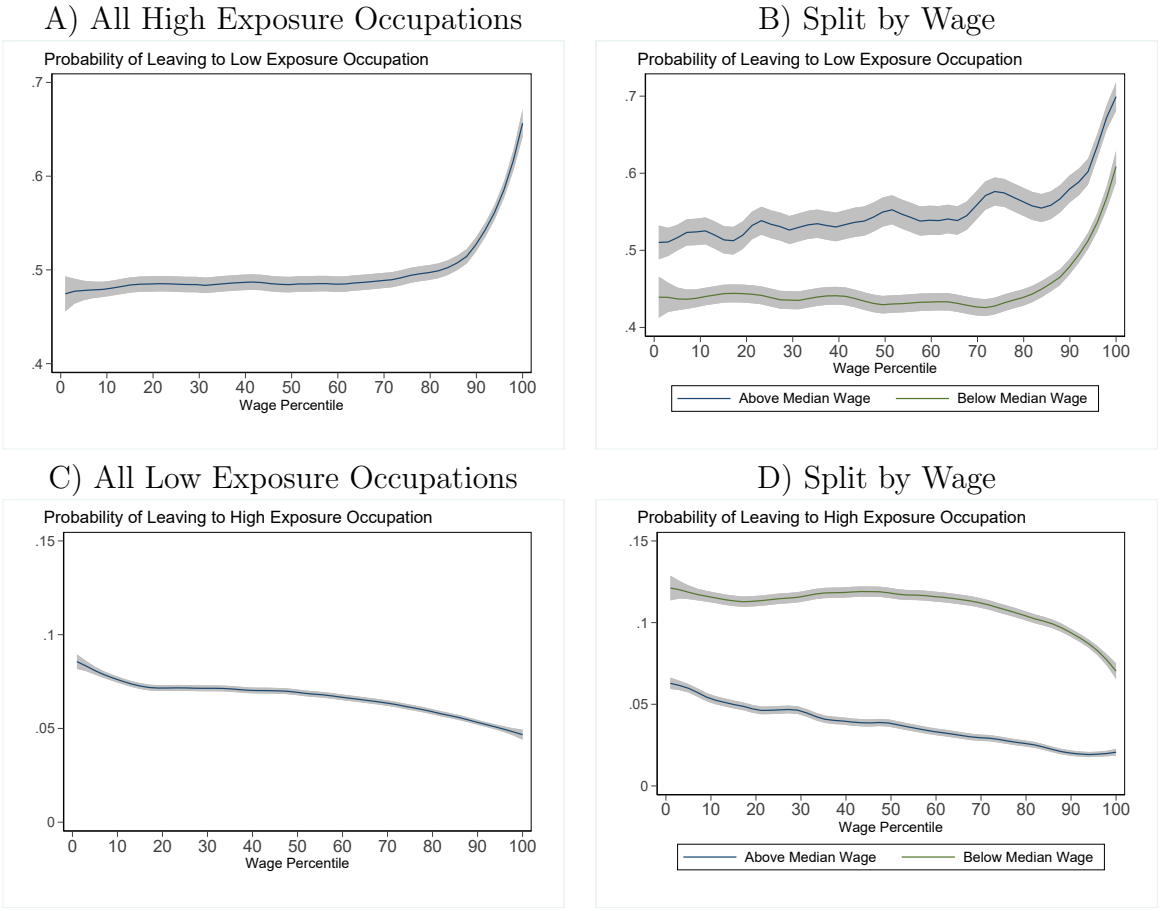


Figure C21: Robustness: Raw Wages (2)

Panel A shows the probability that a worker leaves to a low exposure occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on raw wages. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95% confidence intervals are shaded in gray. The sample is limited to workers leaving a high exposure occupation. Panel B shows the same graphs, with occupations separated into occupations above and below the median of average wages for high exposure occupations. Panel C shows the probability that a worker leaves to a high exposure occupation compared to their position in their within-occupation wage distribution. The sample is limited to workers leaving a low exposure occupation. Panel D shows the same graphs, with occupations separated into occupations above and below the median of average wages for low exposure occupations. The sample in Panel A includes 81,333 observations for 60,363 male workers. The blue line in Panel B (above the median) includes 39,210 observations for 32,176 male workers. The green line in Panel B (below the median) includes 42,123 observations for 34,669 male workers. The sample in Panel C includes 587,146 observations for 316,446 male workers. The blue line in Panel D (above the median) includes 352,054 observations for 200,669 male workers. The green line in Panel D (below the median) includes 235,092 observations for 159,807 male workers.

