

Copyright

by

Seth Neller

2022

The Dissertation Committee for Seth Neller
certifies that this is the approved version of the following dissertation:

**Essays on Health Economics and the Early-Life Determinants of
Adult Outcomes**

Committee:

Marika Cabral, Supervisor

Mike Geruso

Dean Spears

Sheila Olmstead

**Essays on Health Economics and the Early-Life Determinants of
Adult Outcomes**

by

Seth Neller

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2022

Acknowledgments

Pending.

Essays on Health Economics and the Early-Life Determinants of Adult Outcomes

Seth Neller, Ph.D.

The University of Texas at Austin, 2022

Supervisor: Marika Cabral

The three chapters of this dissertation explore topics in health economics, namely how early-childhood health circumstances affect long-run health and economic outcomes, as well as how insurance reimbursement impacts the nature of physician practices.

The first chapter assesses the impact of *in utero* and early-childhood exposure to wildfire smoke on longevity.¹ To identify areas that were exposed to wildfire pollution, we leverage mid-20th century (1930-1969) California wildfires and smoke dispersion modeling. We then combine these wildfire pollution data with comprehensive, restricted-use administrative data. These linked data allow us to measure childhood wildfire smoke exposure for four decades of birth cohorts and to observe a rich set of later-life outcomes. Using these data, we exploit plausibly exogenous variation in smoke exposure—which is a function of fire timing and size as well as wind direction and speed—to identify long-run effects. We find that moving from the 25th to 75th percentile of early-life wildfire smoke exposure results in 1.7 additional deaths before age 55 per 1,000 individuals, conditional on surviving past early childhood. Aggregating these effects across ages 30 to 80 translates to 46 life years lost per 1,000 persons.

The second chapter considers the impact of *in utero* and early-childhood exposure to wildfire smoke on longevity as well as economic achievement, human capital accumulation, and disability in

¹Chapters 1 and 2 are adapted from the manuscript, “Ashes to Ashes: The Lifelong Consequences of Early-Life Wildfire Exposure,” which is jointly authored with Samuel Arenberg. The contribution statement at the conclusion of this dissertation details my contributions to this manuscript.

mid-to-late adulthood.¹ To identify areas that were exposed to wildfire pollution, we leverage mid-20th century (1930-1969) California wildfires and smoke dispersion modeling. We then combine these wildfire pollution data with comprehensive, restricted-use administrative data from the Social Security Administration and Census Bureau. These linked data allow us to measure childhood wildfire smoke exposure for four decades of birth cohorts and to observe a rich set of later-life outcomes. Using these data, we exploit plausibly exogenous variation in smoke exposure—which is a function of fire timing and size as well as wind direction and speed—to identify long-run effects. We find that moving from the 25th to 75th percentile of early-life wildfire smoke exposure results in 1.7 additional deaths before age 55 per 1,000 individuals, conditional on surviving past early childhood. Aggregating these effects across ages 30 to 80 translates to 46 life years lost per 1,000 persons. We further find that smoke exposure results in unfavorable changes to a wide range of later-life outcomes across economic achievement, educational attainment, and disability measures. From these results, we estimate that each child born in California during our sample period sustained, on average, approximately \$22,000 of discounted damages in lost life expectancy and lost earnings due to wildfire smoke. These findings suggest that warming temperatures, which exacerbate the duration and intensity of wildfire seasons, are already meaningfully affecting the life cycles of exposed children through increased smoke exposure.

The third chapter exploits spatial discontinuities in Medicare payment rates to estimate the effect of reimbursements on primary care physicians' choice of organizational structure. I find that a 1 percent increase in Medicare reimbursement leads to a 1.7 to 2.2 percentage point increase in primary care doctors who practice with a small group (defined as 25 providers or fewer). This effect is driven by changes in the tails of the practice size distribution: a 1.8 percentage point increase in physicians who are affiliated with the smallest (1- or 2-provider) practice groups with a corresponding decrease in physicians joining very large practices (≥ 150 providers). I do not, however, detect any evidence of physician sorting or bunching around the boundary in response to differential payment, supporting the underlying assumptions of my regression discontinuity design. Accordingly, my findings suggest that Medicare pricing may be a factor in the trend of consolidation in the physician and clinical services market.

Contents

Acknowledgments	iv
Abstract	v
List of Tables	xi
List of Figures	xii
Chapter 1 The Impact of Early Childhood Wildfire Smoke Exposure on Longevity	1
1.1 Introduction	1
1.2 Background and Conceptual Framework	5
1.2.1 Ambient Pollution and Childhood Development	5
1.3 Data	8
1.3.1 Wildfire Data	8
1.3.2 Smoke Data	10
1.3.3 Outcome Data and Linkages using Restricted Census Data	12
1.3.4 Other Data	15
1.4 Validation of Smoke Measures	15
1.5 Methodology	16
1.6 Long-Run Mortality Effects	18
1.7 Robustness	21
1.7.1 Controlling for Proximity to Fires	21
1.7.2 Changes to Specification	21
1.7.3 Measurement Error	24
1.7.4 Issues with Two-Way Fixed Effect Estimates and Alternative Estimation . .	25
1.8 Supplemental Results	26

1.8.1	Early-Life Exposure to Wildfire Smoke Induces Individuals to Sort into Higher-Mortality Occupations and Neighborhoods	26
1.8.2	Wildfire Shocks Are Not Transmitted to Children Via Local Economic Channels and Long-Run Effects Do Not Meaningfully Vary Across Person and Place Characteristics	27
1.9	Conclusion	28
Chapter 2 The Impact of Early Childhood Wildfire Smoke Exposure on Economic, Education, and Disability Outcomes		36
2.1	Introduction	36
2.2	Background and Conceptual Framework	40
2.2.1	Ambient Pollution and Childhood Development	40
2.3	Data	44
2.3.1	Wildfire Data	44
2.3.2	Smoke Data	46
2.3.3	Outcome Data and Linkages using Restricted Census Data	48
2.3.4	Other Data	50
2.4	Validation of Smoke Measures	51
2.5	Methodology	52
2.6	Long-Run Effects on Economic, Educational, and Disability Outcomes	54
2.6.1	Economic Achievement	54
2.6.2	Educational Attainment	56
2.6.3	Disability	57
2.7	Robustness	58
2.7.1	Controlling for Proximity to Fires	58
2.7.2	Changes to Specification	59
2.7.3	Mortality Selection	62
2.7.4	Measurement Error	63
2.7.5	Issues with Two-Way Fixed Effect Estimates and Alternative Estimation . .	63
2.8	Supplemental Results	64
2.8.1	Effects on Earnings and Disability Grow More Severe with Age	65

2.8.2	Childhood Smoke Exposure Affects the Entire Distribution of Economic Achievement and Earnings	66
2.8.3	Wildfire Shocks Are Not Transmitted to Children Via Local Economic Channels and Long-Run Effects Do Not Meaningfully Vary Across Person and Place Characteristics	67
2.8.4	The Effect of Childhood Wildfire Exposure on Family Formation, Institutionalization, and Other Outcomes	67
2.9	Discussion and Conclusion	68
2.9.1	The Cost and Fiscal Burden of Childhood Wildfire Smoke Exposure	68
2.9.2	Conclusion	70
Chapter 3	How Do Medicare Payments Influence Physician Practice Structure?	81
3.1	Introduction	81
3.2	Background and Data	84
3.2.1	Medicare Payments and Identifying Variation	84
3.2.2	Data	85
3.3	Theoretical Model	87
3.3.1	Setup	87
3.3.2	Comparative Statics	89
3.4	Empirical Framework	90
3.4.1	Identification and Econometric Model	90
3.4.2	Validity of Regression Discontinuity Design	92
3.5	Results	95
3.5.1	Main results	95
3.5.2	Robustness	97
3.5.3	Alternate IV and Full Sample Results	98
3.5.4	Heterogeneity	99
3.6	Discussion	101
Appendix A	Appendix to Chapters 1 and 2: The Effect of Early Life Wildfire Smoke Exposure	115
A	Data Construction	115
A.1	Matching Fire Dates to Meteorologically Similar Days for HYSPLIT Analysis	115

A.2	Place of Birth String Matching Algorithm	116
A.3	Creation of Adjusted Numident Death Rates	116
A.4	Data Construction for Control Variables and Heterogeneity Analysis	118
B	Additional Results and Robustness	120
B.1	Infant Mortality	120
B.2	Comparison of Results with Existing Literature	122
B.3	Bounding Mortality Selection	124
B.4	Migration-Related Measurement Error	125
B.5	Other Supplemental Results	126
Appendix B	Appendix to Chapter 3: How Do Medicare Payments Affect Physician Practice Structure?	160
A	Variations of the Theoretical Model	160
B	Sample Construction	164
C	Additional Figures and Tables	167
Bibliography		188
Contribution Statement		201

List of Tables

2.1	The Cost and Fiscal Burden of Early-Life Wildfire Smoke Exposure	80
3.1	Main Results	112
3.2	Full-Sample and Alternate IV Results	113
A.1	Summary Statistics	156
A.2	Main Results: Table of Estimates	157
A.3	Context for Main Results: Level Estimates	158
A.4	Supplemental Results: Wildfire Smoke Does Not Have Detectable Effects on Contemporaneous Local Economic Conditions or Net Migration	159
B.1	Medicare-Designated Payment Localities	186
B.2	Alternate Results: Treatment Outcomes	187

List of Figures

1.1	Descriptive Statistics: There is Substantial Variation in Year-to-Year Acres Burned, but Intensity has Increased Sharply over the Last 30 Years	29
1.2	Identifying Variation: Smoke Exposure Varies Over Time and Geography	30
1.3	Validation of Treatment: Modeled Wildfire Smoke Exposure Increases Monthly PM _{2.5}	31
1.4	Main Results: Early-Life Wildfire Exposure Reduces Longevity in Adulthood	32
1.5	Robustness: Controlling for Nearby Fires Does Not Affect Estimates, Implying that Smoke Exposure is Driving Effects	33
1.6	Robustness to Additional Controls and Specifications	34
1.7	Supplemental Results: Individuals Exposed to Higher Levels of Early-Life Wildfire Exposure Sort into Higher-Mortality Occupations and Neighborhoods	35
2.1	Descriptive Statistics: There is Substantial Variation in Year-to-Year Acres Burned, but Intensity has Increased Sharply over the Last 30 Years	72
2.2	Identifying Variation: Smoke Exposure Varies Over Time and Geography	73
2.3	Validation of Treatment: Modeled Wildfire Smoke Exposure Increases Monthly PM _{2.5}	74
2.4	Main Results: Early-Life Wildfire Exposure Reduces Long-Run Economic and Educational Achievement; Increases Disability	75
2.5	Robustness: Controlling for Nearby Fires Does Not Affect Estimates, Implying that Smoke Exposure is Driving Effects	76
2.6	Robustness to Additional Controls and Specifications	77
2.7	Supplemental Results: The Effects Early-Life Wildfire Exposure Increase Proportionally with Age for Key Time-Varying Outcomes	78
2.8	Supplemental Results: Early-Life Wildfire Exposure Impacts the Entire Distribution of Economic and Earnings Outcomes	79

3.1	Identifying Variation: Medicare Payment Locality Boundaries	104
3.2	Identifying Variation: Boundary Differences in Medicare Payment	105
3.3	Validation of Research Design: Spatial Provider Density	106
3.4	Validation of Research Design: Smoothness of LASSO-Predicted Outcomes	107
3.5	Main Results: Discontinuities at the Boundary	108
3.6	Distributional Effects on Group Size	109
3.7	Robustness: Bandwidth Sensitivity	110
3.8	Heterogeneous Effects by Border Characteristics	111
A.1	Motivating Figure: Wildfires are a Growing Source of PM _{2.5} Pollution	129
A.2	Illustration of Conceptual Framework: Wildfire Smoke Could Affect Long-Run Health through Multiple Channels	130
A.3	Graphical Illustration of Data Construction for Primary Data Sets	131
A.4	Descriptive Statistics: Location of Fires	132
A.5	Illustration of Modeled Smoke Plumes Under Different Conditions	133
A.6	Descriptive Statistics: Changes in Wind Direction and Intensity Provide Plausibly Exogenous Variation	134
A.7	Major City Borders Are Stable Over Time, Suggesting Little Risk of Measurement Error from Changing Longitude / Latitude Centroids	135
A.8	Descriptive Statistics: Comparison of Adjusted and Unadjusted Death Rates per Numident to Expected Cohort Death Rates	136
A.9	Validation of Treatment: Modeled Wildfire Smoke Exposure Increases Monthly Measures of Various Pollutants	137
A.10	Illustration of Treatment Definition	138
A.11	Illustration of Regional Definitions	139
A.12	Descriptive Statistics: Fire Seasonality	140
A.13	Robustness: Unadjusted and Adjusted Mortality Measures Display Similar Effects .	141
A.14	Additional Results: Wildfire Smoke Exposure Increases Infant Mortality	142
A.15	Additional Results: There is No Detectable Effect of Wildfire Smoke Exposure on Fertility Timing or Overall Fertility	143
A.16	Additional Results: Reductions in Longevity with Infant Mortality Added	144

A.17 Additional Results: Early-Life Wildfire Smoke Exposure Reduces Earnings and High School Completion; Increases Cognitive Difficulty	145
A.18 Additional Results: Early-Life Wildfire Smoke Exposure Increases Use of Social Safety Net; Induces Sorting into Higher-Mortality Industries	146
A.19 Robustness: Residual Plots Demonstrate that Impacts of Wildfire Exposure are Generally Linear	147
A.20 Robustness: Residual Plots Demonstrate that Impacts of Wildfire Exposure are Not Driven by the Tails of the Distribution	148
A.21 Robustness: Adjusting for Mortality Selection Does Not Qualitatively Affect Estimates	149
A.22 Robustness: Adjusting for Migration-Specific Measurement Error Does Not Qualitatively Affect Estimates	150
A.23 Robustness: Using Alternative Estimation Strategies Do Not Result in Materially Different Estimates	151
A.24 Robustness: Multiple Hypothesis Testing Does Qualitatively Affect Conclusions About Statistical Significance for Index Components	152
A.25 Supplemental Results: The Effects of Early-Life Wildfires Smoke Data on Labor Force Participation, Economic Index, and Disability Index are Increasing Proportionally with Age	153
A.26 Supplemental Results: There Are Few Statistically Significant Differences in the Effect of Early-Life Wildfire Smoke Across Groups	154
A.27 Supplemental Results: Early-Life Wildfire Smoke Exposure Does Not Significantly Affect Family Formation, Migration, or Placebo Outcomes, but Does Increase Institutionalization	155
B.1 Validation of Research Design: Alternate Measures of Spatial Provider Density . .	172
B.2 Validation of Research Design: Covariate Balance Tests	173
B.3 Main Results: Without Border FE	174
B.4 Localities in the Dallas-Fort Worth Area (Physician Clusters)	175
B.5 Distributional Effects on Group Size	176
B.6 Distributional Results: Effects on Smallest and Largest Groups	177
B.7 Other Results: Treatment Outcomes	178
B.8 Other Results: Other Practice Structure Outcomes	179

B.9	Border Leave-one-out Analysis	180
B.10	Other Robustness Checks	181
B.11	Main Results: Full Sample	182
B.12	Distributional Effects on Group Size	183
B.13	Heterogeneity by Physician Characteristics	184
B.14	Prediction of Log FTHHI by Group Size Measure	185

Chapter 1

The Impact of Early Childhood Wildfire Smoke Exposure on Longevity

1.1 Introduction

It is now well established that air pollution increases contemporaneous mortality of both the very old and very young, two age groups most sensitive to their environments. However, substantially less is known about a related question, which we discuss in this paper: how do the harms of exposure to air pollution at very young ages manifest in reduced longevity in later adulthood? Given that long-run effects frequently comprise a large fraction of the damages from early-life shocks (Currie and Almond, 2011; Almond et al., 2018), understanding the ramifications of childhood pollution exposure across the life cycle is vital in assessing its total costs. This paper provides the first evidence that increases in ambient air pollution in early life increase mortality at older ages.

To determine the long-run effects of early-life exposure to air pollution, we study wildfires, the only major emissions source of fine particulate matter ($PM_{2.5}$) in the United States that is growing over time and one that has generated approximately 29% of all fine particulate pollution in recent years.¹ Additionally, it is expected that wildfire pollution will *continue* to grow: the United States Forest Service predicts that 1 °C increase in average temperature will result in a 200-600% increase in acres burned in the Western United States (Vose et al., 2012). Given that

¹Source: Author calculations using National Emissions Inventory data from the Environmental Protection Agency. See Appendix Figure A.1 for more detail.

experts predict global temperatures will increase by at least 2 °C over the global baseline by 2050 (IPCC, 2021), the prospect of increased pollution from wildfires is nearly certain.

There are two key challenges to conducting a long-term analysis of childhood air pollution. The first is that pollution has only been reliably measured since the mid-to-late 1960s, with the first meaningful regulation occurring with the Clean Air Act Amendments of 1970. This lack of historical data and sources of exogenous variation has limited researchers' ability to evaluate the ramifications of early-life pollution exposure over the life cycle. We overcome this challenge by examining wildfires occurring in California, a source of pollution which has been tracked reliably since at least 1930. Additionally, wildfire smoke exposure, which is a function of fire timing, acres burned, wind direction, and wind speed, provides potentially exogenous variation for an analysis of long-term outcomes. In order to supplement these wildfire records and generate measures of smoke exposure, we first hand-gather dates in which the wildfires occurred using hundreds of historical newspaper articles. We then utilize wildfire pollution modeling tools to generate historical measures of wildfire smoke exposure. These modeled smoke measures are utilized by the United States Forest Service to predict wildfire smoke and compare well to patterns generated by satellites (Larkin et al., 2009). In this paper, we perform additional exercises to further validate this modeled smoke measure and find that it strongly predicts pollution exposure in a modern context.

The second challenge in assessing the long-run impacts of early-life shocks is linking childhood circumstances to later-life outcomes. To address this difficulty, we obtain restricted-use administrative data from the Social Security Administration and U.S. Census Bureau. In addition to providing comprehensive high-quality microdata, these data sets have two features that are ideally suited for our analysis. First, the administrative data provides exact city and state of birth, which allows us to accurately assign measures of wildfire smoke exposure to individuals. Second, these data also include unique identifiers which we use to link individuals—and their associated early-life exposure measures—to comprehensive death records (as applicable).

We evaluate the long-term effects of exposure using a cohort distributed leads and lags model and find that the effects of childhood pollution exposure—which we define as exposure before age 5—last throughout the entire life cycle. Specifically, we find that moving from the 25th to the 75th percentile of exposure—roughly 5 additional months with *any* wildfire smoke—results in 1.7 additional deaths prior to age 55 per 1,000 persons, conditional on surviving past childhood.²

²Results are conditional on surviving past childhood because it is necessary for an individual to obtain a Social Security number to be included in our data, a process that did not typically occur until adolescence for the cohorts born in our sample. Also of note: we feature survival until age 55 as a main outcome as that is the latest age that

Increases in mortality appear as early as age 30 and dissipate by age 80—summing the total effects between these two ages translates to 46 life years lost per 1,000 persons. These reductions in adult longevity exceed the number of life years lost due to infant mortality attributable to wildfire pollution. This further underscores the fact that quantifying long-run effects is an essential exercise in assessing the costs of childhood harms. Using our estimates, we conservatively estimate that the average child in our sample sustains approximately \$12,500 of damages in lost life years due to wildfire smoke.

To demonstrate that our results are driven by smoke exposure—rather than proximity to the actual wildfires themselves—we control for nearby fires during childhood. Inclusion of these controls has almost no effect on our estimates, suggesting that the impacts that we find relate to pollution damage, rather than long-run effects of trauma or negative economic shocks due to natural disaster exposure. Additionally, we show that our estimates are stable across a range of robustness tests, including adding controls for climate and other time-varying factors as well as different functional forms and specifications. Additionally, we consider alternate estimation methods to evaluate the impact of weighting issues endemic to models that utilize two-way fixed effects (e.g., Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfoeuille, 2020; Goodman-Bacon, 2021a; Sun and Abraham, 2021) and find that results are qualitative similar when estimated using other methods.

To better understand how childhood smoke exposure affects mortality, we conduct a supplemental analysis to understand how occupational and neighborhood characteristics may contribute to reduced longevity. We find that individuals affected by wildfire smoke during childhood select occupations and neighborhoods that have higher mortality risk, providing suggestive evidence that economic factors could be driving mortality results, particularly those in early adulthood.³

Our findings that early-life exposure to wildfire smoke pollution reduces longevity contribute to three distinct areas of economic research. First, this paper contributes to our understanding of how circumstances in early childhood affect adult longevity. To date, nearly all of our understanding in this area comes from studies showing that expanded access to the social safety net while young decreases mortality at later ages (e.g., Aizer et al., 2016; Bailey et al., 2020; Goodman-Bacon, 2021b). To the best of our knowledge, this is the first causal research to assess the longevity impacts of childhood exposure to ambient air pollution from any source.⁴ These longevity effects

we observe all cohorts in our sample.

³Additionally, Chapter 2 of this dissertation also considers the degree to which reduced economic achievement and educational attainment could explain the reductions in longevity that we observe.

⁴Notably, we are also unaware of any research displaying non-causal *associations* between childhood air pollution and adult mortality. This is likely due to the substantial difficulties in locating historical pollution measures and

are important to measure, as they result in a greater number of life years lost than those lost from infant deaths and also comprise over half of the total damages from smoke exposure measured in this paper.

Second, we expand the literature examining the long-run effects of early-life air pollution on adult outcomes. This work has largely focused on the impact of the long-run effects of the Clean Air Act Amendments (“CAAA”) of 1970 on earnings around age 30 (Isen et al., 2017b) and on educational attainment (Colmer and Voorheis, 2020; Voorheis, 2017).⁵ This paper increases our understanding of the long-run effects of childhood pollution exposure by examining a different—and highly relevant—source of pollution. While the pollution from industrial and vehicular sources studied in this literature have been falling for many years, pollution from wildfire smoke has been increasing—and likely will continue to increase—over time.⁶ Given the growing prevalence of wildfires and evidence from medical research suggesting that particulates from wildfires are potentially more harmful to human health than pollution from other sources, understanding the harm done by wildfire pollution is increasingly important.⁷

Third and finally, this study contributes to our rapidly growing understanding of the effects of pollution from fires on human health. Thus far, this literature has concentrated primarily on the contemporaneous effects of fire-related pollution on mortality (Jayachandran, 2009; Miller et al., 2017), infant health (Rangel and Vogl, 2019), and labor market outcomes (Borgschulte et al., 2020). As emphasized above, this paper adds to the existing knowledge of these health effects by quantifying the long-run harms of wildfire smoke, which we find to be a substantial component of the costs of these fires. Furthermore, insofar as pollution from these fires is expected to grow due to global increases in temperature (Moritz et al., 2012; Vose et al., 2012), we view this paper as an important contribution to our understanding of the costs of climate change (e.g. Jacob et al., 2007; Schlenker and Roberts, 2009; Graff Zivin and Neidell, 2014; Barreca et al., 2016; Hsiang et al.,

longitudinally linking them to adult mortality outcomes, as discussed in previous paragraphs. Tangentially related are studies that combine cross-sectional variation in air pollution exposure with relatively short-run mortality measures to estimate the effects of air pollution on life expectancy (see Chen et al., 2013, for a notable example). However, because these papers use mortality measures that typically involve 1-to-5-year follow-up periods, they are still identifying the contemporaneous effects of air pollution, which is fundamentally different than the focus of this paper.

⁵ An early contributor to this literature is Sanders (2012), who studies the long-run effects of recession-induced decreases in pollution on test scores. Additionally, there is also a small literature examining the impacts of the lead air pollution on educational attainment (Grönqvist et al., 2020) and crime/anti-social behavior (Aizer and Currie, 2019; Reyes, 2007, 2015).

⁶ As demonstrated in Appendix Figure A.1, wildfire smoke comprised over 29% of fine particulate matter pollution during the 2017-2020 period, more than the emissions from fuel combustion/generation, industrial activities, and vehicular use *combined*.

⁷ See, for example, Aguilera et al. (2021a) and Aguilera et al. (2021b)

2017).

1.2 Background and Conceptual Framework

1.2.1 Ambient Pollution and Childhood Development

Because this paper estimates the impact of air quality in early life on outcomes in late life, in this section, we first explain the biological influence of ambient pollution on human health. Further, we document several reasons why wildfire smoke may be more harmful than other forms of ambient pollution. Throughout, we focus on the role of fine particulate matter, or PM_{2.5}, because it is widely considered to be the most harmful to human health (Peeples, 2020).

PM_{2.5} is particulate matter, either solid or liquid, that is less than 2.5 microns in diameter (for reference, a human hair is about 70 microns in diameter). It enters the body through inhalation into the lungs. Although some of the matter will be removed by normal clearance mechanisms, some of the matter will reach the bloodstream, where the particulates can cause a multitude of cardiovascular and respiratory problems. Although the precise pathways are not yet fully understood, a leading hypothesis is that many particulates carry or produce reactive oxygen species (“ROS”). Excesses of ROS can interfere with normal cellular processes and cause damage to tissues and organs. Such oxidative stress is associated with onset of diseases ranging from diabetes to atherosclerosis. PM_{2.5} can even enter the gastrointestinal tract, causing imbalances in the intestinal microecology, which is known to affect a host of other systems, including the central nervous system. Further, fine particulate matter has been shown to weaken the immune system and cause increased susceptibility to infections months after initial exposure (Landguth et al., 2020).

The harm from fine particulate matter may be exacerbated in early life: young children breathe in more air relative to their body weight, they tend to spend more time outdoors, and their bodies are growing rapidly. Holm et al. (2021) summarizes the potential connection between wildfire exposure at young ages and outcomes at older ages, “[E]arly-life exposure may act by ‘programming’ fundamental metabolic, structural, and cell signaling mechanisms that may result in lifelong impacts.” Furthermore, early-life and late-life health statuses are strongly correlated, suggesting that harms persist over the life cycle (Case and Paxson, 2010).

Finally, there is an emerging literature that suggests PM_{2.5} from wildfire smoke may be worse for human health than other sources of ambient pollution.⁸ One reason is that wildfire smoke has a

⁸Consult Holm et al. (2021) for a detailed overview of the potential health impacts of wildfire smoke on children.

greater concentration of ultra-fine particulates, which are less than 0.1 microns in diameter (Holm et al., 2021). These particles can penetrate deeper into the body than larger particulates, cross the blood-brain barrier, damage neural tissue, and impair brain development (Schraufnagel, 2020). Another reason is that wildfire smoke contains more polycyclic aromatic hydrocarbons, which induce greater oxidative stress. Wildfire smoke also contains more volatile organic compounds, many of which are understood to be carcinogenic (e.g., formaldehyde and benzene). Additionally, wildfire smoke has shown to contain meaningful amounts of heavy metals, such as lead, that have been linked to adverse long-term outcomes (California Air Resources Board, 2021; Grönqvist et al., 2020). The likelihood that wildfire smoke is more harmful than other forms of ambient pollution has been demonstrated by animal studies (Wegesser et al., 2010) as well as observational analyses of child hospitalizations (Aguilera et al., 2021a; Aguilera et al., 2021b). Further, a recent paper that examines the effects of wildfire smoke pollution on the labor market (Borgschulte et al., 2020) finds effects on extensive-margin labor supply that are 2-7 times larger than a comparison group of papers that utilize variation in other pollution sources. In summary, these studies support the idea that the composition of wildfire smoke particles may be particularly harmful for human health.

Conceptual Framework

To formalize the hypothesized impact of early-life pollution exposure on long-run outcomes and our reduced-form findings discussed in subsequent sections, we present a simple conceptual model.⁹ This framework, which is visually represented by Appendix Figure A.2, models the life cycle in four distinct periods. In the first period, childhood health (h_1) is determined as a function of genetic endowments (h_0) and inputs during early life (I_1), so that $h_1 = h_1(h_0, I_1)$.¹⁰ Given its documented adverse health effects—discussed in detail above—air pollution in childhood is modeled as negatively affecting early-life inputs and health, such that $\partial h_1 / \partial I_1 < 0$.

In the second period, individuals choose an amount of education (e), the quality of which is a function of childhood health: $e = e(h_1) = e(h_0, I_1)$. Because health is seen as a positive input into the educational process, childhood smoke exposure is predicted by the model to reduce educational attainment:

⁹What follows is similar in spirit to the frameworks presented by Isen et al. (2017b) and Colmer and Voorheis (2020).

¹⁰For purposes of this model, health is broadly defined to include physical health and measures of mental well-being, including cognitive ability and non-cognitive skills.

$$\frac{\partial e}{\partial I_1} = \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} < 0.$$

After completion of their education careers, individuals progress into the labor market, the third period of the model. Because total earnings depend on both education and innate ability (physical and mental, both represented by health stock in our model), we model earnings (ℓ) as:

$$\begin{aligned}\ell &= \ell(e(h_1), h_1) \\ &= \ell(e(h_0, I_1), h_0, I_1).\end{aligned}$$

With this structure, the impact of early-life pollution exposure is given by:

$$\frac{\partial \ell}{\partial I_1} = \underbrace{\frac{\partial \ell}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1}}_{\substack{\text{Lower earnings} \\ \text{due to poorer} \\ \text{health}}} + \underbrace{\frac{\partial \ell}{\partial e} \cdot \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1}}_{\substack{\text{Lower earnings due to} \\ \text{impact of reduced} \\ \text{education } (\partial e / \partial I_1)}} < 0.$$

As displayed in the above equation, the impacts of childhood smoke exposure affect earnings through two channels. The first term captures the direct channel: lower health stock could cause lower earnings through poor physical health (which likely manifests through reduced extensive-margin labor supply), lower cognitive ability, or reduced non-cognitive abilities (both of which could manifest as either reduced labor supply or lower wages). The second term captures the indirect effect of educational attainment. Insofar as individuals have lower levels of education due to wildfire smoke exposure, this will translate to lower earnings.

The last stage of the model concerns health in later adulthood and longevity (h_2), which is determined as a function of childhood health (h_1) and later-life health inputs (I_2).¹¹ Adult health inputs are themselves based on an individual's education and earnings, $I_2 = I_2(e, \ell)$. Taken together, this gives:

$$\begin{aligned}h_2 &= h_2(h_1, I_2(e, \ell)) \\ &= h_2(h_0, I_1, I_2(e(h_0, I_1), \ell(e(h_0, I_1), h_0, I_1))).\end{aligned}$$

¹¹It would also be reasonable to utilize a model with many periods, where adult health and earnings are determined contemporaneously within a given period and act as inputs into the next period's outcomes. However, we abstract away from such a scenario in favor of simplicity and expositional clarity.

Thus, the impact of a childhood wildfire shock on later life health is written as:

$$\frac{\partial h_2}{\partial I_1} = \underbrace{\left[\frac{\partial h_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right]}_{\text{Worse health due to direct/sustained effects}} + \underbrace{\left[\frac{\partial I_2}{\partial e} \cdot \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right]}_{\text{Lower investments due to adverse health impact on education (e)}} + \underbrace{\left[\left(\frac{\partial I_2}{\partial \ell} \cdot \frac{\partial \ell}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right) + \left(\frac{\partial I_2}{\partial \ell} \cdot \frac{\partial \ell}{\partial e} \cdot \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right) \right]}_{\text{Lower investments due to adverse health impact on earnings (\ell) and impact of reduced education (\partial e / \partial I_1) on earnings (\ell)}}$$

The first term details the direct effect of early-life health on later-life health. Insofar as reductions in early life health capital are permanent, this term will be negative. The second term relates to the indirect effect of lower educational attainment, which could increase adverse health behaviors. While the evidence on education and later-life health is mixed, in Section Chapter 2 we utilize estimates from the literature to assess the role of this channel. Finally, the last two terms capture the effects of wildfire smoke transmitted through the earnings channel. Lower earnings in life could result in lower health investments. To the extent possible, we evaluate the degree to which our mortality estimates can be explained by lower earnings (Chapter 2 of this dissertation) or other socioeconomic factors, such as occupation and neighborhood choice (Section 1.8.1).

1.3 Data

To estimate the effects of wildfire smoke on long-run outcomes, we construct a longitudinal data set linking childhood measures of wildfire smoke exposure to administrative and large-scale survey data detailing adult measures of longevity, economic achievement, educational attainment, and disability in adulthood. A simplified version of this data construction process, presented in Appendix Figure A.3, displays data sources and key steps used to clean and assemble our data. The sections that follow discuss these items in greater detail.

1.3.1 Wildfire Data

Our analysis utilizes wildfires from California, the state where wildfires are most severe and have been tracked for over a century. Specifically, we obtain “fire atlases” from the California Department of Forest and Fire Protection’s (“CAL FIRE”) Fire Perimeter Database, which is “the most complete digital record of fire history in California” (California Department of Forestry and Fire Protection, 2018). Two characteristics of this database make it ideally-suited for our analysis of

the long-run effects of air pollution from wildfire smoke. First, the database includes comprehensive data on fires for nearly 100 years—and includes some coverage of fires as early as 1878. This consistent and long-running tracking of fires allows to measure differential early-life air pollution exposure for children born as early as 1930. The availability of this historical data is a vital feature for our analysis. Because we can observe children born many years ago, we are able to evaluate longer-term effects—such as longevity and mid-to-late adult outcomes—that were previously unmeasurable with existing data.¹²

Second, the CAL FIRE data is highly detailed—it includes geospatial data for each fire (i.e., the exact area that was burned), the acreage burned, and the date that each fire was discovered. These characteristics are key inputs into the pollution modeling process discussed later in this section and allow us to generate strong predictions of the cohorts who were exposed to smoke pollution. To the best of our knowledge, there is no other source that combines both the number of years and the quality of data present in the fire atlas data.¹³

Although the Fire Perimeter Database has high-quality information regarding Californian wildfires, 42% of fires during our study period do not have exact dates when the fire occurred—only the year of the fire. To complete this missing information, we hand-gathered dates for each fire from historical newspaper archives.¹⁴ Overall, 512 fires were missing exact dates and required hand-dating; of those fires, we were able to assign dates to 458 (88%) with high confidence. The remaining fires were treated as missing. Notably, fires that could not be confidently dated tend to be fires that are smaller and/or are in remote areas and are therefore less likely to result in harmful pollution. For purposes of our analysis, we focus on fires that burned 1,000 acres or more, which comprise approximately 89% of the total acres burned during our study period (1,232 fires in total).

Figure 1.1 displays statistics regarding the severity of Californian wildfires—measured by total acres burned—over time. Panel A focuses on the study period of 1930 through 1969 and

¹²Specifically, most studies of the long-run effect of childhood air pollution utilize pollution monitor data (largely available starting in the mid-to-late 1960s) and leverage exogenous variation from the Clean Air Act Amendments of 1970. Because these data sources are relatively recent, papers that use them are limited to observing adult outcomes relatively early in life.

¹³Notably, the National Interagency Fire Center has detailed wildfire records, but these date back only to 1983. Additionally, the United States Forest Service maintains historical fire perimeter databases for much of the Western United States, but these appear to be incomplete, with many lacking any information on the date of the fire, a key input into the pollution modeling analysis.

¹⁴Fires were located in newspaper records by searching internet archives—newspapers.com and genealogybank.com—for key terms. For instance, a typical search included typical key words, such as “acres burned”, along with information gathered from the Fire Perimeter Database: the year of the fire, the name of the forest (e.g., “Mendocino National Forest”), and a prominent geological feature near to the fire (e.g., “Clear Lake”).

demonstrates that there is substantial temporal variation in the number of burned acres: some years (1937 and 1963) there are fewer than 20,000 acres burned while others (1950) reach nearly 500,000 acres.¹⁵ To provide further context for our sample period and to motivate the growing problem of wildfires, we also include Panel B, which displays the temporal trends of fires from 1930 through 2021. Within the figure, which has a vertical axis eight times the height of Panel A, we see two trends. First, the variation in acres burned was relatively constant from 1930 through 1990. However, there has been a marked increase in the severity of fire seasons over the last 25-30 years. In fact, the four most extreme seasons in California’s recorded history have all occurred within the last five years, emphasizing the potential for these fires to have a growing impact on health and economic activity. In addition to the temporal variation in fires, Appendix Figure A.4 presents the spatial distribution of fires (in 10-year bins), underscoring the fact that wildfire hazard is dispersed throughout the state.

1.3.2 Smoke Data

A central challenge of assessing the health impacts of wildfire pollution is understanding which individuals were exposed to smoke. Recent papers studying contemporaneous impacts of recent wildfire pollution—such as Miller et al. (2017) and Borgschulte et al. (2020)—utilize satellite remote sensing data to identify smoke plumes, which is available for the years 2005 through 2021. However, no such data is available during the sample period in which our cohorts were born. When the ability to directly observe pollution is absent, papers have historically assigned treatment using simple proximity to the pollution source (e.g., Clay et al., 2016) or by utilizing a combination of wind direction and proximity to the pollution source (e.g., Schlenker and Walker, 2015; Rangel and Vogl, 2019). We build on these techniques by utilizing sophisticated pollution modeling tools, joining recent environmental economics papers such as Sullivan (2017) and Grainger and Ruangmas (2018) who integrate these models into their empirical approach. Specifically, we utilize the National Oceanic and Atmospheric Administration’s (“NOAA”) HYSPLIT model to predict dispersion of wildfire smoke (Stein et al., 2015). HYSPLIT, which is used by the U.S. Forest Service to predict wildfire pollution, utilizes the location, date, and size of the fire (in acres) as inputs and combines them with gridded four-dimensional meteorological data and fuel composition data obtained from the U.S. Forest Service’s BlueSky modeling framework (Larkin et al., 2009) to predict

¹⁵A potential concern with historical data is that recorded fires could increase over time solely due to improvements in tracking them. However, we note that there is no statistically detectable trend ($p = 0.58$) in the number of acres burned over the sample period, alleviating this concern.

smoke plumes from each fire that compare favorable to satellite observations. In addition to the satellite comparisons performed by existing research, we also validate the model-generated smoke plumes by assessing their impact on ground-level pollution. This analysis, which is detailed in Section 1.4, finds that modeled smoke exposure is highly predictive of elevated pollution in a modern (2004-2019) context.

Our use of pollution modeling provides several advantages over the proximity or wind direction methods. First, HYSPLIT incorporates the amount of fuel burned, which is a function of both (a) acres burned and (b) the underlying fuel source (i.e., the density of forest or brush in the area of the fire). Given that the sizes of fires in our sample range from 1,000 acres to over 200,000 acres and occur everywhere from brush land to densely wooded forests, this combination is critical in accurately assessing the amount of air pollution from wildfires. Second, pollution modeling utilizes wind dynamics in a more sophisticated manner than is typically observed in existing methods by further incorporating wind speed and wind direction outside of the immediate vicinity of the fire. Wind speed is an important consideration for fires because calm winds can result in smoke patterns that diffuse around the origin point of the fire, while strong winds can send fire pollution for hundreds—if not thousands—of miles in a specific direction. Because wildfire smoke can travel so far from its origin, inclusion of wind dynamics that are not in the immediate vicinity of the fire are also necessary to correctly assign exposure to wildfire smoke. These considerations are illustrated by Appendix Figure A.5, which displays modeled smoke output under three different scenarios with varying fire size and wind intensity. In each scenario, the smoke plumes differ meaningfully, underscoring the importance of incorporating fire size and rich wind dynamics into the treatment definition.

While utilizing HYSPLIT provides several advantages over existing treatment definitions, there are two limitations to the smoke plume output. First, HYSPLIT generates smoke plumes as if the entire acreage is burned within first twenty-four hours from the start of the fire. Because wildfires often last several days, we generate smoke plumes for three separate dates: the date in which the fire began and two subsequent days.¹⁶ Second, the model allows use of meteorological data only back to 1948, which is roughly in the middle of our sample period. To compensate for this, we match each fire date from *before* 1948 with a date *on or after* 1948. This matching process is done by finding the date with the smallest Euclidean distance across several different measures of wind direction and intensity at the ground level as well as multiple levels of the atmosphere.

¹⁶Areas are considered to be treated if they spatially intersect any of the plumes generated by this procedure.

The details of this matching process are covered in more detail in Appendix Section A.1.

The identifying variation for our analysis is displayed in Figure 3.2, which details measures of smoke exposure across birth cohorts and across space. Panel A demonstrates the degree to which birth cohorts are exposed to wildfire smoke during the period from *in utero* through age 4 (“IU-Age 4”).¹⁷ As further discussed in Section 1.5, we focus on this particular age range because that is when children appear to be most sensitive to wildfire shocks. Figure 3.2 Panel A displays the mean, interquartile, and interventile ranges of *in utero* through age 4 smoke by year-of-birth cohort, along with the mean and interquartile range (“IQR”) for the entire sample. The figure illustrates that the amount of exposure varies substantially across California birth cohorts, with peak exposure for cohorts born in the 1950s and lowest exposure for 1930s cohorts.

Even within cohorts, there is substantial range in exposure, as demonstrated by the meaningful differences between the most- and least-intensely treated individuals. This within-cohort variation is more fully illustrated by Panel B, which demonstrates how smoke exposure—defined as the percentage of months in a given period with any smoke coverage—varies over geography in ten-year bins. As displayed in the figure, there are clear geographic differences in exposure to smoke pollution. However, these differences do not come from the same heavily and lightly treated areas for every birth cohort. For instance, in some years, Northern California is more intensely exposed, while in other years Southern California is subjected to higher smoke pollution. As discussed above, this is due in part to fluctuations in fire timing and severity. It is also a function of wind speed and direction, factors that vary meaningfully, as demonstrated by Appendix Figure A.6. As discussed further in the sections to follow, these environmental determinants of smoke exposure provide plausibly exogenous variation in our treatment variable.

1.3.3 Outcome Data and Linkages using Restricted Census Data

To study the long-run effects of childhood exposure to wildfire pollution, we link smoke exposure in early life to three restricted-use datasets from the United States Census Bureau: the Social Security Administration’s (“SSA”) Numident file, the 2000 Decennial Census Long Form (a 1-in-6 sample of households) and the 2005-2019 American Community Surveys (“ACS”).¹⁸ There are

¹⁷While we do not observe actual gestational lengths in our data, an individual is assumed to have *in utero* exposure if there is a smoke plume present within the nine months before their birth. Further, exposure “through age 4” indicates smoke exposure at any time prior to a child’s 5th birthday.

¹⁸The SSA Numident file that we obtain through Census Bureau differs from other versions of the Numident database that are available publicly, such as the Berkeley Unified Numident Mortality Database (Goldstein et al., 2021), primarily in the larger number of cohorts covered by the SSA/Census Numident and the ability to link to

three primary features about these data make them ideal for our analysis. First, the Numident file—which is an administrative database of all individuals who have ever been assigned a Social Security number—includes both a person’s city and state of birth, as well as their exact date of birth. These characteristics allow us to assign measures of smoke exposure to individuals in a very precise manner.¹⁹ Second, in addition to detailed information about birth, the Numident file includes mortality information for all individuals in the database. This mortality data is particularly comprehensive for deaths occurring after the mid-1970s (Chetty et al., 2016; Finlay and Genadek, 2021), which we discuss further below. Third, the restricted datasets can be linked together by using a unique, individual-level identifier (a Protected Identification Key, which is essentially a scrambled Social Security number). This linkage allows us to apply our childhood exposure measures—assigned using an individual’s birth place and birth date information from the Numident data—to adult survey outcomes in the 2000 Census, ACS, and other Census products.

In this chapter, we evaluate the effects of early-life smoke pollution on longevity. Our measure of longevity is conditional upon surviving past childhood, because individuals need to obtain a Social Security number prior to inclusion in the Numident data. Within our sample, individuals obtain their Social Security number at 14.4 years of age, on average.²⁰ We evaluate mortality by age 55 as our primary outcome, as that is the oldest age at which we observe all cohorts in our data, and we also evaluate mortality at ages 30 through 80 to calculate a measure of life years lost due to wildfire smoke.

We limit all sample to individuals who had obtained their Social Security Number (“SSN”) by age 19, which includes 95.6% of individuals in our sampling frame, to ensure that any deaths during adulthood can be captured in the Numident death file.²¹ After applying these restrictions, the number of observations in our core sample includes approximately 6,656,000 individuals. The other data sets (discussed further below). Nonetheless, certain descriptive analyses in this paper will use the Berkeley Unified Numident Mortality Database (henceforth referred to as the “Public-version Numident”) to minimize the amount of restricted Census data that is subject to disclosure.

¹⁹The ability to observe *city* of birth is a substantial improvement over most data sources, which only provide state or country of birth—if they provide any birth place information at all. Having city-level data—rather than only county-level data—is particularly helpful in our setting of California because many counties in California have very large land areas. (The *median* county is nearly the size of Delaware and the largest is almost the size as West Virginia.) Thus, having city of birth detail considerably improves the accuracy of treatment assignment. Additionally, the ability to observe exact date of birth is also a substantial improvement over public data sources, which are typically limited to birth year.

²⁰The date of Social Security number (“SSN”) receipt is proxied by the “initial cycle date” in Numident. It is worth noting that the SSA’s “Enumeration at Birth” program, which allows parents to complete SSN applications for newborns as part of the birth registration process, did not start until 1987. Accordingly, individuals in our sample typically obtained their SSN in late childhood or adolescence.

²¹We include this sample restriction because individuals who obtain Social Security numbers later in life may be subject to survival bias.

summary statistics for mortality outcomes are displayed in Appendix Table A.1.

While the SSA Numident file provides key information necessary for our analysis, there are two challenges to using the data. First, place of birth data is generated by information that individuals enter when completing the application for a Social Security number. These responses are stored as non-standardized string variables, which include spelling errors, transcription errors, use of abbreviations, and/or inclusion of places that do not qualify as cities. For example, an individual born in Los Angeles, California might write their city of birth as “Los Angeles,” “Los Angels,” “Westwood LA” (a neighborhood in Los Angeles), or simply “LA.” In order to standardize these responses, we limited our sample to all individuals who reported their birth state as California and then hand-matched common responses to populated places detailed in the United States Geological Survey’s Geographic Names Information System (“GNIS”) file. The GNIS file includes the county, longitude, and latitude of each location.²² The remaining observations were matched using probabilistic algorithms, which are further discussed in Appendix Section A.2. Overall, we matched 99.5% of individuals to a GNIS location, where 99.1% of these matches either had a perfect text match or were hand-matched.

Second, there are challenges when utilizing Numident to evaluate longevity, the main outcome of our analysis. Because Numident began tracking deaths in the early 1960s—and not comprehensively until the mid-1970s (Finlay and Genadek, 2021)—rates of death-by-age are understated, particularly for our earliest cohorts. To correct for this, we constructed adjusted death rates that scale death rates per the Numident by death rates from the Social Security Administration’s Cohort Life Tables (“SSA Cohort Life Tables”), a process that is discussed in greater detail in Appendix Section A.3. As demonstrated by Appendix Figure A.8, the adjusted death rates provide a much better approximation to expected cohort death rates than the unadjusted values. Additionally, adjusted death rates for earlier cohorts (those born in the 1930s and 1940s) are more comparable with rates for cohorts born from 1950-1960, further suggesting that adjusted rates are more appropriate as the primary outcome. Nonetheless, we demonstrate in our Section 1.6 analysis that results are robust to using unadjusted rates.

²²Longitude and latitude relate to the centroid of each place. A potential concern with utilizing city centroids is that the location of these centroids could drift over time as city boundaries change. However, as demonstrated by Appendix Figure A.7 and discussed in the accompanying notes, this is likely not an issue because boundaries are quite stable over time.

1.3.4 Other Data

In addition to our longitudinal data linking early-life wildfire smoke exposure to long-run mortality outcomes, we also utilize several data sources for the purpose of supplemental and validation analyses, which we briefly outline here.

Environmental Protection Agency (“EPA”) Monitoring Data, 2004-2019. The EPA collects daily monitor readings for a variety of pollutants. In order to relate this monitoring data to smoke exposure, we generate modeled smoke plumes using 2004-2019 wildfire data from CAL FIRE and the techniques discussed in Section 1.3.2. These data are incorporated into the analysis in Section 1.4 which demonstrates that the modeled smoke measures are highly predictive of ground-level pollution.

Control Variable and Heterogeneity Data. We also utilize a variety of other variables—namely county-level aggregates, historical hospital data, and climate measures—to include in our robustness and in heterogeneity analyses (Section 1.7 and Section 1.8, respectively). The data sources from which these variables were derived, and our method for incorporating them into our main analysis data set, are discussed further in Appendix Section A.4.

1.4 Validation of Smoke Measures

Because modeled smoke exposure discussed in Section 1.3.2 will form the basis for our treatment variable, we show in this section that it strongly predicts ground-level air pollution in a modern context.²³ In order to do so, we estimate the following distributed leads-and-lags regression model:

$$Y_{it} = \gamma_{y(t),i} + \eta_{m(t),i} + \sum_{j=-4}^4 \beta_j AnySmoke_{i,t-j} + \varepsilon_{it}, \quad (1.1)$$

where the unit of analysis is a monitor i with associated pollution reading observed in year-month t . To control for time-varying and seasonality-specific trends for each monitor, monitor-by-year fixed effects ($\gamma_{y(t),i}$) and monitor-by-calendar-month fixed effects ($\eta_{m(t),i}$) are included in the regression.

The treatment variable, $AnySmoke_{i,t-j}$, is equal to one a monitor was exposed to smoke during period $t - j$. Specifically, we consider a monitor to be “exposed” if the modeled smoke plume spatially intersects with the monitor’s geocoordinates. We include leads and lags of the treatment variable to (1) account for serial correlation in smoke exposure and (2) trace out the

²³We do not validate our measure in a *historical* context because monitor data does not exist for long-ago periods.

effect of smoke exposure on pollution over time. The coefficients of interest on these leads and lags, $\beta_j \{j \in -4...4\}$, represent the marginal effect of smoke coverage during relative month j on a monthly pollution measure during month t .

The main results of this analysis are displayed in Figure 1.3, which shows the impact of predicted wildfire smoke plumes on *mean* daily PM_{2.5} and *max* daily PM_{2.5} in Panels A and B, respectively. Within the figure, the estimated impact of smoke coverage during relative month j is presented with 95% confidence intervals. The estimated impact on pollution during the month in which the fire occurs (β_0) is presented in the upper-right-hand corner of each panel, along with the associated standard error and baseline mean.