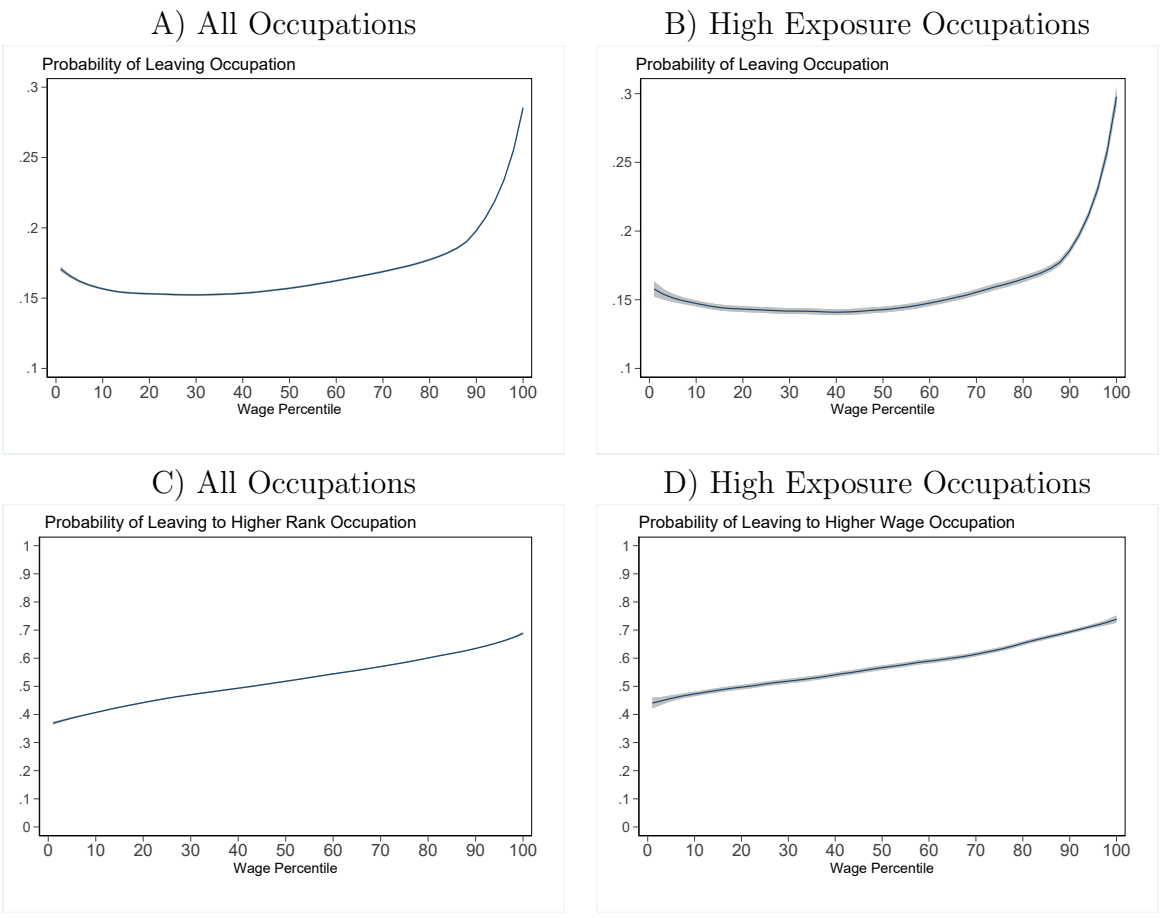


Figure C22: Robustness: All Workers (1)

Panel A shows the probability that a worker leaves their occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on raw wages. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95% confidence intervals are shaded in gray. Panel B is the same as Panel A, but limits the sample to high exposure occupations. Panel C shows the probability of a departing worker changing to an occupation with a higher rank (in the year of the switch) compared to their position in their within-occupation wage distribution. Occupation ranks are determined by comparing year dummy coefficients from equation 4.4. The sample is limited to workers leaving their occupations. Panel D is the same as Panel C, but limits the sample to high exposure occupations. The sample in Panel A includes 6,320,861 observations for 1,343,254 workers, Panel B includes 543,665 observations for 149,909 workers, Panel C includes 1,064,082 observations for 572,755 workers, and Panel D includes 85,286 observations for 63,828 workers.