We find that our modeled smoke exposure predicts meaningful increases in ground-level PM_{2.5} during the month in which the fire occurs: mean daily pollution increases by 1.26 $\mu\text{g}/\text{m}^3$ (13% over the mean) and max daily pollution by 3.58 $\mu\text{g}/\text{m}^3$ (18% over the mean). These effects decay over the two months after the start of the fire, as fine particulates remain in the air for weeks at a time.²⁴ In addition to the impacts on PM_{2.5}, smoke exposure has an impact on a range of other pollutants. As displayed in Appendix Figure A.9, we find increases in mean and maximum for larger particulate matter (PM₁₀) of 9% and 15%, respectively, alongside 2-5% increases over the mean for Carbon Monoxide (CO), Ozone (O₃), and Nitrogen Dioxide (NO₂). Taken together, these results support the assertion that our modeled wildfire smoke measure is an effective predictor of pollution.

1.5 Methodology

To evaluate the long-run effects of early childhood wildfire smoke exposure, we estimate the following cohort distributed lags and leads model:

$$Y_{cb} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \sum_{k=-5}^5 \delta_k Smoke_{c,b+k} + \varepsilon_{cb}, \quad (1.2)$$

where the unit of analysis is a birth cohort born in a city c in year-month b . The treatment variable, $Smoke_{c,b+k}$, and its associated leads and lags detail the number of months with any smoke exposure at age k , where a month of smoke exposure is defined as any spatial intersection between a smoke

²⁴The slow decay of the effects is also likely due to the fact that some fires, particularly in recent years, can last for a very long time (>1 month). However, because wind direction is autocorrelated over short periods, our modeled smoke measure is still able to capture these lagged pollution effects, even though we are unable to explicitly incorporate the duration of the fire into our model.

plume and a child's city of birth during that month (see Appendix Figure A.10 for an example of how treatment is defined). The coefficients of interest on these treatment variables, δ_k ($k \in -5 \dots 5$), provide the marginal effect of an additional smoke month during age k . The δ coefficients for the *in utero* period through age 5 ($\delta_{IU} \dots \delta_5$) allow us to trace out the dynamic effects of additional smoke exposure at various ages. Meanwhile, the coefficients for the years leading up to conception ($\delta_{-5} \dots \delta_{-1}$) measure the impact of smoke exposure that occurs before a child could be conceived. Accordingly, these estimates collectively serve as a placebo test, since no effect is expected from air pollution exposure during these periods.

We include city-of-birth fixed effects (μ_c) to control for time-invariant factors related to individuals born in the same city.²⁵ Additionally, we include year-of-birth \times region-of-birth fixed effects ($\lambda_{y(b),r(c)}$) where region of birth is defined in our baseline specification as Northern or Southern California.²⁶ These fixed effects capture time-varying differences by regions of California that may differ in terms of smoke exposure, climate, and economic activity. Further, because both birth outcomes (Buckles and Hungerman, 2013) and wildfires are highly seasonal, we include calendar-month-of-birth fixed effects, $\alpha_{m(b)}$.²⁷ Standard errors for estimates from this equation are clustered at the county level.²⁸

In addition to our model detailed in Equation 1.2, we also estimate an equation that collapses the effects of childhood wildfire smoke exposure into a single variable:

$$Y_{cb} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \delta \cdot Smoke_c^{IU\text{-Age } 4} + \varepsilon_{cb}. \quad (1.3)$$

The treatment variable, $Smoke_c^{IU\text{-Age } 4}$ is a single continuous variable equal to the number of months with any smoke exposure from the *in utero* ("IU") period through age 4.²⁹ For notational convenience, estimates from this equation are referred to as our "summary estimates," "IU-Age 4 estimates," or " $\delta_{IU\text{-Age } 4}$ " throughout the remainder of this paper.

We focus on the *in utero* through age 4 period primarily for two reasons. First, as discussed

²⁵Cities with fewer than 100 births are grouped together with their nearest neighbors until each combined city group has at least 100 associated births during the sample period.

²⁶The counties grouped into either Northern or Southern California are displayed in Appendix Figure A.11, Panel A. In our robustness checks in Section 1.7, we instead utilize State Economic Areas—precursors to modern-day commuting zones—and find similar results. These areas are illustrated in Appendix Figure A.11, Panel B.

²⁷Seasonality in wildfires is demonstrated Appendix Figure A.12, which shows that 90.3% of acres are burned during the months of July through November.

²⁸This is true for all other regression estimates presented in this paper, unless noted otherwise.

²⁹Because we do not observe actual gestation period in our data, the *in utero* period is assumed to begin nine months prior to birth.

extensively in Currie and Almond (2011) and Almond et al. (2018), the years prior to age 5 are a particularly important time of development and shocks during that period have been shown to have particularly meaningful later-life impacts. Second, as will be demonstrated by the results in Section 1.6, the age-of-exposure effects largely dissipate by age 5 (consistent with the existing literature) and thus we focus on the impact of smoke exposure that occurs before that age.

Our research design relies on the assumption that the intensity of wildfire smoke exposure during the first years of life are uncorrelated with other factors, conditional on our fixed effect controls. Because we control for time-invariant factors specific to a given place of birth (via our city-of-birth fixed effects, μ_c) as well as region-specific, time-varying factors (via $\lambda_{y(b),r(c)}$, our vector of year-of-birth \times region-of-birth fixed effects), potentially confounding factors would need to be time-varying at a sub-regional (e.g., city or county) level. The leads in our distributed leads and lags model (δ_k for $k \in \{-5, \dots, -1\}$ in Equation 1.2) provide a useful diagnostic test as to whether or not such confounding issues exist. If such confounders exist, then these coefficients will likely be economically and statistically different from zero. Additionally, in Section 1.7, we consider different fixed effect and control variable configurations to address the possibility of unobserved confounders that are also conditionally correlated with wildfire smoke exposure and find that results are robust to these alternate specifications.

1.6 Long-Run Mortality Effects

The first outcome we estimate is the long-run impact of childhood wildfire smoke exposure on longevity, specifically survival to age 55, which is the oldest age that we observe all of the cohorts in our sample. (Recall that all analyses are conditional on surviving past childhood, as it is necessary to obtain a Social Security number to be included in the data.) The results of this analysis are displayed in Panel A of Figure 1.4 where each point represents a value of δ_k from Equation 1.2 along with associated 95% confidence intervals. Each point estimate represents the marginal effect of an additional smoke month during age k , where the *in utero* age ($k = \text{IU}$) represents the nine months preceding birth and negative ages ($k \leq -1$) represent years prior to the *in utero* period. Additionally, our measure summarizing the marginal impact of an additional month with exposure during the *in utero* through age 4 period (described by Equation 1.3) is presented in the upper-right-hand corner along with the associated standard error.

The results in Figure 1.4 demonstrate a clear effect of air pollution from wildfires during

early childhood: our summary estimates show that an additional month of smoke exposure during the *in utero* through age 4 period ($\delta_{IU\text{-Age}4}$) results in 0.34 additional deaths before age 55 per 1,000 individuals. Moving from the 25th to the 75th percentile—an increase of 5 months—results in 1.7 additional deaths before age 55 per 1,000 (for conciseness, the effect of moving from the 25th to the 75th percentile will henceforth be denoted as $p_{25} \rightarrow p_{75}$). This effect is a 1.8% increase over the age 55 cumulative death rate within our sample. The leads and lags coefficients, which show the age-specific exposure effects, demonstrate when childhood smoke exposure matters the most. We find that the effects of exposure are largest around age 2 and have dissipated by age 5.³⁰ We do not, however, find any effects for our placebo estimates of smoke exposure before conception (relative periods -5 through -1), supporting the assertions that exposed areas are not differentially trending prior to exposure or subject to changing birth composition in anticipation of future wildfire smoke shocks (see Section 1.7 for further discussion).

While studies of the *contemporaneous* effects of air pollution find the strongest impacts when children are *in utero* or infants, our age-specific effects imply that the most sensitive period for severe, transitory wildfire smoke shocks occurs when children are slightly older (ages 1-2). This could be for two primary reasons. First, slightly older children are likely exposed to much higher levels of wildfire smoke because they spend more time outside and have higher lung capacity. Time outside matters a great deal because smoke exposure is substantially higher outdoors—for instance, Burke et al. (2021) find that outdoor PM_{2.5} can be approximately 10 to 20 times higher than indoor measures during periods of heavy wildfire smoke. Further, because wildfire smoke tends to be transitory, this ratio is much higher than for persistent pollution sources that have time to infiltrate homes (Liang et al., 2021). Accordingly, it is reasonable that children who are still young enough to be vulnerable—but are old enough to spend more time outdoors during periods of elevated air pollution—have more pronounced responses to wildfire smoke. Second, because our result is conditional on surviving past childhood, it excludes deaths that occur very early in life. Accordingly, insofar as individuals die before obtaining a social security number—something which occurs at age 14, on average, in our sample—they will be excluded from this mortality calculation. We consider this by estimating cohort-level infant mortality in Appendix Section B.1 and adding this mortality into our estimates. While these added deaths increase the response to air pollution

³⁰There are some concerns that the attenuation of effects by age 5 could be due to measurement error, since treatment is assigned based on city of birth and thus may incorrectly assigned due to migration. We bound this measurement error using migration data from the 1940 Complete-Count Census (Ruggles et al., 2021) and find that it does not qualitatively affect the results. See Section 1.7 for details.

in early life, our effects are still most pronounced for slightly older children.

We further explore the effects of wildfire air pollution on cumulative mortality across a variety of ages in Figure 1.4, Panel B. Specifically, we estimate summary coefficients ($\delta_{IU\text{-Age}4}$) for regressions where the outcome is cumulative mortality by age a (where $a \in \{30, 35, \dots, 80\}$). The results of these separate regressions are displayed in Panel B of Figure 1.4 along with 95% confidence intervals. We find that small effects manifest relatively early (age 30) and return to zero by age 80.³¹ Note that age effects for *any* estimate of cumulative mortality must mechanically “fade” to zero as all members of the counterfactual group eventually die of old age. In our context, it appears that wildfire pollution shocks do not appreciably affect mortality after age 80, although the estimates are highly imprecise. To aggregate these mortality effects, we calculate the life years lost between ages 30 to 80 and find that an additional month of wildfire smoke exposure results in 9.2 life years lost per 1,000 persons ($p_{25} \rightarrow p_{75} = 46$ per 1,000).³² As discussed in Section 1.3.3, our mortality data are adjusted to better match aggregate death patterns in our sample period. We re-estimate the analyses detailed above using the *unadjusted* rates and obtain similar estimates—8.1 life years lost per additional smoke month compared to 9.2 life years when using adjusted death rates. See Appendix Figure A.13 for more detail.

To further contextualize our mortality estimates, we compare the cohort-level life years lost due to smoke-induced infant mortality to the reduction in life years attributable to decreased adult longevity. As detailed in Appendix Section B.1, we calculate the effects of wildfire pollution on infant mortality using California vital statistics data and a cohort distributed leads-and-leads model in a similar spirit to the one specified in Equation 1.2. From our estimates, we calculate that moving from the 25th to 75th percentile of smoke exposure results in 11.7 life years lost per 1,000 births due to infant mortality, as compared a reduction of 46 years per 1,000 due to shorter adult life spans, as shown above. Decreased longevity results in a larger number of life years primarily because the critical period where wildfire smoke can reduce adult longevity is substantially longer than the period where infants are at high risk of death from wildfire smoke. That is, there are simply more opportunities for air pollution from fires to reduce adult longevity. Given that reductions in infant mortality are a staple of cost-benefit analyses for environmental policy, our finding that early-life air pollution has even larger effects through the channel of reduced adult longevity underscores the importance of incorporating long-run effects of early-life harms into policy calculations.

³¹Tabular data on point estimates and associated standard errors are presented in Appendix Table A.2, Panel A.

³²Standard errors for this value were obtained using bootstrap (100 iterations).

1.7 Robustness

We conduct a variety of robustness tests to ensure that our estimates do not qualitatively change when we (1) evaluate the direct effects of the fires themselves; (2) vary our specification or choose alternative treatment definitions; (3) consider the effects of mortality selection; (4) adjust for migration-related measurement error; or (5) employ alternate estimation methods. We discuss each of these exercises below.

1.7.1 Controlling for Proximity to Fires

This paper has thus far attributed the later-life harms of wildfire exposure to the substantial increases in air pollution that stem from these fires. However, wildfires sometimes cause physical damage to populated lands and may also necessitate evacuation or other responses that could result in childhood trauma. Accordingly, it is worth evaluating the degree to which our treatment may capture effects of wildfire-induced trauma, rather than harms purely from early-life air pollution exposure. To do so, we re-estimate our leads and lags specifications and include controls for fires that are within 10 miles of an individual’s city of birth—specifically, individuals are assigned direct fire exposure if *any* part of the fire is within 10 miles of the city centroid. As shown by Figure 1.5, controlling for proximity to fires does not materially affect our estimates, suggesting that wildfire smoke—and not trauma or economic damage from the direct effects of the fire—are driving results.³³

1.7.2 Changes to Specification

In this subsection, we include a variety of controls and vary the configuration of fixed effects in our baseline specification. Before detailing the results of these modifications, it is worth briefly discussing primary threats to internal validity and how these threats are either addressed by our main specification or can be diagnosed via our leads-and-lags coefficients. The two primary threats

³³Given that loss and displacement from fires are serious events that have been shown to have lasting consequences (Schwank, 2021), it is worth considering why we *don’t* see effects when controlling for these estimates. One simple explanation is that, since fires tend to occur away from population centers, very few fires are sufficiently near to people to cause lasting harm. Anecdotal evidence from historical newspaper records, which were reviewed to supplement missing information on wildfire dates, is consistent with this assertion. We noted only sparing reports of large evacuations or towns that were lost to wildfires, and even these locations tended to be sparsely populated resort or logging towns. However, it is worth noting that in *recent* years there have been several notable incidents where fires directly affected population centers. This is attributable, in part, to the rapid growth of the number of individuals living in the wildland urban interface (Radeloff et al., 2018), which both increases the exposure of populated places to naturally occurring wildfires—such as those caused by lightning strikes—and increases the risk that population-adjacent wildfires will be caused by human activity.

to internal validity would be if: (1) wildfire smoke was spuriously correlated with other characteristics and trends in a given area; or (2) wildfire smoke was co-determined with other weather-related factors—such as temperature, precipitation, or wind—that affect long-run outcomes. Our main specification addresses the first potential threat to identification by including detailed city of birth fixed effects (that finely control for time-invariant factors), along with year-of-birth \times region-of-birth effects (which control for cohort-varying factors that are specific to Northern and Southern California, respectively). However, in the event that these controls do not adequately adjust for spuriously correlated time-invariant or time-varying factors, our lead coefficients ($\delta_{-5}, \delta_{-4}, \dots, \delta_{-1}$)—which capture the marginal effect of an additional month of smoke exposure *prior* to conception—provide useful tests. If these placebo coefficients display pre-trends or are consistently and significantly different than zero, then that would indicate potential issues with our analysis.³⁴ However, we do not observe level shifts or trends in the placebo coefficients displayed in Section 1.6, supporting the idea that time-invariant and time-varying factors are not confounding our analysis.

In addition to the credibility added by our placebo tests, we also further address potential threats to internal validity by adding various controls and modifying our fixed effect specifications. The results of these tests are shown in Figure 1.6, which displays summary estimates for our baseline specification (on the top row) along with summary estimates of alternative specifications (on other rows and described by the in-figure text). We briefly discuss each of these robustness specifications below.

Additional Controls

Our first specification includes controls for climate variables, which have been shown to be both a strong determinant of wildfires and also human health in the short- and long-run (e.g., Barreca et al., 2016; Isen et al., 2017a). Specifically, we include controls for the number of days above 32 degrees Celsius during the *in utero* period through the first six months of life—i.e., the critical period in which temperature strongly affects long-run earnings, as shown by Isen et al. (2017a).

³⁴To provide specific examples, if there are time-varying trends (i.e., if areas with high propensity for smoke exposure are trending in a different manner than low-propensity areas), we would expect to see trends in our placebo coefficients prior to the conception date. Additionally, the lead coefficients will also detect if there is a compositional response in *anticipation* of a wildfire smoke shock. Consider if, for instance, wealthier or more educated parents—who are expected to have wealthier and more educated children—move away from an area just before their child is born. In this case, we would expect to see a preconception trend since the composition of the children born in a given area is changing over time. However, we see no trends that would suggest this is occurring. (Regarding the specific threat of anticipatory migration, we also do not find any evidence that wildfire smoke exposure induces migration, as indicated in the last column of Appendix Table A.4. This is also consistent with findings from Borgschulte et al. (2020) who do not detect smoke-induced migration behavior in a modern context.)

We also control for (a) the mean of monthly precipitation and (b) the fraction of days in cardinal wind direction bins interacted with an individual’s county of birth. Both sets of these controls are calculated over the *in utero* through age 4 period.³⁵ However, as displayed in the Figure 1.6, controlling for these outcomes does not qualitatively affect our estimates.

Next, in order to better control for granular time trends, we follow a standard practice in the early-life determinants literature by interacting baseline county-level characteristics with time trends (recent examples include Bailey et al., 2020, 2021; Hoynes et al., 2016). Specifically, we interact 1930 county-level data with a decade-varying-linear spline in time.³⁶ We find that inclusion of these interacted splines does not meaningfully change our estimates.

Last, we control for the possibility that more-and-less smoke prone areas are trending differently by dividing overall smoke propensity (defined over the entire period) into quartiles and interacting these quartiles with birth year fixed effects. Inclusion of these additional fixed effect controls do not significantly affect our estimates.

Different Fixed Effect Specifications

We next consider different fixed effect specifications that increase the granularity of controls along time and spatial dimensions. First, we replace our year-of-birth by region-of-birth fixed effects—where region is defined as Southern California or Northern California—with more detailed year-month-of-birth by region-of-birth fixed effects. If there are short-lived shocks that are contemporaneous with wildfires—such as regional temperature spikes or other extreme weather—they will be captured by these controls. Next, we instead increased the granularity of our region-of-birth fixed effects to be based on State Economic Areas (precursors to modern-day commuting zones), rather than a South-North distinction.³⁷ These controls will capture relatively local trends that may not be common to the larger region. As displayed in Figure 1.6, modifications to these fixed effects yield highly similar estimates to our baseline specification.

³⁵Precipitation could affect children primarily via agricultural income shocks. Additionally, we control for county-by-wind direction bins because it is possible that wildfire smoke could blow in exclusively from an otherwise “good” or “bad” direction, i.e., a direction that has little or lots of pollution from non-wildfire sources (Deryugina et al., 2019). If this is the case, then we would systematically under (over) state the effects of wildfire smoke on long-run outcomes.

³⁶These characteristics include: (1) the fraction of children in school; (2) the percent immigrant population; (3) the relative share of manufacturing and agricultural employment; (4) the average age at first marriage; and (5) the average family income. All characteristics were obtained from the 1930 Census 100% Sample (Ruggles et al., 2021), except for family income, which was obtained from the 1934 IRS Statements of Income. (No income-related questions were asked during the 1930 Census.)

³⁷See Appendix Figure A.11 for illustration of State Economic Areas.

Different Treatment Definition

Our treatment definition implicitly assumes that each additional month of exposure during a given age range has the same marginal effect—i.e., that the dose-response function is linear in the number of months with *any* smoke exposure from the *in utero* period through age 4. To relax this assumption, we redefine our treatment variable to be an indicator variable equal to one if an individual is exposed to an above-median amount of smoke and zero otherwise. Specifically, for an in-sample median smoke measure \bar{m} , we estimate:

$$Y_{cb} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \tau \cdot \mathbf{I}(Smoke_c^{\text{IU-Age } 4} > \bar{m}) + \varepsilon_{cb}, \quad (1.4)$$

which is the same specification as Equation 1.3 with a transformed treatment variable. To facilitate comparison with our estimated treatment, we also estimate a separate equation with $Smoke_c^{\text{IU-Age } 4}$ as the *dependent* variable to obtain the average number of additional smoke months for those with above-average exposure:

$$Smoke_c^{\text{IU-Age } 4} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)}\kappa \cdot \mathbf{I}(Smoke_c^{\text{IU-Age } 4} > \bar{m}) + \varepsilon_{cb}. \quad (1.5)$$

We then scale the effects from Equation 1.4 ($\hat{\tau}$) by those from Equation 1.5 ($\hat{\kappa}$) to generate a comparable estimate to our main results. The scaled estimates from this exercise, presented in the final row of Figure 1.6, are qualitatively very similar to the effects estimated from our baseline, although they are less precise, consistent with the decreased variation in the treatment variable.

In addition to this exercise, we also provide evidence supporting the linearity of the dose-response function using residualized scatter plots. These plots, which are displayed in Appendix Figure A.19, generally support the choice of a linear functional form for our analysis. See accompanying notes to these figures for further information on how plots were constructed and additional discussion of interpretation.

1.7.3 Measurement Error

Because our estimates use place of birth to assign treatment, our treatment will be mis-measured to some degree for any individuals who migrate away from where they were born within the first four to five years of life. This is of particular concern when interpreting our age-specific estimates, as estimates at older ages will mechanically have higher rates of migration away from their birthplace

and thus have higher degrees of measurement error. To assess the potential impact of measurement error on our main result, we utilize the 1940 Census 100% Sample to determine where children live at age five and how it compares to their place of birth, which is inferred using their 5-year migration status.³⁸ With these data, we restrict our sample to children age five who were born in California and utilize information on their counties of birth and residence to calculate measures of “true” exposure and exposure based solely on place of birth (i.e., assigned in the same way as our main analyses). Then, using these assigned treatments, we calculate how much our estimates have potentially been attenuated due to measurement error, under the assumption that error in assignment is classical. A more detailed description of this approach, as well as support for the assumption of classical measurement error, are discussed in detail in Appendix Section B.4.

The results of this exercise, where we adjust our estimates for the calculated attenuation factors, are displayed in Appendix Figure A.22. Within the figure, we present the attenuation-adjusted leads-and-lags coefficients for our main outcomes from Section 1.6 alongside the original estimates. As demonstrated by the figure, the measurement-error-adjusted estimates are nearly indistinguishable from our baseline estimates, likely due in part to the relatively low rates of migration among small children (only 14% of children had migrated by age 5 within the sample).

1.7.4 Issues with Two-Way Fixed Effect Estimates and Alternative Estimation

This paper utilizes two-way fixed effects (“TWFE”) controls in our estimation of long-term outcomes. However, as discussed in a fast-growing literature (e.g., Callaway and Sant’Anna, 2021; de Chaisemartin and D’Haultfoeuille, 2020; Goodman-Bacon, 2021a; Sun and Abraham, 2021), TWFE models may be difficult to interpret due to the unintuitive way in which they weight different observations across time. This can be particularly problematic if there is staggered timing, time-varying treatment effects, and/or dynamic treatment effects. Recently, several papers—such as Cengiz et al. (2019) and the previously referenced studies—have proposed methods for handling these issues in commonly used difference-in-difference and event-study frameworks where a policy is enacted and remains in effect.

However, the existing literature has not settled on the best way to deal with these weighting

³⁸The 1940 Census is best-suited for this analysis because it asked 5-year migration questions of all individuals in the 100% sample. This contrasts with later Decennial Censuses where migration was asked only of a smaller sub-sample and/or was asked only about 1-year migration. In this context, the larger sample size is particularly useful as we are looking at relatively small sub-sample of individuals, i.e., those who were five years old in 1940 and were also born in California.

issues for continuous treatments that turn on and off, which is the case for our research design.³⁹ To address potential weighting issues, we estimate an alternate regression specification that is separately estimated *for each year-of-birth cohort* and we then combine these estimates by simple average to ensure intuitive weighting. Specifically, we estimate:

$$Y_{cb} = \lambda_{y(b),r(c)} + \alpha_{m(b)} + \delta \cdot Smoke_{cb}^{\text{IU-Age4}} + \bar{Y}_c + f(X_{cb}) + \varepsilon_{cb},$$

where the above equation is very similar to our summary specification (Equation 1.3), but differs in key ways to allow separate estimation for each year-of-birth cohort. Because we are estimating each regression using only within year-of-birth cohort (cross-sectional) variation, we are unable to utilize our city-of-birth fixed effects (μ_c) that are present in Equation 1.3. To increase precision and control for potential differences across places that may be correlated with both wildfire smoke and long-run outcomes, we instead include long-run city-of-birth averages of the outcome variable (\bar{Y}_c) as well as indicators for quartiles of smoke propensity, county-level demographic controls, and climate controls discussed earlier in this section. The results of this exercise are presented in Appendix Figure A.24. Within the figure, the estimates from our alternative estimation are qualitatively similar, though less precise—although our results pertaining to longevity, economic achievement, and educational attainment all exceed traditional levels of statistical significance.⁴⁰

1.8 Supplemental Results

We have established that early-childhood wildfire exposure results in adverse long-run impacts to longevity. In this section, we provide further evidence of potential additional mechanisms for our mortality results.

1.8.1 Early-Life Exposure to Wildfire Smoke Induces Individuals to Sort into Higher-Mortality Occupations and Neighborhoods

As discussed in Section 1.2, there are many ways in which early life air pollution exposure could affect longevity. One potential way in which this could occur is if individuals are induced into occupations or neighborhoods that increase their risk of death via environmental exposures, occu-

³⁹Very recent work by Callaway et al. (2021) provide alternative estimation strategies for difference-in-differences designs with continuous treatment. Future versions of this paper will consider the applicability of these methods to our setting.

⁴⁰We obtain standard errors through 100 cluster bootstrap iterations.

pational safety hazards, increased levels of stress, higher rates of violent crime, or other factors. To further explore this possibility, we examine the additional mortality risk that is associated with individuals' choice of occupation and neighborhood. Specifically, we utilize the full Numident file (i.e., without restricting to Californian births) and link it to both the full American Community Surveys and 2000/2010 Decennial Censuses. Using these data, we then restrict to the same age ranges that are covered within this paper and calculate 10-year mortality rates for each occupation and Census tract.⁴¹ We then assign occupation-specific mortality rates to individuals based on their observed occupation in the ACS and assign tract-specific mortality rates to individuals based on their residence in the Decennial Census.

Having constructed these ten-year mortality measures, we utilize them as outcome variables to assess whether childhood smoke exposure causes individuals to sort into higher-mortality occupations and neighborhoods. The results of these regressions, displayed in Figure 1.7, clearly demonstrate that more exposed individuals sort into occupations and neighborhoods with greater mortality risk. While this result does not conclusively show that longevity is increasing due to occupation and neighborhood choice—after all, lower socioeconomic status exposes individuals to a number of increased mortality risks that could also be correlated with these choices—it does provide suggestive evidence as to the channels that could contribute to premature mortality.⁴²

1.8.2 Wildfire Shocks Are Not Transmitted to Children Via Local Economic Channels and Long-Run Effects Do Not Meaningfully Vary Across Person and Place Characteristics

We perform two other supplemental analysis, both of which are discussed further in Appendix Section B.5 but are briefly summarized here. First, we perform an analysis of contemporaneous economic outcomes (e.g., employment and family income) and net migration to ascertain if wildfire smoke shocks could be transmitted to children via local economic channels (see Appendix Table A.4). We do not find any statistically significant evidence of reductions in economic activity. Further, the reductions that we do find are also economically small and cannot explain a substantial portion of the longevity results that we find.

Second, we consider heterogeneity across individual-level and place-level characteristics,

⁴¹Occupation-specific mortality rates are also within sex and 5-year age bin, while tract-specific mortality rates are within sex.

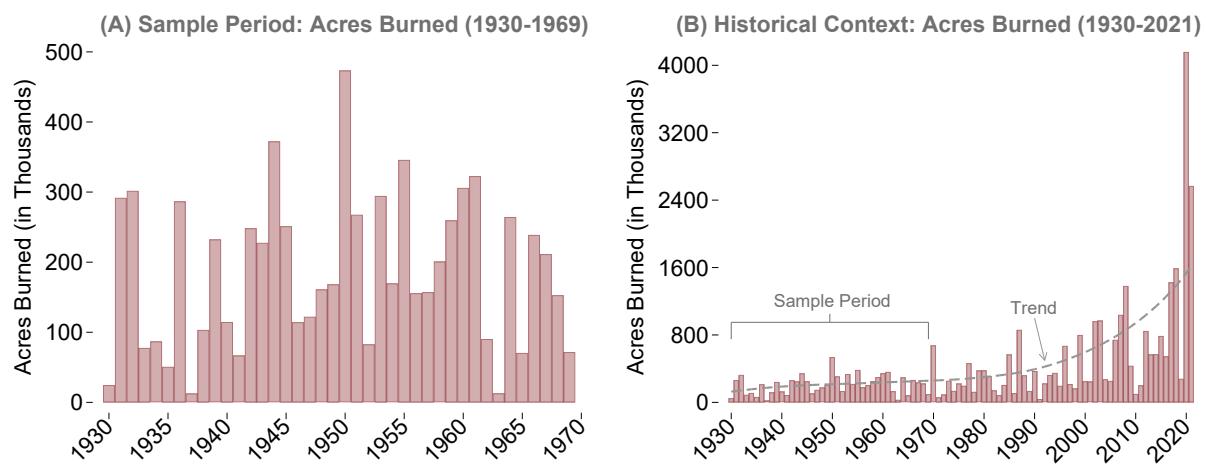
⁴²Appendix Figure A.18, Panel B, also considers the 10-year mortality rate of an individual's industry and finds qualitatively similar findings to those discussed here.

including sex, race/ethnicity, urban status, county-level income, hospital access, and county-level economic dependence on lumber (i.e., places that could be more sensitive to loss of timber due to wildfire). The results of this analysis, shown in Appendix Figure A.26, display little-to-no heterogeneity. Out of our twenty-four regressions (six groups \times four outcomes), there are only three that have statistically different interaction effects.

1.9 Conclusion

In this paper, we contribute to the understanding of the long-term effects of childhood air pollution exposure by leveraging variation in an increasingly important source: smoke from wildfires. Specifically, we examine longevity, a previously unstudied outcome in the literature and find that moving from the 25th to the 75th percentile of smoke exposure results in 46 life years lost per 1,000 persons. In the conclusion of the next chapter, we incorporate these harms with losses due to reductions in economic activity in order to determine the costs of wildfire smoke exposure within our cohorts. Additionally, the next chapter also discusses a variety of potential policy implications for mitigating and ameliorating the effects of early life air pollution exposure from wildfires.

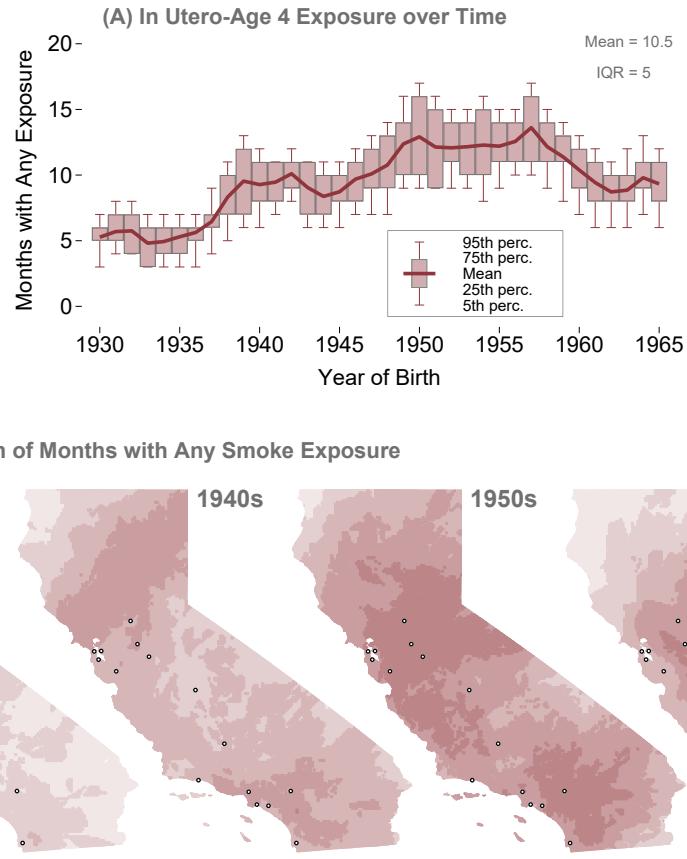
Figure 1.1: Descriptive Statistics: There is Substantial Variation in Year-to-Year Acres Burned, but Intensity has Increased Sharply over the Last 30 Years



Notes: The purpose of this figure is to display the substantial over-time variation in wildfires during the sample period (Panel A) and the stark increase in wildfire acreage burned over the last thirty years (Panel B). See Section 1.3.1 for further discussion.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018) and 2021 fire data from fire.ca.gov/incidents/ (updated through October 11, 2021).

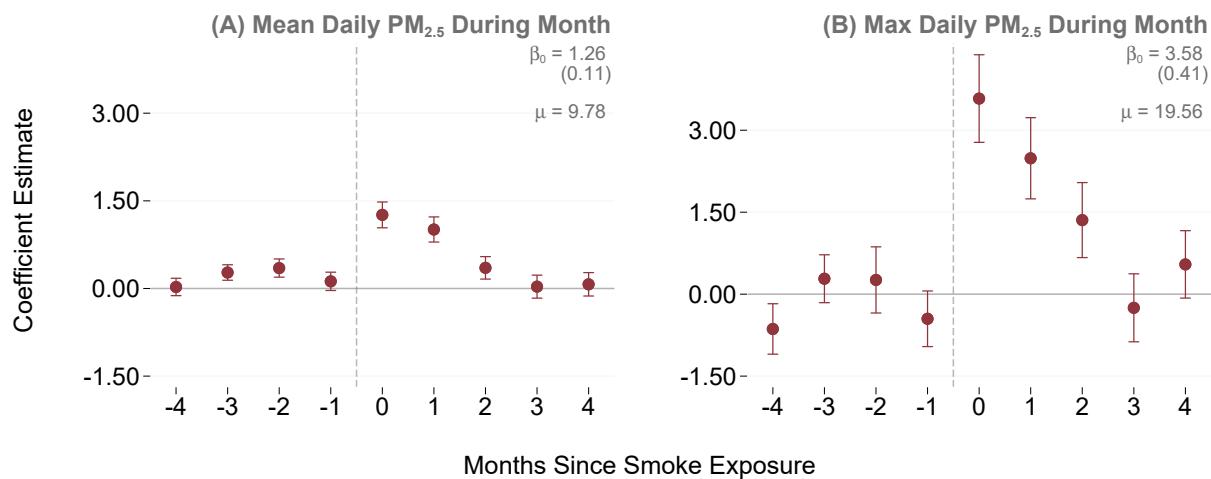
Figure 1.2: Identifying Variation: Smoke Exposure Varies Over Time and Geography



Notes: The purpose of this figure is to display the identifying variation utilized in our analysis. Panel A displays how the birth-weighted months of smoke plume exposure from an individual's *in utero* period through age 4 ("IU-Age 4") varies across birth cohorts. Within this panel, the mean and various quantiles are plotted for each cohort with the overall mean and interquartile range ("IQR") plotted in the upper-right corner. Panel B displays the spatial variation in smoke exposure according to the percentage of months in a given decade that experienced smoke exposure.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), the Public-version NumIdent (Goldstein et al., 2021).

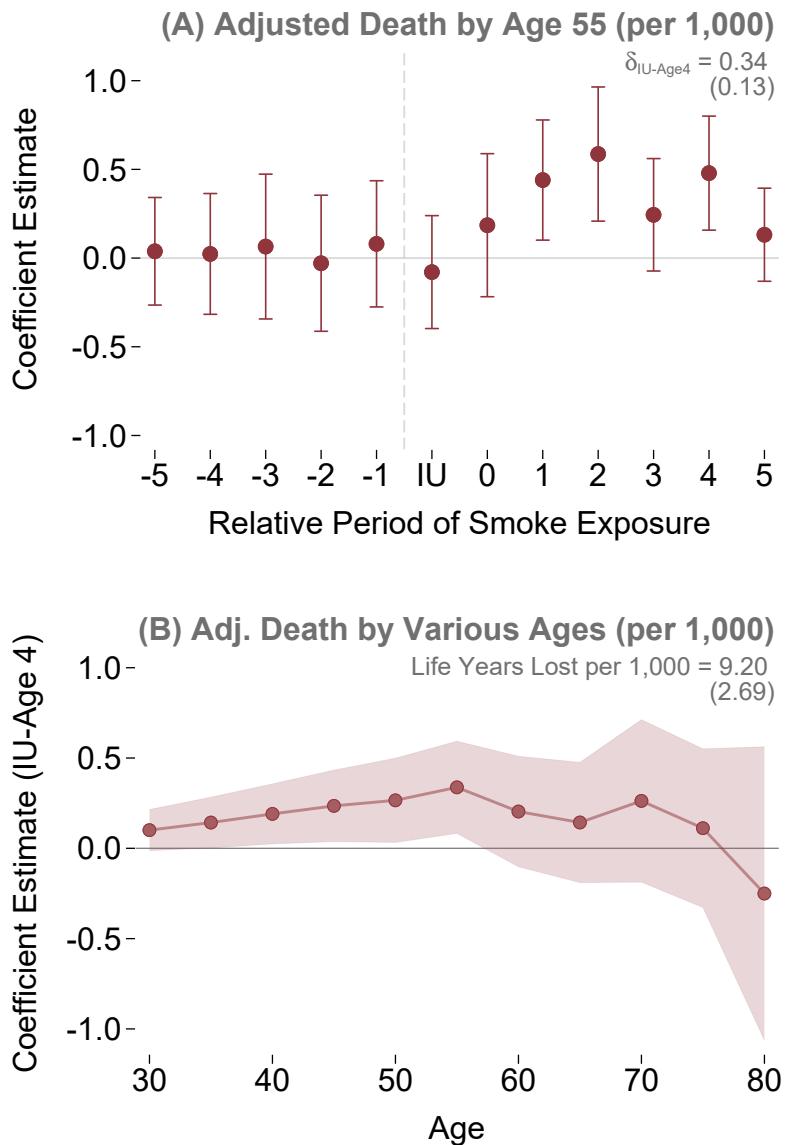
Figure 1.3: Validation of Treatment: Modeled Wildfire Smoke Exposure Increases Monthly PM_{2.5}



Notes: The purpose of this figure is to display the impact of modeled smoke plume coverage on monthly pollution levels. Each point represents a coefficient from the regression detailed by Equation 1.4 with corresponding 95% confident intervals. The point estimate for the impact on pollution during the month in which the fire occurs (β_0), along with the standard error and baseline mean are presented in the upper right-hand corner of the graph.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), and Daily Summary Data Files from the EPA Data Mart.

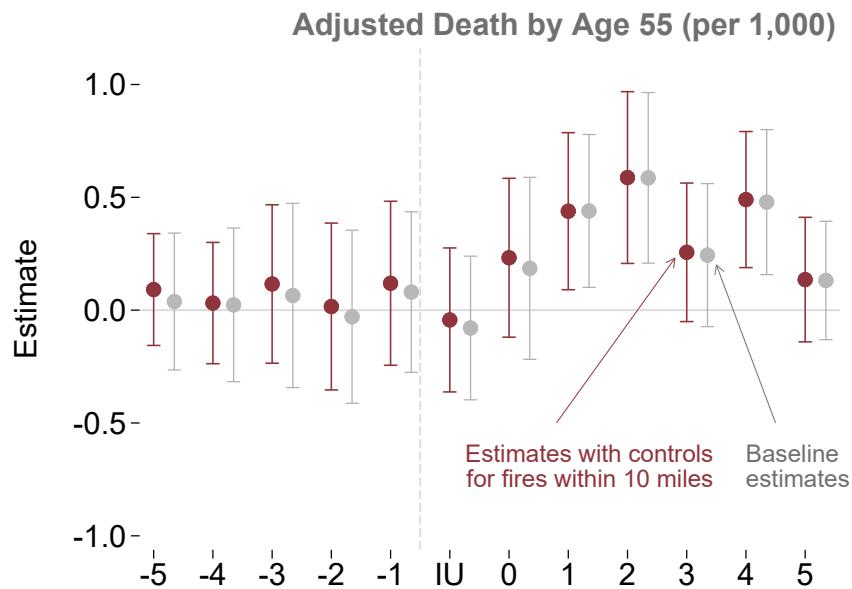
Figure 1.4: Main Results: Early-Life Wildfire Exposure Reduces Longevity in Adulthood



Notes: The purpose of this figure is to display the impact of wildfire smoke exposure on longevity in adulthood. Within Panel A, each point represents a δ_k coefficient from the regression detailed by Equation 1.2 with corresponding 95% confident intervals, representing the marginal effect of an additional month of smoke exposure during the stated period. The upper right-hand corner displays the summary coefficient for an additional month of exposure during the *in utero* period through Age 4 (Equation 1.3), along with the associated standard error. Panel B displays the summary coefficients from Equation 1.3 and associated 95% confidence intervals for the effect of an additional month of exposure on at death by various ages (30 through 80).

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), and Restricted Census Numident.

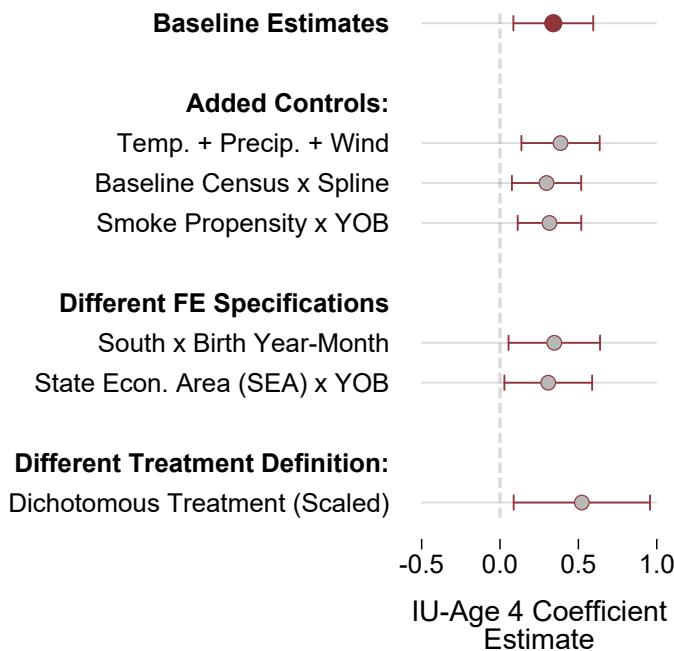
Figure 1.5: Robustness: Controlling for Nearby Fires Does Not Affect Estimates, Implying that Smoke Exposure is Driving Effects



Notes: The purpose of this figure is to demonstrate that smoke exposure, rather than proximity to wildfires themselves, are driving for estimates. To do so, the figure compares our baseline estimates from our leads-and-lags specification to those when we re-estimate the specification with controls for fires that are within 10 miles of an individual's city of birth. See the text of Section 1.7.1 for further discussion.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

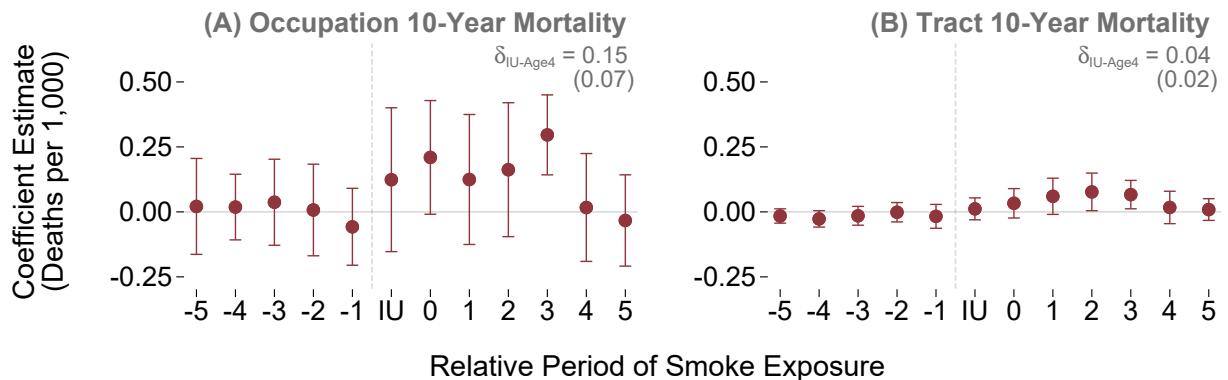
Figure 1.6: Robustness to Additional Controls and Specifications



Notes: The purpose of this figure is to display the robustness of results to differing specifications. Within the figure, our summary coefficients (Equation 1.3) for our preferred specification are presented on the top line, with estimates for differing specifications (due to added controls, changes to fixed effect structure, or modifications of functional form) included in the rows below. See Section 1.7 for further description.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Figure 1.7: Supplemental Results: Individuals Exposed to Higher Levels of Early-Life Wildfire Exposure Sort into Higher-Mortality Occupations and Neighborhoods



Notes: The purpose of this figure is to display the how early childhood wildfire exposure induces individuals to sort into higher-mortality occupations and neighborhoods. See notes to Figure 1.4 for discussion of graph construction and the text of Section 1.8.1 for more detail regarding construction of the outcome variable.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, Restricted American Community Surveys, and Restricted Full-count Decennial Censuses.

Chapter 2

The Impact of Early Childhood Wildfire Smoke Exposure on Economic, Education, and Disability Outcomes

2.1 Introduction

It is now well established that air pollution increases contemporaneous mortality of both the very old and very young, two age groups most sensitive to their environments. However, substantially less is known about a related question, which we discuss in this paper: how do the harms of exposure to air pollution at very young ages manifest in reduced well-being and longevity in later adulthood? Given that long-run effects frequently comprise a large fraction of the damages from early-life shocks (Currie and Almond, 2011; Almond et al., 2018), understanding the ramifications of childhood pollution exposure across the life cycle is vital in assessing its total costs. This paper provides the first evidence—to our knowledge—quantifying the effects of childhood pollution shocks on measures of earnings and disability past early adulthood (through age 59).

To determine the long-run effects of early-life exposure to air pollution, we study wildfires, the only major emissions source of fine particulate matter ($PM_{2.5}$) in the United States that is growing over time and one that has generated approximately 29% of all fine particulate pollution

in recent years.¹ Additionally, it is expected that wildfire pollution will *continue* to grow: the United States Forest Service predicts that 1 °C increase in average temperature will result in a 200-600% increase in acres burned in the Western United States (Vose et al., 2012). Given that experts predict global temperatures will increase by at least 2 °C over the global baseline by 2050 (IPCC, 2021), the prospect of increased pollution from wildfires is nearly certain.

There are two key challenges to conducting a long-term analysis of childhood air pollution. The first is that pollution has only been reliably measured since the mid-to-late 1960s, with the first meaningful regulation occurring with the Clean Air Act Amendments of 1970. This lack of historical data and sources of exogenous variation has limited researchers' ability to evaluate the ramifications of early-life pollution exposure over the life cycle. We overcome this challenge by examining wildfires occurring in California, a source of pollution which has been tracked reliably since at least 1930. Additionally, wildfire smoke exposure, which is a function of fire timing, acres burned, wind direction, and wind speed, provides potentially exogenous variation for an analysis of long-term outcomes. In order to supplement these wildfire records and generate measures of smoke exposure, we first hand-gather dates in which the wildfires occurred using hundreds of historical newspaper articles. We then utilize wildfire pollution modeling tools to generate historical measures of wildfire smoke exposure. These modeled smoke measures are utilized by the United States Forest Service to predict wildfire smoke and compare well to patterns generated by satellites (Larkin et al., 2009). In this paper, we perform additional exercises to further validate this modeled smoke measure and find that it strongly predicts pollution exposure in a modern context.

The second challenge in assessing the long-run impacts of early-life shocks is linking childhood circumstances to later-life outcomes. To address this difficulty, we obtain restricted-use administrative data from the Social Security Administration and U.S. Census Bureau. In addition to providing comprehensive high-quality microdata, these data sets have two features that are ideally suited for our analysis. First, the administrative data provides exact city and state of birth, which allows us to accurately assign measures of wildfire smoke exposure to individuals. Second, these data also include unique identifiers which we use to link individuals—and their associated early-life exposure measures to their responses in the 2000 and 2010 Decennial Censuses and 2005-2019 American Community Surveys.

We evaluate the long-term effects of exposure using a cohort distributed leads and lags

¹Source: Author calculations using National Emissions Inventory data from the Environmental Protection Agency. See Appendix Figure A.1 for more detail.

model and find that the effects of childhood pollution exposure—which we define as exposure before age 5—last throughout the entire life cycle. Because of the wide range of outcomes available to us in our data—and because early-life shocks have been shown to affect nearly every facet of adult life—we standardize outcomes (à la Kling et al., 2007) and combine them into three indices covering economic achievement, educational attainment, and disability. We also specifically consider “headline” outcomes within each group that typify our results and are more easily interpreted. We find that moving from the 25th to 75th percentile of smoke exposure—roughly 5 additional months with *any* wildfire smoke—decreases our index of economic achievement by 1.45% of a standard deviation (%-SD), headlined by an \$890 decrease in annual earned income (defined as wages and self-employment income). Consistent with our economic results, we find that the same level of exposure decreases our educational attainment index by 1.35%-SD, including a 0.40 percentage point decrease in high school completion. Finally, we find an adverse change of 0.80%-SD in our index of disabilities. Our most notable finding within this group of outcomes is that moving from the 25th to 75th percentile of smoke exposure increases the likelihood of difficulty remembering, concentrating, or making decisions by 0.30 percentage points, supporting the idea that early-life pollution exposure impacts cognition and mental health throughout the entire life cycle. Taken together, we conservatively estimate that the average child in our sample sustains approximately \$22,000 of discounted damages in lost earnings and lost life years due to wildfire smoke.

To demonstrate that our results are driven by smoke exposure—rather than proximity to the actual wildfires themselves—we control for nearby fires during childhood. Inclusion of these controls has almost no effect on our estimates, suggesting that the impacts that we find relate to pollution damage, rather than long-run effects of trauma or negative economic shocks due to natural disaster exposure. Additionally, we show that our estimates are stable across a range of robustness tests, including adding controls for climate and other time-varying factors as well as different functional forms and specifications. We also consider the role of mortality selection by including bounded estimates in a similar spirit as Lee (2009) and find that wildfire smoke’s differential effects on longevity do not materially affect estimates of other long-run outcomes. Additionally, we consider alternate estimation methods to evaluate the impact of weighting issues endemic to models that utilize two-way fixed effects (e.g., Callaway and Sant’Anna, 2021; de Chaisemartin and D’Haultfœuille, 2020; Goodman-Bacon, 2021a; Sun and Abraham, 2021) and find that results are qualitative similar when estimated using other methods.

To better understand how childhood smoke exposure affects long-run outcomes, we conduct

a variety of supplemental analyses. First, to further understand how the dynamics of childhood wildfire smoke exposure evolve as individuals grow older, we consider heterogeneity by age at survey. We find economically larger effects—both in levels and in relation to the mean—on earnings and cognitive difficulty at older ages. This finding suggests that evaluations of the impacts of early-life air pollution—and potentially other childhood shocks—should consider later-life dynamics to capture the total costs. Additionally, to better understand how wildfire pollution exposure affects the distribution of outcomes, we consider heterogeneity across the economic index and wage distributions. We find that early-life exposure to wildfire pollution causes in a leftward shift of the entire distribution, suggesting that both extensive-margin labor supply and earnings (conditional on employment) are affected.²

Our findings that early-life exposure to wildfire smoke pollution adversely affects economic achievement, educational attainment, and disability in mid-to-late adulthood contribute to two distinct areas of economic research. First, we expand the literature examining the long-run effects of early-life air pollution on adult outcomes. This work has largely focused on the impact of the long-run effects of the Clean Air Act Amendments (“CAAA”) of 1970 on earnings around age 30 (Isen et al., 2017b) and on educational attainment (Colmer and Voorheis, 2020; Voorheis, 2017).³ This paper increases our understanding of the long-run effects of childhood pollution exposure in two primary ways. First, we causally assess the impacts of early-life air pollution substantially past age 30—in our setting, we examine impacts on ages 35 through 59—and are the first paper to do so.⁴ This is a meaningful contribution because early-life impacts in some settings can be detected at younger ages and then “fade” over time (see Currie and Almond, 2011; Almond et al., 2018, for examples), while it is also possible for childhood harms to *increase* proportionally as individuals age. Understanding that childhood air pollution shocks grow proportionally worse as individuals age—as we find in this paper—is important when assessing the cumulative costs of early-life harm from air pollution and may also have implications for other childhood health shocks. The second way this paper complements existing research is by examining a different—and highly relevant—source

²Additionally, we perform further analyses to understand how local economic shocks from smoke exposure (Borgschulte et al., 2020) could affect long-run outcomes, explore effect heterogeneity by person and place characteristics, and assess the impact of childhood air pollution exposure on other long-run outcomes that do not fall into the categories of economic achievement, educational attainment, or disability.

³An early contributor to this literature is Sanders (2012), who studies the long-run effects of recession-induced decreases in pollution on test scores. Additionally, there is also a small literature examining the impacts of the lead air pollution on educational attainment (Grönqvist et al., 2020) and crime/anti-social behavior (Aizer and Currie, 2019; Reyes, 2007, 2015).

⁴Within the medical literature, Russ et al. (2021) combine modeled pollution measures with longitudinal survey data ($n = 572$) and find an association between childhood air pollution and late-life cognitive decline.

of pollution. While the pollution from industrial and vehicular sources studied in this literature have been falling for many years, pollution from wildfire smoke has been increasing—and likely will continue to increase—over time.⁵ Given the growing prevalence of wildfires and evidence from medical research suggesting that particulates from wildfires are potentially more harmful to human health than pollution from other sources, understanding the harm done by wildfire pollution is increasingly important.⁶

Second, this study contributes to our rapidly growing understanding of the effects of pollution from fires on human health. Thus far, this literature has concentrated primarily on the contemporaneous effects of fire-related pollution on mortality (Jayachandran, 2009; Miller et al., 2017), infant health (Rangel and Vogl, 2019), and labor market outcomes (Borgschulte et al., 2020). As emphasized above, this paper adds to the existing knowledge of these health effects by quantifying the long-run harms of wildfire smoke, which we find to be a substantial component of the costs of these fires. Furthermore, insofar as pollution from these fires is expected to grow due to global increases in temperature (Moritz et al., 2012; Vose et al., 2012), we view this paper as an important contribution to our understanding of the costs of climate change (e.g. Jacob et al., 2007; Schlenker and Roberts, 2009; Graff Zivin and Neidell, 2014; Barreca et al., 2016; Hsiang et al., 2017).

2.2 Background and Conceptual Framework

2.2.1 Ambient Pollution and Childhood Development

Because this paper estimates the impact of air quality in early life on outcomes in late life, in this section, we first explain the biological influence of ambient pollution on human health. Further, we document several reasons why wildfire smoke may be more harmful than other forms of ambient pollution. Throughout, we focus on the role of fine particulate matter, or PM_{2.5}, because it is widely considered to be the most harmful to human health (Peeples, 2020).

PM_{2.5} is particulate matter, either solid or liquid, that is less than 2.5 microns in diameter (for reference, a human hair is about 70 microns in diameter). It enters the body through inhalation into the lungs. Although some of the matter will be removed by normal clearance mechanisms, some of the matter will reach the bloodstream, where the particulates can cause a multitude of

⁵ As demonstrated in Appendix Figure A.1, wildfire smoke comprised over 29% of fine particulate matter pollution during the 2017-2020 period, more than the emissions from fuel combustion/generation, industrial activities, and vehicular use *combined*.

⁶See, for example, Aguilera et al. (2021a) and Aguilera et al. (2021b)

cardiovascular and respiratory problems. Although the precise pathways are not yet fully understood, a leading hypothesis is that many particulates carry or produce reactive oxygen species (“ROS”). Excesses of ROS can interfere with normal cellular processes and cause damage to tissues and organs. Such oxidative stress is associated with onset of diseases ranging from diabetes to atherosclerosis. PM_{2.5} can even enter the gastrointestinal tract, causing imbalances in the intestinal microecology, which is known to affect a host of other systems, including the central nervous system. Further, fine particulate matter has been shown to weaken the immune system and cause increased susceptibility to infections months after initial exposure (Landguth et al., 2020).

The harm from fine particulate matter may be exacerbated in early life: young children breathe in more air relative to their body weight, they tend to spend more time outdoors, and their bodies are growing rapidly. Holm et al. (2021) summarizes the potential connection between wildfire exposure at young ages and outcomes at older ages, “[E]arly-life exposure may act by ‘programming’ fundamental metabolic, structural, and cell signaling mechanisms that may result in lifelong impacts.” Furthermore, early-life and late-life health statuses are strongly correlated, suggesting that harms persist over the life cycle (Case and Paxson, 2010).

Finally, there is an emerging literature that suggests PM_{2.5} from wildfire smoke may be worse for human health than other sources of ambient pollution.⁷ One reason is that wildfire smoke has a greater concentration of ultra-fine particulates, which are less than 0.1 microns in diameter (Holm et al., 2021). These particles can penetrate deeper into the body than larger particulates, cross the blood-brain barrier, damage neural tissue, and impair brain development (Schraufnagel, 2020). Another reason is that wildfire smoke contains more polycyclic aromatic hydrocarbons, which induce greater oxidative stress. Wildfire smoke also contains more volatile organic compounds, many of which are understood to be carcinogenic (e.g., formaldehyde and benzene). Additionally, wildfire smoke has shown to contain meaningful amounts of heavy metals, such as lead, that have been linked to adverse long-term outcomes (California Air Resources Board, 2021; Grönqvist et al., 2020). The likelihood that wildfire smoke is more harmful than other forms of ambient pollution has been demonstrated by animal studies (Wegesser et al., 2010) as well as observational analyses of child hospitalizations (Aguilera et al., 2021a; Aguilera et al., 2021b). Further, a recent paper that examines the effects of wildfire smoke pollution on the labor market (Borgschulte et al., 2020) finds effects on extensive-margin labor supply that are 2-7 times larger than a comparison group of papers that utilize variation in other pollution sources. In summary, these studies support the idea

⁷Consult Holm et al. (2021) for a detailed overview of the potential health impacts of wildfire smoke on children.

that the composition of wildfire smoke particles may be particularly harmful for human health.

Conceptual Framework

To formalize the hypothesized impact of early-life pollution exposure on long-run outcomes and our reduced-form findings discussed in subsequent sections, we present a simple conceptual model.⁸ This framework, which is visually represented by Appendix Figure A.2, models the life cycle in four distinct periods. In the first period, childhood health (h_1) is determined as a function of genetic endowments (h_0) and inputs during early life (I_1), so that $h_1 = h_1(h_0, I_1)$.⁹ Given its documented adverse health effects—discussed in detail above—air pollution in childhood is modeled as negatively affecting early-life inputs and health, such that $\partial h_1 / \partial I_1 < 0$.

In the second period, individuals choose an amount of education (e), the quality of which is a function of childhood health: $e = e(h_1) = e(h_0, I_1)$. Because health is seen as a positive input into the educational process, childhood smoke exposure is predicted by the model to reduce educational attainment:

$$\frac{\partial e}{\partial I_1} = \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} < 0.$$

After completion of their education careers, individuals progress into the labor market, the third period of the model. Because total earnings depend on both education and innate ability (physical and mental, both represented by health stock in our model), we model earnings (ℓ) as:

$$\begin{aligned} \ell &= \ell(e(h_1), h_1) \\ &= \ell(e(h_0, I_1), h_0, I_1). \end{aligned}$$

With this structure, the impact of early-life pollution exposure is given by:

$$\begin{aligned} \frac{\partial \ell}{\partial I_1} &= \underbrace{\frac{\partial \ell}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1}}_{\text{Lower earnings}} + \underbrace{\frac{\partial \ell}{\partial e} \cdot \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1}}_{\text{Lower earnings due to}} < 0. \\ &\quad \text{due to poorer} \qquad \qquad \qquad \text{impact of reduced} \\ &\quad \text{health} \qquad \qquad \qquad \text{education } (\partial e / \partial I_1) \end{aligned}$$

⁸What follows is similar in spirit to the frameworks presented by Isen et al. (2017b) and Colmer and Voorheis (2020).

⁹For purposes of this model, health is broadly defined to include physical health and measures of mental well-being, including cognitive ability and non-cognitive skills.

As displayed in the above equation, the impacts of childhood smoke exposure affect earnings through two channels. The first term captures the direct channel: lower health stock could cause lower earnings through poor physical health (which likely manifests through reduced extensive-margin labor supply), lower cognitive ability, or reduced non-cognitive abilities (both of which could manifest as either reduced labor supply or lower wages). The second term captures the indirect effect of educational attainment. Insofar as individuals have lower levels of education due to wildfire smoke exposure, this will translate to lower earnings. As part of our analysis in Section 2.6.2, we perform back-of-the-envelope calculations to assess how much of the earnings effect is potentially transmitted through this education channel.

The last stage of the model concerns health in later adulthood and longevity (h_2), which is determined as a function of childhood health (h_1) and later-life health inputs (I_2).¹⁰ Adult health inputs are themselves based on an individual's education and earnings, $I_2 = I_2(e, \ell)$. Taken together, this gives:

$$\begin{aligned} h_2 &= h_2(h_1, I_2(e, \ell)) \\ &= h_2(h_0, I_1, I_2(e(h_0, I_1), \ell(e(h_0, I_1), h_0, I_1))). \end{aligned}$$

Thus, the impact of a childhood wildfire shock on later life health is written as:

$$\frac{\partial h_2}{\partial I_1} = \underbrace{\left[\frac{\partial h_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right]}_{\text{Worse health due to direct/sustained effects}} + \underbrace{\left[\frac{\partial I_2}{\partial e} \cdot \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right]}_{\text{Lower investments due to adverse health impact on education (e)}} + \underbrace{\left[\left(\frac{\partial I_2}{\partial \ell} \cdot \frac{\partial \ell}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right) + \left(\frac{\partial I_2}{\partial \ell} \cdot \frac{\partial \ell}{\partial e} \cdot \frac{\partial e}{\partial h_1} \cdot \frac{\partial h_1}{\partial I_1} \right) \right]}_{\text{Lower investments due to adverse health impact on earnings (\ell) and } (\partial e / \partial I_1) \text{ on earnings (\ell)}}$$

The first term details the direct effect of early-life health on later-life health. Insofar as reductions in early life health capital are permanent, this term will be negative. The second term relates to the indirect effect of lower educational attainment, which could increase adverse health behaviors. While the evidence on education and later-life health is mixed, in Section 2.6.2 we utilize estimates from the literature to assess the role of this channel. Finally, the last two terms capture the effects of wildfire smoke transmitted through the earnings channel. Lower earnings in life could result in

¹⁰It would also be reasonable to utilize a model with many periods, where adult health and earnings are determined contemporaneously within a given period and act as inputs into the next period's outcomes. However, we abstract away from such a scenario in favor of simplicity and expositional clarity.

lower health investments. To the extent possible, we evaluate the degree to which our mortality estimates (from Chapter 1) can be explained by lower earnings in Section 2.6.1.

2.3 Data

To estimate the effects of wildfire smoke on long-run outcomes, we construct a longitudinal data set linking childhood measures of wildfire smoke exposure to administrative and large-scale survey data detailing adult measures of longevity, economic achievement, educational attainment, and disability in adulthood. A simplified version of this data construction process, presented in Appendix Figure A.3, displays data sources and key steps used to clean and assemble our data. The sections that follow discuss these items in greater detail.

2.3.1 Wildfire Data

Our analysis utilizes wildfires from California, the state where wildfires are most severe and have been tracked for over a century. Specifically, we obtain “fire atlases” from the California Department of Forest and Fire Protection’s (“CAL FIRE”) Fire Perimeter Database, which is “the most complete digital record of fire history in California” (California Department of Forestry and Fire Protection, 2018). Two characteristics of this database make it ideally-suited for our analysis of the long-run effects of air pollution from wildfire smoke. First, the database includes comprehensive data on fires for nearly 100 years—and includes some coverage of fires as early as 1878. This consistent and long-running tracking of fires allows to measure differential early-life air pollution exposure for children born as early as 1930. The availability of this historical data is a vital feature for our analysis. Because we can observe children born many years ago, we are able to evaluate longer-term effects—such as longevity and mid-to-late adult outcomes—that were previously unmeasurable with existing data.¹¹

Second, the CAL FIRE data is highly detailed—it includes geospatial data for each fire (i.e., the exact area that was burned), the acreage burned, and the date that each fire was discovered. These characteristics are key inputs into the pollution modeling process discussed later in this section and allow us to generate strong predictions of the cohorts who were exposed to smoke

¹¹Specifically, most studies of the long-run effect of childhood air pollution utilize pollution monitor data (largely available starting in the mid-to-late 1960s) and leverage exogenous variation from the Clean Air Act Amendments of 1970. Because these data sources are relatively recent, papers that use them are limited to observing adult outcomes relatively early in life.

pollution. To the best of our knowledge, there is no other source that combines both the number of years and the quality of data present in the fire atlas data.¹²

Although the Fire Perimeter Database has high-quality information regarding Californian wildfires, 42% of fires during our study period do not have exact dates when the fire occurred—only the year of the fire. To complete this missing information, we hand-gathered dates for each fire from historical newspaper archives.¹³ Overall, 512 fires were missing exact dates and required hand-dating; of those fires, we were able to assign dates to 458 (88%) with high confidence. The remaining fires were treated as missing. Notably, fires that could not be confidently dated tend to be fires that are smaller and/or are in remote areas and are therefore less likely to result in harmful pollution. For purposes of our analysis, we focus on fires that burned 1,000 acres or more, which comprise approximately 89% of the total acres burned during our study period (1,232 fires in total).

Figure 2.1 displays statistics regarding the severity of Californian wildfires—measured by total acres burned—over time. Panel A focuses on the study period of 1930 through 1969 and demonstrates that there is substantial temporal variation in the number of burned acres: some years (1937 and 1963) there are fewer than 20,000 acres burned while others (1950) reach nearly 500,000 acres.¹⁴ To provide further context for our sample period and to motivate the growing problem of wildfires, we also include Panel B, which displays the temporal trends of fires from 1930 through 2021. Within the figure, which has a vertical axis eight times the height of Panel A, we see two trends. First, the variation in acres burned was relatively constant from 1930 through 1990. However, there has been a marked increase in the severity of fire seasons over the last 25-30 years. In fact, the four most extreme seasons in California’s recorded history have all occurred within the last five years, emphasizing the potential for these fires to have a growing impact on health and economic activity. In addition to the temporal variation in fires, Appendix Figure A.4 presents the spatial distribution of fires (in 10-year bins), underscoring the fact that wildfire hazard is dispersed throughout the state.

¹²Notably, the National Interagency Fire Center has detailed wildfire records, but these date back only to 1983. Additionally, the United States Forest Service maintains historical fire perimeter databases for much of the Western United States, but these appear to be incomplete, with many lacking any information on the date of the fire, a key input into the pollution modeling analysis.

¹³Fires were located in newspaper records by searching internet archives—newspapers.com and genealogybank.com—for key terms. For instance, a typical search included typical key words, such as “acres burned”, along with information gathered from the Fire Perimeter Database: the year of the fire, the name of the forest (e.g., “Mendocino National Forest”), and a prominent geological feature near to the fire (e.g., “Clear Lake”).

¹⁴A potential concern with historical data is that recorded fires could increase over time solely due to improvements in tracking them. However, we note that there is no statistically detectable trend ($p = 0.58$) in the number of acres burned over the sample period, alleviating this concern.

2.3.2 Smoke Data

A central challenge of assessing the health impacts of wildfire pollution is understanding which individuals were exposed to smoke. Recent papers studying contemporaneous impacts of recent wildfire pollution—such as Miller et al. (2017) and Borgschulte et al. (2020)—utilize satellite remote sensing data to identify smoke plumes, which is available for the years 2005 through 2021. However, no such data is available during the sample period in which our cohorts were born. When the ability to directly observe pollution is absent, papers have historically assigned treatment using simple proximity to the pollution source (e.g., Clay et al., 2016) or by utilizing a combination of wind direction and proximity to the pollution source (e.g., Schlenker and Walker, 2015; Rangel and Vogl, 2019). We build on these techniques by utilizing sophisticated pollution modeling tools, joining recent environmental economics papers such as Sullivan (2017) and Grainger and Ruangmas (2018) who integrate these models into their empirical approach. Specifically, we utilize the National Oceanic and Atmospheric Administration’s (“NOAA”) HYSPLIT model to predict dispersion of wildfire smoke (Stein et al., 2015). HYSPLIT, which is used by the U.S. Forest Service to predict wildfire pollution, utilizes the location, date, and size of the fire (in acres) as inputs and combines them with gridded four-dimensional meteorological data and fuel composition data obtained from the U.S. Forest Service’s BlueSky modeling framework (Larkin et al., 2009) to predict smoke plumes from each fire that compare favorable to satellite observations. In addition to the satellite comparisons performed by existing research, we also validate the model-generated smoke plumes by assessing their impact on ground-level pollution. This analysis, which is detailed in Section 2.4, finds that modeled smoke exposure is highly predictive of elevated pollution in a modern (2004-2019) context.

Our use of pollution modeling provides several advantages over the proximity or wind direction methods. First, HYSPLIT incorporates the amount of fuel burned, which is a function of both (a) acres burned and (b) the underlying fuel source (i.e., the density of forest or brush in the area of the fire). Given that the sizes of fires in our sample range from 1,000 acres to over 200,000 acres and occur everywhere from brush land to densely wooded forests, this combination is critical in accurately assessing the amount of air pollution from wildfires. Second, pollution modeling utilizes wind dynamics in a more sophisticated manner than is typically observed in existing methods by further incorporating wind speed and wind direction outside of the immediate vicinity of the fire. Wind speed is an important consideration for fires because calm winds can result in smoke pat-

terns that diffuse around the origin point of the fire, while strong winds can send fire pollution for hundreds—if not thousands—of miles in a specific direction. Because wildfire smoke can travel so far from its origin, inclusion of wind dynamics that are not in the immediate vicinity of the fire are also necessary to correctly assign exposure to wildfire smoke. These considerations are illustrated by Appendix Figure A.5, which displays modeled smoke output under three different scenarios with varying fire size and wind intensity. In each scenario, the smoke plumes differ meaningfully, underscoring the importance of incorporating fire size and rich wind dynamics into the treatment definition.

While utilizing HYSPLIT provides several advantages over existing treatment definitions, there are two limitations to the smoke plume output. First, HYSPLIT generates smoke plumes as if the entire acreage is burned within first twenty-four hours from the start of the fire. Because wildfires often last several days, we generate smoke plumes for three separate dates: the date in which the fire began and two subsequent days.¹⁵ Second, the model allows use of meteorological data only back to 1948, which is roughly in the middle of our sample period. To compensate for this, we match each fire date from *before* 1948 with a date *on or after* 1948. This matching process is done by finding the date with the smallest Euclidean distance across several different measures of wind direction and intensity at the ground level as well as multiple levels of the atmosphere. The details of this matching process are covered in more detail in Appendix Section A.1.

The identifying variation for our analysis is displayed in Figure 2.2, which details measures of smoke exposure across birth cohorts and across space. Panel A demonstrates the degree to which birth cohorts are exposed to wildfire smoke during the period from *in utero* through age 4 (“IU-Age 4”).¹⁶ As further discussed in Section 2.5, we focus on this particular age range because that is when children appear to be most sensitive to wildfire shocks. Figure 2.2 Panel A displays the mean, interquartile, and interventile ranges of *in utero* through age 4 smoke by year-of-birth cohort, along with the mean and interquartile range (“IQR”) for the entire sample. The figure illustrates that the amount of exposure varies substantially across California birth cohorts, with peak exposure for cohorts born in the 1950s and lowest exposure for 1930s cohorts.

Even within cohorts, there is substantial range in exposure, as demonstrated by the meaningful differences between the most- and least-intensely treated individuals. This within-cohort

¹⁵Areas are considered to be treated if they spatially intersect any of the plumes generated by this procedure.

¹⁶While we do not observe actual gestational lengths in our data, an individual is assumed to have *in utero* exposure if there is a smoke plume present within the nine months before their birth. Further, exposure “through age 4” indicates smoke exposure at any time prior to a child’s 5th birthday.

variation is more fully illustrated by Panel B, which demonstrates how smoke exposure—defined as the percentage of months in a given period with any smoke coverage—varies over geography in ten-year bins. As displayed in the figure, there are clear geographic differences in exposure to smoke pollution. However, these differences do not come from the same heavily and lightly treated areas for every birth cohort. For instance, in some years, Northern California is more intensely exposed, while in other years Southern California is subjected to higher smoke pollution. As discussed above, this is due in part to fluctuations in fire timing and severity. It is also a function of wind speed and direction, factors that vary meaningfully, as demonstrated by Appendix Figure A.6. As discussed further in the sections to follow, these environmental determinants of smoke exposure provide plausibly exogenous variation in our treatment variable.

2.3.3 Outcome Data and Linkages using Restricted Census Data

To study the long-run effects of childhood exposure to wildfire pollution, we link smoke exposure in early life to three restricted-use datasets from the United States Census Bureau: the Social Security Administration’s (“SSA”) Numident file, the 2000 Decennial Census Long Form (a 1-in-6 sample of households) and the 2005-2019 American Community Surveys (“ACS”).¹⁷ There are two primary features about these data make them ideal for our analysis. First, the Numident file—which is an administrative database of all individuals who have ever been assigned a Social Security number—includes both a person’s city and state of birth, as well as their exact date of birth. These characteristics allow us to assign measures of smoke exposure to individuals in a very precise manner.¹⁸ Second, the restricted datasets can be linked together by using a unique, individual-level identifier (a Protected Identification Key, which is essentially a scrambled Social Security number). This linkage allows us to apply our childhood exposure measures—assigned using an individual’s birth place and birth date information from the Numident data—to adult

¹⁷The SSA Numident file that we obtain through Census Bureau differs from other versions of the Numident database that are available publicly, such as the Berkeley Unified Numident Mortality Database (Goldstein et al., 2021), primarily in the larger number of cohorts covered by the SSA/Census Numident and the ability to link to other data sets (discussed further below). Nonetheless, certain descriptive analyses in this paper will use the Berkeley Unified Numident Mortality Database (henceforth referred to as the “Public-version Numident”) to minimize the amount of restricted Census data that is subject to disclosure.

¹⁸The ability to observe *city* of birth is a substantial improvement over most data sources, which only provide state or country of birth—if they provide any birth place information at all. Having city-level data—rather than only county-level data—is particularly helpful in our setting of California because many counties in California have very large land areas. (The *median* county is nearly the size of Delaware and the largest is almost the size as West Virginia.) Thus, having city of birth detail considerably improves the accuracy of treatment assignment. Additionally, the ability to observe exact date of birth is also a substantial improvement over public data sources, which are typically limited to birth year.

survey outcomes in the 2000 Census, ACS, and other Census products.

In this Chapter, we focus primarily on outcomes taken from the 2000 Census Long Form and American Community Surveys. Because these surveys provide such a rich set of measures—and because childhood circumstances can affect such a wide range of adult outcomes, as illustrated in Section 2.2—we organize survey measures into three distinct groups: (1) economic achievement, (2) educational attainment, and (3) Census-defined disability measures. For each group, we create indices of standardized outcomes, following Kling et al. (2007), to mitigate concerns about multiple hypothesis testing. To construct these indices, individual component variables are converted to z-scores by subtracting the mean value across cohorts and dividing by the standard deviation of the outcome. Certain variables—poverty status, social safety net use, and disability variables—are then reverse coded as necessary, so that higher values represent “good” outcomes. Finally, each variable is combined into an index by taking a simple mean of these standardized values, where each variable is assigned an equal weight in the index. While we focus on these index outcomes as our main results, we will also present effects for individual index components.

We limit all samples to only include individuals who are younger than age 60 at the time of the survey in order to isolate the impact of wildfires on individuals during their peak earning years. Additionally, we only include individuals who do not have any imputed or missing outcomes for all index components to ensure consistency across analyses. Finally, we restrict our sample to individuals who had obtained their Social Security Number (“SSN”) by age 19, which includes 95.6% of individuals in our sampling frame, to ensure that any deaths during adulthood can be captured in the Numident death file.¹⁹ After applying these restrictions, the number of observations in our core sample includes approximately 897,000 person-years underlying our analysis of economic, education, and disability indices (and their sub-components). The summary statistics for these index component outcomes are displayed in Appendix Table A.1.

While the SSA Numident file provides key information necessary for our analysis, there is a key challenge to using the data. Namely, place of birth data is generated by information that individuals enter when completing the application for a Social Security number. These responses are stored as non-standardized string variables, which include spelling errors, transcription errors, use of abbreviations, and/or inclusion of places that do not qualify as cities. For example, an individual born in Los Angeles, California might write their city of birth as “Los Angeles,” “Los

¹⁹We include this sample restriction because individuals who obtain Social Security numbers later in life may be subject to survival bias.

Angels,” “Westwood LA” (a neighborhood in Los Angeles), or simply “LA.” In order to standardize these responses, we limited our sample to all individuals who reported their birth state as California and then hand-matched common responses to populated places detailed in the United States Geological Survey’s Geographic Names Information System (“GNIS”) file. The GNIS file includes the county, longitude, and latitude of each location.²⁰ The remaining observations were matched using probabilistic algorithms, which are further discussed in Appendix Section A.2. Overall, we matched 99.5% of individuals to a GNIS location, where 99.1% of these matches either had a perfect text match or were hand-matched.

2.3.4 Other Data

In addition to our longitudinal data linking early-life wildfire smoke exposure to long-run survey outcomes, we also utilize several data sources for the purpose of supplemental and validation analyses, which we briefly outline here.

Restricted Decennial Census Full Count Data, 2000 and 2010. These data are drawn from the “short form” questionnaire that is received by all households in a given Census year and accordingly cover nearly the entire U.S. population. The Full-Count Census files are then linked to early-life smoke exposure measures in the same manner discussed in the previous subsection. We utilize the linked data to perform supplemental analysis of neighborhood characteristics as well as certain family formation and long-run migration analyses, all of which are discussed in Section 2.8.

Environmental Protection Agency (“EPA”) Monitoring Data, 2004-2019. The EPA collects daily monitor readings for a variety of pollutants. In order to relate this monitoring data to smoke exposure, we generate modeled smoke plumes using 2004-2019 wildfire data from CAL FIRE and the techniques discussed in Section 2.3.2. These data are incorporated into the analysis in Section 2.4 which demonstrates that the modeled smoke measures are highly predictive of ground-level pollution.

Control Variable and Heterogeneity Data. We also utilize a variety of other variables—namely county-level aggregates, historical hospital data, and climate measures—to include in our robustness and in heterogeneity analyses (Section 2.7 and Section 2.8, respectively). The data sources from which these variables were derived, and our method for incorporating them into our

²⁰Longitude and latitude relate to the centroid of each place. A potential concern with utilizing city centroids is that the location of these centroids could drift over time as city boundaries change. However, as demonstrated by Appendix Figure A.7 and discussed in the accompanying notes, this is likely not an issue because boundaries are quite stable over time.

main analysis data set, are discussed further in Appendix Section A.4.

2.4 Validation of Smoke Measures

Because modeled smoke exposure discussed in Section 2.3.2 will form the basis for our treatment variable, we show in this section that it strongly predicts ground-level air pollution in a modern context.²¹ In order to do so, we estimate the following distributed leads-and-lags regression model:

$$Y_{it} = \gamma_{y(t),i} + \eta_{m(t),i} + \sum_{j=-4}^4 \beta_j AnySmoke_{i,t-j} + \varepsilon_{it}, \quad (2.1)$$

where the unit of analysis is a monitor i with associated pollution reading observed in year-month t . To control for time-varying and seasonality-specific trends for each monitor, monitor-by-year fixed effects ($\gamma_{y(t),i}$) and monitor-by-calendar-month fixed effects ($\eta_{m(t),i}$) are included in the regression.

The treatment variable, $AnySmoke_{i,t-j}$, is equal to one if a monitor was exposed to smoke during period $t - j$. Specifically, we consider a monitor to be “exposed” if the modeled smoke plume spatially intersects with the monitor’s geocoordinates. We include leads and lags of the treatment variable to (1) account for serial correlation in smoke exposure and (2) trace out the effect of smoke exposure on pollution over time. The coefficients of interest on these leads and lags, $\beta_j \{j \in -4...4\}$, represent the marginal effect of smoke coverage during relative month j on a monthly pollution measure during month t .

The main results of this analysis are displayed in Figure 2.3, which shows the impact of predicted wildfire smoke plumes on *mean* daily PM_{2.5} and *max* daily PM_{2.5} in Panels A and B, respectively. Within the figure, the estimated impact of smoke coverage during relative month j is presented with 95% confidence intervals. The estimated impact on pollution during the month in which the fire occurs (β_0) is presented in the upper-right-hand corner of each panel, along with the associated standard error and baseline mean.

We find that our modeled smoke exposure predicts meaningful increases in ground-level PM_{2.5} during the month in which the fire occurs: mean daily pollution increases by 1.26 $\mu\text{g}/\text{m}^3$ (13% over the mean) and max daily pollution by 3.58 $\mu\text{g}/\text{m}^3$ (18% over the mean). These effects decay over the two months after the start of the fire, as fine particulates remain in the air for weeks at a time.²² In addition to the impacts on PM_{2.5}, smoke exposure has an impact on a range of

²¹We do not validate our measure in a *historical* context because monitor data does not exist for long-ago periods.

²²The slow decay of the effects is also likely due to the fact that some fires, particularly in recent years, can last for a

other pollutants. As displayed in Appendix Figure A.9, we find increases in mean and maximum for larger particulate matter (PM_{10}) of 9% and 15%, respectively, alongside 2-5% increases over the mean for Carbon Monoxide (CO), Ozone (O_3), and Nitrogen Dioxide (NO_2). Taken together, these results support the assertion that our modeled wildfire smoke measure is an effective predictor of pollution.

2.5 Methodology

To evaluate the long-run effects of early childhood wildfire smoke exposure, we estimate the following cohort distributed lags and leads model:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \sum_{k=-5}^5 \delta_k Smoke_{c,b+k} + \varepsilon_{cba}, \quad (2.2)$$

where the unit of analysis is a birth cohort born in a city c in year-month b and observed at survey year a . The treatment variable, $Smoke_{c,b+k}$, and its associated leads and lags detail the number of months with any smoke exposure at age k , where a month of smoke exposure is defined as any spatial intersection between a smoke plume and a child's city of birth during that month (see Appendix Figure A.10 for an example of how treatment is defined). The coefficients of interest on these treatment variables, $\delta_k \{k \in -5 \dots 5\}$, provide the marginal effect of an additional smoke month during age k . The δ coefficients for the *in utero* period through age 5 ($\delta_{IU} \dots \delta_5$) allow us to trace out the dynamic effects of additional smoke exposure at various ages. Meanwhile, the coefficients for the years leading up to conception ($\delta_{-5} \dots \delta_{-1}$) measure the impact of smoke exposure that occurs before a child could be conceived. Accordingly, these estimates collectively serve as a placebo test, since no effect is expected from air pollution exposure during these periods.

We include city-of-birth fixed effects (μ_c) to control for time-invariant factors related to individuals born in the same city.²³ Additionally, we include year-of-birth \times region-of-birth fixed effects ($\lambda_{y(b),r(c)}$) where region of birth is defined in our baseline specification as Northern or Southern California.²⁴ These fixed effects capture time-varying differences by regions of California that

²³Cities with fewer than 100 births are grouped together with their nearest neighbors until each combined city group has at least 100 associated births during the sample period.

²⁴The counties grouped into either Northern or Southern California are displayed in Appendix Figure A.11, Panel A. In our robustness checks in Section 2.7, we instead utilize State Economic Areas—precursors to modern-day commuting zones—and find similar results. These areas are illustrated in Appendix Figure A.11, Panel B.

may differ in terms of smoke exposure, climate, and economic activity. Further, because both birth outcomes (Buckles and Hungerman, 2013) and wildfires are highly seasonal, we include calendar-month-of-birth fixed effects, $\alpha_{m(b)}$.²⁵ Finally, we include year-of-birth by age-at-survey fixed effects ($\psi_{y(b),a}$) to ensure comparisons are made between individuals who were both born and surveyed in the same pair of years. Standard errors for estimates from this equation are clustered at the county level.²⁶

In addition to our model detailed in Equation 2.2, we also estimate an equation that collapses the effects of childhood wildfire smoke exposure into a single variable:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \delta \cdot \text{Smoke}_c^{\text{IU-Age } 4} + \varepsilon_{cba}. \quad (2.3)$$

The treatment variable, $\text{Smoke}_c^{\text{IU-Age } 4}$ is a single continuous variable equal to the number of months with any smoke exposure from the *in utero* (“IU”) period through age 4.²⁷ For notational convenience, estimates from this equation are referred to as our “summary estimates,” “IU-Age 4 estimates,” or “ $\delta_{\text{IU-Age } 4}$ ” throughout the remainder of this paper.

We focus on the *in utero* through age 4 period primarily for two reasons. First, as discussed extensively in Currie and Almond (2011) and Almond et al. (2018), the years prior to age 5 are a particularly important time of development and shocks during that period have been shown to have particularly meaningful later-life impacts. Second, as will be demonstrated by the results in Section 2.6, the age-of-exposure effects largely dissipate by age 5 (consistent with the existing literature) and thus we focus on the impact of smoke exposure that occurs before that age.

Our research design relies on the assumption that the intensity of wildfire smoke exposure during the first years of life are uncorrelated with other factors, conditional on our fixed effect controls. Because we control for time-invariant factors specific to a given place of birth (via our city-of-birth fixed effects, μ_c) as well as region-specific, time-varying factors (via $\lambda_{y(b),r(c)}$, our vector of year-of-birth \times region-of-birth fixed effects), potentially confounding factors would need to be time-varying at a sub-regional (e.g., city or county) level. The leads in our distributed leads and lags model (δ_k for $k \in \{-5, \dots, -1\}$ in Equation 2.2) provide a useful diagnostic test as to whether or not such confounding issues exist. If such confounders exist, then these coefficients will likely be

²⁵Seasonality in wildfires is demonstrated Appendix Figure A.12, which shows that 90.3% of acres are burned during the months of July through November.

²⁶This is true for all other regression estimates presented in this paper, unless noted otherwise.

²⁷Because we do not observe actual gestation period in our data, the *in utero* period is assumed to begin nine months prior to birth.

economically and statistically different from zero. Additionally, in Section 2.7, we consider different fixed effect and control variable configurations to address the possibility of unobserved confounders that are also conditionally correlated with wildfire smoke exposure and find that results are robust to these alternate specifications.

2.6 Long-Run Effects on Economic, Educational, and Disability Outcomes

In this section, we build on the literature that examines the long-run effects of childhood air pollution by exploring economic, educational, and disability outcomes at older ages (35-59) than have previously been considered. The results of our analyses are displayed in Figure 2.4. Within the figure, each panel presents an event study for the main index (to the left) as well as the summary coefficients (from Equation 2.3) for the index and each of its components (to the right). As discussed in Section 2.3.3, these variables have been standardized (z-scored) so that their mean is 0 and standard deviation is equal to 1. This procedure allows us to combine many results into a single index. Further, we scale these coefficients such that the estimates presented can be interpreted as a percent of a standard deviation (henceforth, “%‐SD”). In the following sub-sections, we discuss each index and provide more detail regarding the index components.²⁸

2.6.1 Economic Achievement

The results relating to economic achievement are presented in Panel A of Figure 2.4. Within the figure, we observe a clear decrease in economic achievement in later life, with effects largely persisting through age 3 and subsiding by age 5, a pattern that is qualitatively similar to the mortality results discussed above. Aggregating across the *in utero* through age 4 period, we find that the average impact of an additional month of smoke exposure is equal to 0.29%‐SD ($p_{25} \rightarrow p_{75} = 1.45\%$). Upon examining the index components, we find that these effects are consistent across outcomes: there are decreases in income, extensive-margin employment, home ownership, and home value. Further, there are *increases* in the likelihood of being in poverty and requiring use of the social safety net.²⁹

²⁸Tabular data on point estimates and associated standard errors for this figure are presented in Appendix Table A.2, Panel B.

²⁹Poverty and social safety net variables have been reverse coded so that lower values indicate worse outcomes.

Our “headline” economic achievement outcome—i.e., the outcome that typifies our results—is the decrease in earned income (defined as wages plus self-employment income). As indicated by Panel A of Appendix Figure A.17, we find that an additional month with smoke exposure reduces age 35-59 earned income by \$178 per year ($p_{25} \rightarrow p_{75} = \890).³⁰ We calculate the cumulative impact of these lost earnings and discount them back to birth using a 3% discount rate. After incorporating the evolution of our treatment effects over the life cycle—which are discussed in detail in Section 2.8.1—we find the $p_{25} \rightarrow p_{75}$ effect of childhood wildfire pollution exposure reduces discounted earnings over the age ranges in our sample by \$4,710. This suggests that wildfire pollution can have substantial long-run effects on earnings and economic productivity.³¹ The earnings effects appear to be partially driven by decreases in labor force participation ($p_{25} \rightarrow p_{75} = -0.45$ percentage points). To evaluate the degree to which reductions in labor force participation affect our earnings estimates, we perform a back-of-the-envelope calculation which suggests that the effects on extensive-margin labor supply accounts for roughly 27% of the effect on earned income.³²

We also consider the degree to which the decrease in economic activity, and specifically earnings, may contribute to the reduction in longevity (discussed in Chapter 1) by performing an additional back-of-the-envelope calculation. We utilize data from Chetty et al. (2016) to calculate the association between income and life expectancy, and we find that a \$1,000 increase in income is associated with an increase in life expectancy of 0.030 years.³³ Accordingly, our \$890 decrease in earnings may be associated with a reduction of 0.027 life years per person, which suggests that the reduction in earnings could explain as much as 59% of the reduction in life expectancy that we find.³⁴

In addition to highlighting our earnings outcome, we also briefly discuss the degree to which

³⁰These dollar values, along with all dollar values for the remainder of the paper, are in 2010 dollars.

³¹To provide further context, we contrast these magnitudes with earlier-life estimates from the existing literature in Appendix Section B.2.

³²Specifically, within our sample, a 1 percentage point increase in labor force participation is associated with a \$530 increase in earned income. Using this relationship, we can roughly calculate the fraction of the earned income effect “explained” by labor force participation effects as follows:

$$\frac{0.45\text{pp} \times \$530}{\$890} \approx 26.8\%$$

³³We exclude the top income percentile when calculating this slope, as that is consistent with our winsorized income measures.

³⁴We view this calculation as placing an *upper bound* on the degree to which early life shocks affect mortality through the income channel. This is because the relationship from Chetty et al. (2016) is not causal, and so the degree to which our economic results “explain” life expectancy reductions may be overstated. This overstatement will depend on the degree to which reduced health is driving lower income (rather than vice versa). Nonetheless, in the absence of causal estimates of income on life expectancy, we view bounding the income-mortality mechanism in this way to be a useful exercise.

early childhood wildfire smoke exposure increases use of the social safety net, which is defined as any use of the Supplemental Nutrition Assistance Program, Medicaid, Supplemental Security Income, or public assistance income. We find that an additional month with any exposure increases social safety net use by 0.09 percentage points (0.66% over the mean; $p_{25} \rightarrow p_{75} = 0.45$).³⁵ To quantify the effect of this increased use, we utilize the Current Population Survey provided by IPUMS (Flood et al., 2021), which quantifies the household- or person-level cost of each of these programs. Using these data, we calculate that a person using at least one of these programs increases government expenditures by \$7,958 annually, on average. Multiplying this cost by our estimates and discounting back to birth, we find that moving from the 25th to the 75th percentile of childhood smoke exposure results in an additional \$229 of government expenditures per person.³⁶

2.6.2 Educational Attainment

We next consider the effect of smoke exposure on educational attainment, which is displayed in Figure 2.4, Panel B. As with previous outcomes, there is a pronounced effect on educational attainment: an additional month with smoke exposure decreases the educational attainment index by 0.27%-SD ($p_{25} \rightarrow p_{75} = 1.35\%$). Moreover, there is a similar pattern in the outcome dynamics, with effects peaking at ages 1-2 and effectively trending back to zero by age 5.

We begin our discussion of index components with our “headline” outcome: high school completion. As displayed in Appendix Figure A.17, Panel B, we find that a marginal smoke month results in a 0.08 percentage point decrease in high school completion ($p_{25} \rightarrow p_{75} = 0.40\text{pp}$). These effects generally persist across levels of higher education attainment, with estimated reductions in 4-year college ($p_{25} \rightarrow p_{75} = 0.50\text{pp}$) and graduate degree completion ($p_{25} \rightarrow p_{75} = 0.30\text{pp}$).

When estimating the effects on years of schooling, rather than degree completion, we estimate a $p_{25} \rightarrow p_{75}$ reduction of roughly 0.05 years of education. Using this estimate, we can conduct a back-of-the-envelope exercise to assess how much of the earnings effect (from Section 2.6.1) could potentially be explained by this impact on education. Specifically, we utilize a 10% of return per year of schooling (Card, 1999), which translates into a back-of-the-envelope decrease in earnings of 0.5%, or roughly 25% of the earnings effect.³⁷ This suggests that early-life pollution shocks

³⁵See Appendix Figure A.18, Panel A, for event-study graph.

³⁶Specifically, we calculate the average annual cost as $\$7,958 \times 0.45\text{pp} \approx 36$. We then discount this value back to birth using a 3% discount rate: $\sum_{a=35}^{59} [1.03^{-a} \times 36] = \229

³⁷The wage effects found in the previous subsection translate to 2% of wages. This approximation of the education-earnings pathway is consistent with Isen et al. (2017b) who calculate that educational attainment explains approximately 20% of the earnings increases from the Clean Air Act Amendments of 1970.

may be transmitted to earnings primarily via mechanisms other than education, such as shocks to underlying ability, non-cognitive skills, later-life cognitive decline, and/or physical health.

It is also useful to consider how much of the increases in mortality, discussed in Chapter 1 of this dissertation, could be explained by decreases in education. To do so, we utilize estimates from Clark and Royer (2013) and Meghir et al. (2018)—two papers that leverage substantial, sharp variation in years of schooling and high-quality administrative data. While these studies do not find statistically significant decreases in mortality due to higher education, they able to rule out meaningful effects. Applying their estimated bounds on the education-mortality relationship, we calculate that the estimated decreases in education could explain no more than 30% to 51% of the mortality effects that we find.³⁸

2.6.3 Disability

The last set of “index” outcomes we evaluate includes all disability measures available in the American Community Surveys and 2000 Decennial Census. The leads-and-lags coefficients for our disability index—displayed in Figure 2.4, Panel C—show that early-life wildfire smoke exposure causes individuals to have worse later-life disability outcomes, but that these effects are small and statistically imprecise. (Recall that all disability outcomes are reverse coded so that lower values indicate worse outcomes.) However, as displayed in the summary components on the right-hand side of Panel C, the combined effect for the *in utero* through age 4 period achieves statistical significance ($t = -2.13$) with a 0.16%-SD reduction ($p_{25} \rightarrow p_{75} = 0.80\%$). In addition to our disability index, another potentially useful measure of adult health is whether an individual reports *any* disability. We find that an additional smoke month results in a 0.06 percentage point reduction ($t = 1.66$) in individuals reporting at least one disability, or a 0.4% increase over the mean. However, because

³⁸Clark and Royer (2013) rule out reductions in mortality larger than 5% from a 0.315 to 0.464 increase in years of education. We apply this bound to our estimates to calculate the upper bound of mortality than can be explained by education as:

$$\frac{\frac{0.05}{0.315} \times 5\%}{1.8\%} = 30\% \quad \frac{\frac{0.05}{0.464} \times 5\%}{1.8\%} = 44\%.$$

Where the numerator in each equation scales the years-of-education effect that we find (0.05) by the 5% bound in the Clark and Royer (2013) estimates. This effect is then divided by the percentage reduction in age 55 mortality that we discuss in Chapter 1, Section 1.6. Alternatively, we can use estimates from Meghir et al. (2018), who find than a 0.25 increase in years of education results in at most 1.0-1.4 *month* increase in life expectancy (0.083-0.117 years):

$$\frac{\frac{0.05}{0.25} \times 0.083}{0.046} = 51\% \quad \frac{\frac{0.05}{0.25} \times 0.117}{0.046} = 36\%.$$

Where the numerator in each equation scales the years-of-education effect that we find (0.05) by the life expectancy bound in the Meghir et al. (2018) estimates. This effect is then divided by the decrease in our per-person life expectancy estimate from Chapter 1, Section 1.6.

of the noisiness in the event-study estimates, these results are merely suggestive.

When evaluating individual index components, we find that only one outcome—Difficulty Remembering, Concentrating, or Making Decisions (henceforth, “cognitive difficulty”)—achieves statistical significance at the 95% level. The leads-and-lags graph for cognitive difficulty, displayed in Panel C of Appendix Figure A.17, includes a clearer and more convincing pattern than the overall index, with effects following a similar trajectory as other outcomes. The summary coefficient, $\delta_{IU\text{-Age}4}$, implies a marginal effect of 0.06 percentage points ($p_{25} \rightarrow p_{75} = 0.30\text{pp}$). The finding that childhood air pollution exposure causes adult cognitive difficulty is consistent with the medical literature that finds associations between child neuropsychological issues, such as attention deficit hyperactivity disorder and autism spectrum disorder, and air pollution in pre- and post-natal periods (e.g., Suades-González et al., 2015) as well as recent economic studies linking sustained *later* life pollution to dementia (Bishop et al., 2018). However, to the best of our knowledge, this is the first paper to causally link early-life air pollution exposure to such difficulties in later life.

2.7 Robustness

We conduct a variety of robustness tests to ensure that our estimates do not qualitatively change when we (1) evaluate the direct effects of the fires themselves; (2) vary our specification or choose alternative treatment definitions; (3) consider the effects of mortality selection; (4) adjust for migration-related measurement error; or (5) employ alternate estimation methods. We discuss each of these exercises below.

2.7.1 Controlling for Proximity to Fires

This paper has thus far attributed the later-life harms of wildfire exposure to the substantial increases in air pollution that stem from these fires. However, wildfires sometimes cause physical damage to populated lands and may also necessitate evacuation or other responses that could result in childhood trauma. Accordingly, it is worth evaluating the degree to which our treatment may capture effects of wildfire-induced trauma, rather than harms purely from early-life air pollution exposure. To do so, we re-estimate our leads and lags specifications and include controls for fires that are within 10 miles of an individual’s city of birth—specifically, individuals are assigned direct fire exposure if *any* part of the fire is within 10 miles of the city centroid. As shown by Figure 2.5, controlling for proximity to fires does not materially affect our estimates, suggesting that

wildfire smoke—and not trauma or economic damage from the direct effects of the fire—are driving results.³⁹

2.7.2 Changes to Specification

In this subsection, we include a variety of controls and vary the configuration of fixed effects in our baseline specification. Before detailing the results of these modifications, it is worth briefly discussing primary threats to internal validity and how these threats are either addressed by our main specification or can be diagnosed via our leads-and-lags coefficients. The two primary threats to internal validity would be if: (1) wildfire smoke was spuriously correlated with other characteristics and trends in a given area; or (2) wildfire smoke was co-determined with other weather-related factors—such as temperature, precipitation, or wind—that affect long-run outcomes. Our main specification addresses the first potential threat to identification by including detailed city of birth fixed effects (that finely control for time-invariant factors), along with year-of-birth \times region-of-birth effects (which control for cohort-varying factors that are specific to Northern and Southern California, respectively). However, in the event that these controls do not adequately adjust for spuriously correlated time-invariant or time-varying factors, our lead coefficients ($\delta_{-5}, \delta_{-4}, \dots, \delta_{-1}$)—which capture the marginal effect of an additional month of smoke exposure *prior* to conception—provide useful tests. If these placebo coefficients display pre-trends or are consistently and significantly different than zero, then that would indicate potential issues with our analysis.⁴⁰ However, we neither observe level shifts or trends in the placebo coefficients displayed in Section 2.6 (or their

³⁹Given that loss and displacement from fires are serious events that have been shown to have lasting consequences (Schwank, 2021), it is worth considering why we *don't* see effects when controlling for these estimates. One simple explanation is that, since fires tend to occur away from population centers, very few fires are sufficiently near to people to cause lasting harm. Anecdotal evidence from historical newspaper records, which were reviewed to supplement missing information on wildfire dates, is consistent with this assertion. We noted only sparing reports of large evacuations or towns that were lost to wildfires, and even these locations tended to be sparsely populated resort or logging towns. However, it is worth noting that in *recent* years there have been several notable incidents where fires directly affected population centers. This is attributable, in part, to the rapid growth of the number of individuals living in the wildland urban interface (Radeloff et al., 2018), which both increases the exposure of populated places to naturally occurring wildfires—such as those caused by lightning strikes—and increases the risk that population-adjacent wildfires will be caused by human activity.