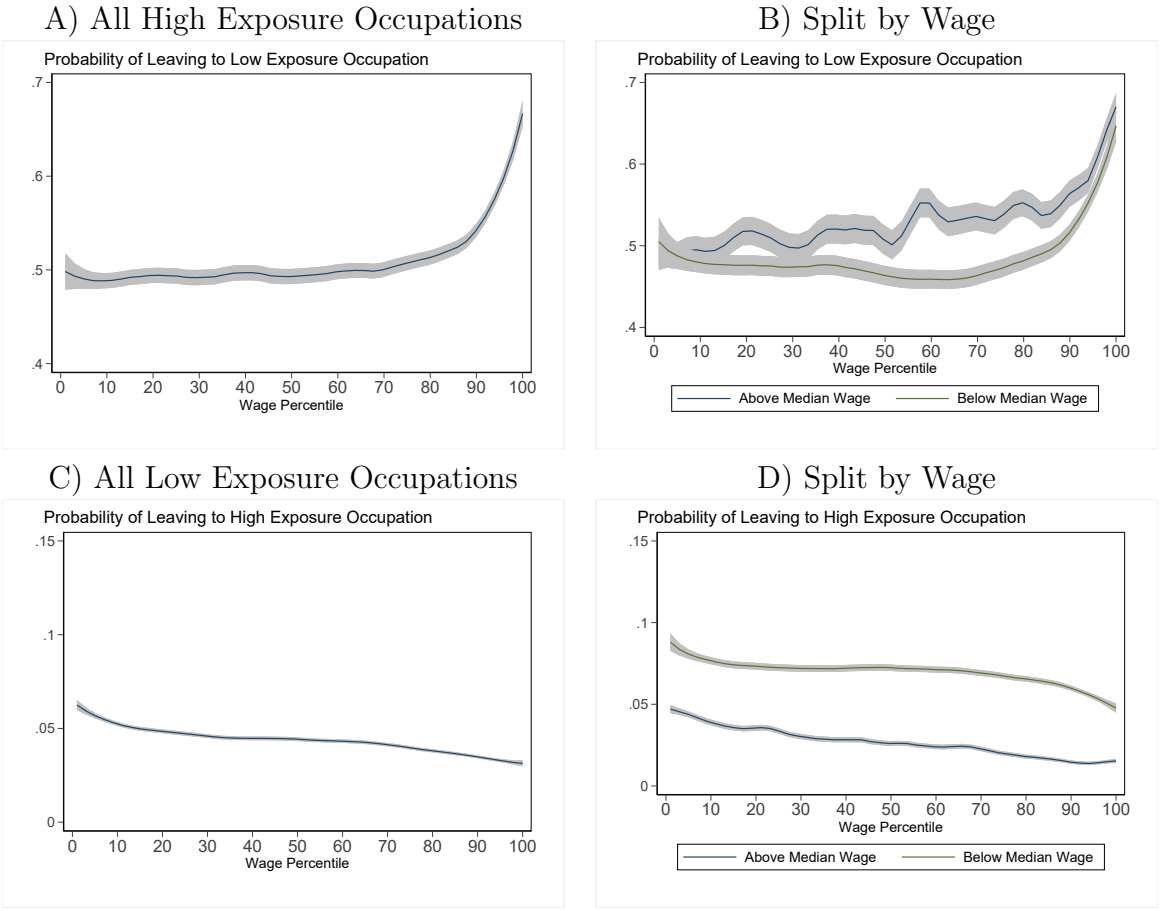


Figure C23: Robustness: All Workers (2)

Panel A shows the probability that a worker leaves to a low exposure occupation compared to their position in their within-occupation wage distribution. The wage percentile is calculated within each occupation and year based on raw wages. The plot is a kernel smoothed local linear regression with a bandwidth of 5 percentiles. 95% confidence intervals are shaded in gray. The sample is limited to workers leaving a high exposure occupation. Panel B shows the same graphs, with occupations separated into occupations above and below the median of average wages for high exposure occupations. Panel C shows the probability that a worker leaves to a high exposure occupation compared to their position in their within-occupation wage distribution. The sample is limited to workers leaving a low exposure occupation. Panel D shows the same graphs, with occupations separated into occupations above and below the median of average wages for low exposure occupations. The sample in Panel A includes 85,286 observations for 63,828 workers. The blue line in Panel B (above the median) includes 45,257 observations for 36,522 workers. The green line in Panel B (below the median) includes 40,029 observations for 33,846 workers. The sample in Panel C includes 978,796 observations for 537,777 workers. The blue line in Panel D (above the median) includes 591,462 observations for 343,914 workers. The green line in Panel D (below the median) includes 387,334 observations for 274,746 workers.

Bibliography

Adhvaryu, A., Kala, N., and Nyshadham, A. (2020). The Light and the Heat: Productivity Co-Benefits of Energy-Saving Technology. *The Review of Economics and Statistics*, 102(October):779–792.

Agbim, C., Araya, F., Faust, K. M., and Harmon, D. (2020). Subjective versus objective energy burden: A look at drivers of different metrics and regional variation of energy poor populations. *Energy Policy*, 144:111616.

Anderson, G. B., Dominici, F., Wang, Y., McCormack, M. C., Bell, M. L., and Peng, R. D. (2013). Heat-related emergency hospitalizations for respiratory diseases in the Medicare population. *American journal of respiratory and critical care medicine*, 187(10):1098–1103.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, 24(2):3–30.

Aroonruengsawat, A. and Auffhammer, M. (2011). Impacts of climate change on residential electricity consumption: evidence from billing data. In Libecap, G. D. and Steckel, R. H., editors, *The Economics of Climate Change: Adaptations Past and Present*, page 311342. University of Chicago Press.

Auffhammer, M., Baylis, P., and Hausman, C. H. (2017). Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States. *Proceedings of the National Academy of Sciences*, 114(8):1886–1891.

Auffhammer, M., Hsiang, S. M., Schlenker, W., and Sobel, A. (2013). Using weather data and climate model output in economic analyses of climate change. *Review of Environmental Economics and Policy*.

Auffhammer, M. and Mansur, E. T. (2014). Measuring climatic impacts on energy consumption: A review of the empirical literature. *Energy Economics*, 46:522–530.

Autor, D. H. and Dorn, D. (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*.

Azam, M. and Kingdon, G. G. (2013). Are Girls the Fairer Sex in India ? Revisiting Intra-Household Allocation of Education Expenditure. *World Development*, 42:143–164.

Azar, J., Marinescu, I., and Steinbaum, M. (2019). Measuring labor market power two ways. In *AEA Papers and Proceedings*, volume 109, pages 317–321. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Bachmann, R., Bechara, P., and Vonnahme, C. (2020). Occupational Mobility in Europe : Extent , Determinants and Consequences. *De Economist*, 168(1):79–108.

Barreca, A., Clay, K., Deschenes, O., Greenstone, M., and Shapiro, J. S. (2016). Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century. *Journal of Political Economy*, 124(1):105–159.

Barreca, A. and Schaller, J. (2020). The impact of high ambient temperatures on delivery timing and gestational lengths. *Nature Climate Change*, 10(January).

Barreca, A. I. (2012). Climate change, humidity, and mortality in the United States. *Journal of Environmental Economics and Management*, 63(1):19–34.

Beatty, T. K., Blow, L., and Crossley, T. F. (2014). Is there a heat-or-eattrade-off in the uk? *Journal of the Royal Statistical Society Series A: Statistics in Society*, 177(1):281–294.

Bednar, D. J. and Reames, T. G. (2020). Recognition of and response to energy poverty in the United States. *Nature Energy*, 5(6):432–439.

Berger, D., Herkenhoff, K., and Mongey, S. (2022). Labor market power. *American Economic Review*, 112(4):1147–1193.

Bhattacharya, J., DeLeire, T., Haider, S., and Currie, J. (2003). Heat or eat? Cold-weather shocks and nutrition in poor American families. *American Journal of Public Health*, 93(7):1149–1154.

Blanchette, A. (2020). *Porkopolis: American Animality, Standardized Life, and the Factory Farm*. Duke University Press.

Bolotova, Y. V. (2022). Competition issues in the us beef industry. *Applied Economic Perspectives and Policy*, 44(3):1340–1358.

Borusyak, K. and Hull, P. (2021). Non-Random Exposure to Exogenous Shocks: Theory and Applications. *NBER Working Paper*.

Bosch, G. and Weinkopf, C. (2017). Reducing Wage Inequality: The Role of the State in Improving Job Quality. *Work and Occupations*, 44(1) 68-8.

Bozio, A., Guillot, M., and Bozio, A. (2020). The Contribution of Payroll Taxation to Wage Inequality in France. *IZA Discussion Paper*, (13317).

Brockway, A. M. and Dunn, L. N. (2020). Weathering adaptation: Grid infrastructure planning in a changing climate. *Climate risk management*, 30:100256.

Brown, M. A., Soni, A., Lapsa, M. V., Southworth, K., and Cox, M. (2020). High energy burden and low-income energy affordability: Conclusions from a literature review. *Progress in Energy*, 2(4):042003.

Burillo, D., Chester, M. V., Ruddell, B., and Johnson, N. (2017). Electricity demand planning forecasts should consider climate non-stationarity to maintain reserve margins during heat waves. *Applied Energy*, 206:267–277.

Burke, M., Dykema, J., Lobell, D. B., Miguel, E., and Satyanath, S. (2015). Incorporating climate uncertainty into estimates of climate change impacts. *Review of Economics and Statistics*, 97(2):461–471.

Burke, M. and Emerick, K. (2016). Adaptation to climate change: Evidence from us agriculture. *American Economic Journal: Economic Policy*, 8(3):106–140.

Burke, P. J. and Abayasekara, A. (2018). The price elasticity of electricity demand in the United States: A three-dimensional analysis. *The Energy Journal*, 39(2):123–146.

Cachon, G. P., Gallino, S., and Olivares, M. (2012). Severe weather and automobile assembly productivity. *Columbia Business School Research Paper*, (12/37).

Caldwell, S. and Danieli, O. (2024). Outside options in the labor market. *Review of Economic Studies*, page rdae006.

California Energy Commission (2022). Building Climate Zones by Zip Code.

California Public Utilities Commission (2009). Decision Adopting Settlements on Marginal Cost, Revenue Allocation, and Rate Design.

Carleton, T., Jina, A., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Kopp, R. E., Mccusker, K. E., Nath, I., Rising, J., Rode, A., Seo, H. K., Viaene, A., Yuan, J., and Zhang, A. T. (2022). Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits. *The Quarterly Journal of Economics*, pages 2037–2105.

Carleton, T. A. and Hsiang, S. M. (2016). Social and economic impacts of climate. *Science*, 353(6304):aad9837.

Carpena, F. (2019). How Do Droughts Impact Household Food Consumption and Nutritional Intake? A Study of Rural India. *World Development*, 122(226057):349–369.

Carpenter, C. and Dobkin, C. (2009). The effect of alcohol consumption on mortality: regression discontinuity evidence from the minimum drinking age. *American Economic Journal: Applied Economics*, 1(1):164–182.

Castells-Quintana, D., Lopez-U, Ribe, P., and Mcdermott, T. K. J. (2018). Adaptation to climate change : A review through a development economics lens. *World Development*, 104:183–196.

Chaparro, C. M. and Suchdev, P. S. (2019). Anemia epidemiology, pathophysiology, and etiology in low- and middle-income countries. *Annals of the New York Academy of Sciences*, 1450(1):15–31.

Chen, M., Sanders, K. T., and Ban-Weiss, G. A. (2019). A new method utilizing smart meter data for identifying the existence of air conditioning in residential homes. *Environmental Research Letters*, 14(9):094004.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

Choi, C.-Y. and Chudik, A. (2019). Estimating impulse response functions when the shock series is observed. *Economics Letters*, 180:71–75.

Chuang, Y., Delmas, M. A., and Pincetl, S. (2022). Are residential energy efficiency upgrades effective? An empirical analysis in Southern California. *Journal of the Association of Environmental and Resource Economists*, 9(4):641–679.

Clemens, J. (2021). How do firms respond to minimum wage increases? understanding the relevance of non-employment margins. *Journal of Economic Perspectives*, 35(1):51–72.

Coffman, D. J. (2022). The excluded workers: The nlra, farm laborers, and a lineage of exploitation. *Drake J. Agric. L.*, 27:85.

Copernicus Climate Change Service (1979-2012). ECMWF Reanalysis v5 (ERA5). Accessed April 26, 2021.

Cortes, G. M. and Gallipoli, G. (2017). The Costs of Occupational Mobility: an Aggregate Analysis. 16(2):275–315.

Cullen, J. B., Friedberg, L., and Wolfram, C. (2004). Consumption and changes in home energy costs: How prevalent is the heat or eatdecision? *University of California, Berkeley manuscript*.