⁴⁰To provide specific examples, if there are time-varying trends (i.e., if areas with high propensity for smoke exposure are trending in a different manner than low-propensity areas), we would expect to see trends in our placebo coefficients prior to the conception date. Additionally, the lead coefficients will also detect if there is a compositional response in *anticipation* of a wildfire smoke shock. Consider if, for instance, wealthier or more educated parents—who are expected to have wealthier and more educated children—move away from an area just before their child is born. In this case, we would expect to see a preconception trend since the composition of the children born in a given area is changing over time. However, we see no trends that would suggest this is occurring. (Regarding the specific threat of anticipatory migration, we also do not find any evidence that wildfire smoke exposure induces migration, as indicated in the last column of Appendix Table A.4. This is also consistent with findings from Borgschulte et al. (2020) who do not detect smoke-induced migration behavior in a modern context.)

related appendix figures), supporting the idea that time-invariant and time-varying factors are not confounding our analysis.

In addition to the credibility added by our placebo tests, we also further address potential threats to internal validity by adding various controls and modifying our fixed effect specifications. The results of these tests are shown in Figure 2.6, which displays summary estimates for our baseline specification (on the top row) along with summary estimates of alternative specifications (on other rows and described by the in-figure text). We briefly discuss each of these robustness specifications below.

Additional Controls

Our first specification includes controls for climate variables, which have been shown to be both a strong determinant of wildfires and also human health in the short- and long-run (e.g., Barreca et al., 2016; Isen et al., 2017a). Specifically, we include controls for the number of days above 32 degrees Celsius during the *in utero* period through the first six months of life—i.e., the critical period in which temperature strongly affects long-run earnings, as shown by Isen et al. (2017a). We also control for (a) the mean of monthly precipitation and (b) the fraction of days in cardinal wind direction bins interacted with an individual’s county of birth. Both sets of these controls are calculated over the *in utero* through age 4 period.⁴¹ However, as displayed in the Figure 2.6, controlling for these outcomes does not qualitatively affect our estimates.

Next, in order to better control for granular time trends, we follow a standard practice in the early-life determinants literature by interacting baseline county-level characteristics with time trends (recent examples include Bailey et al., 2020, 2021; Hoynes et al., 2016). Specifically, we interact 1930 county-level data with a decade-varying-linear spline in time.⁴² We find that inclusion of these interacted splines does not meaningfully change our estimates.

Last, we control for the possibility that more-and-less smoke prone areas are trending differently by dividing overall smoke propensity (defined over the entire period) into quartiles and

⁴¹Precipitation could affect children primarily via agricultural income shocks. Additionally, we control for county-by-wind direction bins because it is possible that wildfire smoke could blow in exclusively from an otherwise “good” or “bad” direction, i.e., a direction that has little or lots of pollution from non-wildfire sources (Deryugina et al., 2019). If this is the case, then we would systematically under (over) state the effects of wildfire smoke on long-run outcomes.

⁴²These characteristics include: (1) the fraction of children in school; (2) the percent immigrant population; (3) the relative share of manufacturing and agricultural employment; (4) the average age at first marriage; and (5) the average family income. All characteristics were obtained from the 1930 Census 100% Sample (Ruggles et al., 2021), except for family income, which was obtained from the 1934 IRS Statements of Income. (No income-related questions were asked during the 1930 Census.)

interacting these quartiles with birth year fixed effects. Inclusion of these additional fixed effect controls do not significantly affect our estimates.

Different Fixed Effect Specifications

We next consider different fixed effect specifications that increase the granularity of controls along time and spatial dimensions. First, we replace our year-of-birth by region-of-birth fixed effects—where region is defined as Southern California or Northern California—with more detailed year-*month*-of-birth by region-of-birth fixed effects. If there are short-lived shocks that are contemporaneous with wildfires—such as regional temperature spikes or other extreme weather—they will be captured by these controls. Next, we instead increased the granularity of our region-of-birth fixed effects to be based on State Economic Areas (precursors to modern-day commuting zones), rather than a South-North distinction.⁴³ These controls will capture relatively local trends that may not be common to the larger region. As displayed in Figure 2.6, modifications to these fixed effects yield highly similar estimates to our baseline specification.

Different Treatment Definition

Our treatment definition implicitly assumes that each additional month of exposure during a given age range has the same marginal effect—i.e., that the dose-response function is linear in the number of months with *any* smoke exposure from the *in utero* period through age 4. To relax this assumption, we redefine our treatment variable to be an indicator variable equal to one if an individual is exposed to an above-median amount of smoke and zero otherwise. Specifically, for an in-sample median smoke measure \bar{m} , we estimate:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \tau \cdot \mathbf{I}(Smoke_c^{\text{IU-Age } 4} > \bar{m}) + \varepsilon_{cba}, \quad (2.4)$$

which is the same specification as Equation 2.3 with a transformed treatment variable. To facilitate comparison with our estimated treatment, we also estimate a separate equation with $Smoke_c^{\text{IU-Age } 4}$ as the *dependent* variable to obtain the average number of additional smoke months for those with above-average exposure:

$$Smoke_c^{\text{IU-Age } 4} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \kappa \cdot \mathbf{I}(Smoke_c^{\text{IU-Age } 4} > \bar{m}) + \varepsilon_{cba}. \quad (2.5)$$

⁴³See Appendix Figure A.11 for illustration of State Economic Areas.

We then scale the effects from Equation 2.4 ($\hat{\tau}$) by those from Equation 2.5 ($\hat{\kappa}$) to generate a comparable estimate to our main results. The scaled estimates from this exercise, presented in the final row of Figure 2.6, are qualitatively very similar to the effects estimated from our baseline, although they are less precise, consistent with the decreased variation in the treatment variable.

In addition to this exercise, we also provide evidence supporting the linearity of the dose-response function using residualized scatter plots. These plots, which are displayed in Appendix Figure A.19, generally support the choice of a linear functional form for our analysis. See accompanying notes to these figures for further information on how plots were constructed and additional discussion of interpretation.

2.7.3 Mortality Selection

Because Chapter 1 of this dissertation demonstrates mortality effects from early childhood wildfire smoke, a potential concern, particularly for the outcomes discussed in Section 2.6, is that mortality selection could lead to bias in observed long-run outcomes. Put differently, it is possible that more highly exposed children will not be present in later-life surveys because they died before the survey date—either as infants or in adulthood—creating potential compositional deviations from the implied counterfactual at later ages.

To evaluate the degree to which this selection may affect estimates, we perform a bounding exercise in the same spirit as Lee (2009). This exercise is explained in detail in Appendix Section B.3, but we discuss the general method here. First, we estimate the death rate at a given age a using the same methods employed in Chapter 1 of this dissertation. Then, using these estimates, we predict the number of deaths for each birth cohort (defined by city and year-month of birth) by age a .⁴⁴ We then assign extreme values—equal to either the within-sample 1st percentile of the outcome distribution (for a lower bound) or the 99th percentile of the outcome distribution (for an upper bound)—to impute the outcome of the deceased individuals as of the survey date. We then create a new outcome variable by averaging these values into the values of other individuals observed in that birth cohort.⁴⁵ The results of this exercise are displayed in Appendix Figure A.21.

⁴⁴These estimates also include infant mortality, discussed in Appendix Section B.1.

⁴⁵To give a simplified example, suppose we observe 1,000 individuals of a given cohort at age 55 with an earnings average of \$44,000. Additionally, assume that, based on their childhood wildfire exposure, an additional two individuals died before age 55 who would have survived had they not been exposed to wildfire smoke. We then bound the effects of these two missing individuals by assuming that they were at the 1st percentile of the earnings distribution (\$0) or at the 99th percentile of the earnings distribution (\$286,000). Then the new cohort averages are as follows:

$$\text{Lower : } \frac{(1,000 \times \$44,000) + (2 * \$0)}{(1000 + 2)} = \$43,920 \quad \text{Upper : } \frac{(1,000 \times \$44,000) + (2 * \$286,000)}{(1000 + 2)} = \$44,483$$

As demonstrated by the figure, our estimates are not sensitive to mortality selection.

2.7.4 Measurement Error

Because our estimates use place of birth to assign treatment, our treatment will be mis-measured to some degree for any individuals who migrate away from where they were born within the first four to five years of life. This is of particular concern when interpreting our age-specific estimates, as estimates at older ages will mechanically have higher rates of migration away from their birthplace and thus have higher degrees of measurement error. To assess the potential impact of measurement error on our main result, we utilize the 1940 Census 100% Sample to determine where children live at age five and how it compares to their place of birth, which is inferred using their 5-year migration status.⁴⁶ With these data, we restrict our sample to children age five who were born in California and utilize information on their counties of birth and residence to calculate measures of “true” exposure and exposure based solely on place of birth (i.e., assigned in the same way as our main analyses). Then, using these assigned treatments, we calculate how much our estimates have potentially been attenuated due to measurement error, under the assumption that error in assignment is classical. A more detailed description of this approach, as well as support for the assumption of classical measurement error, are discussed in detail in Appendix Section B.4.

The results of this exercise, where we adjust our estimates for the calculated attenuation factors, are displayed in Appendix Figure A.22. Within the figure, we present the attenuation-adjusted leads-and-lags coefficients for our main outcomes from Section 2.6 alongside the original estimates. As demonstrated by the figure, the measurement-error-adjusted estimates are nearly indistinguishable from our baseline estimates, likely due in part to the relatively low rates of migration among small children (only 14% of children had migrated by age 5 within the sample).

2.7.5 Issues with Two-Way Fixed Effect Estimates and Alternative Estimation

This paper utilizes two-way fixed effects (“TWFE”) controls in our estimation of long-term outcomes. However, as discussed in a fast-growing literature (e.g., Callaway and Sant’Anna, 2021; de Chaisemartin and D’Haultfœuille, 2020; Goodman-Bacon, 2021a; Sun and Abraham, 2021),

These upper and lower bound averages are then used as outcome variables in our regressions.

⁴⁶The 1940 Census is best-suited for this analysis because it asked 5-year migration questions of all individuals in the 100% sample. This contrasts with later Decennial Censuses where migration was asked only of a smaller sub-sample and/or was asked only about 1-year migration. In this context, the larger sample size is particularly useful as we are looking at relatively small sub-sample of individuals, i.e., those who were five years old in 1940 and were also born in California.

TWFE models may be difficult to interpret due to the unintuitive way in which they weight different observations across time. This can be particularly problematic if there is staggered timing, time-varying treatment effects, and/or dynamic treatment effects. Recently, several papers—such as Cengiz et al. (2019) and the previously referenced studies—have proposed methods for handling these issues in commonly used difference-in-difference and event-study frameworks where a policy is enacted and remains in effect.

However, the existing literature has not settled on the best way to deal with these weighting issues for continuous treatments that turn on and off, which is the case for our research design.⁴⁷ To address potential weighting issues, we estimate an alternate regression specification that is separately estimated *for each year-of-birth cohort* and we then combine these estimates by simple average to ensure intuitive weighting. Specifically, we estimate:

$$Y_{cba} = \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \delta \cdot Smoke_{cb}^{\text{IU-Age4}} + \bar{Y}_c + f(X_{cb}) + \varepsilon_{cba},$$

where the above equation is very similar to our summary specification (Equation 2.3), but differs in key ways to allow separate estimation for each year-of-birth cohort. Because we are estimating each regression using only within year-of-birth cohort (cross-sectional) variation, we are unable to utilize our city-of-birth fixed effects (μ_c) that are present in Equation 2.3. To increase precision and control for potential differences across places that may be correlated with both wildfire smoke and long-run outcomes, we instead include long-run city-of-birth averages of the outcome variable (\bar{Y}_c) as well as indicators for quartiles of smoke propensity, county-level demographic controls, and climate controls discussed earlier in this section. The results of this exercise are presented in Appendix Figure A.24. Within the figure, the estimates from our alternative estimation are qualitatively similar, though less precise—although our results pertaining to longevity, economic achievement, and educational attainment all exceed traditional levels of statistical significance.⁴⁸

2.8 Supplemental Results

We have established that early-childhood wildfire exposure results in adverse long-run impacts to economic achievement, educational attainment, and disability in later life. In this section,

⁴⁷Very recent work by Callaway et al. (2021) provide alternative estimation strategies for difference-in-differences designs with continuous treatment. Future versions of this paper will consider the applicability of these methods to our setting.

⁴⁸We obtain standard errors through 100 cluster bootstrap iterations.

we provide further evidence on (1) how effects evolve as individuals age, (2) what extent smoke exposure reshapes the distribution of outcomes, and (3) other heterogeneity and outcomes that may be affected by childhood air pollution exposure.

2.8.1 Effects on Earnings and Disability Grow More Severe with Age

As discussed earlier in this paper, a key advantage of our setting is that we are able to observe the effects of childhood air pollution shocks for adults in later life, which enables us to better understand how early-life harms evolve throughout the life cycle. It is not *ex-ante* clear how treatment effects evolve as individuals age: the impact of childhood shocks may “fade” when transitioning from early adulthood to mid-life if, for instance, they simply accelerate the manifestation of an underlying condition that would have otherwise impacted an individual a few years later. On the other hand, it is possible that the impacts of early-life harms, mitigated by the higher “health stock” of younger individuals, *increase* proportionally as individuals age.

To assess which of these scenarios holds empirically, we evaluate the two “headline” outcomes with the potential to vary as individuals age: earnings and cognitive difficulty. Specifically, we estimate the following equation:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \sum_{j \in \mathcal{J}} \delta_j \cdot (\text{AgeBin}_j \times \text{Smoke}_c^{\text{IU-Age } 4}) + \varepsilon_{cba}, \quad (2.6)$$

where \mathcal{J} is the set of age bins in 5-year increments. The results of this equation are displayed in Figure 2.7. Within the figure, each point represents a separate estimate of δ_j with associated 95% confidence intervals. For further context, the effect as the percent of the age-specific mean is also presented, along with a line of best fit for the coefficients.

While the individual coefficients from this exercise are somewhat imprecise, taken together they tell a clear story: the adverse impact of childhood air pollution exposure on earnings and cognitive difficulty is growing with age, both in levels and proportionally. The earnings effect (as a percent of the age-specific mean) for the youngest age bin is less than one third the effect for the oldest age bin; the impact of cognitive difficulty is less than one half.⁴⁹ As discussed in Section

⁴⁹ Appendix Figure A.25 presents qualitatively similar results for labor force participation, economic index, and disability index outcomes. Notably, our impact on labor force participation is almost completely driven by the effects in the oldest age bin: a $p_{25} \rightarrow p_{75} = 1.07$ percentage points. When we assess how much of the later-life earnings could be attributable to labor force participation (performing a back-of-the-envelope in the same manner as was discussed in Section 2.6.1), we find that reduced LFP “explains” 43.7% of the late-life earnings effect.

2.6.1, we utilize these life-cycle earnings dynamics to arrive at a discounted earnings loss of \$4,710 in response to a $p_{25} \rightarrow p_{75}$ increase in wildfire smoke exposure. To underscore the importance of incorporating these age-specific effects into our earnings loss computation, we calculate the discounted loss of earnings using an alternative method. Due to data constraints, the existing literature has assumed that the reductions in early-adulthood earnings persist proportionally throughout the life cycle (i.e., that a 1% reduction in age 30 earnings will translate to a 1% reduction in earnings at ages 35, 45, 55, etc.). We follow this method by using the proportional reduction in earnings that we observe at our earliest age bin (aged 40 and under; $p_{25} \rightarrow p_{75} = 1.0\%$) and assume that this effect is constant as people age. Such an exercise would result in a discounted cost of air pollution that is 44% lower than our estimate—a substantial understatement. Indeed, these findings also suggest that there could also be considerable value in investigating longer-run impacts of early-life shocks from sources aside from air pollution.

2.8.2 Childhood Smoke Exposure Affects the Entire Distribution of Economic Achievement and Earnings

While we establish that childhood exposure to smoke reduces economic achievement and earnings *on average* (Section 2.6.1), it is not clear how this exposure is reshaping the earnings distribution. To assess this, we calculate the in-sample quintiles of the economic index and earnings distribution. We then estimate separate regressions where the outcome variable is equal to one if an individual is in a given quintile bin and zero otherwise.

The results of this exercise are displayed in Figure 2.8 where each bar represents the coefficient from a separate regression and associated 95% confidence intervals. Both Panels A and B tell a similar story: early-life exposure to wildfire pollution meaningfully increases the probability that individuals are in the worst quintile of the economic achievement and earnings distribution ($p_{25} \rightarrow p_{75} = 0.6$ to 0.7 percentage points) and that impact is almost completely offset by reductions in the best quintile ($p_{25} \rightarrow p_{75} = 0.5$ to 0.6 pp). These results are consistent with a left shift of the distribution of outcomes. Our findings also draw an interesting contrast with those of Isen et al. (2017b) who perform a very similar exercise to ascertain the distributional impact of the Clean Air Act Amendments of 1970 (“CAAA”). While they find substantial improvements to the bottom of the income distribution (which is analogous to our findings), their effect is largely displaced by increases in mass among the middle quantiles.

2.8.3 Wildfire Shocks Are Not Transmitted to Children Via Local Economic Channels and Long-Run Effects Do Not Meaningfully Vary Across Person and Place Characteristics

We perform two other supplemental analysis, both of which are discussed further in Appendix Section B.5 but are briefly summarized here. First, we perform an analysis of contemporaneous economic outcomes (e.g., employment and family income) and net migration to ascertain if wildfire smoke shocks could be transmitted to children via local economic channels (see Appendix Table A.4). We do not find any statistically significant evidence of reductions in economic activity. Further, the reductions that we do find are also economically small and cannot explain a substantial portion of the results that we find.

Second, we consider heterogeneity across individual-level and place-level characteristics, including sex, race/ethnicity, urban status, county-level income, hospital access, and county-level economic dependence on lumber (i.e., places that could be more sensitive to loss of timber due to wildfire). The results of this analysis, shown in Appendix Figure A.26, display little-to-no heterogeneity. Out of our twenty-four regressions (six groups \times four outcomes), there are only three that have statistically different interaction effects.

2.8.4 The Effect of Childhood Wildfire Exposure on Family Formation, Institutionalization, and Other Outcomes

In addition to the wide set of outcomes explored in previous sections, we briefly discuss certain outcomes here that do not fall into one of our three indices but are of interest. These outcomes, along with their scaled coefficients (as a percent of the mean), are presented in Appendix Figure A.27. As demonstrated by the figure, the effects that we estimate for family formation outcomes (whether a person has ever been married and the number of children they have living with them at the time of survey) and migration are either statistically insignificant, economically insignificant, or both. Likewise, we estimate the effect of early-childhood wildfire smoke on two placebo outcomes—sex and race/ethnicity—which should not be affected by air pollution. As expected, we do not detect any impact on these outcomes.

We do, however, see increases in the number of individuals who are held in institutional group quarters—i.e., correctional or nursing facilities—as of the 2010 Decennial Census.⁵⁰ Specifically, we

⁵⁰This variable is not available in the Restricted 2000 Decennial Census. Note that correctional and nursing facilities are grouped together because the Census Bureau did not permit RDC researchers to disclose results using

find that moving from the 25th to 75th percentile of exposure increases institutionalization by 0.025 percentage points, a 2.1% increase over the mean. This finding is consistent with the literature finding early-life-exposure impacts of lead (e.g., Reyes, 2007; Aizer and Currie, 2019; Grönqvist et al., 2020) and other particulate matter (Voorheis, 2017) on incarceration and behavioral issues. Further, it underscores the ability of wildfire pollution exposure to cause the sort of degradation in health and non-cognitive skills that ultimately leads to institutionalization.

2.9 Discussion and Conclusion

2.9.1 The Cost and Fiscal Burden of Childhood Wildfire Smoke Exposure

We end our analysis by calculating the cost of wildfires to children in our sample by multiplying our estimates by the average exposure of each child, which gives the per-person benefit of “shutting off” wildfire pollution. Additionally, we also quantify the fiscal burden by estimating the amount of lost tax revenue as well as increases in social safety net costs and incarceration costs as a result of wildfires.

The results of this exercise are presented in Table 2.1. The left half of the table (Panel A) details our estimates of the individual cost of wildfire pollution in two areas: earnings and life years lost in adulthood (from Chapter 1). Using the same discounting method discussed in Sections 2.6.1 and 2.8.1, we calculate the discounted losses for the average amount of *in utero* through age 4 wildfire smoke exposure within our sample (10.5 months) and find a reduction of approximately \$9,900 per person. Added to this amount is the estimated cost of lost life years in adulthood, calculated in Chapter 1. Specifically, we estimate the number of lost life years ($\frac{9.2 \text{ life years lost}}{1,000 \text{ persons}} \times 10.5 \text{ smoke months}$) by the estimated value of a statistical life year of \$130,000 to obtain a per-person mortality cost of roughly \$12,600.⁵¹ Taken together, these estimates imply per-person losses of \$22,450 due to wildfire smoke exposure. To provide further perspective on the magnitude of these losses, we multiply them by 500,000 persons, which is roughly the size of a single California birth cohort. The exercise yields losses of \$11.22 billion dollars for every single “cohort”

sub-categories of the institutional group quarters variable during the time period in which these results were disclosed. However, as approximately 85.2% of institutionalized persons in our sample age range were in correctional facilities, it is reasonable to interpret these effects as primarily pertaining to incarceration.

⁵¹The value of a statistical life year (“VLSY”) was obtained from Cutler and Richardson (1999) and inflated to 2010 dollars. Alternatively, we could utilize the methods employed by Carleton et al. (2020) and Bailey et al. (2020). To determine the VLSY, these studies use the *age-invariant* value of statistical life estimates from the Environmental Protection Agency and divide them by median life expectancy in the United States to arrive an approximate value of \$200,000 (in 2010\$). If we instead utilized these estimates, the life years lost would be valued at \$19,320.

subjected to this average level of exposure. These estimates, while large, conservatively omit several components. First, because of this paper’s focus on *later* life outcomes, we omit any earnings losses before age 35 and any losses due to infant mortality. While we do not find statistically significant early-adult earnings effects, incorporation of infant mortality costs would increase the estimate by \$1.6 or \$2.5 billion dollars. Additionally, we do not attempt to quantify the reduced quality of life from increased disability and the lost non-pecuniary benefits due to lower educational attainment (Oreopoulos and Salvanes, 2011).

In addition to the direct, person-level costs of wildfire smoke exposure, there are also increased fiscal burdens on governments via lower tax revenue, increased use of the social safety net, and increased incarceration. These fiscal burdens are displayed in Panel B of Table 2.1. The first row, which is concerned with lower tax revenue, takes the decreased earnings (from Panel A) and multiplies them by the average federal marginal tax rate from 2000-2019 of 13.2% (Tax Policy Center, 2021) to arrive at the lost tax revenue.⁵² The cost impact of increased social safety net use is taken from Section 2.6.1 and scaled by the average exposure of 10.5 months to arrive at a discounted burden of \$482 per person. Finally, we utilize the increases in incarceration from the previous section to calculate the burden of housing inmates. Using costs from Vera Institute of Justice (2015, 2012), Bureau of Justice Statistics (2016), and Sawyer and Wagner (2020), we calculate the per-year cost of incarceration and discount back to birth.⁵³ This exercise yields an average discounted cost of \$99 per person. When combined, we calculate that these factors impose an additional cost of \$1,887 per person, or \$0.94 billion per 500,000 individuals.

These impacts were estimated using historical birth cohorts from the 1930s through 1960s. However, it is worth considering whether these effects are expected for children born in most recent years. On one hand, there are substantial advances in medical technology available to modern children that were not available during our sample period, and there is an increased ability to avoid air pollution due to indoor air filtration. However, the degree to which these technologies are able to offset the lasting effects of early-life shocks is unknown: insofar as impacts to childhood health

⁵²This exercise abstracts away from losses in state and local tax revenue, but these could also be substantial as lower earnings likely also translate into reduced sales taxes via decreased consumption.

⁵³Specifically, we gather incarceration costs for local jails (Vera Institute of Justice, 2015), state prisons (Vera Institute of Justice, 2012), and federal prisons (Bureau of Justice Statistics, 2016), converting them to 2010 dollars. We then weight them by the relative number of individuals in each level of incarceration (Sawyer and Wagner, 2020) to obtain the cost per incarceration year. We further adjust this estimate to reflect that we use the “institutionalized group quarters” variable from the Census, which also includes individuals in nursing facilities. (As discussed in the previous section, we cannot separately disclose incarceration results due to disclosure restrictions.) Accordingly, we down-weight our cost estimate by the fraction of institutionalized 35-59 year-olds that were in prison (85.2%). We then use this down-weighted incarceration cost measure and discount back to birth, using the conservative assumption that individuals are equally likely to be incarcerated at any given year in this age range.

are not salient—which might be the case for mild cognitive impairment or exposures that increase long-run cancer risk, for example—it is unclear whether modern technology provides substantial protection from long-run harms. Furthermore, it is possible that any improvements in our ability to lessen or avoid the impacts of childhood wildfire smoke may be more than offset by the increasing severity of wildfire seasons (Figure 2.1), which subject children to more frequent and more severe smoke exposure.⁵⁴ Because the net effect of these countervailing factors are unclear, we believe that more research on the early childhood impacts of wildfire smoke in a modern context is needed. A potential advantage in studying wildfire pollution is that, unlike industrial pollution sources such as coal power plants, the emissions profile has likely not substantially changed over time. Accordingly, any studies that determine the medium-run effects of early-life smoke exposure within a more recent context (using outcomes such as education or early-adult earnings) will be able to compare findings to estimates in this paper. Such comparisons may improve our ability to understand how the impacts of these childhood shocks have evolved over time.

2.9.2 Conclusion

In this paper, we contribute to the understanding of the long-term effects of childhood air pollution exposure by leveraging variation in an increasingly important source: smoke from wildfires. Specifically, we provide a deeper understanding of the impact of early-life exposure across the life cycle by evaluating economic, education, and disability outcomes at later ages. Specifically, we find that moving from the 25th to the 75th percentile of smoke exposure reduces annual earnings by \$890. We further find that this level of exposure reduces the probability of completing high school (by 0.40 percentage points) and increases the probability of cognitive difficulty (by 0.30 percentage points). Further investigation of these later-life harms demonstrates that they grow proportionally worse as individuals age and that failure to incorporate these later-life dynamics can potentially cause large understatements in the discounted damages from early-life shocks.

Because the findings in this paper and the previous chapter demonstrate that childhood exposure to wildfire pollution has substantial long-term costs, there are potentially larger benefits than previously understood from policies which abate, ameliorate, or defend against the harms from wildfire smoke. We conclude by briefly discussing these policies. First, because worsening wildfires—and the amounts of related air pollution—are exacerbated by the warming climate (Vose

⁵⁴Future iterations of this paper will quantify the degree to which present-day children are more exposed to wildfire smoke than their historical counterparts in our sample period.

et al., 2012; IPCC, 2021; Zhuang et al., 2021), any policy that slows or reverses the trend of higher temperatures will have an associated benefit of improved health from decreased wildfire smoke. However, absent a reversal of warming temperatures, it is possible that improved forest management practices could reduce the air pollution from wildfires. A substantial amount of resources in fire prevention and containment are devoted to protecting homes and other structures (Plantinga et al., 2020); however, the costs of wildfire pollution suggest that minimizing the biomass burned is important even when fires occur in remote areas.

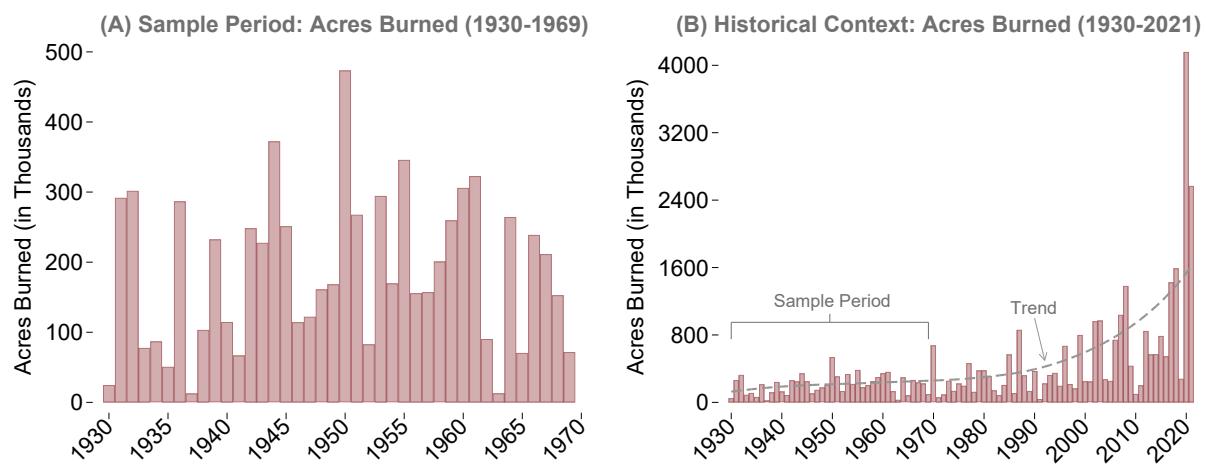
One potential policy that could abate the occurrence of large wildfires is an increased use of prescribed burns, a technique that has potentially been underutilized in California and other western states (Boxall, 2021; Sommer, 2021). Traditionally, local and state authorities have had substantial responsibility for enacting prescribed burns. However, because wildfire smoke can create meaningful interstate health externalities, it is possible that Federal regulators should take a larger role in encouraging forest management practices that prevent large fires. Schultz et al. (2019) suggests that adequate funding and incentives are key barriers to sufficient use of prescribed burns, two areas that could benefit from an increased role by the Federal government.⁵⁵

While abatement of wildfires could be an important policy lever, it is not the only option for reducing harms of childhood pollution. Because the impacts of wildfire smoke on indoor air quality are much lower (Burke et al., 2021), policies that encourage indoor behavior and improve indoor air filtration in homes, schools, and workplaces may be beneficial. Although more research on the effectiveness of filtration against wildfire smoke is needed, existing studies suggest that improving general air filtration, especially in schools, is very cost-effective (e.g., Gilraine, 2020; Stafford, 2015).⁵⁶ Finally, we note that it may be possible to mitigate injury from childhood pollution shocks even after children are exposed (Billings and Schnepel, 2018). Absent interventions to reduce or avoid wildfire smoke through other methods, understanding ways to reduce the harms of air pollution through medical treatment, social programs, or other means is an important area for future research.

⁵⁵Of course, pollution from controlled burns can be harmful as well. However, controlled burns consume much smaller amounts of biomass and can be planned on days when pollution would be blown away from densely-populated areas.

⁵⁶Less is known about the impact of air filtration in homes. Burke et al. (2021) note that there is substantial dispersion in the degree of smoke infiltration among different homes, but that household characteristics, such as income, do not explain much of the variation.

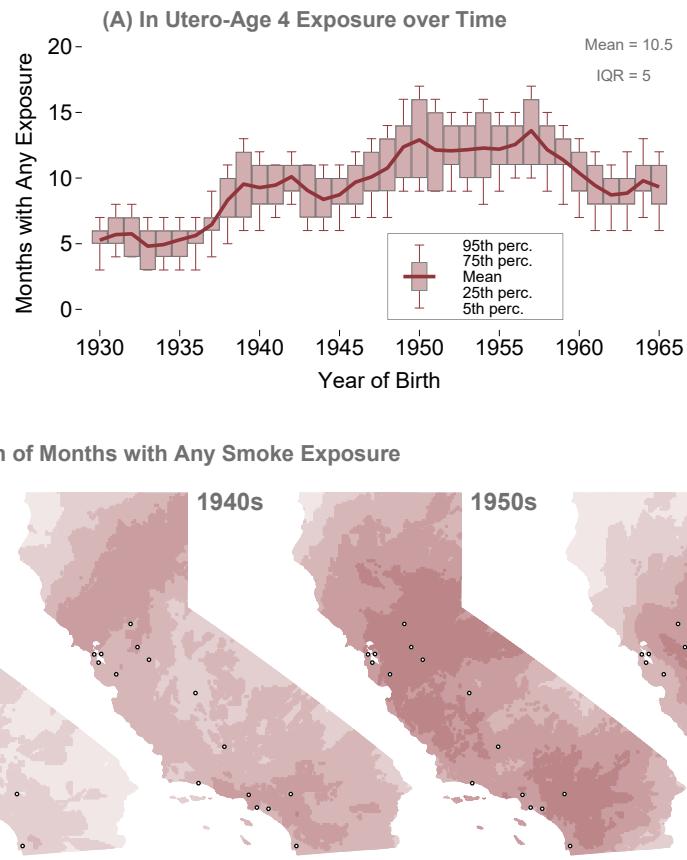
Figure 2.1: Descriptive Statistics: There is Substantial Variation in Year-to-Year Acres Burned, but Intensity has Increased Sharply over the Last 30 Years



Notes: The purpose of this figure is to display the substantial over-time variation in wildfires during the sample period (Panel A) and the stark increase in wildfire acreage burned over the last thirty years (Panel B). See Section 2.3.1 for further discussion.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018) and 2021 fire data from fire.ca.gov/incidents/ (updated through October 11, 2021).

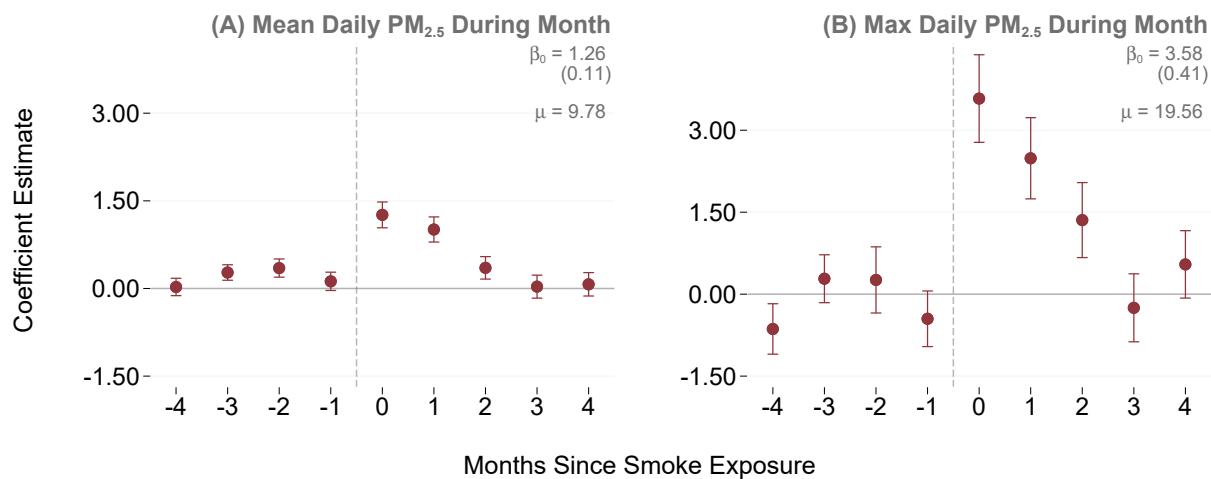
Figure 2.2: Identifying Variation: Smoke Exposure Varies Over Time and Geography



Notes: The purpose of this figure is to display the identifying variation utilized in our analysis. Panel A displays how the birth-weighted months of smoke plume exposure from an individual's *in utero* period through age 4 ("IU-Age 4") varies across birth cohorts. Within this panel, the mean and various quantiles are plotted for each cohort with the overall mean and interquartile range ("IQR") plotted in the upper-right corner. Panel B displays the spatial variation in smoke exposure according to the percentage of months in a given decade that experienced smoke exposure.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), the Public-version NumIdent (Goldstein et al., 2021).

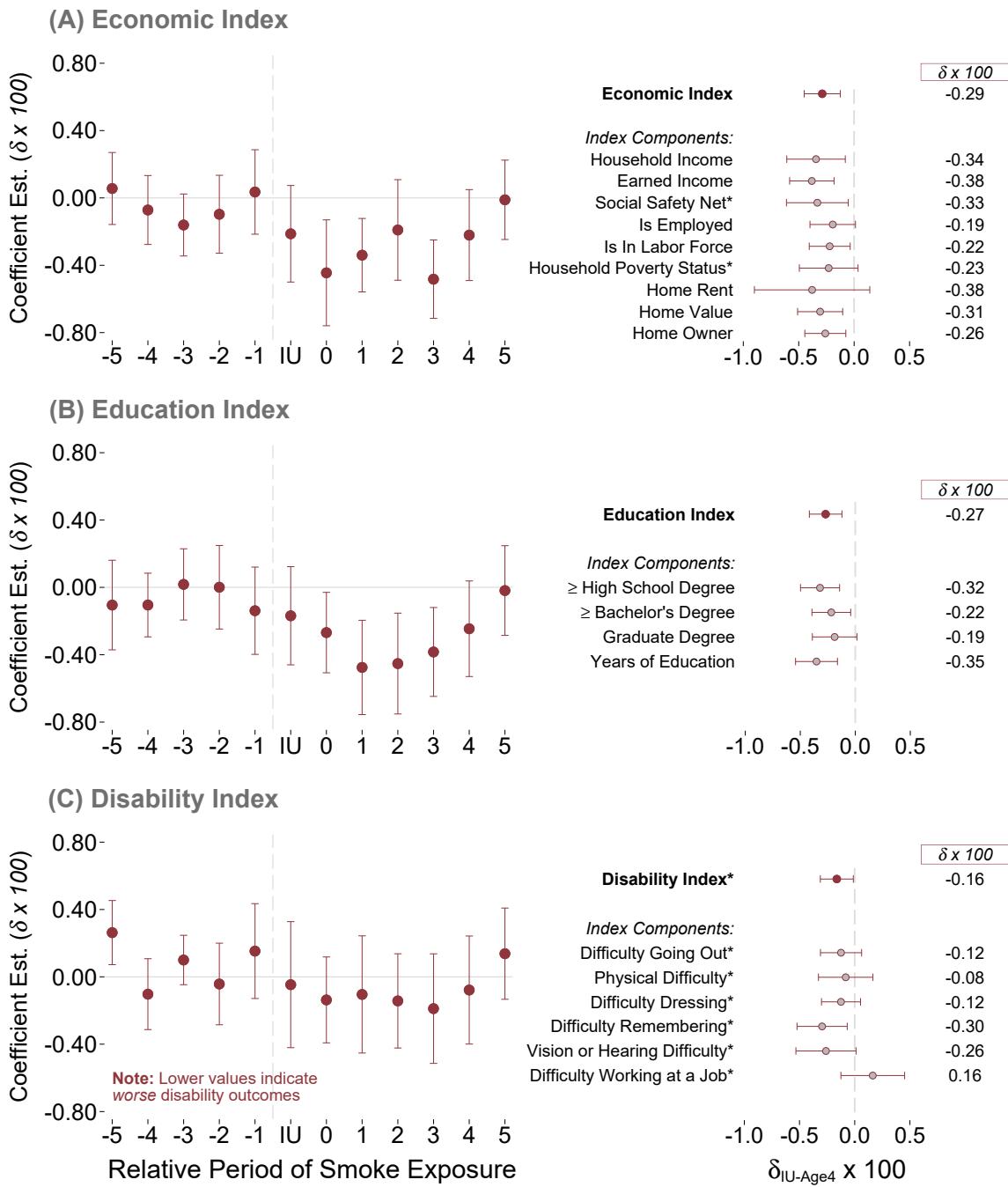
Figure 2.3: Validation of Treatment: Modeled Wildfire Smoke Exposure Increases Monthly PM_{2.5}



Notes: The purpose of this figure is to display the impact of modeled smoke plume coverage on monthly pollution levels. Each point represents a coefficient from the regression detailed by Equation 2.4 with corresponding 95% confident intervals. The point estimate for the impact on pollution during the month in which the fire occurs (β_0), along with the standard error and baseline mean are presented in the upper right-hand corner of the graph.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), and Daily Summary Data Files from the EPA Data Mart.

Figure 2.4: Main Results: Early-Life Wildfire Exposure Reduces Long-Run Economic and Educational Achievement; Increases Disability

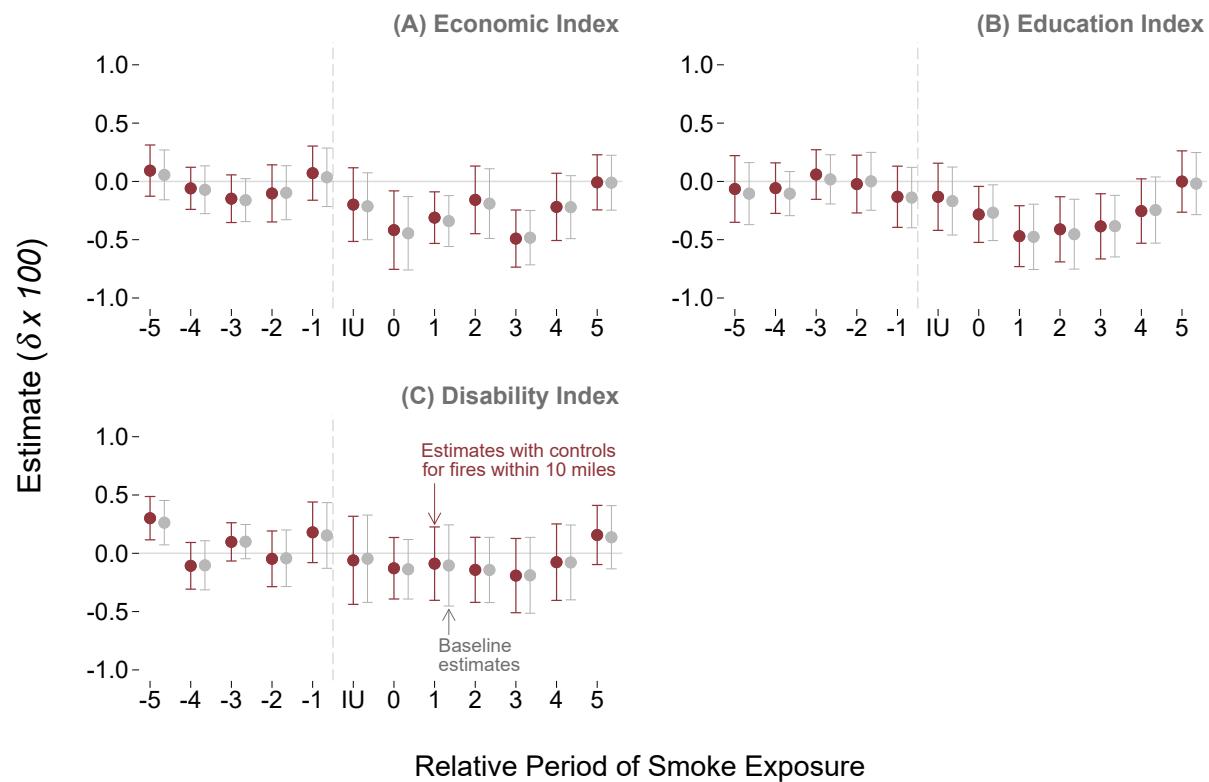


* Indicates that measure has been reverse coded so that higher values correspond with better outcomes (and v.v.).

Notes: The purpose of this figure is to display the impact of wildfire smoke exposure on long-run measures of economic achievement, educational attainment, and disability. Within the figure, each panel represents a different index of outcomes with leads-and-lags coefficients (from Equation 2.2) plotted on the left side of each panel, while summary estimates (from Equation 2.3) of the effect over the *in utero* through Age 4 period are plotted on the right. These summary coefficients are present for the applicable index and each of its components and are accompanied by the actual coefficient values (in the “ $\delta \times 100$ ” column). All estimates in this figure should be interpreted as the percentage-of-a-standard-deviation change in response to an additional month of smoke exposure during a given period.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

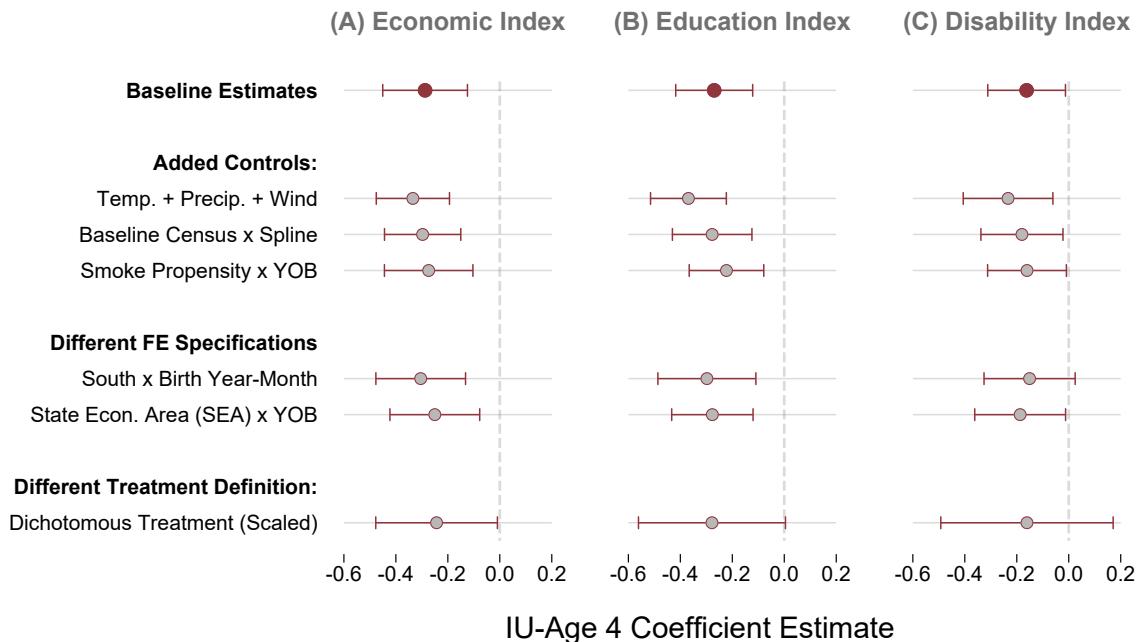
Figure 2.5: Robustness: Controlling for Nearby Fires Does Not Affect Estimates, Implying that Smoke Exposure is Driving Effects



Notes: The purpose of this figure is to demonstrate that smoke exposure, rather than proximity to wildfires themselves, are driving for estimates. To do so, the figure compares our baseline estimates from our leads-and-lags specification to those when we re-estimate the specification with controls for fires that are within 10 miles of an individual's city of birth. See the text of Section 2.7.1 for further discussion.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

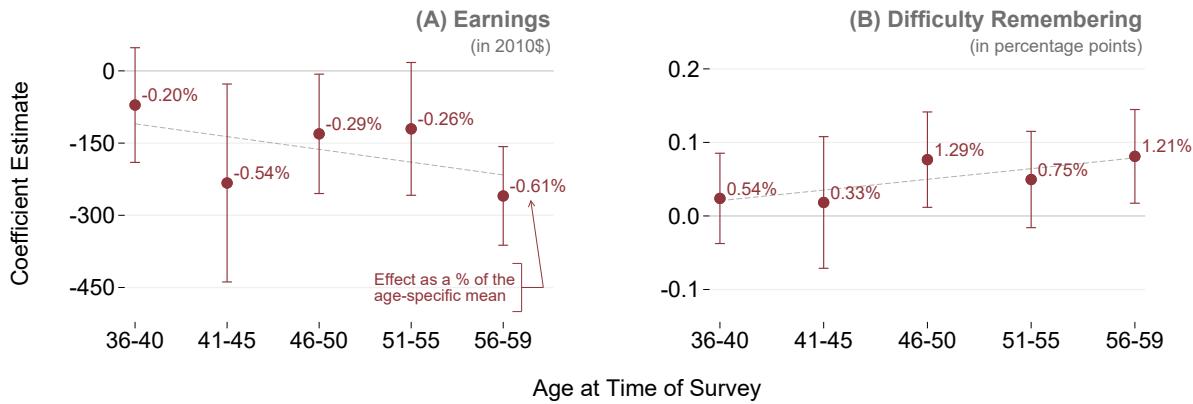
Figure 2.6: Robustness to Additional Controls and Specifications



Notes: The purpose of this figure is to display the robustness of results to differing specifications. Within the figure, our summary coefficients (Equation 2.3) for our preferred specification are presented on the top line, with estimates for differing specifications (due to added controls, changes to fixed effect structure, or modifications of functional form) included in the rows below. See Section 2.7 for further description.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Figure 2.7: Supplemental Results: The Effects Early-Life Wildfire Exposure Increase Proportionally with Age for Key Time-Varying Outcomes



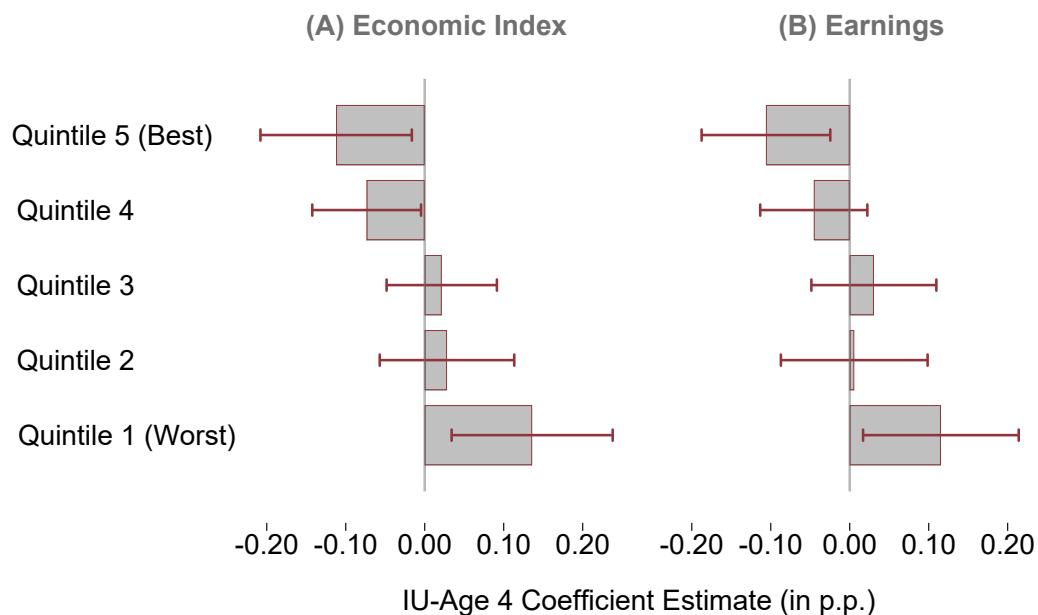
Notes: The purpose of this figure is to display how the effects of childhood wildfire smoke exposure evolve as individuals age. Within the figure, each point represents a separate estimate of δ_j with associated 95% confidence intervals, from the following equation:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \sum_{j \in \mathcal{J}} \delta_j \cdot (AgeBin_j \times Smoke_c^{IU-Age})^4 + \varepsilon_{cba},$$

For further context, the effect as a percent of the age-specific mean is also presented, along with a line of best fit for the coefficients. As demonstrated by the figure, the effects of early life wildfire pollution shocks increase proportionally with age, a finding which is discussed further in Section 2.8.1.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Figure 2.8: Supplemental Results: Early-Life Wildfire Exposure Impacts the Entire Distribution of Economic and Earnings Outcomes



Notes: The purpose of this figure is to display the distributional impacts of wildfire smoke exposure. Within the figure, each bar represents the coefficient from a separate regression where the outcome is whether or not an individual falls into that quintile (see Section 2.8.2 for more detail). Estimates are associated with 95% confidence intervals.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Table 2.1: The Cost and Fiscal Burden of Early-Life Wildfire Smoke Exposure

(A) Lost Earnings and Life Years		(B) Fiscal Burden	
Earnings Loss (Discounted)	\$ 9,892	Lost Federal Tax Revenue (Discounted)	\$ 1,306
Life Years Lost (Longevity)	\$ 12,558	Additional Social Safety Net Costs	\$ 482
		Additional Incarceration Costs	\$ 99
Loss per Person	\$ 22,450	Fiscal Cost per Person	\$ 1,887
Loss per 500,000 Persons (in \$B)	\$ 11.22	Fiscal Cost per 500,000 Persons (in \$B)	\$ 0.94

Notes: The purpose of this table is to translate our estimates from Chapter 1, Section 1.6; and Chapter 2, Sections 2.6 and 2.8 into a measure of the total costs. Within the figure, there are two panels. Panel A summarizes the total per-person costs of lost earnings and lost life years, while Panel B summarizes the total per-person fiscal burden from lost tax revenue and government expenditures. Additionally, both panels scale these results by 500,000 persons, which is the approximate size of a single California birth cohort. See Section 2.9 for further discussion of how this table was calculated.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, Restricted American Community Surveys, and Restricted Full-count Decennial Censuses.