Currie, J. (2006). The take-up of social benefits. In Auerbach, A. J., Card, D., and Quigley, J. M., editors, *Public Policy and the Income Distribution*, pages 80–148. Russell Sage Foundation.

Davis, L. W. and Gertler, P. J. (2015). Contribution of air conditioning adoption to future energy use under global warming. *Proceedings of the National Academy of Sciences*, 112(19):5962–5967.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. World Bank Publications.

Deaton, A. (2024). Rethinking my economics.

Deaton, A. and Drèze, J. (2009). Food and nutrition in India: Facts and interpretations. *Economic and Political Weekly*, 44(7):42–65.

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic literature*, 52(3):740–798.

Deryugina, T., MacKay, A., and Reif, J. (2020). The long-run dynamics of electricity demand: Evidence from municipal aggregation. *American Economic Journal: Applied Economics*, 12(1):86–114.

Deschênes, O. and Greenstone, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American economic review*, 97(1):354–385.

Deschênes, O. and Greenstone, M. (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics*, 3(4):152–185.

Deser, C., Phillips, A., Bourdette, V., and Teng, H. (2012). Uncertainty in climate change projections: the role of internal variability. *Climate dynamics*, 38:527–546.

Dillender, M. (2021). Climate Change and Occupational Health : Are There Limits to Our Ability to Adapt? *The Journal of Human Resources*.

Dillon, A., McGee, K., and Oseni, G. (2015). Agricultural Production, Dietary Diversity and Climate Variability. *Journal of Development Studies*, 51(8):976–995.

Dimick, M. (2023). Conflict of laws? tensions between antitrust and labor law. *The University of Chicago Law Review*, 90(2):379–436.

Dutta, G. and Mitra, K. (2017). A literature review on dynamic pricing of electricity. *Journal of the Operational Research Society*, 68:1131–1145.

Eichhorst, W., Marx, P., and Wehner, C. (2017). Labor market reforms in Europe: towards more flexicure labor markets ? *Journal for Labour Market Research*.

Energy Information Administration (2018). One in Three U.S. Households Faces a Challenge in Meeting Energy Needs.

Energy Insights USA, D. (2020). 2019 California Residential Appliance Saturation Study.

Evans, G. W. (2016). Childhood poverty and adult psychological well-being. *Proceedings of the National Academy of Sciences*, 113(52):14949–14952.

Farbotko, C., Kitara, T., Dun, O., and Evans, C. (2022). A climate justice perspective on international labour migration and climate change adaptation among Tuvaluan workers. *Oxford Open Climate Change*, 2(October 2021):1–15.

Fournier, E. D., Federico, F., Porse, E., and Pincetl, S. (2019). Effects of building size growth on residential energy efficiency and conservation in California. *Applied energy*, 240:446–452.

Frank, D. A., Neault, N. B., Skalicky, A., Cook, J. T., Wilson, J. D., Levenson, S., Meyers, A. F., Heeren, T., Cutts, D. B., Casey, P. H., et al. (2006). Heat or eat: the low income home energy assistance program and nutritional and health risks among children less than 3 years of age. *Pediatrics*, 118(5):e1293–e1302.

Garg, T., Gibson, M., and Sun, F. (2020a). Extreme temperatures and time use in china. *Journal of Economic Behavior & Organization*, 180:309–324.

Garg, T., Jagnani, M., and Taraz, V. (2020b). Temperature and human capital in india. *Journal of the Association of Environmental and Resource Economists*, 7(6):1113–1150.

Gathmann, C. and Schönberg, U. (2010). How General is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, 28(1).

Georgieff, M. K. (2020). Iron deficiency in pregnancy. *American Journal of Obstetrics and Gynecology*, 223(4):516–524.

Goldenson, N., Leung, L. R., Mearns, L. O., Pierce, D. W., Reed, K. A., Simpson, I. R., Ullrich, P., Krantz, W., Hall, A., Jones, A., et al. (2023). Use-inspired, process-oriented gcm selection: Prioritizing models for regional dynamical downscaling. *Bulletin of the American Meteorological Society*, 104(9):E1619–E1629.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–2526.

Gopalan, C., Rama Sastri, V., Balasubramanian, S., Narasinga Rao, B., Deosthale, Y., and Pant, K. (1989). Nutritive Value of Indian Foods.

Graff Zivin, J. and Neidell, M. (2014). Temperature and the Allocation of Time : Implications for Climate Change. *Journal of Labor Economics*, 32(1).

Graff Zivin, J. and Neidell, M. (2014). Temperature and the allocation of time: Implications for climate change. *Journal of Labor Economics*, 32(1):1–26.

Greening, L. A., Greene, D. L., and Difiglio, C. (2000). Energy efficiency and consumptionthe rebound effecta survey. *Energy policy*, 28(6-7):389–401.

Groes, F., Kircher, P., and Manovskii, I. (2015). The U-Shapes of Occupational Mobility. *Review of Economic Studies*, 82(October 2014):659–692.

Harrigan, J., Reshef, A., and Toubal, F. (2021). The March of the Techies: Job Polarization Within and Between Firms. *Research Policy*, 50(7):104008.

Harrison, C. and Popke, J. (2013). "Because you got to have heat": the networked assemblage of energy poverty in eastern north carolina. In *The New Geographies of Energy*, pages 248–260. Routledge.

Hatton, E. (2020). *Coerced: Work Under Threat of Punishment*. Univ of California Press.

Heflin, C., London, A. S., and Scott, E. K. (2011). Mitigating material hardship: The strategies low-income families employ to reduce the consequences of poverty. *Sociological Inquiry*, 81(2):223–246.

Heilmann, K., Kahn, M. E., and Tang, C. K. (2021). The urban crime and heat gradient in high and low poverty areas. *Journal of Public Economics*, 197:104408.

Hensvik, L. and Nordström, O. (2023). The Skill-Specific Impact of Past and Projected Occupational Decline. *Labour Economics*, 81(January 2023).

Hernández, D. (2016). Understanding 'energy insecurity' and why it matters to health. *Social science & medicine*, 167:1–10.

Hernández, D. and Bird, S. (2010). Energy burden and the need for integrated low-income housing and energy policy. *Poverty & public policy*, 2(4):5–25.

Hernandez, T. and Gabbard, S. (2019). Findings from the national agricultural workers survey (naws) 2015–2016: a demographic and employment profile of united states farmworkers. *Dep Labor Employ Train Adm Wash Dist Columbia*.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Holmes, S. M. (2013). *Fresh Fruit, Broken Bodies: Migrant Farmworkers in the United States*. California Series in Public Anthropology.

Hou, X. (2010). Can Drought Increase Total Calorie Availability? The Impact of Drought on Food Consumption and the Mitigating Effects of a Conditional Cash Transfer Program. *Economic Development and Cultural Change*.

ICMR, I. C. o. M. R. (2010). Nutrient Requirements and Recommended Dietary Allowances for Indians.

IIPS and ICF (2017). National Family Health Survey (NFHS-4), 2015-16: India.

Isen, A., Rossin-Slater, M., and Walker, R. (2017). Relationship between season of birth, temperature exposure, and later life wellbeing. *Proceedings of the National Academy of Sciences*, 114(51):13447–13452.

Jayasinghe, M., Chai, A., Ratnasiri, S., and Smith, C. (2017). The power of the vegetable patch : How home-grown food helps large rural households achieve economies of scale & escape poverty. *Food Policy*, 73(April):62–74.

Jessel, S., Sawyer, S., and Hernández, D. (2019). Energy, poverty, and health in climate change: a comprehensive review of an emerging literature. *Frontiers in public health*, 7:470168.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182.

Kaul, T. (2018). Intra-household allocation of educational expenses : Gender discrimination and investing in the future. *World Development*, 104:336–343.

Ke, X., Wu, D., Rice, J., Kintner-Meyer, M., and Lu, N. (2016). Quantifying impacts of heat waves on power grid operation. *Applied energy*, 183:504–512.

Kim, J., Lee, A., and Rossin-Slater, M. (2021). What to expect when it gets hotter: the impacts of prenatal exposure to extreme temperature on maternal health. *American Journal of Health Economics*, 7(3):281–305.

Kishiyama, M. M., Boyce, W. T., Jimenez, A. M., Perry, L. M., and Knight, R. T. (2009). Socioeconomic disparities affect prefrontal function in children. *Journal of cognitive neuroscience*, 21(6):1106–1115.

Kotchen, M. J. (2017). Longer-run evidence on whether building energy codes reduce residential energy consumption. *Journal of the Association of Environmental and Resource Economists*, 4(1):135–153.

Kuang, J. (2024a). California tries again to protect workers from indoor heat - except in prisons.

Kuang, J. (2024b). California workers must wait even longer for indoor heat protections.

Kumar, A. (2020). *Monopsony capitalism: Power and production in the twilight of the sweatshop age*. Cambridge University Press.

Labandeira, X., Labeaga, J. M., and López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy policy*, 102:549–568.

Lane, D. J. R. and Richardson, D. R. (2014). The active role of vitamin C in mammalian iron metabolism: much more than just enhanced iron absorption! *Free Radical Biology and Medicine*, 75:69–83.

Le Barbanchon, T. and Rizzotti, N. (2020). The Task Content of French Jobs. *SSRN*.

Liu, M., Shamdasani, Y., and Taraz, V. (2023). Climate change and labor reallocation: Evidence from six decades of the indian census. *American Economic Journal: Economic Policy*, 15(2):395–423.

Longden, T., Quilty, S., Riley, B., White, L. V., Klerck, M., Davis, V. N., and Frank Jupurrurla, N. (2022). Energy insecurity during temperature extremes in remote Australia. *Nature Energy*, 7(1):43–54.

LoPalo, M. (2020). Temperature, Worker Productivity, and Adaptation: Evidence from Survey Data Production. *SSRN*, pages 1–67.

Luckstead, J. and Devadoss, S. (2019). The importance of h-2a guest workers in agriculture. *Choices*, 34(1):1–8.

Macaluso, C. (2023). Skill remoteness and post-layoff labor market outcomes.

Maestas, N., Mullen, K. J., Powell, D., Von Wachter, T., and Wenger, J. B. (2017). Working Conditions in the United States: Results of the 2015 American Working Conditions Survey. *RAND Labor and Population*.

Maestas, N., Mullen, K. J., Powell, D., Von Wachter, T., and Wenger, J. B. (2023). The value of working conditions in the united states and implications for the structure of wages. *American Economic Review*, 113(7):2007–2047.

Mahoney, E. (2020). Farm workers face double threat: Wildfire smoke and covid-19.

Marcus, H., Schauer, C., and Zlotkin, S. (2021). Effect of Anemia on Work Productivity in Both Labor- and Occupations : A Systematic Narrative Synthesis. *Food and Nutrition Bulletin*, 42(2):289–308.