Chapter 3

How Do Medicare Payments Influence Physician Practice Structure?

3.1 Introduction

In its role as the largest purchaser of clinical services in the United States, the Centers for Medicare and Medicaid Services (“CMS” or “Medicare”) yields tremendous influence on physician decisions through its price-setting behavior. The research considering the impact of these prices has largely focused the effects of reimbursement on intensive-margin treatment decisions: i.e. how physicians modify the degree of resource utilization in response to changes in prices. However, an emerging literature has shown the impacts of Medicare payments reach much further, from technology adoption (Finkelstein, 2007; Acemoglu and Finkelstein, 2008; Clemens and Gottlieb, 2014) to private-market price-setting (Clemens and Gottlieb, 2017).

This paper builds on this literature by examining how Medicare policy shapes the way in which physician practices are structured. Specifically, I examine how primary care physicians change their choice of practice structure—as indicated by the size of medical group that they are affiliated with—in response to differential Medicare payment. The structure of physician practices is relevant to both patient health and healthcare costs. Larger multispecialty practice groups may have positive impacts on health outcomes (Epstein et al., 2010; Baker et al., 2019); however, they may also gain bargaining power with commercial insurers, leading to higher prices, premiums, and out-of-pocket costs in the private market (Dunn and Shapiro, 2014; Baker et al., 2014).¹ Given the

¹Others, such as Scott et al. (2017) and Short and Ho (2019) find no effects or negative effects of vertical integration on quality measures.

size of the market for privately financed physician and clinical services (\$287 billion in 2016), even limited increases in commercial prices could have large cost ramifications.

To assess the effect of differential reimbursements on these outcomes, this paper exploits spatial discontinuities in Medicare fee-for-service payments within several metropolitan areas across the United States. Medicare has differential payments for providers located in certain “localities,” which correspond to select metropolitan areas.² Providers practicing outside of these metropolitan areas are assigned a “Rest of State” payment rate, which is typically less generous. Using a boundary regression discontinuity design, I exploit these sharp differences in payment to assess the plausibly exogenous impact of different prices on providers practicing near these borders who serve the same patient populations and face the same input costs.

I find that a 1 percent increase in Medicare reimbursement leads to a 1.7 to 2.2 percentage point (“p.p.”) increase in the share of primary care physicians (“PCPs”) practicing in smaller medical groups (defined as practices with 25 or fewer providers). This finding is consistent with a theoretical model of physician choice, in which doctors balance financial considerations with non-financial preferences over practice structure. Comparative statics from this model predict that income from more generous reimbursement allows clinicians to adopt practice structures that they find are more suited to their non-financial tastes but that are otherwise less profitable.

To more completely understand how across-the-board changes in Medicare reimbursement affect the nature of medical practices, I examine how differential reimbursement influences the distribution of group sizes. Within my preferred sample, I find that the distribution on the higher-payment side of the boundary stochastically dominates the low-payment distribution. Further analysis underscores that the effects on group size manifest early in the distribution: in response to Medicare reimbursement that is 1 percent more generous, the fraction of primary care doctors practicing in the smallest (1-2 provider) groups increases by 1.8 percentage points. These effects persist through the distribution until they are offset by reductions in PCPs affiliated with groups of 150 or greater. Given that groups of this size are generally comprised of hospital systems or physician groups that are affiliated with a particular hospital system, these findings suggest that higher overall Medicare prices may decrease vertical integration in healthcare markets.

My empirical findings are robust across specifications, bandwidth choices, and sample restrictions. Furthermore, to perform tests of my identifying assumptions, I conduct a variety of

²These localities often closely follow county boundaries, though they are defined by CMS on the zip-code level. The full list of designated payment localities can be found in Appendix Table B.1

bunching tests across different sub-samples and find no evidence that physicians endogenously respond to differential payment by relocating to the side of the border with higher reimbursement. Additionally, I find that observable physician characteristics do not predict whether they receive higher reimbursement, further underscoring that selection is not a driving force of the results.

Finally, I perform a back-of-the-envelope calculation to assess the degree to which these changes in practice structure could result in increases in private-market prices through changes in bargaining power. When combining my estimates with those from Dunn and Shapiro (2014), I find that a 1% *decrease* in Medicare prices may result in a 0.15% to 0.50% *increase* in private prices. Given the relative size of the private market to Medicare, it is possible that a substantial fraction—if not all—of savings from decreased Medicare reimbursement are simply offset by higher costs to participants in the privately financed market.

My findings contribute to the literature that analyzes how Medicare pricing affects market-wide behavior. Acemoglu and Finkelstein (2008) show that a reform in the Medicare Prospective Payment System, which increased the relative cost of labor by moving from full to partial cost reimbursement, increased adoption in new medical technologies in the hospital setting.³ Additionally, Clemens and Gottlieb (2017) employ a geographic consolidation of payment areas and sharp adjustment to the relative price of surgical versus medical care to investigate how Medicare's prices influence those set by the bargaining processes between providers and private insurers. They find that changes in private prices closely follow those made by Medicare. Finally, the research that is most closely related to this paper is Dranove and Ody (2019), who find that increasing the relative reimbursement for hospital-employed physicians leads to an increase in the share of doctors that vertically integrated with hospital systems (and therefore could participate in these higher reimbursement rates). I build upon their work by finding that, in addition to the directly incentivized effects that they estimate, broad pricing pressure can affect market consolidation and integration as well.

This paper also contributes to the body of research studying the impact of prices on physician location. Much of the physician location literature, summarized by Nicholson and Propper (2012), focuses on broad physician location choice in response to financial incentives and the urban-rural divide in provider access and concludes that financial incentives have at most modest effects

³Clemens and Gottlieb (2014) also study technological adoption using a consolidation of geographic adjustments in the Medicare Physician Fee Schedule. They find that increased reimbursement is associated with higher rates of MRI utilization, particularly by non-radiologists. Another paper related to Medicare-induced technology adoption and market structure change—although not directly due to price changes—is Finkelstein (2007), who studies the aggregate effects of the rollout of Medicare.

on location choice. This paper extends this literature by examining whether financial incentives affect location choice within urban areas, as indicated by bunching analyses and other metrics of spatial concentration. Consistent with the literature on broad location choice, I find that financial incentives have no detectable effects.

The remainder of this paper proceeds as follows: Section 3.2 discusses the background information on Medicare geographic adjustment factors, which provide the identifying variation for my analysis, and the data sources used. Section 3.3 presents a stylized theoretical model and resulting comparative statics to motivate the expected effect of Medicare payments on practice structure. Section 3.4 details the identification strategy, econometric specification, and tests of validity for my research design. Section 3.5 discusses main estimates, robustness checks, and heterogeneity analyses. Section 3.6 discusses the implications of my findings and concludes.

3.2 Background and Data

3.2.1 Medicare Payments and Identifying Variation

CMS reimburses physician services under Medicare Part B through use of the Physician Fee Schedule. In a given year, providers in payment locality a are reimbursed for service j based on three different components:

$$Payment_{aj} = ConversionFactor \times \mathbf{RVU}'_j \times \mathbf{GAF}_a, \quad (3.1)$$

where *ConversionFactor* is an annually updated baseline for payment and **RVU** represents a vector of three Relative Value Unit (“RVU”) components, which are intended to capture the resources required for the labor, practice expense and malpractice premiums, respectively, for a given service (and are therefore a measure of quantity). The **GAF** term represents geographic adjustment factors that account for the fact that labor, practice expense, and malpractice premiums vary in price across various regions.

Medicare-designated metropolitan areas, which have not changed since 1997, receive a GAF that is set based on locality-specific costs, while the remaining areas of the state, including non-designated cities, are assigned a GAF based on costs outside of the designated localities.⁴ Figure 3.1 illustrates the geographic differences in payment areas. The top panel of the figure shows

⁴The GAF is calculated using physician wages (labor); employee wages, costs for contracted services, rent and supplies (practice expense); as well as liability insurance premiums (malpractice).

how different payment localities are dispersed throughout the United States. There are 17 states with multiple geographic adjustment factors and 37 designated metropolitan areas within those states (see Appendix Table B.1 for a list of these localities). The bottom panel provides a more detailed view of two localities in the Dallas-Fort Worth Metroplex to illustrate how the resulting payment differences vary over the geography of urban areas. Providers in the Dallas and Fort Worth areas receive more generous reimbursements than providers of the same services in surrounding suburban areas (ex. Plano or Flower Mound). These sharp, localized, *within-state* differences in payments provide the variation used for my analysis. To be clear, I do not utilize across-state differences in geographic adjustment factors in my analysis, since it is likely that many factors (e.g., Medicaid/private insurance environment, licensing requirements, etc.) change across these boundaries.

The magnitude of the differences across geographic boundaries is illustrated by Figure 3.2, where the unit of analysis is a locality boundary.⁵ As evidenced by Panel A, cross-sectional discontinuities between payment areas can be quite large—in some cases over 10%. Furthermore, for providers that face the same costs, as is the case for those close to the boundary for these payments, this additional Medicare revenue translates directly to profit, increasing its potential to impact provider decisions. Meanwhile, these discontinuities are quite stable over time, as detailed in the Panel B. I view this stability as an advantage—Clemens and Gottlieb (2014) demonstrate that physician responses to changes in Medicare prices evolve over several years. Due to the lack of variation in prices over time, I believe the effects that I identify are indicative of long-run behavior.⁶

3.2.2 Data

The data used for the empirical analysis is primarily gathered from the Centers for Medicare and Medicaid Services. The Medicare Physician and Other Supplier Public Use File supplies provider-level data on addresses, summary information on patients treated, Medicare utilization, and characteristics such as certification (e.g., Medical Doctor, Doctor of Osteopathic Medicine) and primary specialty type. The CMS Physician Compare data set is used to gather additional provider-level information on group size and affiliations, medical school attended, and year graduated.⁷

⁵For instance, the boundary between Dallas and Rest of Texas payment localities is one such locality boundary, as is the border between Dallas and Fort Worth, both of which are displayed in Figure 3.1(b).

⁶Additionally, boundary discontinuity designs readily lend themselves to the identification of long-term effects, as demonstrated by Ambrus et al. (2020) and Dell (2010), both of which identify long-run impacts from geographic discontinuities.

⁷The medical school rankings in this paper were obtained from Schnell and Currie (2018).

Within the Physician Compare data, providers are considered to be part of a group practice if they have filed a Medicare claim associated with that group's taxpayer identification number within the last twelve months. Because non-physician providers also bill Medicare, the group size measures in this data will differ slightly from those who only measure the number of physicians associated with a given group.

The geographic adjustment factors and relative value units are obtained from the Physician Fee Schedule for reimbursements under Medicare Part B. The American Community Survey (“ACS”) five-year (2011-2015) estimates (Ruggles et al., 2019) were used to obtain demographic information on the Census-block-group level, which includes variables on age, educational attainment, insurance coverage, household income, home values, and population. Finally, geographic shape files were obtained from the U.S. Census Bureau. All data used is from 2015 unless otherwise noted.⁸ Further information on the data used and sample construction is detailed in Appendix B, Section B.

To construct the sample of physicians used for analysis, I first geocode all providers in states that have multiple payment localities.⁹ Next, using the longitude and latitude for each physician address, I calculated the straight-line distance between providers and nearby locality borders, keeping all doctors within 10 miles.¹⁰ I exclude from my analysis boundaries where physicians on one side of the boundary do not provide a good counterfactual for those on the other side. This includes boundaries that straddle state lines (because providers on each side face a different regulatory environment, as well as different private and state-level public insurance), as well as those that are defined by rivers (because providers may face different demographic environments). Additionally, I remove borders where there are less than five providers within the bandwidth on either side of the border and remove providers that practice in a different locality than either side of the boundary they would otherwise be assigned to.¹¹

My final sample is further limited to only include physicians who specialize in primary care; accordingly, all subsequent references to doctors or physicians throughout the remainder of the paper focus on primary care physicians. I focus on primary care physicians for two reasons. First,

⁸While earlier years are available (specifically 2012-2014), I focus only on the cross-section due to the lack of overtime variation in my treatment, as shown in Figure 3.2(b), and to avoid confounding my results with after-effects of major legislation, namely the Patient Protection and Affordable Care Act, passed in 2010.

⁹This process was done using Texas A&M GeoServices.

¹⁰Locality borders are defined by adjacent zip codes that are assigned different geographic adjustment factors.

¹¹Consider Figure 3.1(b) for an example of this. A doctor practicing in northwest Coppell could be within a bandwidth-distance of the Fort Worth-Rest of Texas locality border and thus eligible for assignment to that border, even though she practices in the Dallas locality. Accordingly, restrictions are made to ensure that this does not occur.

non-primary care doctors are much more likely than to have multiple practice locations (Xierali, 2018), making treatment assignment based on a single location problematic and likely to introduce measurement error. Second, relative to medical or surgical specialists, doctors in primary care are relatively homogeneous—only two specialties, Family Practice and Internal Medicine, make up roughly 95% of PCPs.¹² This homogeneity readily facilitates analysis within the regression discontinuity framework discussed in Section 3.4.

Finally, in my preferred sample, I maintain only boundaries that have payment differences of 2 percent or more (I refer to this as my “High-Impact” sample). Unless specifically noted, this sample will be used for all analyses. However, as shown in Section 3.5.3, my results are robust to including all observations.

3.3 Theoretical Model

3.3.1 Setup

While theoretical models of physician behavior have considered how differential prices might induce doctors to change treatment intensity or patient composition between Medicare and privately insured patients—McGuire and Pauly (1991) and Glied and Zivin (2002) are two well-cited examples—there is a lack of theory illustrating how different Medicare payments might induce variation in medical practice size. Accordingly, I present a stylized model of structure choice that contrasts physician profit-seeking behavior with non-financial considerations. However, before doing so, it is useful to understand the considerations that underlie the choices that physicians make when selecting different group sizes.¹³ The significant dispersion in medical group size suggests that there are a number of advantages and disadvantages to joining a larger medical group. These advantages and disadvantages are summarized by Casalino et al. (2003), who analyze physician assessments of the advantages of larger group practice—which include bargaining leverage with health plans, economies of scale, and lifestyle factors—as well as the disadvantages—lack of coordination, lack of physician leadership and other problems stemming from group incentives. These issues, along with loss of autonomy over patient care, are echoed in popular literature on the subject (e.g., Lipton 2012).

¹²In contrast, the top two specialist disciplines (Emergency Medicine and Diagnostic Radiology) comprise less than 17% of all specialty providers.

¹³I focus specifically on physicians as decision makers, as they provide the majority of health care services and substantially influence decisions in the clinical setting. This is true even if professional managers make decisions as agents of the clinicians.

With these factors in mind, consider a highly simplified model where physicians choose practice structure, s , at the beginning of each period to maximize total utility, u .¹⁴ Total utility is additively separable into utility from the physician's share of profits, $v(\pi(s))$, and utility from non-financial characteristics, $w(s)$. In such a case, the physician solves:

$$\max_s u(s) = v(\pi(s)) - w(s), \quad (3.2)$$

where both $v(\cdot)$ and $w(\cdot)$ are increasing and concave in s .¹⁵ The assumptions on $v(\cdot)$ are standard, while those on $w(\cdot)$ are motivated by the aforementioned literature—e.g. disutility from coordination and lack of autonomy is increasing in the size of the group practice but exhibits diminishing disutility in larger practices.

Within this model, the physician's profit as a function of group size s is:

$$\pi(s) = p^o(s) \cdot q^o + p^m \cdot q^m - mc(s) \cdot (q^m + q^o), \quad (3.3)$$

where p^x and q^x denote the price and quantity of patients covered by insurer type $x = \{\text{Medicare } (m), \text{ Other private } (o)\}$, and $mc(s) \cdot (q^m + q^o)$ represents a cost function that is linear in quantity and the same for publicly and privately insured patients. The marginal cost function, $mc(s)$, is assumed to be decreasing and convex to capture efficiency costs associated with larger groups. Finally, adapting the conceptual framework from Clemens and Gottlieb (2017), I parameterize private prices, p^o , as:

$$p^o(s) = (1 - \theta(s)) \cdot p^{ins} + \theta(s) \cdot p^m, \quad (3.4)$$

where $\theta(s) \in [0, 1]$ is the relative bargaining strength of the insurer. The term p^{ins} is the highest price the insurer is willing to pay the physician and p^m is the Medicare price, which acts as the physician's reservation price.¹⁶ Lower values of $\theta(s)$ represent physician groups with substantial

¹⁴This framework makes two substantial simplifications for the purposes of clarity and tractability. First, the model treats the physician's location as exogenous. As discussed in Section 3.4.2, this assumption appears to be empirically supported, as there is little-to-no evidence of spatial response to differential payments. Second, quantity could also increase in tandem with higher prices due to physician-induced demand, as demonstrated by Clemens and Gottlieb (2014) or patient composition of Medicare vs. privately insured patients could change. Appendix B, Section A investigates how incorporating quantity responses affects model outcomes. I find that, under similar assumptions, predictions are largely the same as discussed below. Additionally, the results discussed further in Appendix B, Section C suggest minimal extensive-margin response to differential prices, supporting the assumption that patient composition is unchanging.

¹⁵Note that, even though the function $-w(s)$ is convex, it also satisfies the conditions for quasi-concavity based on the strict monotonicity of its first derivative. Accordingly, u is strictly quasi-concave and therefore can achieve a maximum.

¹⁶Note that, in any equilibrium where bargaining has resulted in an agreement between rational parties, it must

bargaining power and are able negotiate prices closer to the insurer's reservation price. Conversely, high values of $\theta(s)$ represent doctors with lower bargaining power and are therefore paid rates closer to the reservation price. To reflect larger groups' ability to elicit higher prices (Dunn and Shapiro, 2014), $\theta(s)$ is assumed to be decreasing and is additionally assumed convex to reflect diminishing bargaining power as s becomes large.

3.3.2 Comparative Statics

To assess the impact of differential Medicare prices on selected group size, consider the first order conditions of the model specified by Equations 3.2 through 3.4:

$$\frac{\partial u(s)}{\partial s} = 0 \Leftrightarrow v_{\pi\pi}\pi_s - w_s = 0 = \phi(s), \quad (3.5)$$

where subscripts denote derivatives with respect to the subscripted argument. Then, totally differentiating to obtain the comparative static of Medicare prices, p^m , on group size, s , yields:

$$\frac{ds}{dp^m} = -\frac{\partial\phi(s)/\partial p^m}{\partial\phi(s)/\partial s} = -\frac{v_{\pi\pi}\pi_s\pi_{p^m} + v_{\pi\pi}s p^m}{u_{ss}}. \quad (3.6)$$

Utilizing the properties of the model specified in Section 3.3.1, the sign of Equation 3.6 is:

$$\frac{ds}{dp^m} < 0. \quad (3.7)$$

Thus, the model predicts that group size responds negatively to higher Medicare prices, as the income effect of higher prices affords physicians the luxury of choosing group sizes that more closely align with their non-financial preferences. This conclusion holds even when price or marginal costs are assumed to be unresponsive to group size (i.e. $\theta_s = 0$ or $mc_s = 0$). I next consider whether this conclusion is empirically true in the following sections.

be that $p^{ins} \geq p^m$.

3.4 Empirical Framework

3.4.1 Identification and Econometric Model

To identify the effects of Medicare payments on practice structure, I estimate the following model for all providers within the neighborhood of a locality border:

$$Y_{ip} = \delta \cdot High_p + \phi_{b(p)} + f(Z_p) + h(X_{ip}) + \varepsilon_{ip}, \quad (3.8)$$

where i denotes a primary care physician practicing in location p . The main outcome of interest, Y_{ip} , is a variable indicating whether a PCP is affiliated with a small practice group, defined as a medical group practice with 25 or fewer providers. The threshold of 25 providers was chosen because that is the group size generally used by CMS to distinguish small and large practice groups.¹⁷

The variable $High$ is a binary variable equal to one if a provider is on the side of the border with higher payment rates. The coefficient on $High$, δ , is the parameter of interest. To account for locality-specific characteristics of providers, I also include border fixed effects, $\phi_{b(p)}$. In certain specifications, $\phi_{b(p)}$ represents cluster-by-border fixed effects, wherein nearby physicians assigned to the same border are grouped together into clusters (using k -means clustering) in order to better control for spatial characteristics of the localities that I study. Next, $f(Z_{ip})$ represents a polynomial in distance from the border that varies on each side of the cutoff. In my preferred specification, I utilize a first-order polynomial. Finally, $h(X_{ip})$ is comprised of a flexible set of local, patient-population, and provider controls.¹⁸

The identifying variation from my analysis comes from sharp changes in geographic adjustment factors over space. I assign GAFs to providers based on their practice location and specialty.

¹⁷See, for instance, Centers for Medicare and Medicaid Services (2015). This measure was chosen as my primary measure, rather than a measure of average group size selected, because moderate shifts in the *distribution* of group sizes can have very large effects on the *average* group size that are difficult to interpret and may seem implausible at first glance. To be concrete: in the context I study, if 5% of primary care physicians in a market moved from the 75th percentile of group size (roughly 250 providers) to the 25th percentile of group size (roughly 2 providers), this would result in a reduction of ~ 0.50 log points (or a $\sim 40\%$ decrease) in the average log group size chosen by physicians in that market. Nonetheless, I also consider the effects of this payment discontinuity on average (log) group size and find that the effects are consistent with the other results presented in this paper. See Appendix B, Section C for more detail.

¹⁸Local controls include percentage of population by (a) age bins, (b) educational attainment bins, and (c) insurance coverage categories; median household income, median real-estate value, and density controls. Block group characteristics are assigned to individual physicians by calculating an inverse-distance-weighted average of characteristics within 2 miles of a physician's location, where distance is calculated from the physician location to block-group population centroid. Patient population controls include the average age and hierarchical condition category score for a physician's patient population. Finally, physician characteristics include quantiles of medical school rank and experience, as well as controls for gender.

Specifically,

$$GAF_{ip} = \sum_{\kappa \in work, pe, mp} gaf_{\kappa(p)} \cdot rvu_share_{\kappa(i)}, \quad (3.9)$$

$gaf_{\kappa(p)}$ is the geographic adjustment factor component (work, professional expense, or malpractice premiums) that varies based on physician zip code and $rvu_share_{\kappa(i)}$ is the share of relative value units in each component for physician i 's specialty, calculated at the national level.¹⁹ The GAF is constructed using national RVU shares—rather than individual shares—to avoid any endogenous change in service mix in response to differential payment.

The interpretation of the coefficient of interest, δ , in Equation 3.8 is the change in the outcome in response to higher payment. To facilitate interpretation as the effect of a 1 percent change in reimbursement, I also estimate a regression discontinuity-instrumental variables (“RD-IV”) design as follows:

$$Y_{ip} = \delta \cdot \widehat{GAF}_{ip} + \phi_{b(p)} + f(Z_p) + h(X_{ip}) + \varepsilon_{ip}, \quad (3.10)$$

$$GAF_{ip} = \theta \cdot High_p + \phi_{b(p)} + f(Z_p) + h(X_{ip}) + \varepsilon_{ip}. \quad (3.11)$$

The coefficient θ in the first stage (Equation 3.11) describes the average differential in reimbursement between the high- and low-payment sides of the borders in my sample. In my main analysis, the coefficient of interest, δ , is equal to the percentage point change in the number of PCPs practicing in small groups in response to a 1 percent increase in Medicare generosity. For notational convenience (and to differentiate from estimates from other equations), I refer to this coefficient as ” δ_{RD-IV} ” throughout the paper. This estimate could alternatively be recovered by separately estimating the first-stage regression (Equation 3.11) and the reduced-form equation (Equation 3.8) and subsequently dividing the reduced-form estimate (δ_{RF}) by the first-stage estimate ($\delta_{RD-IV} = \delta_{RF}/\theta$).

The estimation of my reduced-form analysis (Equation 3.8) is performed using ordinary least squares, while RD-IV analyses are estimated using two-stage least squares. All analyses utilize a rectangular kernel in distance from the border and standard errors clustered at the census tract-level to account for spatial correlation. Alternatives to both kernel and clustering choices are considered in Appendix B, Section C, with highly similar results.

¹⁹To provide a concrete example: if 53%, 43%, and 3% of the Family Practice specialty's total RVUs come from work, professional expense, and malpractice RVUs, respectively, then these percentages will be used as weights for the locality-specific work, professional expense, and malpractice GAFs to arrive at the final geographic adjustment factor for that specialty.

The selected bandwidth is based on the travel distance to obtain medical services. Specifically, I set bandwidth equal to the median travel distance in the core-based statistical areas that contain my borders of interest, as determined by the National Household Transportation Survey (“NHTS”). The resulting bandwidth is equal to 4 miles. Robustness to alternate bandwidth choices are considered in Section 3.5.2.

3.4.2 Validity of Regression Discontinuity Design

The primary identifying assumption underlying my regression discontinuity design is that the unobservable determinants of practice structure evolve smoothly across the boundary (i.e. that they are continuous with respect to the running variable). However, in the context studied in this paper, there are clear financial incentives for physicians to migrate to higher payment areas. Accordingly, a threat to my research design is manipulation of the running variable, which would be evidenced by significant high-payment side bunching immediately after the boundary, as providers relocate their practices to more profitable locations.

I consider such behavior in Figure 3.3. Panel A displays the results of a McCrary (2008) bunching test, which shows no evidence of manipulation around the boundary. However, the figure does present a distinct bi-modal pattern, with peaks on either side of the boundary (with a higher-density mass on the high-payment side), which is largely due to changes in population density. Accordingly, Panel B presents a similar analysis to Panel A using the total number of primary care physicians per thousand individuals. Within the figure, each dot represents physician-to-population ratios in one-tenth-of-a-mile increments, along with local linear regression lines and associated 90 percent confidence intervals fit to either side of the cutoff. This figure displays no detectable discontinuity at the cutoff. Notably, there is a statistically insignificant *decrease*—rather than the hypothesized increase—in the physician-to-population ratio on the high-payment side of the boundary.

A further concern regarding manipulation of the running variable is that, due to spatial frictions, physicians might relocate several miles away from the cutoff, rather than immediately to the high-payment side. In this case, it would be possible that there would be no evidence of bunching at the cutoff, but there would still be mass displaced from the low-payment side to the high-payment side. If this were the case, then Figure 3.3B would demonstrate a consistently higher mass of primary care physicians per thousand individuals on the high-payment sides of the boundary, which I do not see. To further extend the analysis, Panels C and D of Figure 3.3 examine

the location choices of PCPs who moved within the last three years, on the basis that physician-movers may be subject to fewer optimization frictions as non-moving physicians. Panel D further limits these movers to those that have moved within hospital referral regions (“HRRs”) and thus may be more informed about local differences in reimbursement than other movers. Neither of these “mover” analyses provide any evidence to suggest that primary care physicians are sorting around these payment boundaries, even when optimization and informational frictions are reduced. Finally, Appendix Figure B.1 considers additional tests of bunching by re-performing two different variants of the analysis in Figure 3.3B: (i) using the log (rather than level) of the physician-to-population ratio and (ii) using the number of practice *locations* (rather than primary care physicians). The results of these analyses closely mirror those in Figure 3.3B, further supporting a lack of sorting around the boundary.

It is worth considering why substantial bunching is not observed at the cutoff, given the impact of differential payments on other aspects of provider practice. There are several possible reasons for this. First, geographic location plays a significant role in patient demand for health care services (e.g., Dunn and Shapiro, 2014; Gowrisankaran et al., 2015; Kessler and McClellan, 2000). Accordingly, relocation of a practice could result in significant patient turnover and thus could be very costly. A second and closely related factor is the impact of spatial proximity on competition between medical practices. To the degree that patients choose providers based on distance, providers will be incentivized to avoid increases in competition—and the resulting patient acquisition costs and/or loss of volume—that would result from bunching. Finally, while the impact of these payments on physicians’ profit is substantial, the actual borders that determine these price differences may not be salient for physicians. This is particularly true for physicians (and groups) who practice in a single locality, as they may not have a readily available counterfactual of the differential payment they would receive if practicing elsewhere. Additionally, while physician groups with multi-locality practices may be more aware of geographic payment differences, they are disincentivized to spatially adjust, because doing so might lead to competition among physicians in the same practice.

While I do not detect any manipulation of the running variable, it is theoretically possible that physician sorting across the boundary is occurring in a way that is not detected by my bunching analysis. While I cannot test for sorting along unobservable dimensions, I can evaluate whether discontinuous changes in observable physician characteristics exist. To evaluate discontinuities in observable characteristics, I apply the least absolute shrinkage and selection operator (“LASSO”)

method to predict (i) receipt of treatment and (ii) choice of small group practice. Employing a LASSO design in this context allows me to utilize a wide range of physician characteristics and their interactions to best predict my outcomes of interest and examine their evolution over the cutoff.²⁰ Insofar as sorting on either side of the boundary occurs, I expect the LASSO-predicted outcomes to also display discontinuous changes at the cutoff.

Figure 3.4 displays the results of this analysis, where Panel A shows the results when the LASSO-predicted GAF is the outcome variable in Equation 3.8 and Panel B demonstrates the estimated effect when the outcome is the LASSO-predicted fraction of primary care physicians in a small practice group. Within the panels, each point represents the means of the outcome in 0.5-mile increments, with the size of the point varying by the number of observations included. All means presented have been regression-adjusted to net out locality-border fixed effects. Each side of the cutoff—where negative distances represent low-payment sides and positive distances are associated with high-payment sides—are fit with linear regression lines and associated 90% confidence intervals. Estimated coefficients from Equation 3.8 along with associated standard errors are presented in the upper-left corner. (All regression discontinuity plots displayed in the remainder of this paper are formatted in this way.)

As displayed in the figure, there is no evidence of sorting across the cutoff. Not only do the point estimates fail to achieve any traditional level of statistical significance, but they are also economically small: a 0.1 percent difference in the predicted geographic adjustment factor and a 0.6 percentage point difference in predicted small group participation, both of which are 1-2 orders of magnitude smaller than the results discussed in the next section. Taken in total, the analyses suggest no physician sorting that would bias my results. As an additional test of robustness, I further examine the sensitivity of my estimates to inclusion of provider and block-group controls in Section 3.5 and find that my results are not sensitive to inclusion of covariates.²¹

²⁰Physician characteristics include indicator variables for specialty type, gender, credential, medical school ranking quantile, experience quantile, and quantiles for a wide array of patient characteristics. Interacting these indicators to allow a highly flexible specification equation yields 1,931 covariates for the LASSO procedure to select from.

²¹As further discussed in Appendix B, Section C and displayed in Appendix Figure B.2, I also examine how individual characteristics vary across the cutoff, both on the physician- and block-group-level. These results also do not suggest any physician sorting across the boundary.

3.5 Results

3.5.1 Main results

The central results of my analysis are graphically displayed in Figure 3.5. Panel A demonstrates the impact of the border discontinuity on physician reimbursements: among borders included in my High-Impact Sample, the discontinuity results in a 5.1 percent increase in reimbursement on average. Panel B displays the effect of this discontinuous increase in payments on physician practice structure. Consistent with the comparative statistics of the model described in Section 3.3, I find that primary care physicians organize into smaller groups in response to more generous payment. Specifically, I estimate that there is an 11.3 percentage point increase in primary care physicians that practice in small groups—defined as groups with 25 or fewer providers—on higher-payment sides of the boundary. The estimates in this figure are mirrored in the first column of Table 3.1, Panels A and B, respectively. An unusual feature of the regression discontinuity plot in Figure 3.5, Panel B worth noting is that the point immediately to the right of the cutoff has a substantially higher fraction of PCPs in small group practices than would be expected based on the regression line. To alleviate concerns that this outlier point is driving the results of my analysis, I perform a “donut-hole” test in Section 3.5.2, wherein I re-perform my analyses after omitting observations within a specified distance of the cutoff. I find that omission of this unusual point does not meaningfully change my results.

To facilitate the interpretation of the effects that I find in my reduced-form analysis, I estimate the RD-IV specification (Equations 3.10 and 3.11) in Panel C of Table 3.1. These estimates imply that a 1 percent increase in the reimbursement rate is associated with a 2.2 p.p. increase in the fraction of primary care doctors affiliated with small practice groups. As illustrated in Columns 2 and 3 of Table 3.1, these results are stable when control variables are added. Furthermore, when I replace my border fixed effects with more spatially fine-grained cluster-by-border fixed effects (Column 4), I obtain similar results.²² As noted in Section 3.4, all specifications utilize a rectangular kernel in distance from the boundary and are clustered at the Census-tract level. Because the results are highly similar across all specifications, I consider Column 1 to be my preferred specification, as it is the most parsimonious.

²²See Appendix Figure B.4 for illustration of these physician clusters and Appendix B, Section B for a discussion of how they were constructed using k -means clustering. To further assess the role of spatial characteristics, Appendix Figure B.3 presents estimates and associated regression discontinuity plots of a specification *without* any spatial fixed effects and finds results that are qualitatively similar, though slightly larger in magnitude.

In addition to my “small group” outcome, I also examine the effects of differential reimbursement across the entire distribution of selected group sizes. Panel A of Figure 3.6 displays the raw cumulative distributions of selected practice sizes separately for the high- and low-reimbursement sides of the border. (Note that the x -axis utilizes a logarithmic scale of base 2, which is employed because the group sizes exhibit substantial positive skewness.) Meaningfully, the distribution on the high-payment side of the border stochastically dominates the low-payment distribution, indicating that a much higher fraction of primary care physicians in higher-payment areas choose smaller practices. A Kolmogorov-Smirnov test easily rejects that the two sides have equal distributions ($p < 0.01$).

To evaluate this within the context of my regression discontinuity framework, I estimate separate RD-IV regressions (Equations 3.10 and 3.11) for various points in the cumulative distribution. This is done by utilizing an outcome of $Y_{ip} = 1(\text{GroupSize} \leq g)$, where g represents a point in the cumulative distribution.²³ The estimates from this analysis, along with their associated 90% confidence intervals are presented in Figure 3.6, Panel B.²⁴ This figure supports the conclusions drawn from Panel A: in response to higher payment, providers are more likely to sort into the small group practices and less likely to be employed by the largest groups.

The effects on practice size manifest very early in the distribution of group sizes. The RD-IV estimate implies that, in response to reimbursement that is 1 percent more generous, there is a 1.8 p.p. increase ($p = 0.050$) in primary care physicians who are affiliated with practice groups with 1 or 2 providers. The effects on the smallest group persist throughout the distribution until they are offset by reductions in groups of 150 greater ($\delta_{RD-IV} = -1.9$, $p = 0.075$).²⁵ Groups of this size are generally comprised of hospital systems or physician groups that are affiliated with a particular hospital system. Accordingly, to frame my results another way, pricing *pressure* might force providers who would have non-financial preferences to practice in a small group to select larger organizations, with the net effect of increasing primary care physicians employed by large, vertically integrated health systems and their associated physician employment groups. More generally, these results may suggest that diminished profitability is a driving force behind the recent trend of consolidation of small physician practices into larger ones (e.g., Capps et al., 2017; Kane, 2019; Muhlestein and Smith, 2016).

²³This analysis is performed using similar methodology as Barcellos et al. (2019), who examine the distributional effects of education on health within a regression discontinuity framework.

²⁴For an analogous exercise focusing on the probability distribution, rather than cumulative distribution, see Appendix B, Section C and Appendix Figure B.5.

²⁵Regression discontinuity plots for these figures are displayed in Appendix Figure B.6.

Finally, I consider two other measures of practice structure, both of which are presented in Appendix Figure B.8. The first measure, presented in Panel A, details effects of differential effects on practice *location* size. Specifically, the outcome for this analysis is the log number of providers at a given practice location, defined as an address-group combination. I do not detect any effects on practice size, suggesting that providers are not responding by *spatially* consolidating into locations with more physicians. This is consistent with the lack of spatial response detailed at length in Section 3.4.2. The second measure, presented in Panel B, details the impacts of differential reimbursement on average log group size. This measure is not presented as a primary outcome because it is challenging to interpret and because moderate shifts in the *distribution* of group sizes can have very large effects on *average* group size. The results in Panel B underscore this: the discontinuity displayed implies a decrease of 0.54 log points (roughly a 41% decrease) in group size. While this may appear implausible at first, approximately the same result would be achieved if 5% of the primary care providers in a market moved from the 75th percentile of group size (roughly 250 providers) to the 25th percentile of group size (roughly 2 providers). Based on the differences in cumulative distributions displayed in Figure 3.6, effects of roughly this magnitude appear to be present, supporting the assertion that the effects on average size are reasonable.²⁶

3.5.2 Robustness

Aside from the assumptions that unobservable characteristics do not vary over the boundary and that there is no meaningful sorting by providers, the choice of bandwidth is the most critical to my analysis. To consider sensitivity of my results to this assumption, I re-estimate my RD-IV results (Equations 3.10 and 3.11) for various multiples of bandwidth, as shown in Figure 3.7 Panel A. Overall, the coefficients are highly consistent, although they lose precision at narrower bandwidths, which is consistent with the bias-variance trade-off.²⁷

Additionally, motivated by the outlier point immediately to the right of the cutoff in Figure 3.5, I perform a “donut” hole analysis (Figure 3.7, Panel B), by estimating separate regressions on samples that omit observations within a given distance from the cutoff, ranging from 0 to 1 mile(s). As displayed in the figure, the RD-IV estimates are highly consistent over the size of the

²⁶I also consider two other more frequently examined outcomes from the physician payment literature: per-patient resource utilization (i.e. the intensive-margin treatment decision) and the number of Medicare patients treated (i.e. the extensive-margin treatment decision). See Appendix B, Section C and Appendix Figure B.7 for further discussion.

²⁷An alternate approach is to utilize methods developed by Calonico et al. (2014) to non-parametrically select bandwidth and estimate results within a local linear regression specification. When applying these methods within a fuzzy regression discontinuity framework (comparable to my RD-IV specification), I obtain highly similar results to my baseline analysis ($\delta = 2.205$, $p = 0.016$).

omitted sample, although my estimates do predictably fail to attain traditional levels of statistical significance when increasingly large number of observations are omitted. In addition to measuring sensitivity to outliers, this test accomplishes two goals. First, insofar as there is sorting *immediately* around the cutoff that is not otherwise detected by my tests in Section 3.4.2, physicians who have manipulated the running variable (and therefore unobservably differ across the cutoff) will be omitted from my sample. Second, these analyses will reduce measurement error in my treatment variable from inaccurately geocoded physician locations, to the degree that such measurement error exists. The consistency of the estimates over the range of “donut hole” sizes supports the idea that my estimates are not meaningfully biased by endogenous sorting or measurement error.

I perform a variety of other robustness checks, which are detailed and discussed further in Appendix B, Section C. However, I will briefly summarize them here. First, I consider whether results are being disproportionately influenced by individual locality borders by re-running my preferred specification while omitting a given border from the analysis. The results of this analysis, which are detailed in Appendix Figure B.9, demonstrate that my results are highly consistent when omitting any given border from my sample. Next, I consider sensitivity of my estimates to different choices of clustering strategy, kernel, and controls (i.e., using quadratic and border-varying linear distance controls as well as granular physician specialty controls). I also consider a wide variety of sample restrictions to test sensitivity of my estimates to sample construction choices and spatial idiosyncrasies. The results of these tests are summarized in Appendix Figure B.10, where I find that my point estimates are consistent regardless of specification choice.

3.5.3 Alternate IV and Full Sample Results

The results so far are based the on reduced-form and RD-IV analyses discussed in Section 3.4. While these methods have the advantage of transparency (i.e. both the first and second stages can be clearly displayed graphically), they do not utilize all the plausibly exogenous variation that is generated by these border discontinuities. Specifically, the RD-IV design utilizes the *average* discontinuity across all borders, rather than exploiting discontinuities specific to each border. To address this issue, I estimate an alternative design, where the second stage is unchanged from Equation 3.10, but the first-stage equation utilizes multiple discontinuities and allows the θ coefficient

to vary across borders:

$$Y_{ip} = \delta \cdot \widehat{GAF}_{ip} + \phi_{b(p)} + f(Z_p, b(p)) + \varepsilon_{ip}, \quad (3.12)$$

$$GAF_{ip} = \sum_{k \in \mathcal{B}} \theta_k \cdot High_p \times 1[b(p) = k] + \phi_{b(p)} + f(Z_p, b(p)) + \varepsilon_{ip}. \quad (3.13)$$

The instruments in Equation 3.13 are created by interacting (i) a binary variable equal to 1 if a provider is located on the high-payment side of the border and (ii) an indicator variable for assignment to a given border. I also allow trends in the running variable to vary on each side of the cutoff *and* by each border, so the first-stage equation is effectively estimating a separate regression discontinuity for each of the borders in my sample. I refer to this estimation strategy as the “RD-Border” specification.

I apply this strategy, along with my RD-IV specification, to both my High Impact Sample (i.e., those borders with substantial payment differences) and the full sample of borders. The results are displayed in Table 3.2. The first column repeats the estimates from my preferred specification (Table 3.1) for comparison. When the RD-IV estimation is applied to the full sample—as shown in Column 2—the estimate decreases meaningfully and also suffers from a high degree of imprecision. This is consistent with a first stage that is roughly half as large as when the High-Impact Sample is used (Appendix Figure B.11). Columns 3 and 4 repeat the analyses of Columns 1 and 2 when applying the “RD-Border” estimation strategy to potentially increase the precision of my estimates. When this method is applied to the High-Impact Sample (Column 3), estimates are similar to my RD-IV analysis (although somewhat smaller, they are not statistically different). When I apply the RD-Border estimation to the full sample, the results—displayed in Column 4—yield nearly identical estimates to those from the High-Impact sample. Accordingly, these analyses demonstrate that the results I find are not driven by my choice of sample selection.²⁸

3.5.4 Heterogeneity

I conclude this section by conducting two heterogeneity analyses to better understand the effect of Medicare pricing on physician practice choice. First, because the financial impact of Medicare reimbursement policy will likely be larger for physicians practicing in areas with a higher share of Medicare-eligible population (those aged 65 and over), I test whether PCPs practicing in high-share

²⁸I also re-perform my distributional analyses from the previous section using the full sample and RD-Border estimation. These results, which are displayed in Appendix Figure B.12, are consistent with those from my High-Impact Sample analysis.

areas respond more strongly to differential Medicare reimbursement. To do so, I divide my High-Impact Sample into two groups: primary care physicians assigned to borders with above-median shares of Medicare-eligible population and PCPs assigned to borders with below-median shares.²⁹ The results of this analysis are displayed in Figure 3.8, Panels A and B, which mirror regression discontinuity plots presented in this paper, but with additional information about the first stage (θ_{FS}) and RD-IV estimates (δ_{IV}) presented in the bottom-right corner. As expected, estimates are substantially larger for primary care doctors who practice in areas where there are more Medicare-eligible individuals, and who are therefore more dependent on Medicare reimbursement. While it is possible that these effects are driven by other factors that are correlated with both larger over-age-65 population shares and higher degrees of responsiveness to payment differentials, this analysis provides further suggestive evidence that the differences in practice choice are driven by Medicare's reimbursement decisions.³⁰

Next, I investigate how variation in private insurer bargaining power (as proxied by market concentration) may affect physician responses to differential Medicare reimbursement. The effects of insurer bargaining power are *ex-ante* ambiguous. In the context of my model (Equation 3.4), powerful insurers are more likely to push negotiated private prices, $p^o(s)$, closer to the physician's reserve price, p^m (in other words, $\theta(s) \rightarrow 1$). This would imply that physicians in more highly concentrated private insurance markets would be more responsive to differential Medicare reimbursement, as Medicare differentials translate more directly to private prices. However, as noted in Barrette et al. (2020), countervailing market power by insurance companies may offset the increases in price that come from higher concentration in the provider markets. This means that the financial returns from selecting a larger group may be lower in markets with powerful insurance companies, and thus it is possible that physician practice choice may be *less* responsive in High-HHI insurance markets.³¹ To test which effect dominates, I divide my High-Impact Sample into

²⁹I choose to divide providers into groups based on border-level, rather than individual-level, characteristics because it is possible that the number of Medicare patients treated is endogenous.

³⁰In a related analysis, I also separately consider responses for the two largest primary care specialties: Internal Medicine and Family Practice. Given that Internal Medicine physicians generally have a larger share of Medicare beneficiaries (roughly 49% more per the 2010-2015 National Ambulatory Medical Care Surveys), it is expected that they would exhibit a stronger response to differential reimbursement. The results of the analyses displayed in Appendix Figure B.13 Panels A and B, confirm this: the RD-IV point estimate is approximately 61% larger for Internal Medicine physicians than for Family Practice doctors (although it should be noted that these estimates are not statistically different at traditional levels). I also consider other types of physician heterogeneity (Panels C through H of Figure B.13) and find little other evidence of differential response. See Appendix B, Section C for further discussion.

³¹This would be formally characterized by $\theta_s^{High\ HHI} \geq \theta_s^{Low\ HHI}$, where values of θ_s that are *more* negative represent larger shifts toward the insurers' reservation price.

physicians who are assigned to borders with above- and below-median market concentration and perform separate regressions on each sample.³² The results, displayed in Figure 3.8 Panels C and D, show that physicians in highly concentrated markets are more responsive to Medicare pricing policy, suggesting that the tighter relationship of private and Medicare prices in these markets may dominate any loss in incremental bargaining power gains from consolidation.

3.6 Discussion

This paper finds that Medicare payments affect practice structure choice among primary care physicians. Specifically, I find that a 1 percent higher payment rate leads to a 1.7-2.2 p.p. increase in the fraction of primary care doctors affiliated with small practice groups. Further investigation of practice size distribution finds that a 1 percent higher reimbursement rate increases the share of doctors in 1-2 provider groups by 1.8 percentage points and *decreases* the share in very large groups (≥ 150 providers) by 1.9 percentage points.

While addressing the welfare impact of these changes is beyond the scope of this paper, it is worthwhile to discuss the implications of larger group sizes on private prices for medical services, achieved through higher levels of bargaining power.³³ To do so, I relate select measures of group size to the Herfindahl-Hirschman Index (“HHI”), the most common metric used to study market power.³⁴ I begin by constructing a measure of HHI by using a similar method to Dunn and Shapiro (2014), who use fixed travel-time HHI (“FTHHI”) to assess market power in physician markets. I calculate the market share for each Census tract by first weighting each provider based on a 20-minute travel-time distance from the population-weighted centroid of that Census tract.³⁵ I then aggregate providers within the same group, and divide by the total number to obtain each group’s

³²Market concentration is determined using CBSA-level measures of HHI obtained from American Medical Association (2016). These CBSA-level estimates are mapped to the border level via a weighted average based on the number of physicians within a bandwidth-distance of the border that are in a given CBSA.

³³Given the changes in physician responses, welfare analysis would require an evaluation of the health impacts, if any, stemming from those responses. As noted in the introduction, the evidence on the health benefits of larger and vertically integrated practices is mixed, with some (Epstein et al., 2010; Baker et al., 2019) arguing that larger practices have health benefits, while others do not find any evidence of health improvements (e.g., Scott et al., 2017) or negative health effects (Short and Ho, 2019).

³⁴I am unable to directly assess the impact of differential Medicare payments on HHI because it is a market-level measure, whereas my treatment and control providers are all in the same market by construction.

³⁵Weighting for provider j is done according to the formula, $w_j = \max\{1 - (d_j/\bar{d}), 0\}$, where d_j is the distance of the provider from the centroid and \bar{d} is the 20-minute travel-time distance. The equivalent distance for a trip of 20 minutes was calculated using data from the NHTS.

Census-tract level market share:

$$s_g = \frac{\sum_{j \in J(g)} w_j}{\sum_{f \in \mathcal{F}} \sum_{j \in J(f)} w_j}, \quad (3.14)$$

where j indicates a provider and w_j the associated weight; $J(g)$ represents the set of providers affiliated with group g ; and \mathcal{F} is the set of all firms in the market. These shares are then used to calculate the FTHHI for each Census tract, which form the basis of the population-weighted average FTHHI for each locality border.