Marinescu, I. and Hovenkamp, H. (2019). Anticompetitive mergers in labor markets. *Ind. LJ*, 94:1031.

Marinescu, I., Ouss, I., and Pape, L.-D. (2021). Wages, hires, and labor market concentration. *Journal of Economic Behavior and Organization*, 184:506–605.

Marks, A. B. (2022). Essential but ignored: Covid-19 litigation and the meatpacking industry. *NEULR*, 14:47.

McKinnon, K. A. and Deser, C. (2018). Internal variability and regional climate trends in an observational large ensemble. *Journal of Climate*, 31(17):6783–6802.

Mullainathan, S. and Shafir, E. (2013). *Scarcity: The new science of having less and how it defines our lives*. Times Books, Henry Holt.

Myers, E. (2020). Asymmetric information in residential rental markets: Implications for the energy efficiency gap. *Journal of Public Economics*, 190:104251.

Naidu, S., Posner, E. A., and Weyl, G. (2018). Antitrust remedies for labor market power. *Harvard law review*, 132(2):536–601.

National Consumer Law Center (2020). The Need for Utility Reporting of Key Credit and Collections Data Now and After the Covid-19 Crisis.

National Sample Survey Office (2003-2012). National Sample Survey, Schedule 1, Rounds 59-64, 66, 68. Accessed October 12, 2022.

Nord, M. and Kantor, L. S. (2006). Seasonal variation in food insecurity is associated with heating and cooling costs among low-income elderly Americans. *The Journal of nutrition*, 136(11):2939–2944.

Novan, K., Smith, A., and Zhou, T. (2022). Residential building codes do save energy: Evidence from hourly smart-meter data. *Review of Economics and Statistics*, 104(3):483–500.

NSS, N. S. S. O. (2014). Nutritional intake in India, 20112012 (Number 560).

OSHA (2024). Heat Standards.

Pailler, S. and Tsaneva, M. (2018). The effects of climate variability on psychological well-being in India. *World Development*, 106:15–26.

Park, R. J., Goodman, J., Hurwitz, M., and Smith, J. (2020). Heat and learning. *American Economic Journal: Economic Policy*, 12(2):306–39.

Park, R. J., Pankratz, N., and Behrer, A. P. (2021). Temperature, Workplace Safety, and Labor Market Inequality. *IZA Discussion Paper*, (14560).

Pavanello, F., De Cian, E., Davide, M., Mistry, M., Cruz, T., Bezerra, P., Jagu, D., Renner, S., Schaeffer, R., and Lucena, A. F. (2021). Air-conditioning and the adaptation cooling deficit in emerging economies. *Nature communications*, 12(1):6460.

Rahimi, S., Huang, L., Norris, J., Hall, A., Goldenson, N., Krantz, W., Bass, B., Thackeray, C., Lin, H., Chen, D., et al. (2024). An overview of the western united states dynamically downscaled dataset (wus-d3). *Geoscientific Model Development*, 17(6):2265–2286.

Randell, H., Gray, C., and Shayo, E. H. (2022). Climatic conditions and household food security : Evidence from Tanzania. *Food Policy*, 112(September):102362.

Rebecca Smith, J. (2012). Immigrant workers and worker's compensation: the need for reform. *American Journal of Industrial Medicine*, 55(6):537–544.

Reich, M., Gordon, D. M., and Edwards, R. C. (1973). A Theory of Labor Market Segmentation. *The American Economic Review*, 63(2):359–365.

Reiss, P. C. and White, M. W. (2005). Household electricity demand, revisited. *The Review of Economic Studies*, 72(3):853–883.

Reséndez, A. (2016). *The Other Slavery: The Uncovered Story of Indian Enslavement in America*. Mariner Books.

Riley, K., Delp, L., Cornelio, D., and Jacobs, S. (2012). From agricultural fields to urban asphalt: the role of worker education to promote california's heat illness prevention standard. *New solutions: a journal of environmental and occupational health policy*, 22(3):297–323.

Rode, A., Baker, R. E., Tamma Carleton, A. L., D'Agostino, M. D., Foreman, T., Gergel, D. R., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., et al. (2023). Is workplace temperature a valuable job amenity? implications for climate change. *Working Paper*.

Rode, A., Carleton, T., Delgado, M., Greenstone, M., Houser, T., Hsiang, S., Hultgren, A., Jina, A., Kopp, R. E., McCusker, K. E., et al. (2021). Estimating a social cost of carbon for global energy consumption. *Nature*, 598(7880):308–314.

Rosen, S. (1986). The theory of equalizing differences. *Handbook of labor economics*, 1:641–692.

Sathaye, J. A., Dale, L. L., Larsen, P. H., Fitts, G. A., Koy, K., Lewis, S. M., and de Lucena, A. F. P. (2013). Estimating impacts of warming temperatures on California's electricity system. *Global Environmental Change*, 23(2):499–511.

Schlosser, E. (2006). Foreword. In Sinclair, U., editor, *The Jungle*, pages vii–xv. Penguin Books.

Schofield, H. and Venkataramani, A. S. (2021). Poverty-related bandwidth constraints reduce the value of consumption. *Proceedings of the National Academy of Sciences*, 118(35):e2102794118.

Schubert, G., Stansbury, A., and Taska, B. (2022). Employer Concentration and Outside Options. *SSRN Working Paper*.

Schwarzwald, K. and Lenssen, N. (2022). The importance of internal climate variability in climate impact projections. *Proceedings of the National Academy of Sciences*, 119(42):e2208095119.

Shah, M. and Millett Steinberg, B. (2017). Drought of Opportunities : Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital Manisha Shah. *Journal of Political Economy*, 125(2).

Sharma, H., Singh, S., and Srivastava, S. (2020a). Socio-economic inequality and spatial heterogeneity in anaemia among children in india: Evidence from nfhs-4 (2015–16). *Clinical Epidemiology and Global Health*, 8(4):1158–1171.

Sharma, H., Singh, S. K., and Srivastava, S. (2020b). Socio-Economic Inequality and Spatial Heterogeneity in Anaemia Among Children in India: Evidence from NFHS-4 (2015-16). *Clinical Epidemiology and Global Health*, 8(4):1158–1171.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., et al. (2018). Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. *Climatic change*, 151:555–571.

Smith, P. L. (1979). Splines as a useful and convenient statistical tool. *The American Statistician*, 33(2):57–62.

Snell, C., Lambie-Mumford, H., and Thomson, H. (2018). Is there evidence of households making a heat or eat trade off in the UK? *Journal of Poverty and Social Justice*, 26(2):225–243.

Somanathan, E., Somanathan, R., Sudarshan, A., and Tewari, M. (2021). The impact of temperature on productivity and labor supply: Evidence from indian manufacturing. *Journal of Political Economy*, 129(6):1797–1827.

Southern California Edison (2012a). Schedule D-CARE California Alternate Rates for Energy Domestic Service.

Southern California Edison (2012b). Schedule D Domestic Service.

Southern California Edison (2012c). Schedule D-FERA Family Electric Rate Assistance.

Southern California Edison (2013a). 2015 General Rate Case (SCE-04 Volume 02, Part 02).

Southern California Edison (2013b). Schedule D-CARE California Alternate Rates for Energy Domestic Service.

Southern California Edison (2013c). Schedule D Domestic Service.

Southern California Edison (2013d). Schedule D-FERA Family Electric Rate Assistance.

Southern California Edison (2014a). California Alternate Rates for Energy (CARE) Program Plan and Budgets Proposal for the 20152017 Program Cycle.

Southern California Edison (2014b). Schedule D-CARE California Alternate Rates for Energy Domestic Service.

Southern California Edison (2014c). Schedule D Domestic Service.

Southern California Edison (2014d). Schedule D-FERA Family Electric Rate Assistance.

Southern California Edison (2015a). Schedule D-CARE California Alternate Rates for Energy Domestic Service.

Southern California Edison (2015b). Schedule D Domestic Service.

Southern California Edison (2015c). Schedule D-FERA Family Electric Rate Assistance.

Southern California Edison (2016a). Schedule D-CARE California Alternate Rates for Energy Domestic Service.

Southern California Edison (2016b). Schedule D Domestic Service.

Southern California Edison (2016c). Schedule D-FERA Family Electric Rate Assistance.

Southern California Edison (2017a). Schedule D-CARE California Alternate Rates for Energy Domestic Service.

Southern California Edison (2017b). Schedule D Domestic Service.

Southern California Edison (2017c). Schedule D-FERA Family Electric Rate Assistance.

Southern California Edison (2017d). Southern California Edison Companys 2018 Annual Report for 2017 Low Income Programs.

Southern California Edison (2018). Southern California Edison Companys (U 338-E) Responses to Administrative Law Judges Ruling Requiring Data From Respondent Utilities.

Southern California Edison (2019). Southern California Edison Companys (U 338-E) Fifteenth Quarterly Report on Progress of Residential Rate Reform.

Southern California Edison (2020). Discontinuance and Restoration of Service (Rule 11).

Stockdale, K. E. (2012). H-2a migrant agricultral workers: Protected from employer exploitation on paper, not in practice. *Creighton L. Rev.*, 46:755.

Taraz, V. (2018). Can farmers adapt to higher temperatures? Evidence from India. *World Development*, 112(December):205–219.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American meteorological Society*, 93(4):485–498.

Thomas, D., Frankenberg, E., Friedman, J., Habicht, J.-P., Hakimi, M., Ingwersen, N., Jaswadi, Jones, N., McKelvey, C., Pelto, G., Sikoki, B., Seeman, T., Smith, J., Sumantri, C., Suriastini, W., and Wilopo, S. (2006). Causal Effect of Health on Labor Market Outcomes : Experimental Evidence. *Working Paper*, (September).

Tiwari, A. (2020). India State and County Shapefile.

Walton, D. and Hall, A. (2018). An assessment of high-resolution gridded temperature datasets over california. *Journal of Climate*, 31(10):3789–3810.

White, C. (2017). The dynamic relationship between temperature and morbidity. *Journal of the Association of Environmental and Resource Economists*, 4(4):1155–1198.

World Bank (2024). Rural population - India.

Xiao, R. J. (2023a). Climate-Induced Labor Risk and Firm Investments in Automation. *SSRN Working Paper*, (April).

Xiao, R. J. (2023b). Climate-Induced Labor Risk: Labor Market Consequences, Firm Labor Adaptation Strategies, and Firm Performance. *SSRN Working Paper*.

Zhu, X., Li, L., Zhou, K., Zhang, X., and Yang, S. (2018). A meta-analysis on the price elasticity and income elasticity of residential electricity demand. *Journal of Cleaner Production*, 201:169–177.