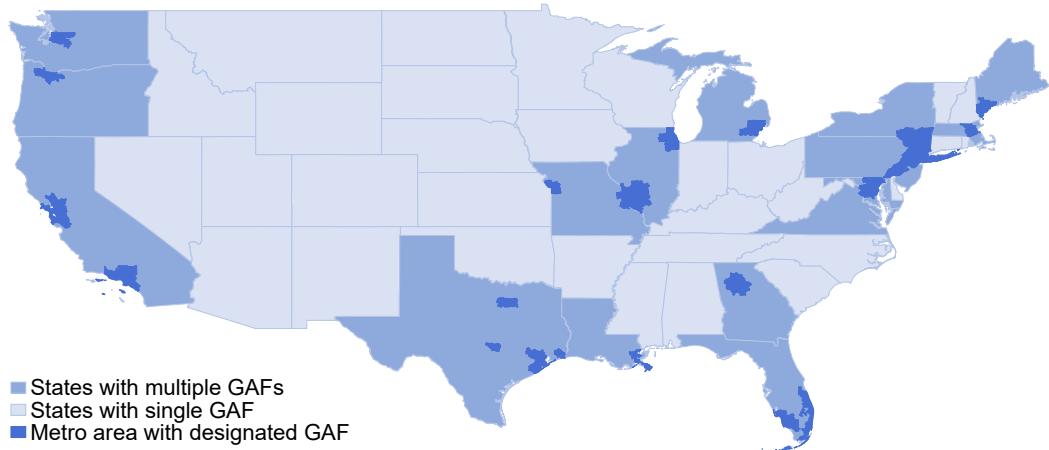
I then relate the mean FTHHI to two measures: (i) the market-level fraction of providers in small groups (the main outcome); and (ii) the market-level fraction of providers in groups of 1-2. The latter measure was chosen because it is the best predictor of FTHHI among my statistically significant distributional results (see Appendix Figure B.14). There is a strong linear relationship between group size measures and FTHHI: a 1 p.p. increase in the fraction of small groups corresponds to a 2.2% *decrease* in FTHHI, whereas a 1 p.p. increase in the fraction of 1-2 provider groups corresponds to a 5.4% decrease in FTHHI. Using these relationships, along with the estimated price increase from a 1% increase in FTHHI from Dunn and Shapiro (2014) and the estimates from Section 3.5, I calculate that a 1% *decrease* in Medicare prices induces a 0.15% - 0.50% increase in private market prices due to increased market power.³⁶ Given the size of the privately financed portion of the primary care services market (\$60 billion), this translates to a \$90 - \$290 million annual transfer from insurance companies to medical providers, compared to \$280 million in cost savings for Medicare, before accounting for other margins of response to Medicare payment.³⁷ Additionally, Baker et al. (2014) demonstrate that increased vertical integration is associated with higher hospital prices when hospitals own physician practices. Accordingly, to the extent that my estimated effects on physicians employed by the largest groups represents a change in hospital ownership of physician practices, price decreases could also have substantial cost spillovers in the inpatient market as well.³⁸

³⁶Dunn and Shapiro (2014) estimate that a 1% increase in 20-minute FTHHI results in a 0.04% to 0.05% increase in private prices.

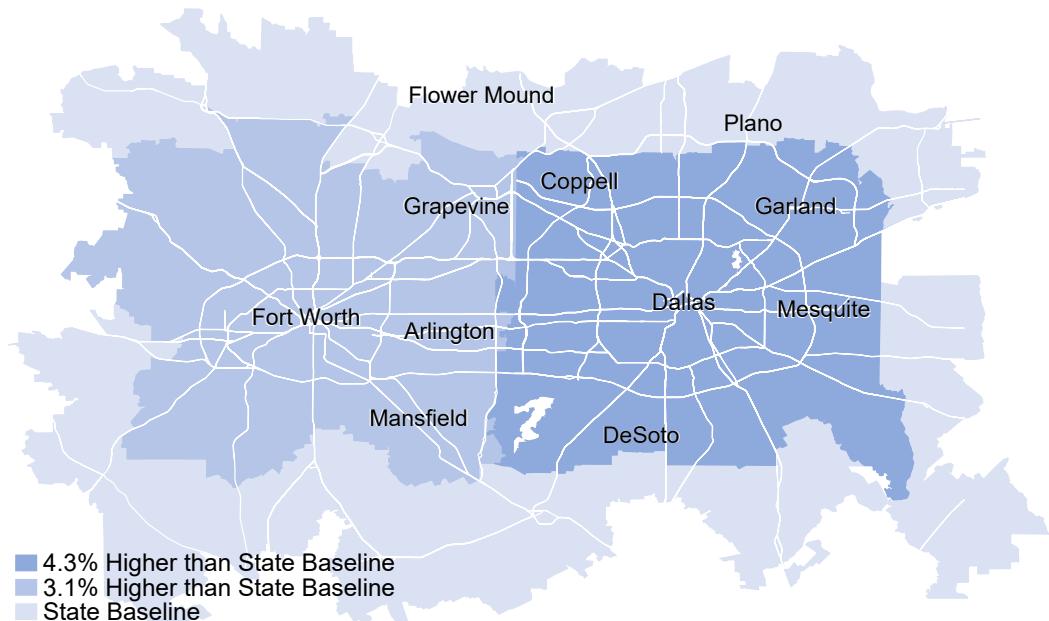
³⁷The size of the private and Medicare markets were derived by multiplying total spending on personal health care ((**alias?**)) by the estimated fraction of total spending attributable from primary care (Jabbarpour et al., 2019). In addition to holding other responses fixed, this calculation also makes the critical assumption that my results are generalizable to a nationwide decrease in Medicare payments. It is also possible that the transfer is overstated, as my calculations do not account for lower costs from larger groups that may be passed onto insurers.

³⁸Given the size of the private hospital market (\$427B per (**alias?**)), even small effects on vertical integration could impose large spillovers in relation to cost savings by Medicare. For instance, given that a 1 p.p. increase in vertical integration results in 0.137% increase in price (Appendix Table 2 of Baker et al., 2014), only a 0.47 p.p. increase in vertically integrated physicians would be required to match the naive Medicare cost savings from a 1% decrease in Medicare prices. Such an estimate is well within the range of possibility, based on the results displayed in Figure 3.6.

In summary, changes in practice structure stemming from lower Medicare prices have the potential to lead to higher private-market prices for both physician and hospital services. Depending on the incidence of these price changes, a substantial degree of savings from cuts in Medicare reimbursement could be borne by privately insured individuals in the form of higher out-of-pocket medical costs and health insurance premiums.



(a) Payment Localities Across the United States



(b) Payment Localities in the Dallas-Fort Worth Area

Notes: The purpose of this figure is to illustrate geographic variation in Medicare payment. Panel A displays differential payments for designated metropolitan areas. There are 17 states that have within-state variation from geographic adjustment by Medicare, which is provided by 37 designated metropolitan areas (indicated by the darkest colors). Light colors indicate states that lack intra-state variation in geographic reimbursement. Panel B illustrates differential geographic adjustments for a large metropolitan area. Primary care providers in the Dallas area and Fort Worth area have payment rates that are 4.3% and 3.1% higher, respectively, than providers of the same health care services in the surrounding suburban areas. Percent differences of geographic adjustment factors are obtained by multiplying each GAF component by the national average of work, practice expense and malpractice RVUs to provide the appropriate weighting of each component. Identifying variation is provided based on the sharp differences in these payments close to the boundaries. Note that the white lines in Panel B represent major roads and highways to provide points of reference.

Source: CMS Physician Fee Schedule.

Figure 3.1: Identifying Variation: Medicare Payment Locality Boundaries

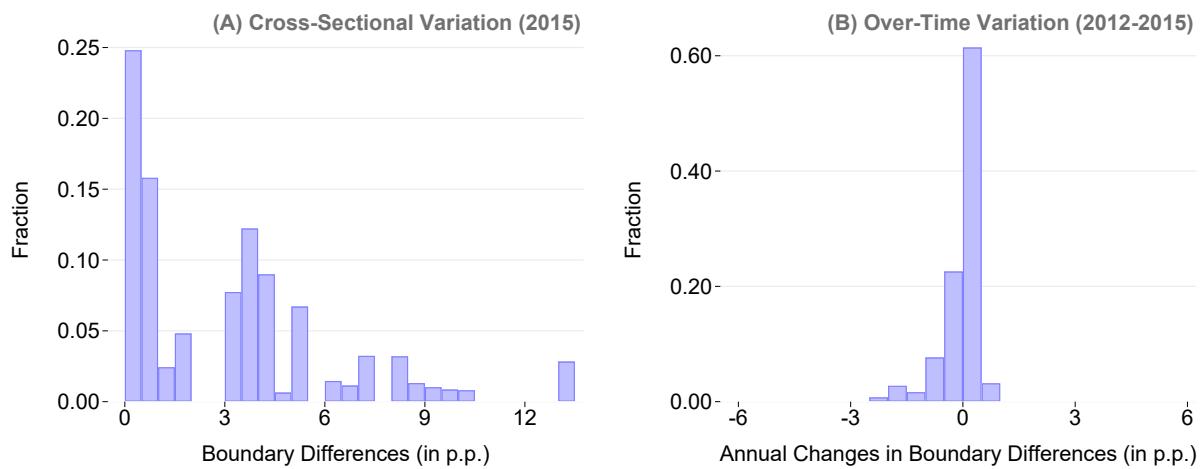


Figure 3.2: Identifying Variation: Boundary Differences in Medicare Payment

Notes: The purpose of this figure is to illustrate the dispersion of boundary differences in geographic adjustment factors and the relatively stability of those differences over time. Panel A shows the distribution of cross-sectional variation in geographic adjustment across locality borders, where a unit difference is interpreted as a 1% premium in a higher-paying adjacent locality. Panel B demonstrates that these differences are relatively constant over time. Percentage differences of geographic adjustment factors are obtained by multiplying each GAF component by the national average of work, practice expense and malpractice RVUs to provide the appropriate weighting of each component. Density calculations are weighted according to the number of primary care physicians within 4 miles of the boundary.

Source: CMS Physician Fee Schedule and CMS Physician and Other Supplier Public Use File.

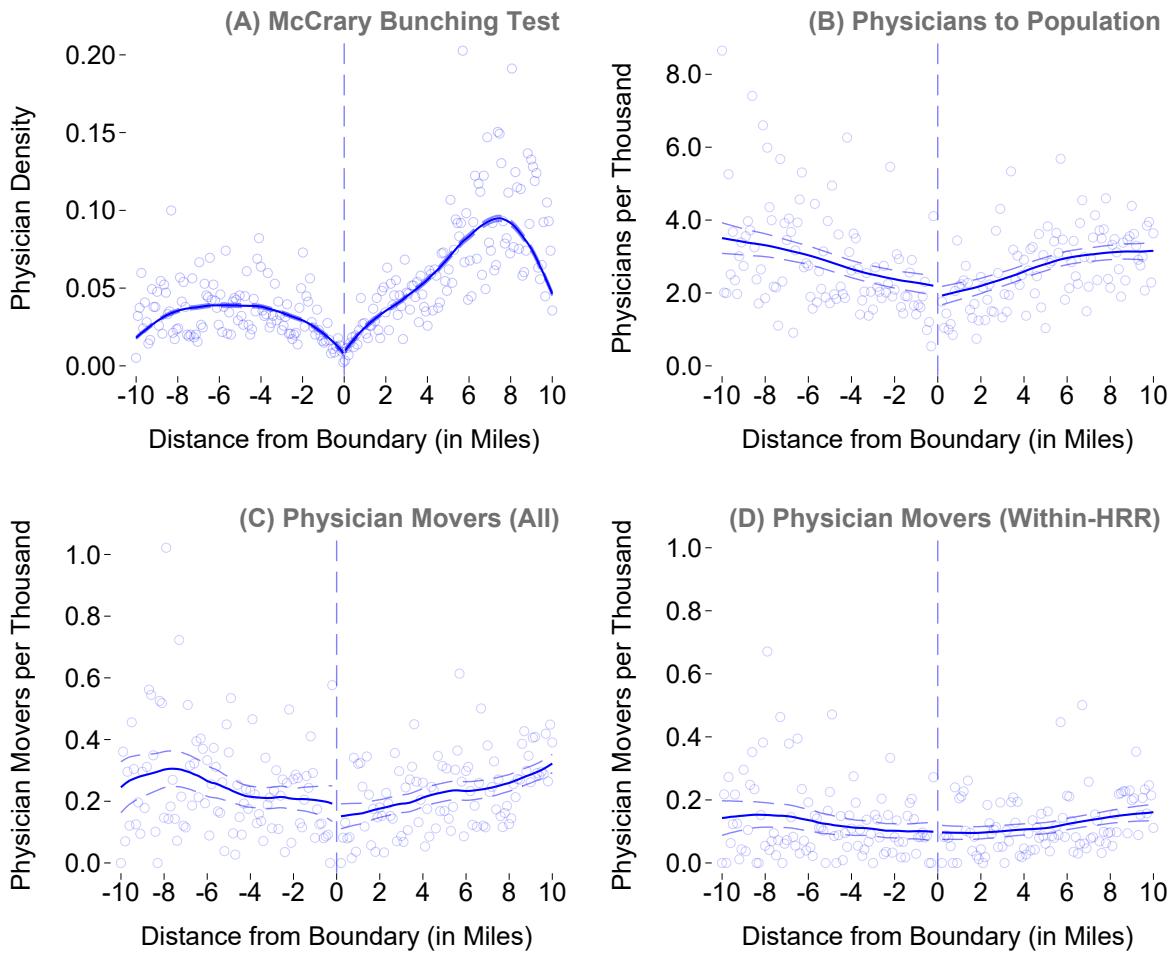


Figure 3.3: Validation of Research Design: Spatial Provider Density

Notes: The purpose of this figure is to illustrate the absence of bunching around locality boundaries. Panel A performs a McCrary (2008) analysis to assess the degree of bunching, if any, around the cutoff. Note that negative values indicate distances in the relatively lower payment areas, and positive values indicate distances in the higher payment areas. Panel B shows the spatial distribution of primary care physicians per 1,000 individuals by their distance to the nearest treatment boundary. Panels C and D examine physicians who moved to their current location during the years 2012-15, with Panel C examining all such movers and Panel D displaying results for movers who stayed within the same Hospital Referral Region (“HRR”). All local linear regression lines are presented with their 90 percent confidence intervals. For presentation purposes, observations are winsorized at the 99th percentile. See also Appendix Figure B.1 for other manipulation tests. This analysis utilizes the High-Impact Sample discussed in Section 3.2.2.

Source: Author calculations using CMS Physician Fee Schedule and CMS Physician and Other Supplier Public Use File.

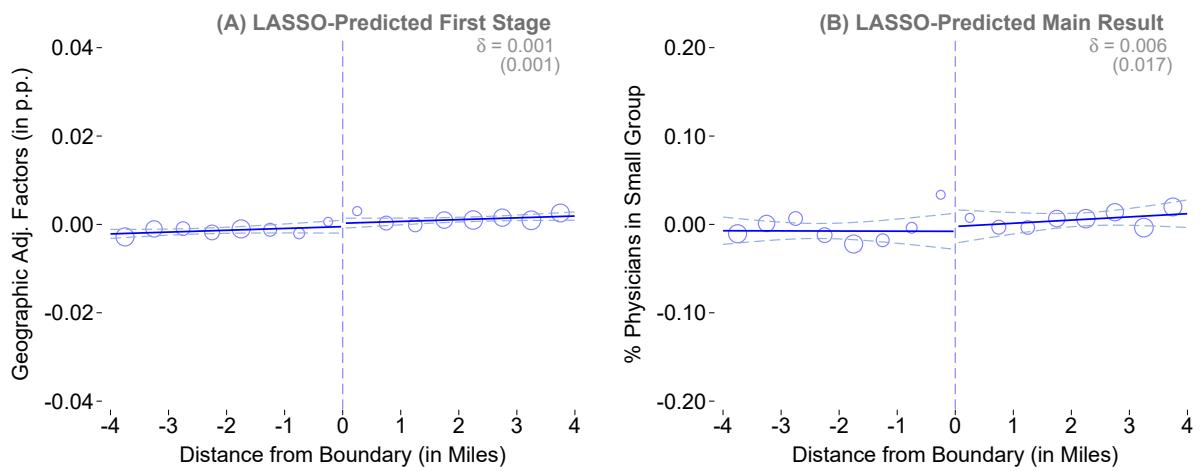


Figure 3.4: Validation of Research Design: Smoothness of LASSO-Predicted Outcomes

Notes: The purpose of this figure is to illustrate the smoothness of covariate-predicted outcomes through the cutoff. As further discussed in the text, each panel represents a LASSO-generated prediction using physician characteristics (and their interactions) to establish that observable characteristics are not predictive of either the treatment (Panel A) or main outcome (Panel B). Within the panels, each point represents the means of the outcome variable (displayed on the y -axis) in 0.5-mile increments, with the size of the point varying by the number of observations included. All means presented have been regression-adjusted to net out locality-border fixed effects. Each side of the cutoff—where negative distances represent low-payment sides and positive distances are associated with high-payment sides—are fit with linear regression lines and associated 90% confidence intervals. Estimated coefficients from Equation 3.8 along with associated standard errors are presented in the upper-left corner. This analysis utilizes the High-Impact Sample discussed in Section 3.2.2.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

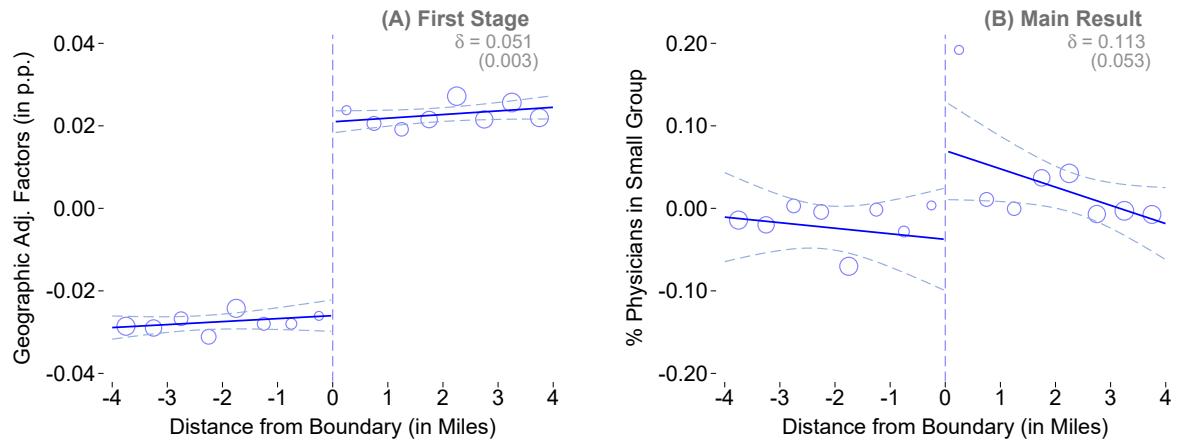


Figure 3.5: Main Results: Discontinuities at the Boundary

Notes: The purpose of this figure is to display the main results of my analysis: the first-stage impact of the border discontinuity (Panel A) and the second-stage impact on the percentage of primary care physicians practicing in small groups (i.e. those with 25 or fewer providers, displayed in Panel B). The coefficients and associated standard errors displayed in each panel are estimated using Equation 3.8. All means presented have been regression-adjusted to net out locality-border fixed effects. Each side of the cutoff—where negative distances represent low-payment sides and positive distances are associated with high-payment sides—are fit with linear regression lines and associated 90% confidence intervals. Estimated coefficients from Equation 3.8 along with associated standard errors are presented in the upper-left corner. Finally, for context, note that the pre-cutoff raw means corresponding to Panels A and B are 1.00 and 0.44, respectively. This analysis utilizes the High-Impact Sample discussed in Section 3.2.2.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, and CMS Physician and Other Supplier Public Use File.

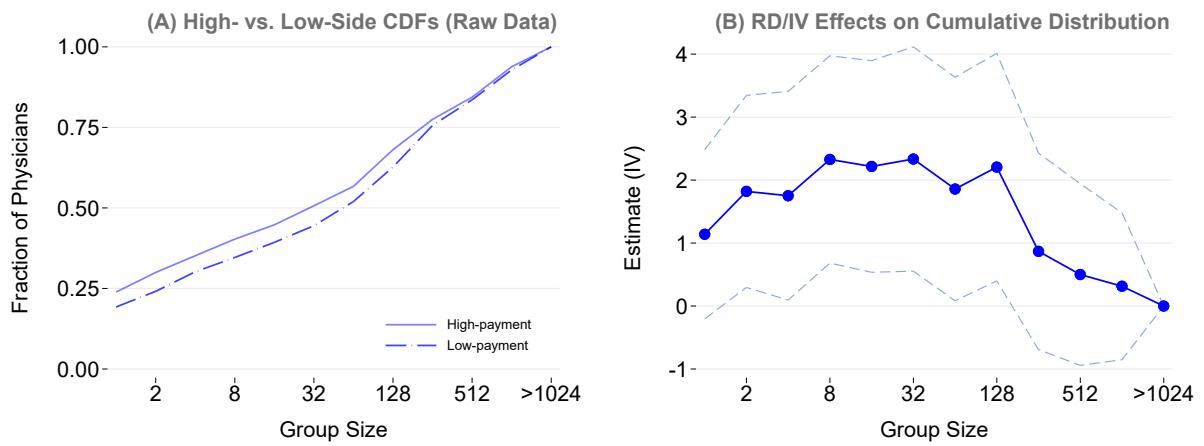


Figure 3.6: Distributional Effects on Group Size

Notes: The purpose of this figure is to display the distributional effects of differential Medicare reimbursement on group size. Panel A separately displays raw cumulative distribution functions of physicians by practice group size for high- and low-reimbursement sides of the border and within the bandwidth. Panel B displays point estimates—and associated 90% confidence intervals—for the RD-IV specification in separate regressions, where the outcome $Y_{ip} = 1(\text{GroupSize} \leq g)$ is generated for each value of group size g on the x -axis. This analysis utilizes the High-Impact Sample discussed in Section 3.2.2.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, and CMS Physician and Other Supplier Public Use File.

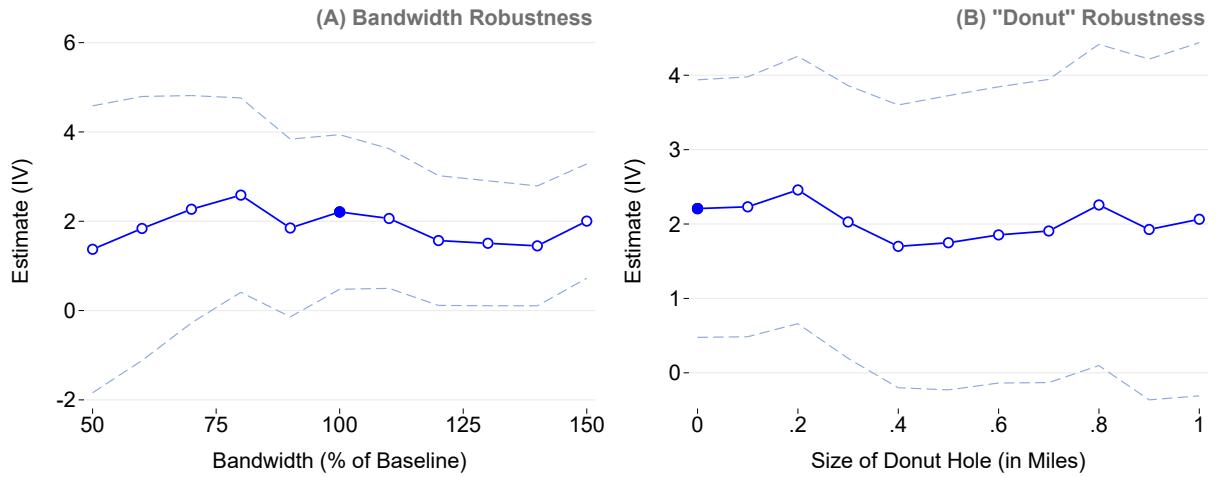


Figure 3.7: Robustness: Bandwidth Sensitivity

Notes: The purpose of this figure is to display robustness tests of the main result displayed in Table 3.1. Panel A displays point estimates for the RD-IV specification for separate regressions utilizing varying bandwidths (ranging from 50% to 150% of the baseline bandwidth, or 2 to 6 miles). Panel B shows estimates from the RD-IV specification for separate “donut” regressions on samples that omit observations within a given distance from the cutoff, ranging from 0 to 1 mile(s). In both panels, 90% confidence intervals are displayed using dashed lines and the preferred estimate is shaded dark blue. These analyses utilize the High-Impact Sample discussed in Section 3.2.2.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

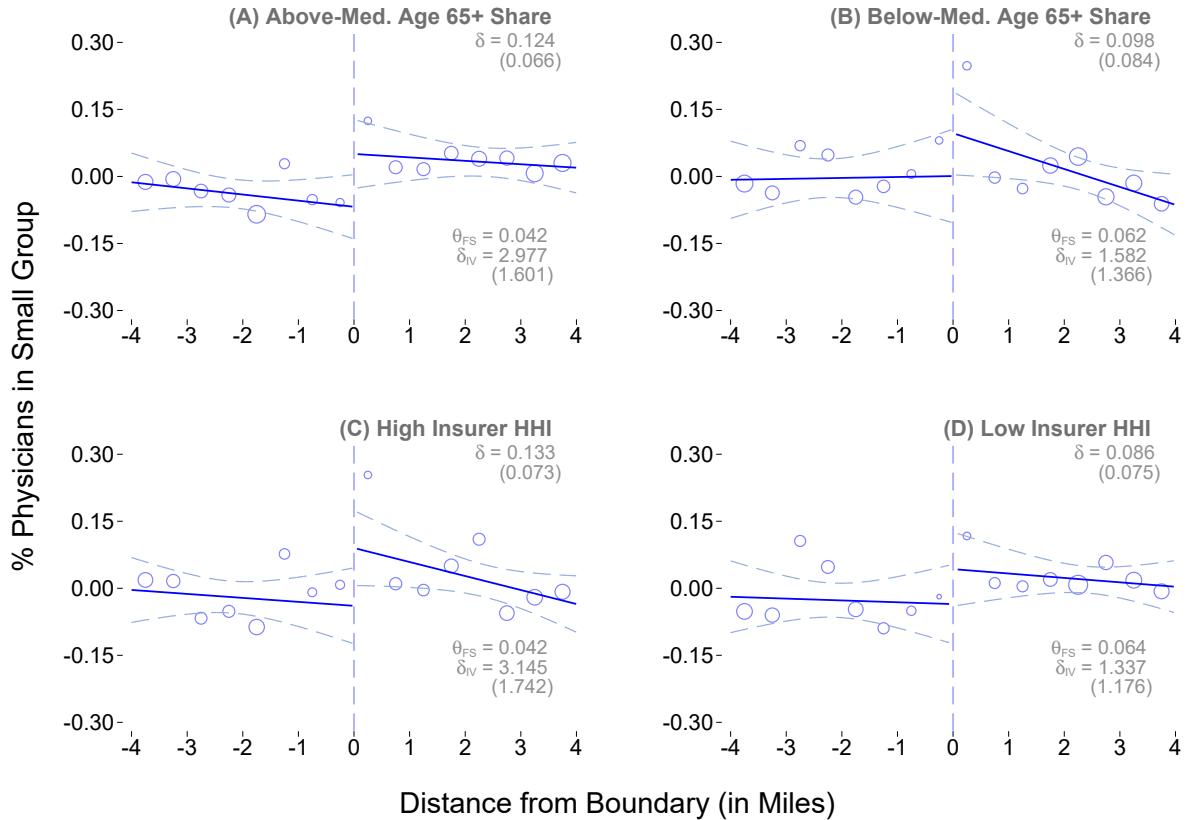


Figure 3.8: Heterogeneous Effects by Border Characteristics

Notes: The purpose of this figure is to display heterogeneity of results by physicians' local environment. Within each panel, reduced-form estimates (Equation 3.8) and associated standard errors are in the upper-right corner, while the first-stage (Equation 3.11) and RD-IV estimate (δ_{IV} , from Equation 3.10) and associated standard errors are in the bottom-right corner. The first set of estimates, displayed in Panels A and B, splits the sample according to borders that have above- and below-median shares of over Medicare-eligible adults, respectively. The second set of estimates, displayed in Panels C and D, splits the sample according to borders that have above- and below-median levels of health insurer market concentration, as measured by the Herfindahl-Hirschman Index ("HHI").

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File, Census Block Group 5-year ACS Estimates (Ruggles et al., 2019), American Medical Association Market Share Calculations.

Table 3.1: Main Results

Outcome	(1)	(2)	(3)	(4)
<i>Panel A: First Stage</i>				
Geographic Adjustment Factors	0.051 (0.003)	0.051 (0.003)	0.051 (0.002)	0.049 (0.002)
<i>Panel B: Reduced Form</i>				
% of Physicians in Small Groups	0.113 (0.053)	0.093 (0.051)	0.098 (0.052)	0.101 (0.051)
<i>Panel C: Instrumental Variables</i>				
% of Physicians in Small Groups	2.208 (1.052)	1.819 (1.009)	1.963 (1.026)	2.055 (1.046)
Controls:				
Distance controls	X	X	X	X
Border FE	X	X	X	
Provider characteristics		X	X	
Local controls			X	
Cluster FE				X
Dep. var. mean	0.456	0.456	0.456	0.456
Observations	5,065	5,065	5,065	5,065
First stage F-statistic	399	396	505	541
Bandwidth (in miles)	4	4	4	4

Notes: This table reports the results of the analysis for different specifications. Panel A describes the estimation of the first stage (i.e. when GAF_{ip} is an outcome) using Equation 3.8. Panel B estimates the effect of the discontinuity on physician practice structure using Equation 3.8, where the coefficients are interpreted as the reduced-form effect of crossing the payment boundary. Panel C displays coefficients when estimating the RD-IV strategy (Equations 3.10 and 3.11), wherein coefficients are interpreted as the effect (in p.p.) of a 1 percent increase in reimbursement. Columns 1 through 4 present the baseline specification and specifications with differing controls. All specifications utilize a rectangular kernel in distance from the boundary and are clustered at the Census-tract level. This analysis utilizes the High-Impact Sample discussed in Section 3.2.2. Significance at 10%, 5%, and 1% levels is represented by *, **, and ***, respectively.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File, Census Block Group 5-year ACS Estimates (Ruggles et al., 2019).

Table 3.2: Full-Sample and Alternate IV Results

Outcome	(1)	(2)	(3)	(4)
% of Physicians in Small Groups	2.208 ** (1.052)	1.339 (1.709)	1.703 ** (0.849)	1.673 ** (0.850)
Sample	High Impact	Full	High Impact	Full
Instrumental variable	RD-IV	RD-IV	RD-Border	RD-Border
Observations	5,065	9,685	5,065	9,685
First stage F-statistic	399	148	7,468	5,920

Notes: This table reports results for different samples and instrumental variable strategies. Columns 1 and 2 display coefficients and associated standard errors when estimating using the RD-IV (Equations 3.10 and 3.11), whereas Columns 3 and 4 display results for the RD-Border IV (Equations 3.12 and 3.13). Columns 1 and 3 utilize the High-Impact Sample discussed in Section 3.2.2 while Columns 2 and 4 use all available borders. All specifications utilize a rectangular kernel in distance from the boundary and are clustered at the Census-tract level. Significance at 10%, 5%, and 1% levels is represented by *, **, and ***, respectively.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

Appendices

Appendix A

Appendix to Chapters 1 and 2: The Effect of Early Life Wildfire Smoke Exposure

A Data Construction

A.1 Matching Fire Dates to Meteorologically Similar Days for HYSPLIT Analysis

As discussed in Section 1.3.1, the available dates for HYSPLIT smoke generation only extend back to 1948, which is in the middle of our sample period. Accordingly, for each fire date, we selected a proxy date that exhibited similar meteorological characteristics as the original day of the fire. Specifically, the following algorithm was performed:

1. Wind vectors— u -wind and v -wind—were obtained from the NOAA/CIRES/DOE 20th Century Reanalysis dataset (“NCDR data”), which covers the period from 1836-1980, and thus covers days from our original sample as well as days covered by HYSPLIT.¹ Specifically, we obtained data for 3-hour increments of surface wind as well as daily wind vectors for at differing heights above the surface (50 meters, 100 meters, 200 meters, and 500 meters). Days that match along these characteristics are likely to have similar plume dynamics, since both surface and higher-elevation winds determine particulate trajectories. The set of all wind vectors is denoted by V , where $v(d)$ is the value of an arbitrary wind vector on a given day d .

¹Wind vectors can be utilized to calculate both the angle at which the wind is blowing, as well as wind speed.

2. The NCDR data is on a 1-degree latitude by 1-degree longitude grid. To assign more weight to nearby fires, inverse distance weighting was utilized with observations further than 100 miles from the fire centroid receiving no weight. These weights are denoted by ω .
3. Finally, for some fire date f , the proxy date d^* is chosen from the set of dates between 1948 and 1980 (the dates of NCDR and HYSPLIT overlap, denoted by D) to minimize the distance-weighted Euclidean distance between the set of wind vectors (V):

$$d^* = \operatorname{argmin}_{d \in D} \left[\sum_{v \in V} \omega(v(f) - v(d))^2 \right]^{1/2}$$

4. In order to have consistency across the sample of fires, this procedure is performed for *all* fires, not just those occurring prior to 1948.

A.2 Place of Birth String Matching Algorithm

As discussed in Section 1.3.3, the vast majority (over 99%) were hand-matched to a GNIS location or had a perfect text match. Hand-matching was performed for any unique place of birth string that had more than 3 responses in the Public-version Numident or more than 10 responses in the restricted Numident. For the remaining observations, we used a bigram matching method to compare the place of birth string to both (a) names of GNIS locations and (b) POB strings that had already been hand-matched, where the best match was retained. Match similarity was calculated using Jaccard similarity scores; only matches with a score of at least 0.5 were retained. To further validate our algorithmic match, we utilized the place-of-birth-string-to-*county*-of-birth crosswalk from Bailey et al. (2020) and Isen et al. (2017b) (see those papers for more details). Then, using our hand-matched sample, we estimated a predicted probability of correct match based on the Jaccard similarity score and whether the county of the algorithmically matched GNIS location agreed with the county per the Bailey et al. (2020) and Isen et al. (2017b) crosswalk. Matches that had a predicted correct probability of 80% or more were retained in the final sample. As noted in Section 1.3.3, we matched 99.5% of individuals to a GNIS location.

A.3 Creation of Adjusted Numident Death Rates

Because the Numident omits potential deaths during the 1930s through mid-1970s, adjustment of the Numident death rates is necessary for our analysis of longevity. To create adjusted rates, the

following procedure was implemented:

1. Conditional death rates—e.g., the rate of death by age 75, conditional on surviving to age 14—were calculated using the SSA Cohort Life Tables. These rates are denoted as r_{bfsd}^{ssa} where b denotes the year of birth, f denotes a person’s sex, s denotes the “survived-to” age (14 in our example), and d notes the “death-by” age (75 in our example).
2. Similar rates were calculated for all individuals born in the United States from 1930-1969 that are included in the Numident (not just those born in California). For these cohorts, the age at which they received their SSN was used as the “survived-to” age, s . These rates are denoted by r_{bfsd}^{num} .
3. Since Numident includes a number of individuals with improperly missing deaths, we also utilized the 2000 and 2010 Decennial Full-Count Censuses and the 2005-2019 American Community Surveys to identify individuals that were (a) missing death information in the Numident and (b) did not appear in any of these Census products. Given that an individual should theoretically appear in each Decennial and has roughly a 22.5% chance of appearing in at least 1 American Community Survey, individuals who are not included are at a high risk of having an unrecorded death.² Let the measure of missingness for each cohort-sex-SSN age cell be denoted as η_{bfs} .
4. Then, using this national-level information, an “adjustment factor” to scale cohort death rates is as follows:

$$abfsd = \frac{r_{bfsd}^{ssa}}{r_{bfsd}^{num} + \eta_{bfs}}.$$

In words, this adjustment rate scales up the deaths per the Numident (plus the number of individuals likely missing death information) to match national birth cohort-by-sex-by-survival age-by-death age moments per the SSA cohort tables.

5. Finally, this rate is applied to death rates for given cells in our underlying data, such that the adjusted death rate at age d for cohorts in a given birthdate-city of birth-race-sex-SSN age cell will be:

²While it is possible that there are other reasons for exclusion (e.g., migration out of the United States, homelessness), omitted death is likely the most plausible explanation, particularly for older cohorts.

$$r_{bcfrsd}^{adj} = r_{bcfrsd}^{unadj} \times a_{bfsd}.$$

While these rates are used for the primary mortality analyses in this paper, as discussed in Section 1.6, use of the unadjusted dates results in very similar estimates of life years lost due to smoke exposure.

A.4 Data Construction for Control Variables and Heterogeneity Analysis

As discussed in Section 1.3.4, we utilize a variety of datasets to obtain control variables (used in Section 1.7) as well as variables for supplemental and heterogeneity analysis (Section 1.8). These additional datasets are described below:

Aggregate County-Level Data, 1930-1970. County-level data are obtained from three primary sources: (1) 1930 and 1940 Full-Count Decennial Censuses (Ruggles et al., 2021), (2) County Books from 1944-1970 (U.S. Census Bureau, 2012)—used to provide county averages for post-1940 outcomes—and (3) the 1934 Internal Revenue Service Statistics of Income (“IRS SOI”). We digitized the IRS SOI data to obtain information on family incomes, which is otherwise not available on the county level until the 1950 Census.³

American Hospital Association Data, 1947-1973 Hospital-level data are obtained from historical American Hospital Association data surveys which were transcribed by Finkelstein (2007).⁴ We standardized hospital names over time and also mapped city and state location strings to features in the GNIS database with associated geocoordinates. We then generated measures of hospital capacity by calculating the number of beds in general hospitals, pediatric hospitals, and labor and delivery hospitals within the 25 miles of an individual’s city and state of birth, which were divided by the number of births within 25 miles (as proxied by the number of individuals in the Numident file). Bed counts were taken from 1947, which is both the earliest year of data and roughly the mid-point for our analysis period.

Climate Data, 1930-1969. These data are obtained from three data sources made available by the NOAA: (1) the Global Historical Climatology Network Daily (“GHCND”) files, which provides daily temperature variables at the observation station-level; (2) the NOAA-CIRES-DOE Twentieth Century Reanalysis file, which provides daily gridded (1° latitude \times 1° longitude) wind

³The 1940 Census includes only wage income, which is top coded at \$5,000 (roughly \$78,000 in 2010 dollars).

⁴Many thanks to Amy Finkelstein for generously providing the data.

direction and speed information; and (3) the University of Delaware Precipitation file (Willmott and Matsuura, 2001), which provides monthly gridded (0.5° latitude \times 0.5° longitude) precipitation data. Station-level GHCND temperature data was aggregated to the county level by first creating tract-level temperature variables (using inverse-distance weighting to stations within 100 kilometers of the tract) and then aggregating population-weighted tract-level measures to the county level. Gridded wind and temperature data were aggregated to the county level via a simple average of all grid points within a county or within 1° of the county's border.

B Additional Results and Robustness

B.1 Infant Mortality

As discussed in Sections 1.6 and 1.7, we perform analyses on infant mortality to compare magnitudes against the reductions in adult longevity and to assess the degree of mortality selection. In the paragraphs that follow, we briefly discuss the data, methods, and regression results for our infant mortality analysis.

Data

Our primary data for the analysis of infant mortality come from birth and death indexes provided by the California Department of Public Health. The birth index provides full name, date of birth, and county of birth, while the death index provides full name, date of birth, date of death, and county. These data cover our entire sample period of 1930-1969.

To create infant mortality rates, we generated counts of births and deaths for each year-month birth cohort (defined based on date and county of birth) from the underlying vital statistics data. One challenge in using these data is that the death records did not include date of birth for deaths occurring prior to 1940. To adjust for this issue, we matched observations from the birth data using name and county information to construct cohort-level rates for these periods. These rates were then adjusted so that annual *period* infant mortality matched aggregated annual vital statistics tabulations from Bailey et al. (2018). However, our results are robust to omitting observations from the 1930s that were constructed in this manner.

Methods

We estimate cohort infant mortality using a method that is similar in spirit to our long-run cohort leads-and-lags estimation (Equation 1.2):

$$IMR_{Cb} = \mu_C + \lambda_{r(C),b} + \sum_{j=-18,-15,\dots,9} \delta_j Smoke_{C,b+j} + f(X_{Cb}) + \varepsilon_{Cb},$$

where the unit of analysis is a birth cohort born in a county C in year-month b . The treatment variable, $SmokeExposure_{C,b+j}$, and its associated leads and lags detail the number of months with any smoke exposure in three-month bins relative to birth. As in our long-term analyses, the coefficients on the leads and lags give us insight into the critical periods of pollution exposure

(for $j \in \{-9, -6, \dots, 9\}$) and also provide placebo estimates to support our research design (for $j \in \{-18, -15, -12\}$).

We include county-of-birth fixed effects (μ_C) and include region-of-birth \times birth cohort fixed effects ($\lambda_{r(C),b}$) where region of birth is defined as Northern or Southern California. Within our controls, $f(X_{Cb})$, we also include county by birth cohort trends to increase precision. All standard errors for this analysis are clustered at the county level.

Regression Results

The results of this analysis are displayed in Appendix Figure A.14. Within the figure, each point (with associated 95% confidence intervals) represents the marginal effect of an additional month of smoke exposure within a given three-month bin (e.g., the coefficient at “0” represents the marginal effect in the three months following birth). Based on the leads and lags coefficients, the critical period for exposure appears to be in the three months preceding and following birth (indicated by δ_{-3} and δ_0 , respectively). The average coefficient across these two bins, along with the standard error, is presented in the upper-right corner, while the mean is presented in the lower right for context. We find that an additional month smoke exposure in the last trimester or first three months of life results in 0.36 additional infant deaths per 1,000 births.

To facilitate comparisons between the results discussed in Section 1.6, we calculate the life years lost due to these additional infant deaths when moving from the 25th to 75th percentile of smoke exposure during the critical period. As discussed in the main text, moving from the 25th to 75th percentile of smoke exposure during the *in utero* through age 4 period results in 5 additional months of wildfire pollution exposure. A similar increase of smoke exposure would result in an increase of 0.43 smoke months during the much shorter critical period for infant mortality. Assuming that the typical infant death would have counterfactually lived to age 75 (roughly the life expectancy at birth for cohorts in our sample), then a $p_{25} \rightarrow p_{75}$ increase in exposure results in 11.7 life years lost per 1,000 persons. For comparison, a single month of exposure—which is close to the *average* level exposure within our sample (0.91 months)—results in 27.0 years of life lost per 1,000. See Section 1.6 for further discussion of how this compares to reductions in adult longevity.

Finally, in Section 1.6, we discuss the fact that our adult longevity estimates omit childhood mortality, since they are conditional on obtaining a Social Security Number. (The average age of SSN receipt was 14.4 in our sample.) Accordingly, Appendix Figure A.16 adds the infant mortality results into the coefficient estimates from Figure 1.4, Panel A to provide further context.

B.2 Comparison of Results with Existing Literature

In this section, we discuss how our findings compare to results from papers that examine the long-run effects of the Clean Air Amendments of 1970 (“CAA”), which substantially *reduced* total suspended particulate (“TSP”) emissions. These papers find that, in response to a reduction of $10 \mu\text{g}/\text{m}^3$ in TSPs, there was an increase of approximately 1.4% in wage earnings by ages 28-32 (Isen et al., 2017b) and a 0.7 to 0.9 to percentage point increase in high school completion (Voorheis, 2017). These papers measure effects primarily in TSPs, which do not translate directly to changes in fine particulate matter pollution ($\text{PM}_{2.5}$), which we measure in this paper, and which is generally seen as a better indicator of the potential of air pollution to harm human health.⁵ However, based on comparisons for monitors that have contemporaneous measurements of both TSPs and $\text{PM}_{2.5}$, we note that a $1 \mu\text{g}/\text{m}^3$ increase in $\text{PM}_{2.5}$ is associated with a 4-5 $\mu\text{g}/\text{m}^3$ increase in TSP. Accordingly, the estimates from the CAAA papers stem from a reported decrease that is equivalent to $\sim 2 \mu\text{g}/\text{m}^3$ of $\text{PM}_{2.5}$. Given that we find that a single month of wildfire exposure increases mean *monthly* $\text{PM}_{2.5}$ by $1.26 \mu\text{g}/\text{m}^3$ (with some effects in the following months), our stated first stage—even for the 5-month ($p_{25} \rightarrow p_{75}$) effects presented in most of the paper—is substantially smaller. Despite this, we find larger effects on earnings ($p_{25} \rightarrow p_{75} \approx 2\%$) than Isen et al. (2017b) but smaller effects on high school completion ($p_{25} \rightarrow p_{75} = 0.4\text{pp}$) than Voorheis (2017). However, as discussed below, while these studies provide compelling causal evidence of the reduced-form impact of the CAAA, comparing their instrumental variable effects (noted above) to the results that we find is challenging for several reasons.

The first set of challenges relates to contextual differences for these papers that may result in meaningfully different results. We begin by noting that the pollution sources that we study are compositionally different. While there is no direct evidence (absent this paper) that childhood wildfire pollution shocks result in worse long-term outcomes, there is a growing body of evidence focusing on contemporaneous outcomes that suggest wildfire pollution could cause larger adverse impacts than an equivalent amount of pollution from other sources. See Section 1.2 for further discussion. Next, as detailed in Section 2.8.1, our effects grow substantially over time: the effect for our oldest age group is three times as large (in percentage terms) as the youngest. If we had instead estimated effects using only our youngest age group, we would find percentage wage reductions closer to 1% for our $p_{25} \rightarrow p_{75}$ effects, which are smaller than those found by Isen et al. (2017b) (though they are still larger relative to our measured reductions in pollution). These age

⁵Measurements of $\text{PM}_{2.5}$ were not available for the study periods in these papers.

effects also help to explain why we find smaller effects on educational outcomes than Voorheis (2017) (since education is largely determined by age 30 and invariant thereafter), whereas we find larger effects than Isen et al. (2017b) (because we focus on later ages). It is also worth noting that, as discussed in Section 2.8.2, wildfire smoke appears to have different effects across the outcome distribution than the CAAA did. Whereas we find that childhood wildfire smoke exposure shifts the entire earnings distribution to the left, Isen et al. (2017b) finds that the left tail is compressed into the middle of the distribution.

The next set of challenges concerns the measured pollution effects of the CAAA and the excludability of the Clean Air Act Amendments as an instrument. First, as acknowledged by Isen et al. (2017b) and Voorheis (2017), the CAAA induced decreases in economic activity (Greenstone, 2002; Walker, 2013). While these economic shocks are widely regarded as small, they could meaningfully down-bias the long-run impacts of air pollution, especially if they are concentrated among vulnerable populations. Next, since the completion of Isen et al. (2017b) and Voorheis (2017), there have been several studies that critically examine the role of “gaming” pollution regulations. One such paper (Gibson, 2019) notes that pollution reductions from later Clean Air Act Amendments are limited to industrial activity around non-attainment monitors, suggesting that the average county-wide pollution reductions stemming from the CAAA of 1970 could be meaningfully lower than previously thought. Gibson (2019) additionally finds that regulation from later Clean Air Act Amendments led to increases in water pollution due to the substitution of firms to different pollution types. Given that industrial water pollution has been shown to negatively impact infant health (e.g., Flynn and Marcus, 2021), this will further offset the benefits found by the CAAA. Lastly, recent research has shown that the monitors readings themselves are subject to manipulation behavior, and so observed reductions in pollution may not reflect true improvements in air quality. For instance, Zou (2021) finds that firms displace their emissions from days where monitors are active to days in which monitoring does not occur. This displacement results in a 7% reduction in *measured* pollution (relative to baseline levels) for counties near regulatory thresholds. Additionally, Grainger et al. (2017) find that regulators in counties near or above regulatory thresholds strategically place new monitors in “clean” areas to reduce the average monitored pollution in these counties.

Taken together, these factors make it highly difficult to make reasonable comparisons between our findings in this paper and other papers which utilize Clean Air Act Amendment variation to identify long-run effects. Accordingly, further research is needed to understand the causal chain

that leads from early-life air pollution exposure in later life outcomes.

B.3 Bounding Mortality Selection

As discussed in Section 1.7, we implement a procedure to bound the impact of mortality selection on our estimates. While that exercise is described generally in the text, we discuss it in detail here. We begin by estimating age-of-exposure effects for death by age a using the techniques described in the text. Explicitly, for $a \in \{30...59\}$, we estimate:

$$M_{cb}^a = \mu_c + \lambda_{y(b),r(c)} + \sum_{k=-5}^5 \delta_k^a \text{Smoke}_{c,b+k} + \varepsilon_{cb},$$

where M^a is the cumulative mortality by age a , conditional on surviving past childhood.

Using the estimates recovered from the previous step, we predict the fraction of individuals that would have died by the age we observe at a given survey (r^a) for each cohort (defined on the city \times year-month of birth level):

$$r_{cb}^a = \sum_{k=-5}^5 \hat{\delta}_k^a \times \text{Smoke}_{c,b+k}$$

This exercise also incorporates the infant mortality estimates from Appendix Section B.1, which are calculated by multiplying the number of smoke months within the 3 months on either side of birth by the summary coefficient over that period (0.36 per 1,000).

We then use the death-by-age rate to construct an upper and lower bound. To do so, we create new cohort-level outcome measures by combining the observed average of the outcome (\bar{y}_{cba})—weighted by the observed individuals in that cohort (n_{cba})—with an extreme value (g) that is weighted by the estimated number of individuals who died before the survey date (calculated using r_{cb}^a):

$$y_{cba}^g = \frac{(\bar{y}_{cba} \times n_{cba}) + (g \times \frac{r_{cb}^a}{1-r_{cb}^a} \times n_{cba})}{n_{cba} \times \frac{1}{1-r_{cb}^a}}.$$

In the equation above, the term $\frac{r_{cb}^a}{1-r_{cb}^a} \times n_{cba}$ captures the number of individuals who are predicted to have died before the survey, while the denominator ($n_{cba} \times \frac{1}{1-r_{cb}^a}$) equals the number of individuals who are observed, plus those who have died. The value of g that is selected is equal to either the 99th in-sample percentile of the outcome distribution (for the upper bound) or the 1st in-sample

percentile (for the lower bound).

Finally, using these revised outcome measures, we re-estimate the following equation for the upper and lower values of g to obtain our bounded estimates:

$$y_{cba}^g = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \sum_{k=-5}^5 \delta_k Smoke_{c,b+k} + \varepsilon_{cba}.$$

The results of this procedure are presented in Appendix Figure A.21. As displayed by the figure, the bounded estimates are highly similar to our original estimates.⁶ We further note that, of the two bounds presented in the figure, the lower bound is likely more representative of the “true” estimate, because individuals who are in poor health (and thus are more likely to die before we can observe them) also tend to have poor outcomes in other areas as well.

B.4 Migration-Related Measurement Error

We perform exercises to assess the impact of measurement error on our estimates, particularly those in the leads-and-lags specification detailed by Equation 1.2. While this procedure is discussed generally in Section 1.7, we describe it in greater detail here. We begin by utilizing the 1940 Census 100% Sample. These data are used because they ask the entire population what county they lived in on April 1, 1935, and thus allow us to accurately assign both county of birth and county of residence—and their corresponding smoke exposure measures—for 5-year-olds in our sample.⁷ Additionally, the sample size allows us to restrict to children aged 5 that were also born in California.

Using this sample of children, we then simulate smoke exposure by fixing a child’s birth place and residence at age 5 but assigning smoke as-if that child was born in different periods in our sample. (This exercise implicitly assumes that the migratory behavior of children in 1940 is applicable to all cohorts.) Using the measures of smoke based on place of birth and place of age 5 residence, we then construct “true” measures at each age, where the correct measure of smoke exposure in the *in utero* period and the first year of life (age 0) is equal to the county-of-birth measure. For other ages (k), we calculate the “true” smoke measure as:

⁶A careful review of the figure will show that, while the upper and lower bounds for our economic and education outcomes are generally symmetrical around the baseline estimate, this is not the case for the disability index. This is because the best possible outcome for the disability index is to have no disabilities at all, which is the case for 84% of our sample. Thus, the upper bound is essentially identical to the baseline and only the lower bound differs.

⁷Children moving out of state were assumed to have no smoke exposure.

$$TrueSmoke_k = \frac{[Smoke_{birth} \times (5 - k)] + [Smoke_{res} \times k]}{5} \quad \text{for } k \leq 5,$$

where $Smoke_{birth}$ and $Smoke_{res}$ represent smoke measures based on an individual's county of birth and residence at age 5, respectively. Note that this exercise that effectively assumes equal likelihood of migration in any given year. Using this true smoke measure, we can calculate the amount of bias under the assumption that the measurement error is classical—an assumption that we revisit below.

For each age, we calculate the measurement error as $u_k = Smoke_{birth} - TrueSmoke_k$ to represent the difference in exposure calculated based on place of birth versus our best estimate of the true exposure calculated above. We then compute the variance of the measurement error ($\sigma_{u,k}^2$) and the “true” age-specific smoke measure ($\sigma_{smoke,k}^2$) to calculate the measurement-error adjusted estimate for each exposure age:

$$\delta_k^{adj} = \frac{\delta_k}{1 - \frac{\sigma_u^2}{\sigma_{smoke,k}^2 + \sigma_u^2}}.$$

As displayed in Appendix Figure A.22 these adjusted estimates are almost identical to our baseline estimates. We conclude this section by briefly revisiting our assumption of classical measurement error. While we can't validate all the conditions necessary for this assumption, we do note that, within our data, $\frac{1}{n} \sum_i u_{k,i} \approx 0$ and $\frac{1}{n} \sum_i TrueSmoke_{k,i} \times u_{k,i} \approx 0$. Accordingly, the use of classical measurement error assumptions appears to be appropriate.

B.5 Other Supplemental Results

In this appendix section, we discuss further supplemental analyses that were only briefly referenced in the text. Specifically, we consider: (1) the degree to which our main results—particular our results concerning mortality—could be explained instead by local economic conditions; and (2) the degree to which there is treatment effect heterogeneity across individual- and place-specific characteristics.

The Effects of Wildfire Smoke Do Not Appear to Be Primarily Transmitted through Local Economic Shocks

As demonstrated by Borgschulte et al. (2020), wildfire smoke has the ability to affect labor market conditions, particularly for older workers. Accordingly, it is possible that the impacts of wildfire smoke on later life outcomes are due, in part, to degraded local economic conditions which could then lead to lower levels of childhood investment. To investigate this possibility, we utilize county-level data from 1930-1970 Decennial Censuses and 1934 IRS Statements of Income Data (see Appendix Section A.4) to estimate the following specification:

$$\Delta Y_{Ct} = \alpha + \tau \cdot Smoke_C^{t',t} + \lambda_t + \varepsilon_{Ct}, \quad (B1)$$

where the unit of analysis is a county, C , observed at time t . The outcome variable is a long-difference (typically of a decade) between time t and the previous period, t' . The use of a long-difference specification effectively nets out any time-invariant characteristics, while the time fixed effect, λ_t , controls for changes common to all areas of California. The treatment variable, $Smoke_C^{t',t}$ represents the number of months with any smoke between time t' and t . In addition to our interest in local economic conditions, we also utilize these data to assess the impact of wildfire air pollution on net migration, which we define as:

$$Migrate_{Ct} = \frac{Pop_{Ct} - Pop_{Ct'} - Births_C^{t',t} + Deaths_C^{t',t}}{Pop_{Ct'}}$$

The results from this equation are displayed in Appendix Table A.4. Within the table, which presents the marginal effect of an additional smoke month during the period, we do not find any statistically significant results. We also note that results are economically small, suggesting that contemporaneous economic do not play a large role in the long-run effects we find and do not induce migratory behavior. However, to provide further context to these results, we relate our (statistically insignificant) effect on family income to estimates from Aizer et al. (2016), who estimate the impact of large ($\sim 25\%$) increases in family income on children's longevity. To do so, we calculate the $p_{25} \rightarrow p_{75}$ effect on income as a percent of the underlying mean (0.4%) and use this estimate to scale the Aizer et al. (2016) finding of a 1-year increase in life expectancy. This procedure yields an estimate of approximately 16 life years lost per thousand individuals ($\frac{0.4\%}{25\%} \times 1 \text{ life year} \times 1000$), or 35% of our estimate. We should further note that this figure is likely an upper bound of the degree to which income shocks could affect longevity, because the effects that

Aizer et al. (2016) estimate are within the context of very poor mothers with dependent children. Accordingly, the returns on income they find are likely much larger than would be true for the population study.⁸

The Effects of Wildfire Smoke Have Little Heterogeneity Across Groups

Next, we consider how the effects of wildfire smoke varies across groups. Specifically, for each group, we estimate:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \delta_0 \cdot \text{Smoke}_c^{\text{IU-Age } 4} + \delta_1 \cdot [\text{Smoke}_c^{\text{IU-Age } 4} \times \mathbf{I}(g(i) = g)] + \varepsilon_{cba}, \quad (\text{B2})$$

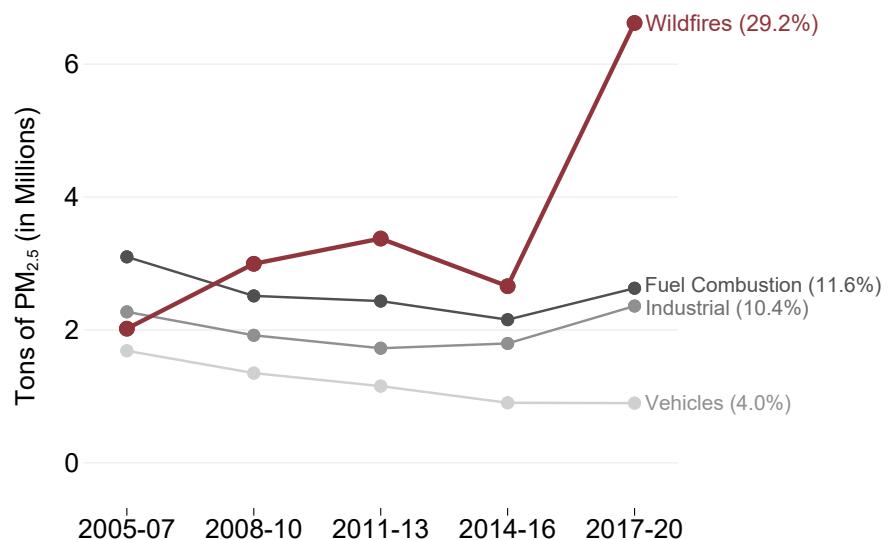
where the indicator variable, $\mathbf{I}(g(i) = g)$, is equal to one if an individual is in the interaction group and zero otherwise. Accordingly, the coefficient on the interaction term, δ_1 , captures the additional effect on a month of wildfire pollution relative to the omitted group.

We present results from this equation in Appendix Figure A.26. Within the figure, there are two sets of estimates presented for each regression, the baseline effect (δ_0) and the *combined* effect ($\delta_0 + \delta_1$) which captures the total effect on the interaction category. Both the baseline effect and combined effect are presented with their 95% confidence intervals. Additionally, estimate pairs that are statistically distinct from each other—i.e., where the interaction term, δ_1 , achieves statistical significance—are denoted with one or more “*” indicating the level of that significance.

As demonstrated by the figure, there is very little heterogeneity among groups: only three of the twenty-four regressions have a statistically significant effect. Accordingly, we conclude that wildfire smoke generally has homogeneous effects across a wide variety of groups. See Appendix Section A.4 for further discussion on how these heterogeneity groups were constructed.

⁸However, we utilize these estimates because they are, to the best of our knowledge, the only causal estimates of an early-life income shock on longevity.

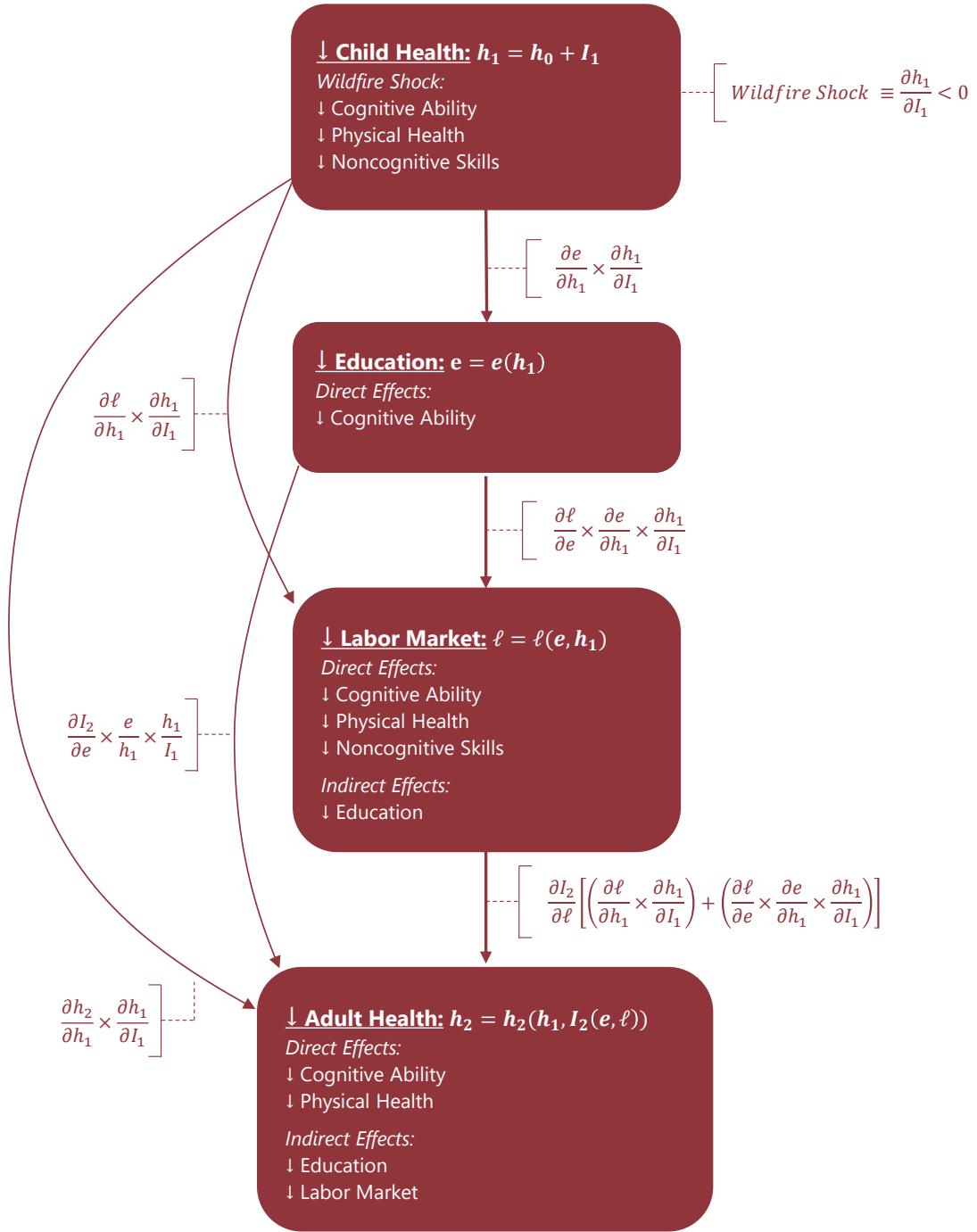
Figure A.1: Motivating Figure: Wildfires are a Growing Source of PM_{2.5} Pollution



Notes: The purpose of this figure is to detail trends in the share of fine particulate matter (PM_{2.5}) attributable to wildfires and other major sources tracked by the Environmental Protection Agency (“EPA”). The connected points represent trends in total tons of pollution emitted while the percentages to the right of the graph represent the share of PM_{2.5} attributable to each source in the most recent period. Percentages do not add up to 100% due to omission of the EPA’s “miscellaneous” category. Years are grouped into bins because the EPA only includes certain pollutant measures as multi-year averages within these data.

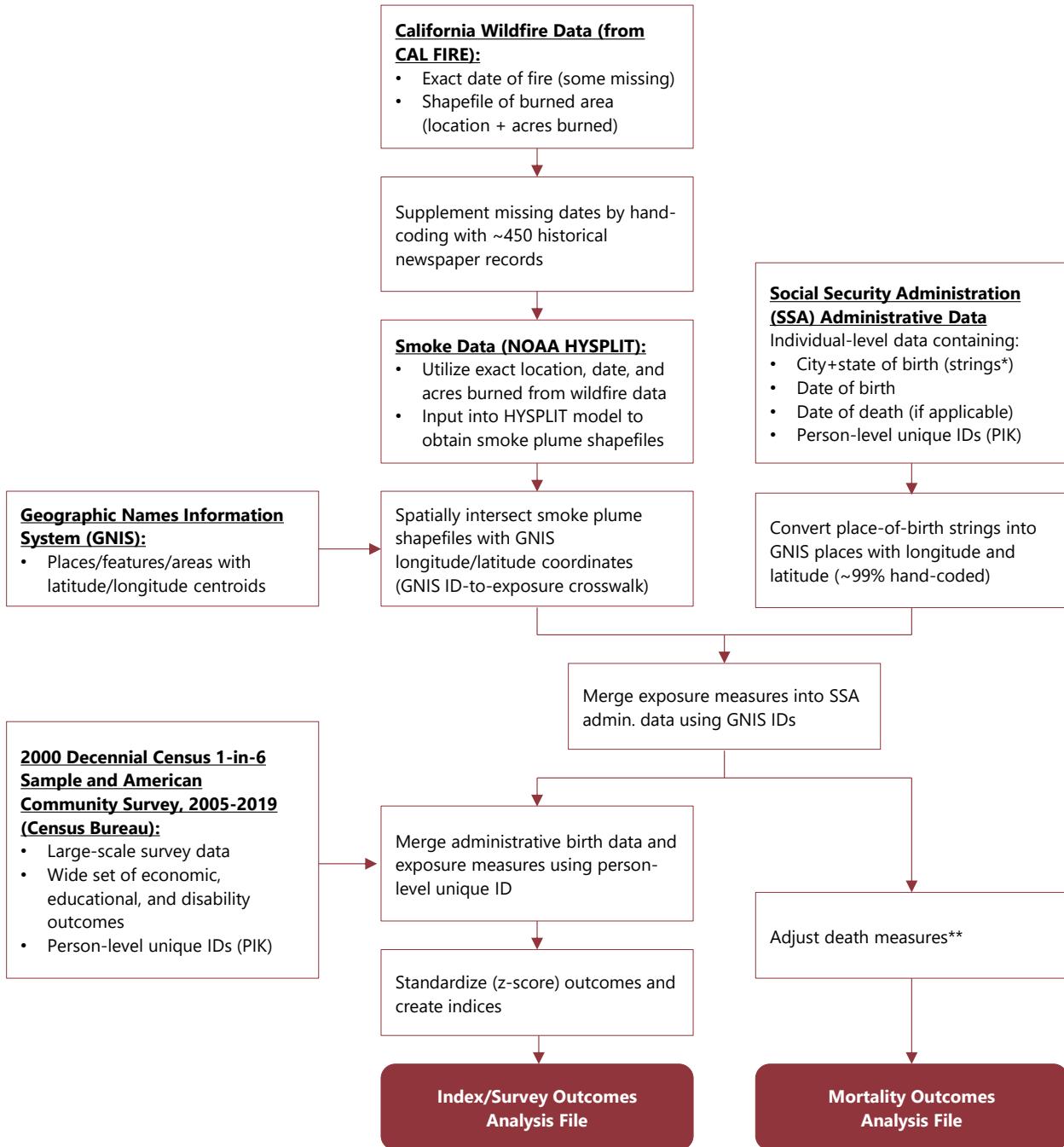
Source: Author calculations using National Emissions Inventory Data from the Environmental Protection Agency.

Figure A.2: Illustration of Conceptual Framework: Wildfire Smoke Could Affect Long-Run Health through Multiple Channels



Notes: The purpose of this figure is to illustrate the channels by which wildfire smoke exposure during childhood could affect long-run outcomes. Each box in the figure represents a different part of the life cycle, as discussed in Sections 1.2 and 2.2, with the text inside the box describing the ways in which childhood pollution could affect outcomes during that stage. The red text associated with each arrow (and the first box) represent the applicable components of the mathematical model described in Sections 1.2 and 2.2—i.e., they describe how the effects are transmitted from one stage to the next.

Figure A.3: Graphical Illustration of Data Construction for Primary Data Sets

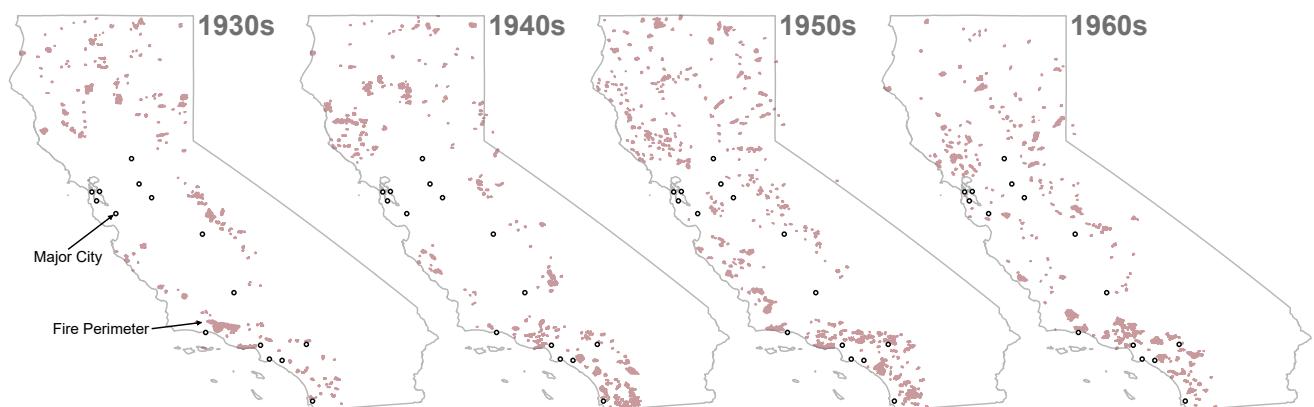


Notes: The purpose of this figure is to illustrate the process by which final datasets constructed, which is further discussed in Sections 1.3 and 2.3. Note that this flow chart is not intended to account for all data sources constructed in this paper, only the primary analysis data sets.

* As discussed in Sections 1.3.3 and 2.3.3, the place of birth data in the administrative SSA data (the Numident file) has not been standardized and is instead presented as a string variable entailing what was originally written on an individual's Social Security application (e.g., an individual born in San Francisco, California might write their city of birth as "San Francisco," "San Fran," or simply "SF"). Accordingly, it was necessary to convert into standardized GNIS places or features.

** The SSA Numident is a nearly comprehensive source of deaths occurring after the mid-1970s but may be missing deaths from earlier years. As further discussed in Sections 1.3.3 and 2.3.3 and Appendix Section A, we therefore adjusted death rates to account for Numident's propensity to undercount these early deaths.

Figure A.4: Descriptive Statistics: Location of Fires



Notes: The purpose of this figure is to display spatial trends in wildfire occurrence. Wildfires occur across the state, although the exact location where they occur varies substantially from year to year.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018).

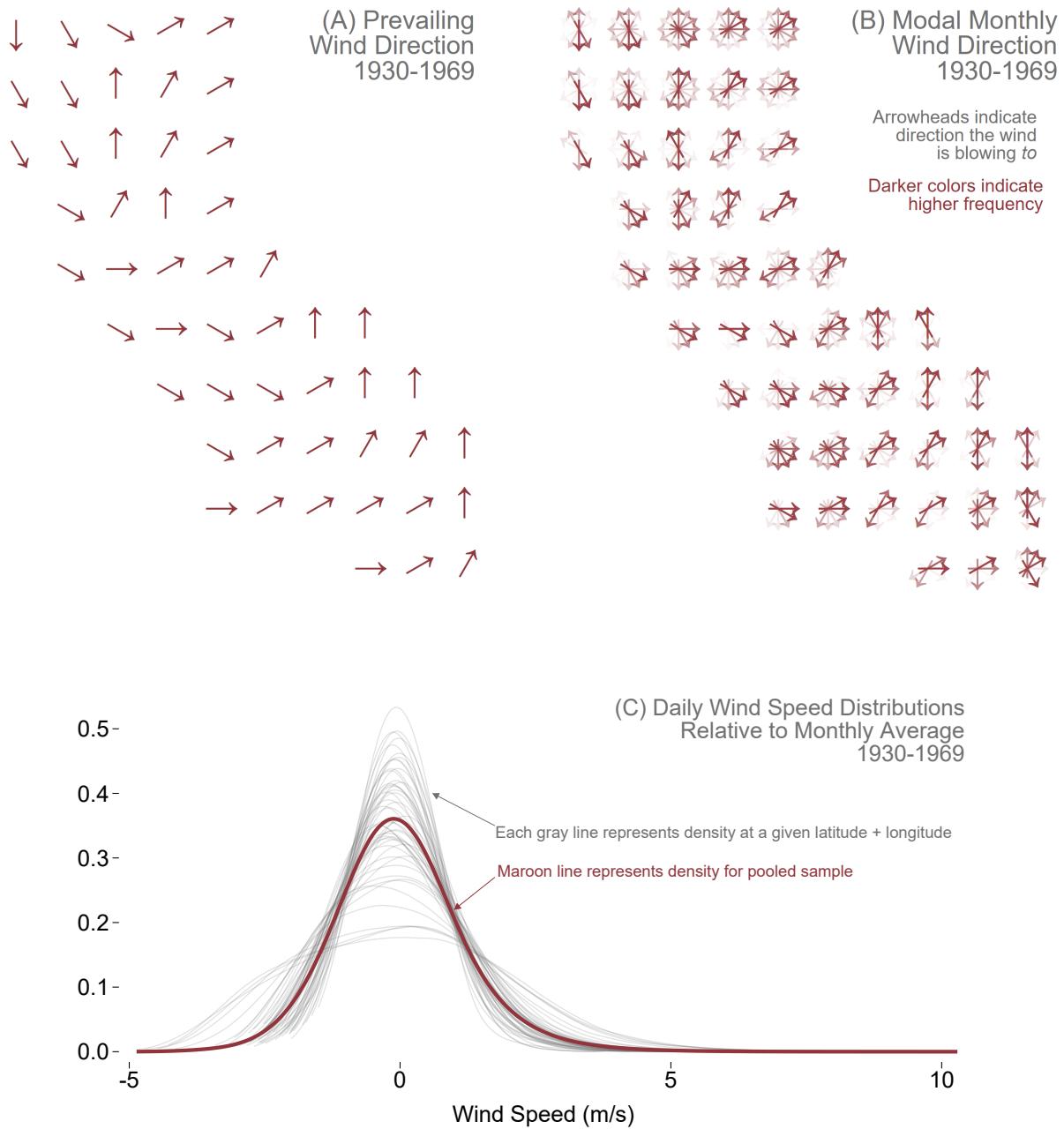
Figure A.5: Illustration of Modeled Smoke Plumes Under Different Conditions



Notes: The purpose of this figure is to illustrate different smoke plumes generated by HYSPLIT under different conditions. Panel A illustrates a smoke plume from a hypothetical large (10,000-acre) fire on a day when there are strong winds (around the 90th percentile of wind speed). Under such conditions, there is a substantial amount of smoke generated, and it can travel for several hundred miles. Panel B illustrates a smoke plume from an alternate small (1,000-acre) fire in the same location and under the same wind conditions. In contrast with Panel A, there is a substantially lower amount of smoke predicted. Finally, Panel C illustrates a large fire under calmer wind conditions (around the 10th percentile of wind speed). While there is a substantial amount of smoke generated, as is the case in Panel A, it is spatially distributed in a diffuse mass surrounding the fire. These figures underscore the importance of incorporating both fire size and wind speed into the modeling of wildfire pollution.

Source: Author simulations using HYSPLIT model output (Stein et al., 2015).

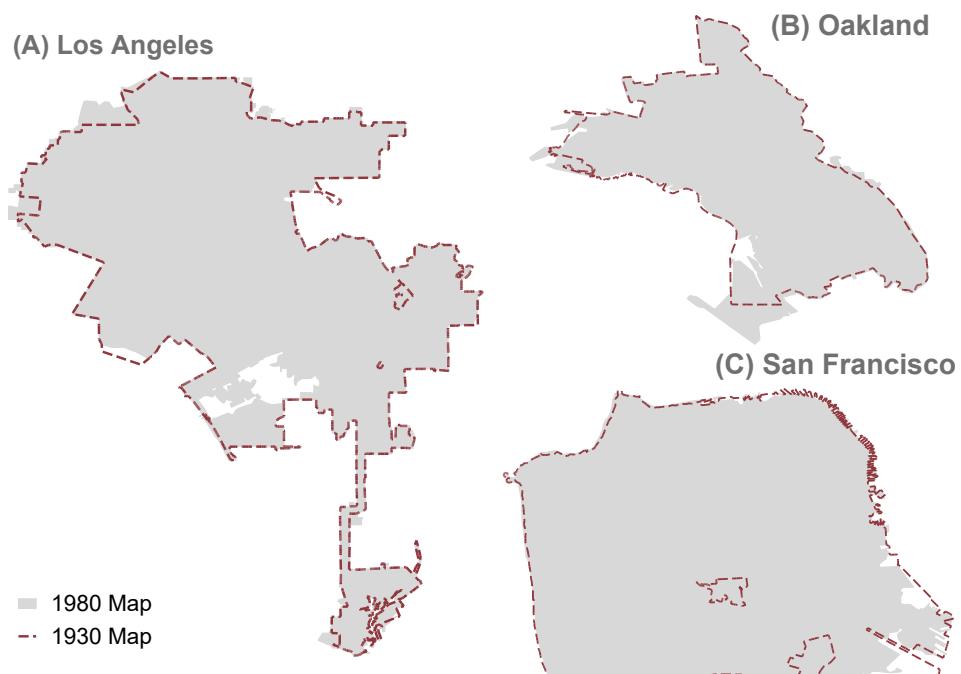
Figure A.6: Descriptive Statistics: Changes in Wind Direction and Intensity Provide Plausibly Exogenous Variation



Notes: The purpose of this figure is to illustrate the variation in wind direction and speed. Within the figure, Panel A displays the prevailing wind direction from 1930-1969 (i.e., the modal wind direction calculated over the entire period) for various points of longitude and latitude in California (state boundary lines have been omitted for visual clarity). Panel B performs a similar exercise, except that the modal direction is calculated for each year-month and plotted in the figure using semi-transparent arrows, so that darker arrows indicate more frequent modal directions. Panel C presents the kernel densities of wind speeds for California (maroon line) and for each of the longitude/latitude points plotted in Panels A and B (semi-transparent gray lines). All speeds are relative to the average speed at that point in that month (calculated over the 1930-1969 period) to emphasize variation in average wind speeds. All panels restrict observations to July through November, as that is when the overwhelming majority of fires occur (and therefore when wind direction is relevant) – see Appendix Figure A.12 for more detail.

Source: Author calculations using the NOAA-CIRES-DOE Twentieth Century Reanalysis File.

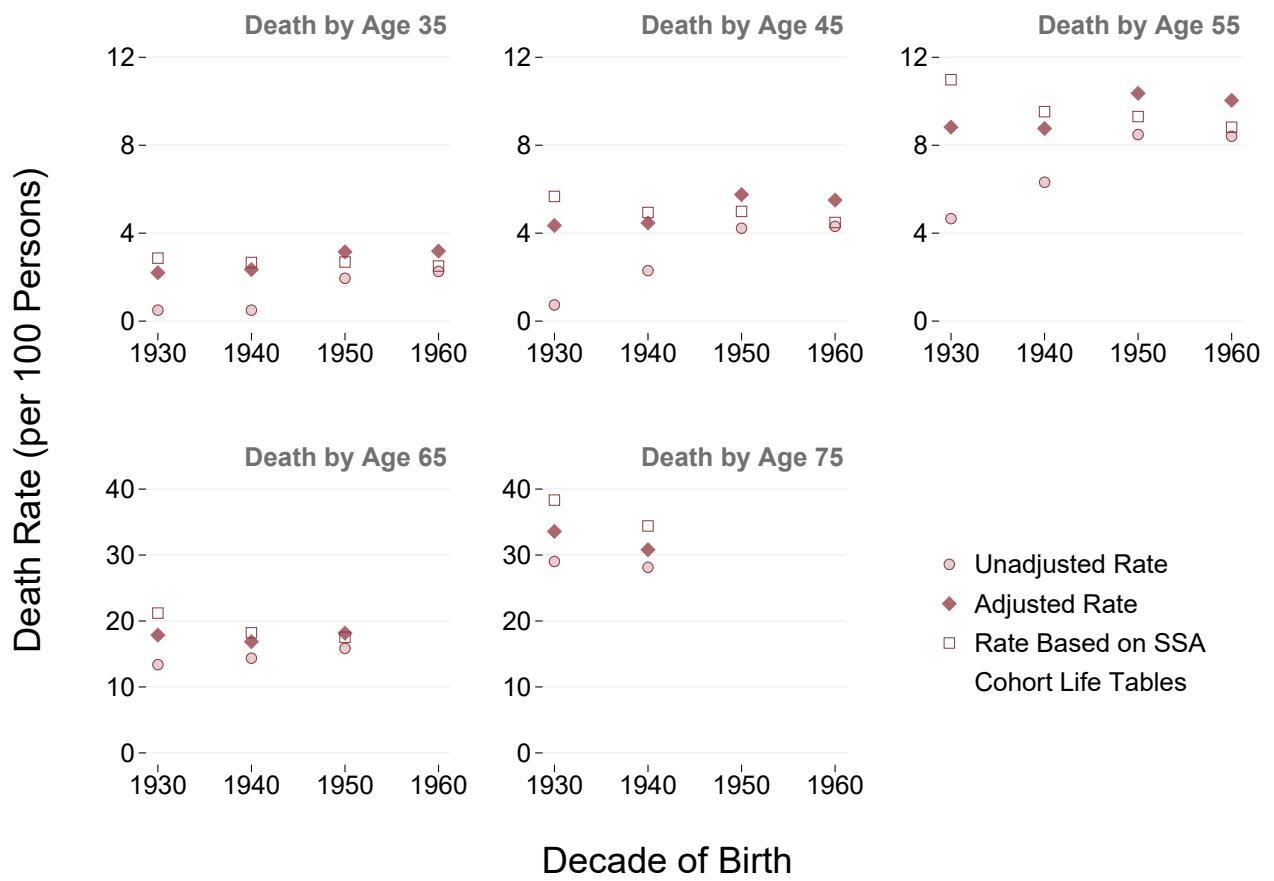
Figure A.7: Major City Borders Are Stable Over Time, Suggesting Little Risk of Measurement Error from Changing Longitude / Latitude Centroids



Notes: The purpose of this figure is to provide evidence that borders for major cities are not changing substantially over time. Given that the longitude and latitude of each birth place is assigned based on city/place centroids, substantially changing borders over time could introduce time-varying measurement error via centroids that are inaccurate for certain birth cohorts. However, this concern is mitigated by the fact that these city borders—presented for all California cities with available GIS data dating back to 1930—are highly similar. (Note that 1980 was chosen as a comparison year as it is the closest year to the end of our sample period for which there was place shapefile data). Instead of changing borders of principal cities, as these major metropolitan areas grew, they simply created new cities and municipalities with their own associated centroids. Accordingly, time-varying measurement error due to changing centroids is likely not an issue.

Source: Author calculations using shapefile data obtained from the Urban Transition Historical GIS Project (Logan and Zhang, 2020) and IPUMS NHGIS (Manson et al., 2021).

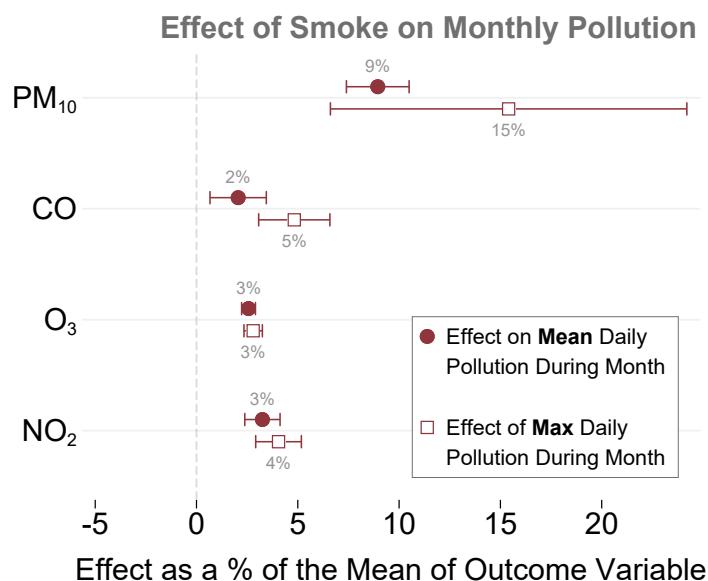
Figure A.8: Descriptive Statistics: Comparison of Adjusted and Unadjusted Death Rates per Numident to Expected Cohort Death Rates



Notes: The purpose of this figure is to contrast how unadjusted and adjusted death rates per the Numident file compare with death rates from SSA Cohort Life Tables. As displayed in the figure, adjusted death rates are much closer to expected values and are more consistent over different cohorts. All values less than 0.5 are bottom coded for disclosure avoidance purposes. See Section 1.3.3 and Appendix Section A for further discussion.

Source: Author calculations using Restricted Census Numident file and SSA Cohort Life Tables.

Figure A.9: Validation of Treatment: Modeled Wildfire Smoke Exposure Increases Monthly Measures of Various Pollutants



Notes: The purpose of this figure is to display the impact of modeled smoke plume coverage on monthly pollution levels across a range of pollutants (described on the *y*-axis). Each point represents the β_0 coefficient from the regression detailed by Equations 1.4 and 2.4 which has been scaled by the mean of the given pollutant for comparability across different pollution types. Each scaled estimate is presented with corresponding 95% confident intervals.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), and Daily Summary Data Files from the EPA Data Mart.

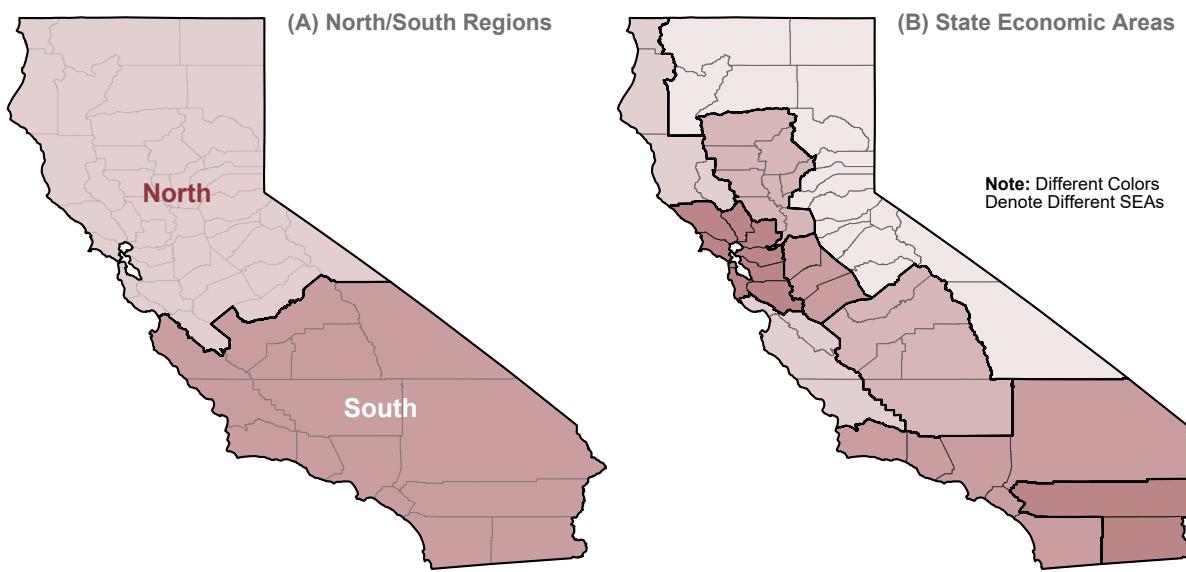
Figure A.10: Illustration of Treatment Definition



Notes: The purpose of this figure is to illustrate how treatment is assigned. A hypothetical/modeled smoke plume is presented in the figure. The cities that are considered to be treated—those that spatially intersect with the smoke plume—are colored maroon, while the unexposed cities—those that are outside of the smoke plume—are colored in black/gray.

Source: Author simulations using HYSPLIT model output (Stein et al., 2015).

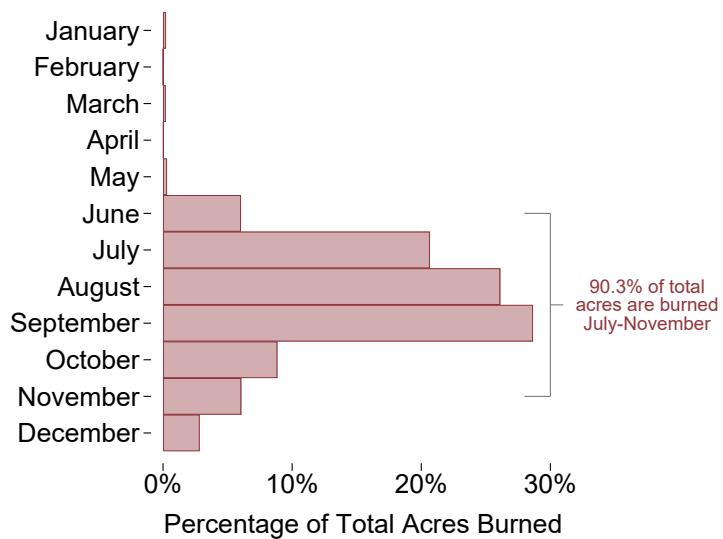
Figure A.11: Illustration of Regional Definitions



Notes: The purpose of this figure is to display different regional definitions used in our analysis. As discussed in Sections 1.5 and 2.5 our core specification utilizes year-of-birth by region (Northern and Southern California) fixed effects. The counties included in Northern and Southern California are included in Panel A. In our robustness checks (Section 1.7 and 2.7), we instead utilize year-of-birth by State-Economic-Area fixed effects. The counties included in the different State Economic Areas are included in Panel B.

Source: Author calculations using data from the U.S. Census Bureau and Ruggles et al. (2021).

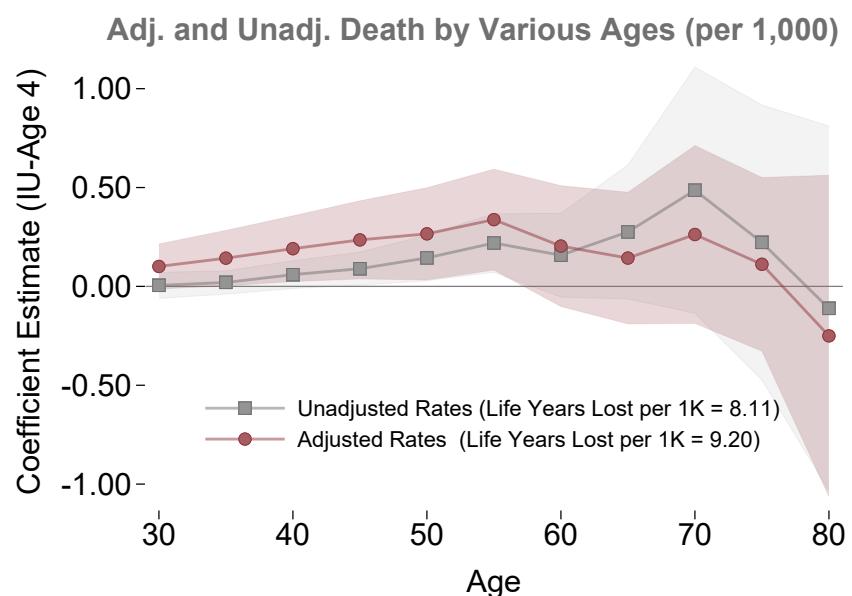
Figure A.12: Descriptive Statistics: Fire Seasonality



Notes: The purpose of this figure is to display seasonality in wildfire acres burned.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018).

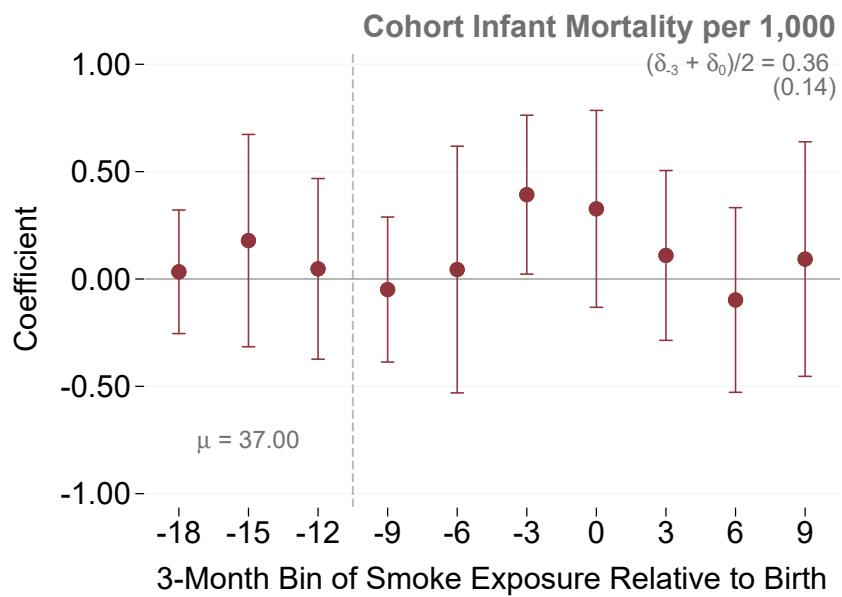
Figure A.13: Robustness: Unadjusted and Adjusted Mortality Measures Display Similar Effects



Notes: The purpose of this figure is to show similarities in death-by-age results when using unadjusted and adjusted measures. See Figure 1.4 and accompanying text for further information on figure construction and discussion of results.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), and Restricted Census Numident.

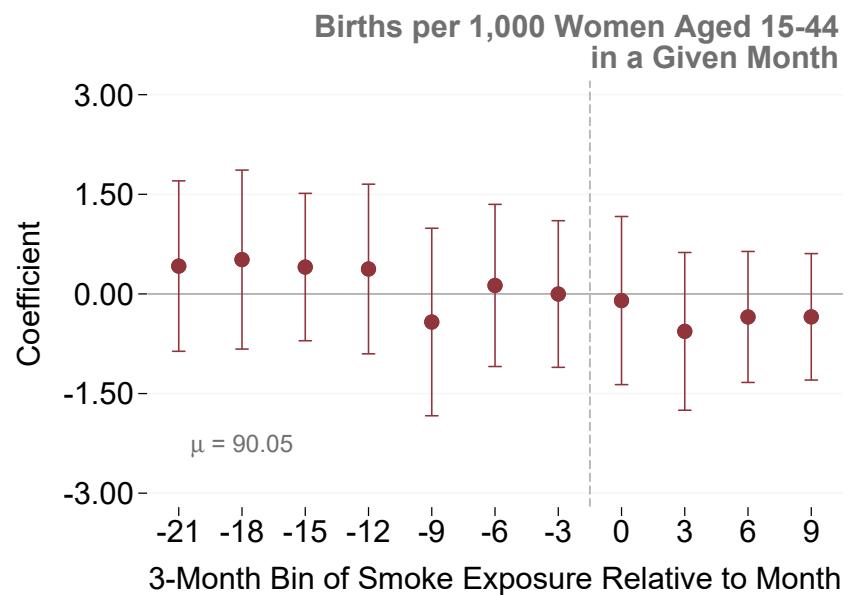
Figure A.14: Additional Results: Wildfire Smoke Exposure Increases Infant Mortality



Notes: The purpose of this figure is to assess the impact of wildfire smoke on cohort infant mortality. Within the figure, each coefficient displays the effect of an additional smoke month within three-month bins, along with 95% confidence intervals. In the upper right-hand corner, the average effect during the critical period—considered to be the three months before (δ_{-3}) and after (δ_0) birth—is presented alongside its standard error. The mean is presented in the lower left-hand corner for reference.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), California Department of Health Vital Statistics, and Natality/Mortality data from Bailey et al. (2018).

Figure A.15: Additional Results: There is No Detectable Effect of Wildfire Smoke Exposure on Fertility Timing or Overall Fertility



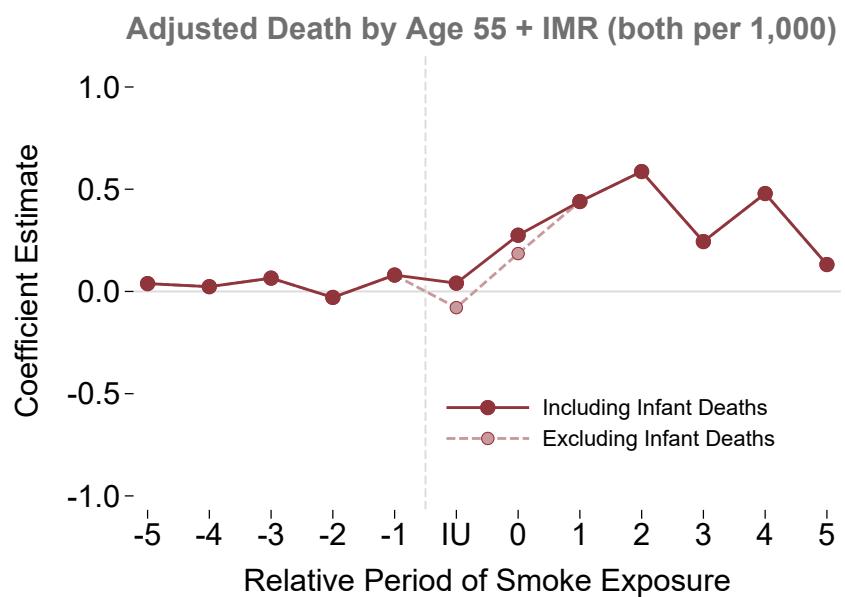
Notes: The purpose of this figure is to estimate the effect of wildfire smoke exposure on fertility, as measured by births per 1,000 women aged 15-44. To do so, we estimate the following equation:

$$BirthsPer1000_{Cb} = \mu_C + \lambda_{r(C),b} + \sum_{j=-21, -18, \dots, 9} \delta_j Smoke_{C,b+j} + \varepsilon_{Cb},$$

which is very similar to the equation used to estimate infant mortality in Appendix A, Section B.1 (see text of that section for more details). Within the figure, coefficients from this equation are plotted with accompanying 95% confidence intervals. The mean of the outcome variable is presented in the bottom-left corner for reference. As displayed by the figure, there do not appear to be any statistically-detectable effects on fertility timing or overall fertility.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), California Department of Health Vital Statistics, and Natality/Mortality data from Bailey et al. (2018).

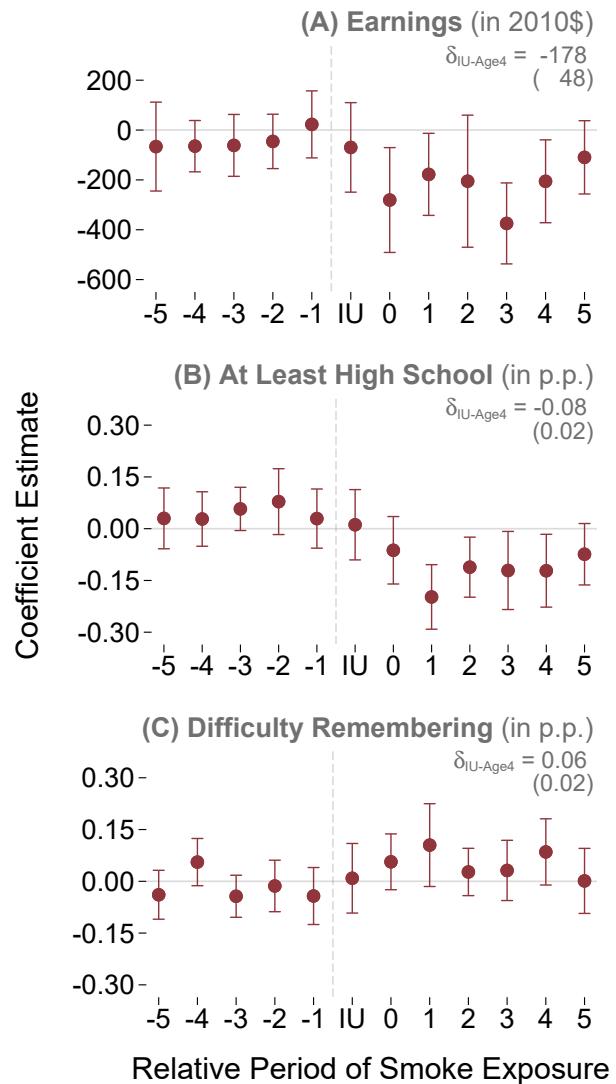
Figure A.16: Additional Results: Reductions in Longevity with Infant Mortality Added



Notes: The purpose of this figure is to display the coefficients from Figure 1.4 Panel A with infant mortality results (discussed in Appendix A, Section B.1) added in. Confidence intervals are omitted from this figure for clarity of presentation.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, California Department of Health Vital Statistics, and Natality/Mortality data from Bailey et al. (2018).

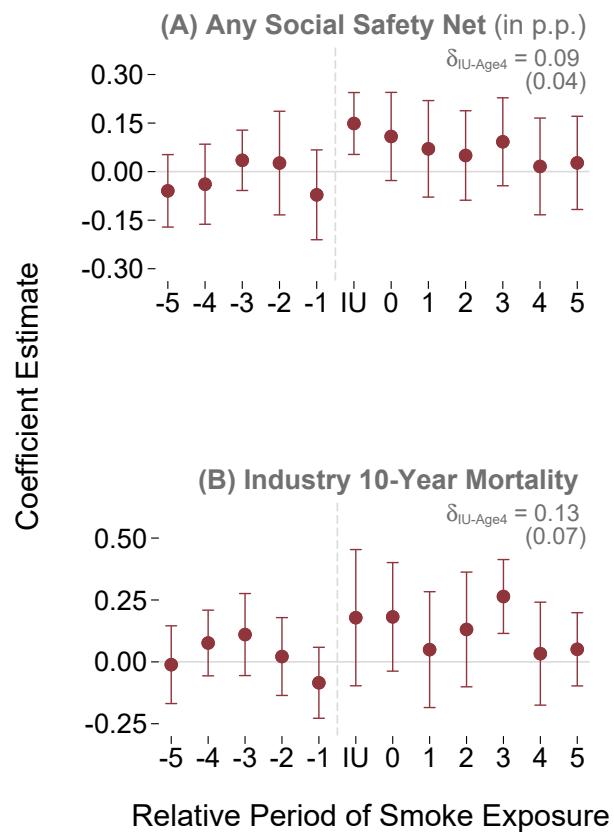
Figure A.17: Additional Results: Early-Life Wildfire Smoke Exposure Reduces Earnings and High School Completion; Increases Cognitive Difficulty



Notes: The purpose of this figure is to display the impact of childhood air pollution exposure from wildfires on a selected group of long-run outcomes. See notes to Figure 2.4 for more detail on graph construction and Section 2.6 for further discussion.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

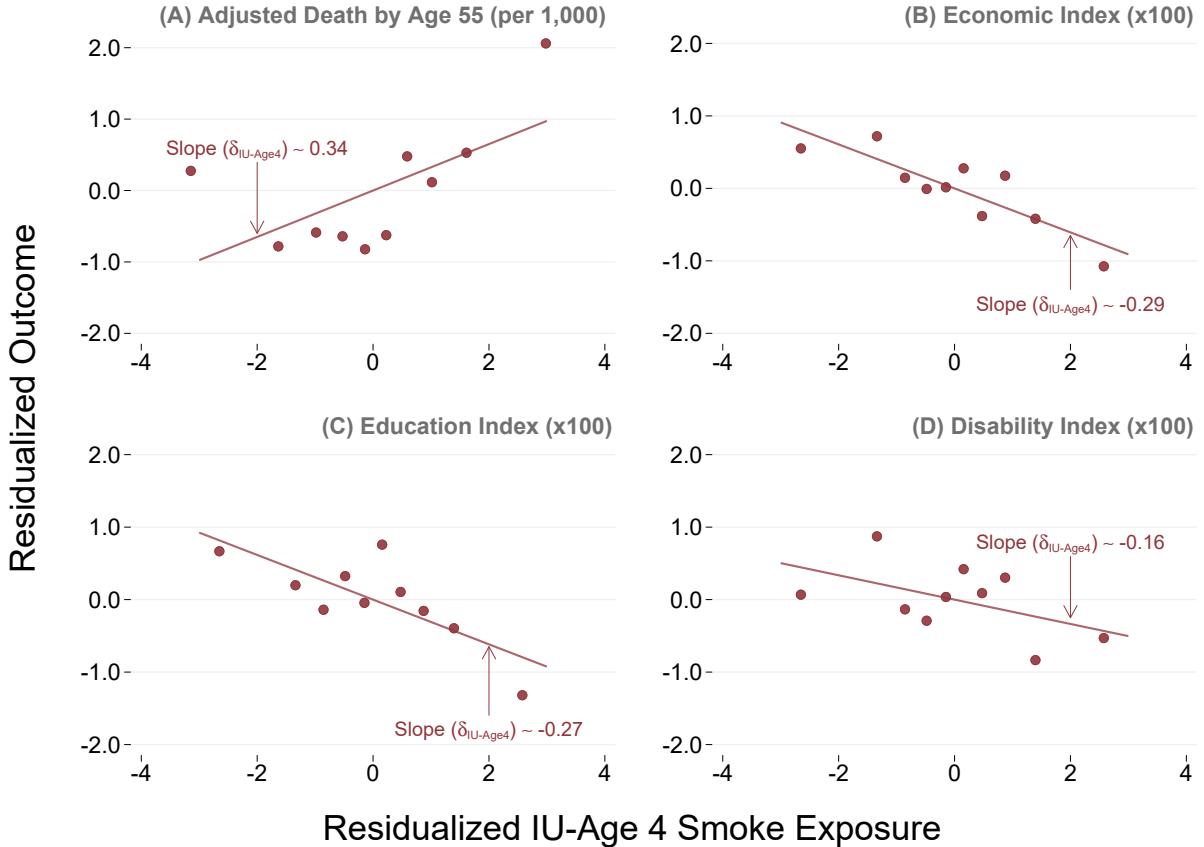
Figure A.18: Additional Results: Early-Life Wildfire Smoke Exposure Increases Use of Social Safety Net; Induces Sorting into Higher-Mortality Industries



Notes: The purpose of this figure is to display the impact of childhood air pollution exposure from wildfires on a selected group of long-run outcomes. See notes to Figure 2.4 for more detail on graph construction and Section 2.6 for further discussion.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, Restricted American Community Surveys, and Restricted Full-count Decennial Censuses.

Figure A.19: Robustness: Residual Plots Demonstrate that Impacts of Wildfire Exposure are Generally Linear



Notes: The purpose of this figure is to show how each of our main outcomes varies with our treatment. To do so, we residualize our baseline fixed effect controls from both the outcome and our summary treatment measure as follows:

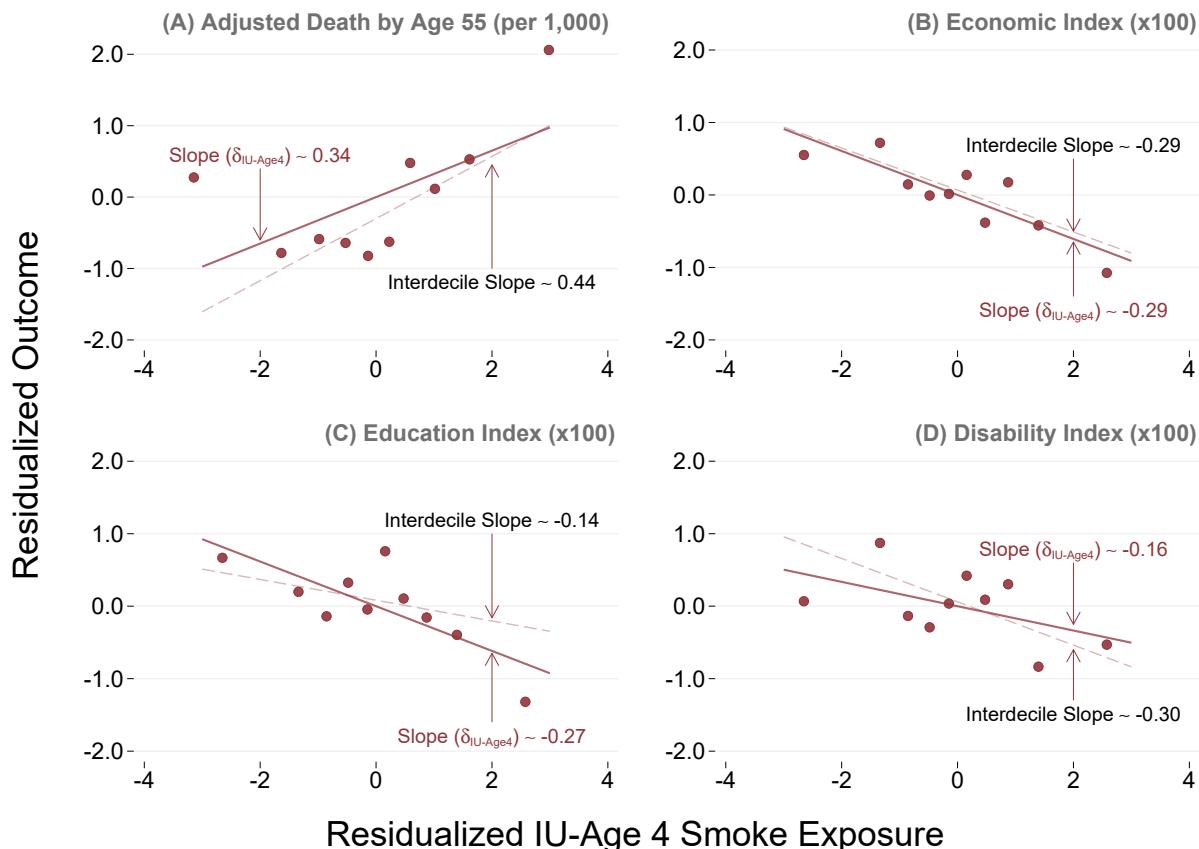
$$y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \epsilon_{cba}$$

$$Smoke_{cb}^{IU\text{-Age}4} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \xi_{cba}.$$

We then collapse the residualized summary treatment ($\tilde{\xi}$) into deciles and plot it against the residualized outcome ($\tilde{\varepsilon}$). The fitted line between $\tilde{\xi}$ and $\tilde{\varepsilon}$ gives the coefficient for our treatment. The general shape of the points within the figure supports our choice of functional form—the dots generally follow a linear pattern, underscoring that the dose-response for our treatment is linear in months of smoke exposure. The only meaningful deviations from this trend occur at the very tails of the distribution. However, as illustrated by Appendix Figure A.20, removal of these points does not materially change the estimates.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

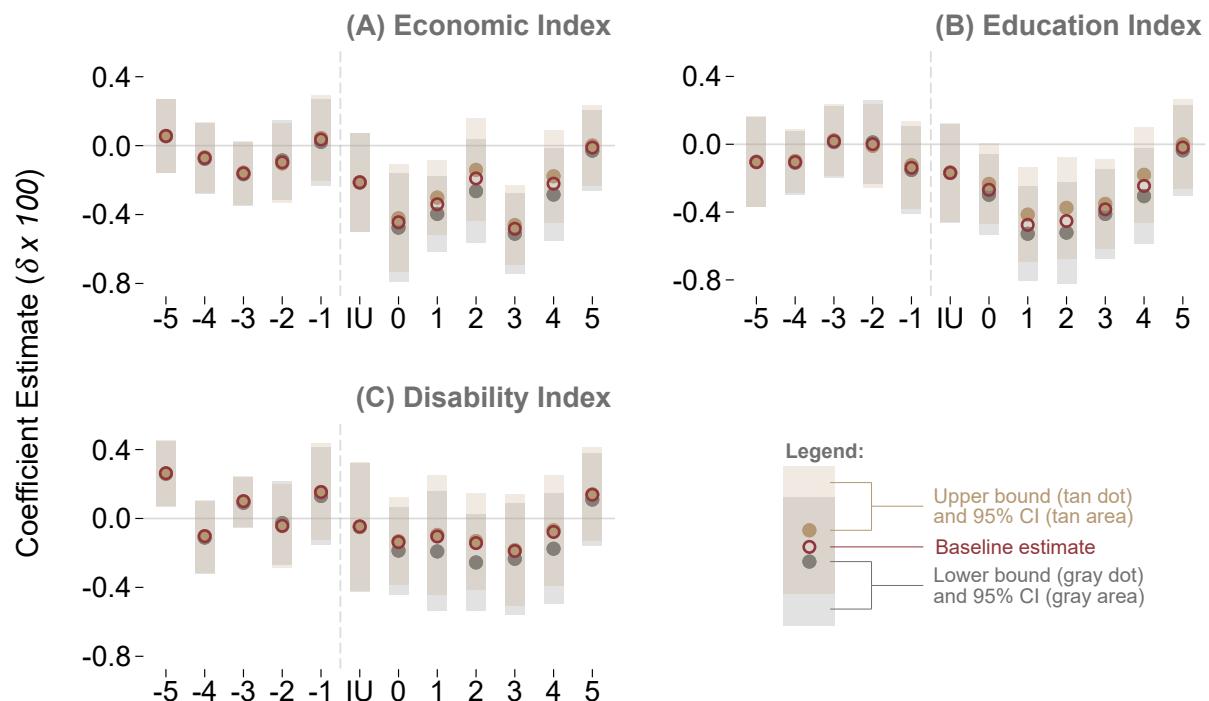
Figure A.20: Robustness: Residual Plots Demonstrate that Impacts of Wildfire Exposure are Not Driven by the Tails of the Distribution



Notes: The purpose of this figure is to further support the analysis displayed in Appendix Figure A.19 and discussed in the accompanying notes. Specifically, we modify the figure to fit lines and calculate slopes using only the points in the 2nd through 9th deciles (the “interdecile slope”) to demonstrate that our estimates are not driven by extreme points in the distribution. See Appendix Figure A.19 for more detail, including information regarding the construction of the graphs.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

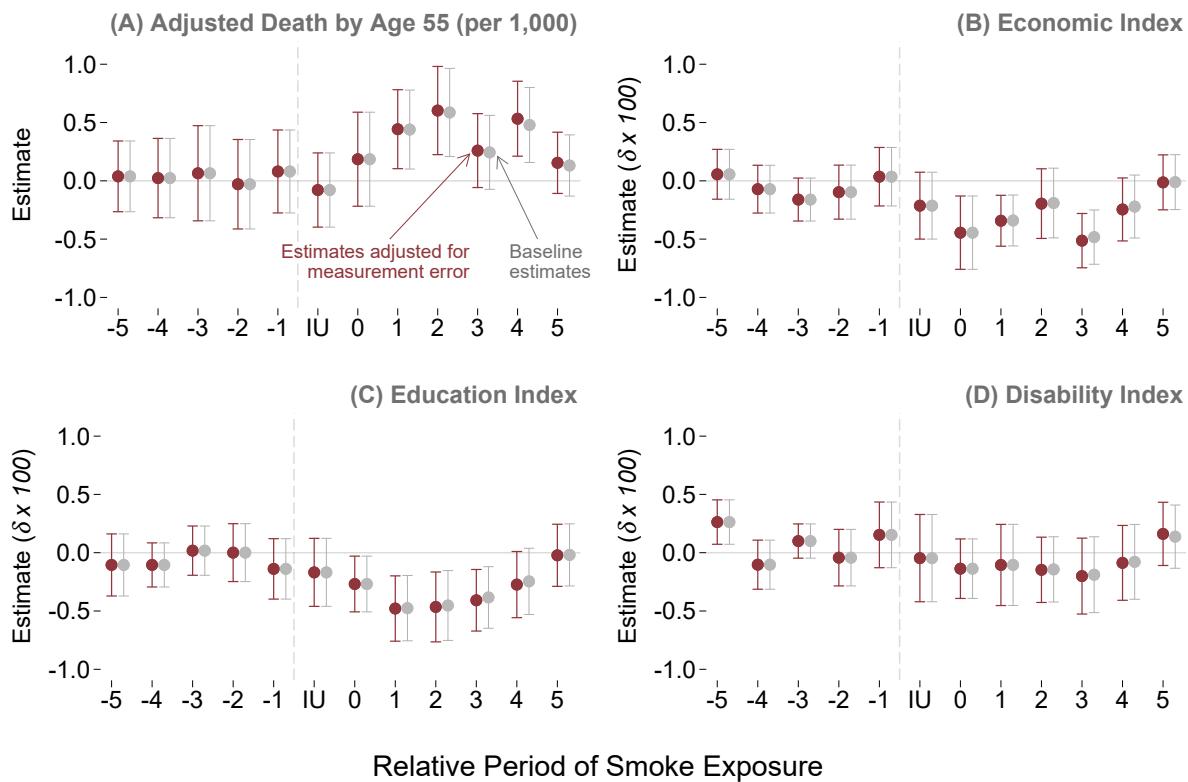
Figure A.21: Robustness: Adjusting for Mortality Selection Does Not Qualitatively Affect Estimates



Notes: The purpose of this figure is to display the bounded estimates for mortality selection discussed in Section 2.7 and Appendix Section B.3. Within the figure, upper- and lower-bounds of the potential impact of mortality selection are presented, alongside baseline estimates (i.e., those from Figure 2.4). See the text of the referenced sections for further description.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, Restricted American Community Surveys, California Department of Health Vital Statistics, and Natality/Mortality data from Bailey et al. (2018).

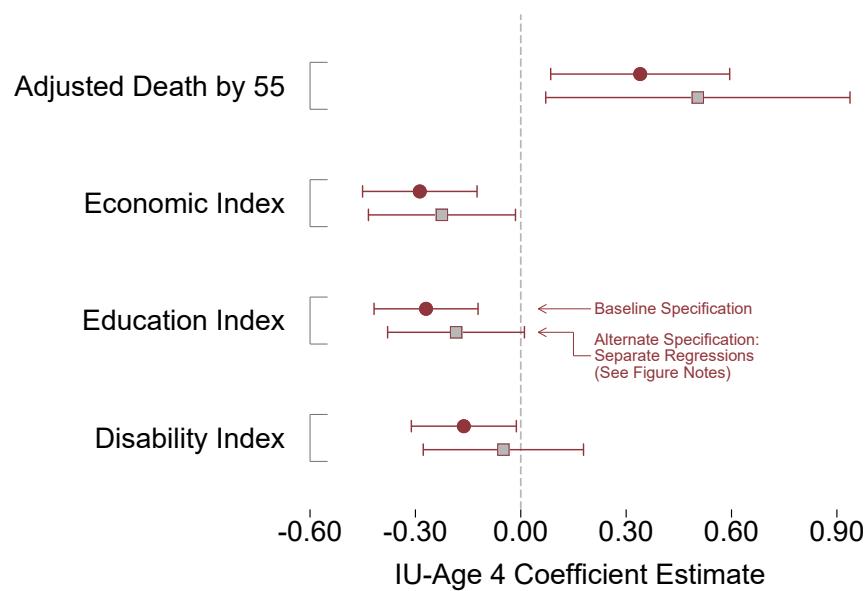
Figure A.22: Robustness: Adjusting for Migration-Specific Measurement Error Does Not Qualitatively Affect Estimates



Notes: The purpose of this figure is to compare our original estimates with those that have been adjusted for migration-specific measurement error. Within the figure, the gray estimates display the baseline estimates whereas the maroon estimates display the estimates when they have been adjusted for measurement error using the techniques described in Sections 1.7 and 2.7, as well as Appendix A, Section B.4. See the text of those sections for further detail.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys, 1940 Census 100% Sample (Ruggles et al., 2021).

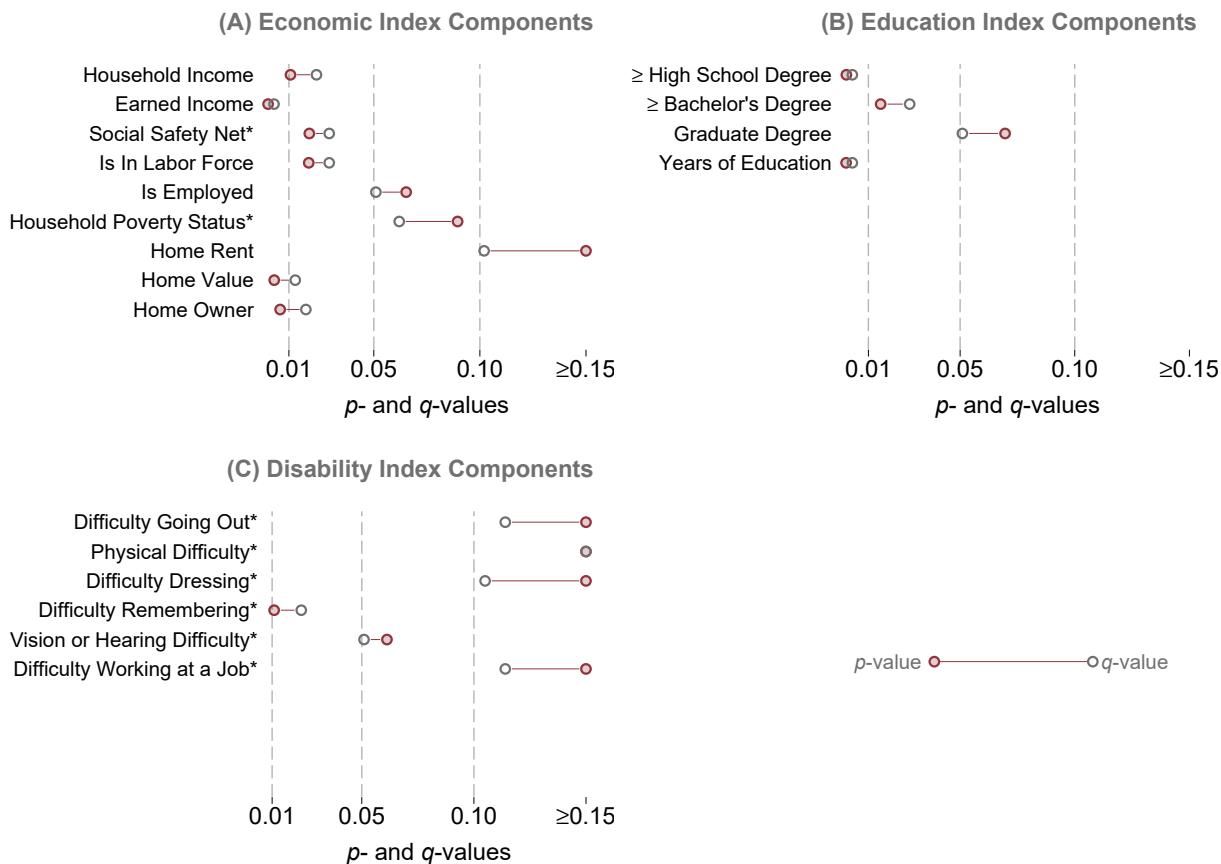
Figure A.23: Robustness: Using Alternative Estimation Strategies Do Not Result in Materially Different Estimates



Notes: The purpose of this figure is to display our estimates from the alternative estimation strategy discussed in Sections 1.7 and 2.7 as compared to our original estimates. This alternative estimation strategy involves separately estimating wildfire smoke effects for each year-of-birth cohort and combining these estimates via simple average. Standard errors are calculated using 100 cluster bootstrap iterations. As further discussed in the text, estimates are qualitatively similar, although the alternative estimates are substantially less precise.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

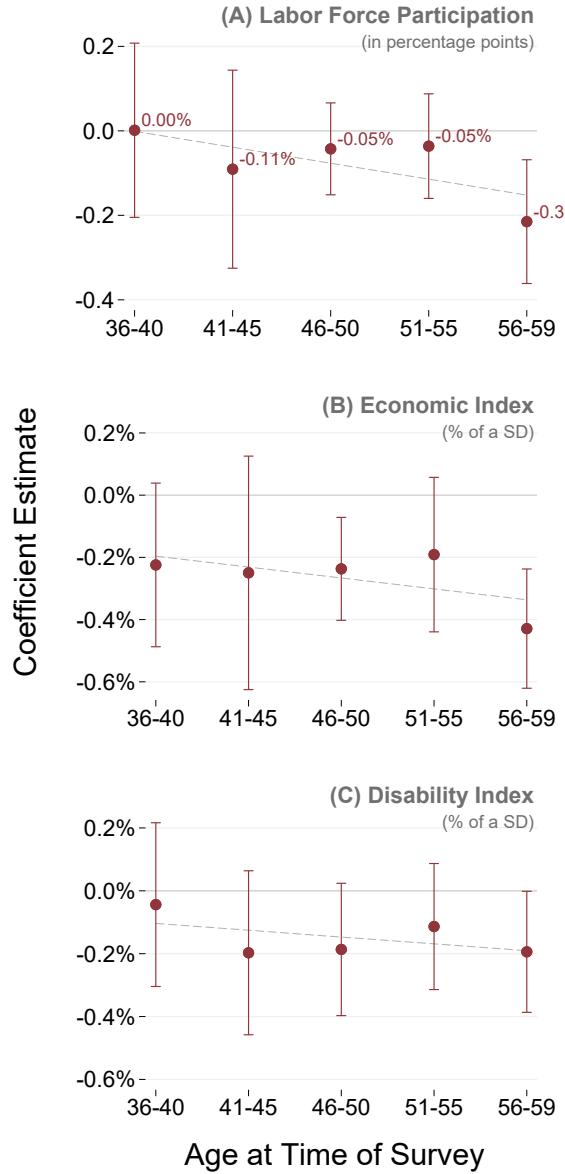
Figure A.24: Robustness: Multiple Hypothesis Testing Does Qualitatively Affect Conclusions About Statistical Significance for Index Components



Notes: The purpose of this figure is to contrast the p -values for our estimates with sharpened q -values (Benjamini and Yekutieli, 2001; Benjamini et al., 2006; Anderson, 2008) that adjust for multiple hypothesis testing. As noted by Anderson (2008), the q -values can be *smaller* than unadjusted p -values when there are many true rejections of the null hypothesis. In this case, then the method can “tolerate” some false rejections as well. Note that this tends to occur for estimates with large p -values, such as those in Panel C, which do not achieve traditional levels of statistical significance even after adjustment.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Figure A.25: Supplemental Results: The Effects of Early-Life Wildfires Smoke Data on Labor Force Participation, Economic Index, and Disability Index are Increasing Proportionally with Age



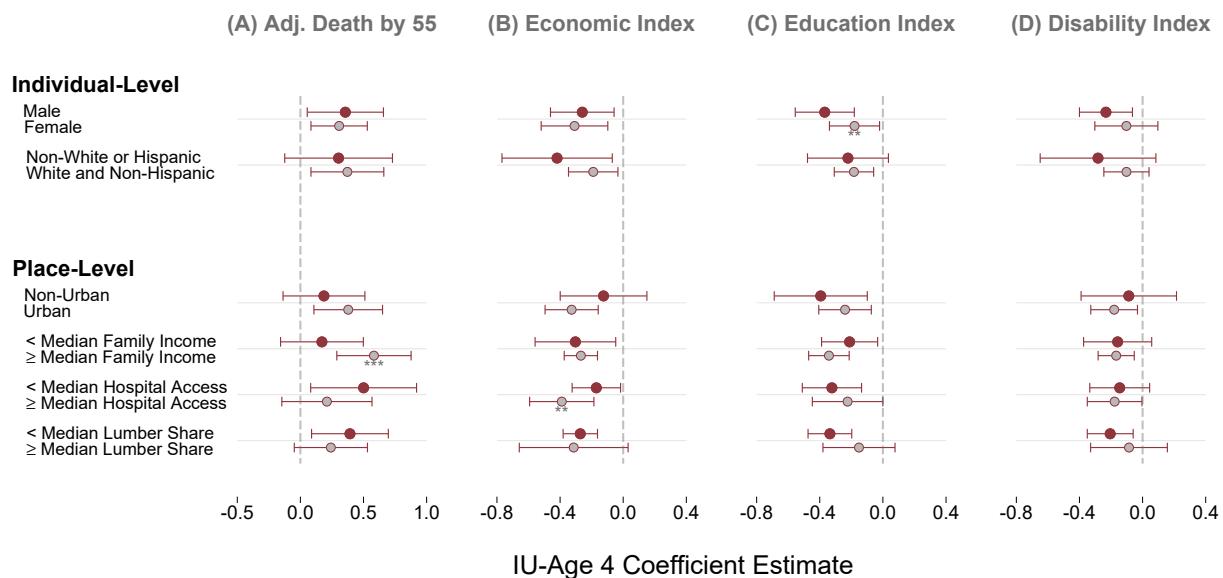
Notes: The purpose of this figure is to display how the effects of childhood wildfire smoke exposure evolve as individuals age. Within the figure, each point represents a separate estimate of δ_j with associated 95% confidence intervals, from the following equation:

$$Y_{cba} = \mu_c + \lambda_{y(b),r(c)} + \alpha_{m(b)} + \psi_{y(b),a} + \sum_{j \in J} \delta_j \cdot (AgeBin_j \times Smoke_c^{IU-Age 4}) + \varepsilon_{cba},$$

For further context, a line of best fit for the coefficients, along the effect as a percent of the age-specific mean is also presented. (Note that the percent of the age-specific mean is *not* presented for the indexes, as they are created to have a mean of zero in the aggregate.)

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Figure A.26: Supplemental Results: There Are Few Statistically Significant Differences in the Effect of Early-Life Wildfire Smoke Across Groups

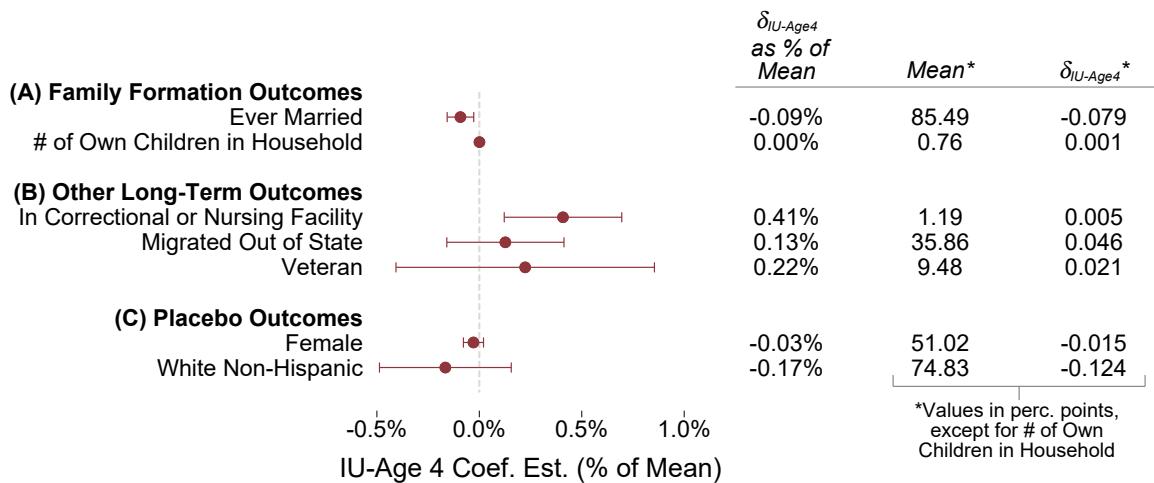


Notes: The purpose of this figure is to display the results of our heterogeneity analysis across groups. Within the figure, there are two sets of estimates presented for each regression detailed by Appendix A, Equation B2: the baseline effect (δ_0) and the *combined* effect ($\delta_0 + \delta_1$) which captures the total effect on the interaction category. Both the baseline effect and combined effect are presented with their 95% confidence intervals. Additionally, estimate pairs that are statistically distinct from each other—i.e., where the interaction term, δ_1 , achieves statistical significance—are denoted as follows:

- * Significant at 90% level
- ** Significant at 95% level
- *** Significant at 99% level

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, Restricted American Community Surveys, 1930 Census 100% Sample (Ruggles et al., 2021), and American Hospital Association Data (Finkelstein, 2007).

Figure A.27: Supplemental Results: Early-Life Wildfire Smoke Exposure Does Not Significantly Affect Family Formation, Migration, or Placebo Outcomes, but Does Increase Institutionalization



Notes: The purpose of this figure is to display the impact of childhood air pollution exposure from wildfires on outcomes that do not fall into the index categories discussed in Section 2.6. Within the figure, estimates—scaled to represent the effect as a percent of the mean—are presented alongside 95% confidence intervals. The tabular effect as a percent of the mean is presented to the first column on the right-hand side of the figure, along with the mean and level (non-scaled) estimate in the second and third columns, respectively.

The outcomes considered fall into three groups. First, we consider family formation outcomes, i.e., whether an individual was ever married (source: ACS) and the number of own children in the household at the time of survey (source: 2000 and 2010 Decennial Censuses). The outcomes in this group are either statistically or economically insignificant. The next group includes outcomes includes whether an individual lives out of state in adulthood (source: 2000 and 2010 Decennial Censuses), their veteran status (source: ACS), or whether they reside in institutional group quarters, the majority of which are correctional or nursing facilities (source: 2010 Decennial Census only; this variable is not available in the Restricted 2000 Decennial Census). Note that correctional and nursing facilities are grouped together because the Census Bureau did not permit RDC researchers to disclose results using sub-categories of the institutional group quarters variable during the time period in which these results were disclosed. The final group includes placebo outcomes—i.e., outcomes which should not be affected by early-childhood wildfire smoke. We do not find any effect on these outcomes, as expected.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, Restricted American Community Surveys, and Restricted Full-count Decennial Censuses.

Table A.1: Summary Statistics

Variable	Mean	Standard Deviation
<i>Economic Outcomes (American Community Survey, Ages 35-59)</i>		
Household Income	\$ 99,367	\$ 87,043
Earned Income*	\$ 44,087	\$ 54,488
Social Safety Net**	14%	34%
Is In Labor Force	76%	43%
Is Employed	71%	45%
Household Poverty Status	11%	31%
Home Rent	\$ 1,176	\$ 614
Home Value	\$ 418,807	\$ 340,128
Home Owner	73%	45%
<i>Education Outcomes (American Community Survey, Ages 35-59)</i>		
At Least High School Degree	94%	24%
At Least Bachelor's Degree	29%	45%
Graduate Degree	10%	30%
Years of Education	14.80	2.62
<i>Disability Outcomes (American Community Survey, Ages 35-59)</i>		
No Disabilities***	84%	37%
Difficulty Going Out (Doing Errands Alone)	5%	23%
Physical Difficulty (Walking or Climbing Stairs)	10%	30%
Difficulty Dressing or Bathing	3%	18%
Difficulty Remembering, Concentration or Making Decisions	6%	24%
Vision or Hearing Difficulty	5%	21%
Difficulty Working at a Job	11%	31%
Number of Disabilities***	0.32	0.85
<i>Mortality per 1,000 (Numident)</i>		
Adjusted Deaths by Age 35	27.22	
Adjusted Deaths by Age 45	50.19	
Adjusted Deaths by Age 55	94.96	
Adjusted Deaths by Age 65	176.27	
Adjusted Deaths by Age 75	321.90	

Notes: The purpose of this table is to detail means and standard deviations for key outcome variables included in our analysis.

* Variable is equal to wages + self-employment income.

** Variable is equal to 1 if an individual has any utilization of SNAP, Medicaid, SSI, or public assistance.

*** These variables are not included in the disability index and are only provided for context.

Source: Public-Use American Community Surveys (Ruggles et al., 2021) and SSA Numident.

Table A.2: Main Results: Table of Estimates

Outcome	Estimate	Standard Error	t-Statistic
<i>Panel A: Mortality Outcomes (per 1,000)</i>			
Adjusted Death by Age 30	0.10	(0.06)	1.73
Adjusted Death by Age 35	0.14	(0.07)	1.99
Adjusted Death by Age 40	0.19	(0.08)	2.25
Adjusted Death by Age 45	0.24	(0.10)	2.33
Adjusted Death by Age 50	0.27	(0.12)	2.23
Adjusted Death by Age 55	0.34	(0.13)	2.59
Adjusted Death by Age 60	0.20	(0.16)	1.31
Adjusted Death by Age 65	0.14	(0.17)	0.84
Adjusted Death by Age 70	0.26	(0.23)	1.14
Adjusted Death by Age 75	0.11	(0.22)	0.50
Adjusted Death by Age 80	-0.25	(0.41)	-0.60
<i>Panel B: Index Outcomes ($\delta \times 100$)</i>			
Economic Index	-0.29	(0.08)	-3.46
Household Income	-0.34	(0.13)	-2.55
Earned Income	-0.38	(0.10)	-3.73
Social Safety Net*	-0.33	(0.14)	-2.33
Is In Labor Force	-0.22	(0.09)	-2.34
Is Employed	-0.19	(0.10)	-1.84
Household Poverty Status*	-0.23	(0.14)	-1.70
Home Rent	-0.38	(0.27)	-1.43
Home Value	-0.31	(0.10)	-2.96
Home Owner	-0.26	(0.09)	-2.75
Education Index	-0.27	(0.08)	-3.57
At Least High School Degree	-0.32	(0.09)	-3.54
At Least Bachelor's Degree	-0.22	(0.09)	-2.42
Graduate Degree	-0.19	(0.10)	-1.81
Years of Education	-0.35	(0.10)	-3.63
Disability Index*	-0.16	(0.08)	-2.13
Difficulty Going Out (Doing Errands Alone)*	-0.12	(0.10)	-1.30
Physical Difficulty (Walking or Climbing Stairs)*	-0.08	(0.13)	-0.65
Difficulty Dressing or Bathing*	-0.12	(0.09)	-1.38
Difficulty with Remembering/Concentrating/Decisions*	-0.30	(0.12)	-2.55
Vision or Hearing Difficulty*	-0.26	(0.14)	-1.87
Difficulty Working at a Job*	0.16	(0.15)	1.11

Notes: The purpose of this table is to detail coefficient estimates presented in Figures 1.4 and 2.4, along with standard errors and t-statistics.

* Indicates that variable has been reverse coded so that higher coefficient values correspond with “better” outcomes.

Source: Author calculations using the Fire Perimeter Database (California Department of Forestry and Fire Protection, 2018), HYSPLIT model output (Stein et al., 2015), Restricted Census Numident, Restricted 2000 Decennial Census 1-in-6 Sample, and Restricted American Community Surveys.

Table A.3: Context for Main Results: Level Estimates

Outcome	Estimate	Mean	Effect as % of Mean	Unit	Note
<i>Economic Outcomes</i>					
Household Income	-\$280	\$99,367	-0.28%	Dollars	*
Earned Income	-\$178	\$44,087	-0.40%	Dollars	*
Social Safety Net	0.09	13.62	0.66%	Percentage Points	*
Is In Labor Force	-0.10	76.01	-0.13%	Percentage Points	*
Is Employed	-0.09	71.40	-0.12%	Percentage Points	
Household Poverty Status	0.07	10.73	0.66%	Percentage Points	
Home Rent	-\$2	\$1,176	-0.20%	Dollars	
Home Value	-\$1,044	\$418,807	-0.25%	Dollars	
Home Owner	-0.12	72.78	-0.16%	Percentage Points	
<i>Education Outcomes</i>					
At Least High School Degree	-0.08	93.94	-0.09%	Percentage Points	*
At Least Bachelor's Degree	-0.10	28.57	-0.34%	Percentage Points	
Graduate Degree	-0.06	10.35	-0.55%	Percentage Points	
Years of Education	-0.01	14.80	-0.06%	Years	
<i>Disability Outcomes</i>					
Difficulty Going Out	0.03	5.48	0.51%	Percentage Points	
Physical Difficulty	0.02	9.65	0.25%	Percentage Points	
Difficulty Dressing	0.02	3.24	0.68%	Percentage Points	
Difficulty Remembering	0.06	6.23	1.01%	Percentage Points	*
Vision or Hearing Difficulty	0.06	4.70	1.17%	Percentage Points	
Difficulty Working at a Job	-0.05	10.81	-0.47%	Percentage Points	

Notes: The purpose of this table is to provide level coefficients for the components of the indices displayed in Figure 2.4 (which displays coefficients that are scaled to be in terms of a percentage of the outcome's standard deviation). Within this table, most level estimates are obtained by multiplying the coefficient estimate by the applicable standard deviation. However, outcomes denoted with a “*” are outcomes for which the level effect has been separately estimated and disclosed.

Source: Author calculations using the Restricted and Public-Use (Ruggles et al., 2021) American Community Surveys.

Table A.4: Supplemental Results: Wildfire Smoke Does Not Have Detectable Effects on Contemporaneous Local Economic Conditions or Net Migration

	(1)	(2)	(3)	(4)	(5)	(6)
	% in Labor Force	% Employed	% Home Owners	Home Value	Family Income	Net Migration
<i>Smoke</i> ^{t',t}	0.00 (0.04)	0.02 (0.04)	-0.04 (0.06)	-\$125 (106)	-\$35 (51)	-0.07 (0.63)
95% CI: Upper Bound	0.09	0.09	0.09	\$83	\$65	1.16
95% CI: Lower Bound	-0.08	-0.06	-0.16	-\$332	-\$135	-1.31
Outcome Mean	60.60	55.90	54.40	\$93,056	\$43,619	24.30
Observations	232	232	172	170	169	232

Notes: The purpose of this table is to demonstrate the impact of an additional smoke month on a variety of outcomes that are *contemporaneous* to the smoke exposure. Each column in the figure represents a different outcome variable estimated from Appendix A, Equation B1, with associated standard errors clustered at the county level. Note that columns 1-3 and 6 are in percentage points, while columns 4 and 5 are denominated in 2010 dollars. See the text of Appendix A, Section B.5 for more detail.

Source: Author calculations using the 1930-1970 Decennial Census (Ruggles et al., 2021), Census County Book Data (U.S. Census Bureau, 2012), and 1934 IRS Statistics on Income.

Appendix B

Appendix to Chapter 3: How Do Medicare Payments Affect Physician Practice Structure?

A Variations of the Theoretical Model

The model in Section 3.3 makes several simplifications for purposes of clarity and tractability. Accordingly, the purpose of this Appendix Section is to test the model's sensitivity to these assumptions, particularly those regarding intensive-margin quantity responses to differential prices. This will be done in two ways: (1) by maintaining the original model, with an additional assumption that Medicare quantity responds positively to price differences and (2) deriving a model with joint choices of practice size and Medicare quantity.

Preliminaries. Recall the following equations from Section 3.3:

$$\max_s u(s) = v(\pi(s)) - w(s) \text{ subject to} \quad (\text{A1})$$

$$\pi(s) = p^o(s) \cdot q^o + p^m \cdot q^m - mc(s) \cdot (q^m + q^o) \quad (\text{A2})$$

$$p^o(s) = (1 - \theta(s)) \cdot p^{ins} + \theta(s) \cdot p^m, \quad (\text{A3})$$

which produce the comparative static of Medicare prices, p^m , on group size, s :

$$\frac{ds}{dp^m} = -\frac{v_{\pi\pi}\pi_s\pi_{p^m} + v_{\pi}\pi_{sp^m}}{u_{ss}} < 0 \quad (\text{A4})$$

The sign of the equation above is based the increasing and concave nature of utility, as well as the following conditions derived from the model:

$$\pi_s = p_s^o q^o - mc_s(q^m + q^o) \quad (\text{A5})$$

$$= -\theta_s(p^{ins} - p^m)q^o - mc_s(q^m + q^o) > 0$$

$$\pi_{p^m} = \theta(s)q^o + q^m > 0 \quad (\text{A6})$$

$$\pi_{sp^m} = \theta_s q^o < 0. \quad (\text{A7})$$

Original Model + Increasing Quantity Assumption. To consider how quantity responses might affect model predictions, first I allow Medicare quantities to respond positively to Medicare prices (formally, $q_p^m > 0$), as shown empirically in Clemens and Gottlieb (2014).¹ Such an assumption would leave $v_{\pi\pi}$, v_{π} , u_{ss} , and π_s unchanged. The affected terms would now include:

$$\pi_{p^m} = \pi_{p^m}^{orig} + (p^m - mc)q_p^m > 0 \quad (\text{A8})$$

$$\pi_{sp^m} = \pi_{sp^m}^{orig} - mc_s q_p^m \gtrless 0, \quad (\text{A9})$$

where $\pi_{p^m}^{orig}$ and $\pi_{sp^m}^{orig}$ represent the values in Equations A6 and A7, respectively. The sign for π_{p^m} is the same as in the original model, under the weak assumption that the marginal Medicare patient is profitable ($p^m - mc > 0$). However, the sign for π_{sp^m} is now technically ambiguous, as the term $-mc_s q_p^m$ is positive. Despite the ambiguous term, it is unlikely that this term is sufficiently influential as to change model predictions due to its small probable magnitude. Empirical evidence on cost savings from mergers and organizational size increases suggest (at best) modest cost savings (see, for instance, Craig et al. 2018). When these marginal cost savings are interacted with known intensive-margin quantity responses (with elasticities of approximately 1-2%), they are unlikely to result in decision-altering quantities. Accordingly, the predictions of the augmented model are likely unchanged from the baseline.

Model with Endogenous Choice of Quantity. I next expand the model to allow physicians

¹Private quantities are assumed to be unresponsive to Medicare prices for clarity of exposition.

to choose the intensive-margin quantity of services delivered to patients (which implicitly assumes that physicians induce demand as agents of patients). In response to changing quantity, physicians also incur non-financial disutility, represented by $h(q)$, due to the cost of inducement. Formally, the physician solves:

$$\max_{s,q^m} u(s, q^m) = v(\pi(s, q^m)) - w(s) - h(q^m) \text{ subject to} \quad (\text{A10})$$

$$\pi(s, q^m) = p^o(s) \cdot q^o + p^m \cdot q^m - mc(s) \cdot (q^m + q^o) \quad (\text{A11})$$

$$p^o(s) = (1 - \theta(s)) \cdot p^{ins} + \theta(s) \cdot p^m, \quad (\text{A12})$$

where the assumptions from the previous model remain, except that $w(s)$ is assumed to be linear (for simplicity) and $h(q)$ is assumed to be increasing in q .² The first order conditions of this model are:

$$\frac{\partial u(s)}{\partial s} = 0 \Leftrightarrow v_\pi \pi_s - w_s = 0 \quad (\text{A13})$$

$$\frac{\partial u(s)}{\partial q^m} = 0 \Leftrightarrow v_\pi \pi_{q^m} - h_{q^m} = 0. \quad (\text{A14})$$

Combining first order conditions to obtain:

$$\phi(s, q^m) = w_s \pi_{q^m} - h_{q^m} \pi_s \quad (\text{A15})$$

and totally differentiating to obtain the comparative static of Medicare prices, p^m , on group size, s , yields:

$$\frac{ds}{dp^m} = -\frac{\partial \phi(s, q^m)/\partial p^m}{\partial \phi(s, q^m)/\partial s} = -\frac{w_s \pi_{q^m} p^m - h_{q^m} \pi_{sp^m}}{w_{ss} \pi_{q^m} + w_s \pi_{sq^m} - h_{q^m} \pi_{ss}}. \quad (\text{A16})$$

Based on the assumptions of the model, signs of each component of the comparative static are as

²As before, private quantities are fixed for clarity of exposition.

follows:

$$w_s, h_{q^m} > 0 \quad (\text{A17})$$

$$w_{ss} = 0 \quad (\text{A18})$$

$$\pi_{q^m p^m} = 1 \quad (\text{A19})$$

$$\pi_{sp^m} = \theta_s q^o < 0 \quad (\text{A20})$$

$$\pi_{sq^m} = -MC_s > 0 \quad (\text{A21})$$

$$\pi_{ss} = p_s^o q^o - mc_{ss}(q^m + q^o) \quad (\text{A22})$$

$$= -\theta_{ss}(p^{ins} - p^m)q^o - mc_{ss}(q^m + q^o) > 0. \quad (\text{A23})$$

Accordingly, the sign of Equation A16 is:

$$\frac{ds}{dp^m} < 0, \quad (\text{A24})$$

which is the same prediction given by the original model.

B Sample Construction

The primary datasets used for my analysis are as follows:

1. **Medicare Physician and Other Supplier Public Use File.** This dataset supplies provider-level data for all providers who treat more than 10 Medicare patients in a given year. The data includes information detailing the exact addresses, summary information on patients treated—including average age and average hierarchical condition category score, which measures patient health—and provider characteristics such as certification (ex. Medical Doctor, Doctor of Osteopathic Medicine) and primary specialty type. The detail file also includes service-level detail regarding procedures by the provider in a given year. Relative value units were assigned to each service by merging in the 2015 Physician Fee Schedule’s Relative Value files obtained from CMS. The source data for utilization and patient-population characteristics are derived from the CMS Chronic Conditions Warehouse, a database with 100% of Medicare enrollment and fee-for-service claims data. Provider characteristics and address are obtained from the National Plan and Provider Enumeration System (“NPPES”). According to NPPES guidelines, any address changes must be communicated by providers within 30 days of the change.
2. **Medicare Physician Compare Data.** This dataset is used to gather physician-level information regarding group practice affiliation, the size of associated medical groups, medical school attended, and year of medical school graduation. Medical school rankings were obtained using the methodology detailed in Schnell and Currie (2018), and provider experience variables were generated by subtracting the current year (2015) from the date of medical school graduation. The source data for Physician Compare variables is the Provider Enrollment, Chain, and Ownership System (“PECOS”), which is Medicare’s enrollment and revalidation system. Group practice affiliation is determined based on the tax identification number that is associated with a given physician’s NPI. Accordingly, it is possible for providers to have multiple affiliations in the Physician Compare data, which is true for roughly 14% of providers in my sample. I assign providers to their smallest practice under the assumption that larger practices are more likely to be included only due to practice privileges and are therefore not the providers’ main practice. As discussed in Section C, this assumption does not meaningfully affect my estimates.

3. Medicare Physician Fee Schedule. This dataset provides detail regarding geographic adjustment factors by locality, RVUs-by-HCPCS code, and zip-to-locality crosswalks. I use the zip-to-locality crosswalks to form locality borders, defined by adjacent zip codes that are assigned different geographic adjustment factors. Physicians are paid using the geographic adjustment factors that correspond to their practice address.

4. Census Data. I obtained local data on the Census-block-group level using 2011-2015 ACS 5-year estimates from the U.S. Census Bureau and IPUMS (Ruggles et al., 2019). This data includes the following variables, which were used as “local” controls in my analysis:

- (a) Percentage of population in given age bins (less than 65, 65-69, 70-74, 75-79, 80-84, and 85-plus);
- (b) Percentage of population by education level (less than high school, high school, some college, college or more);
- (c) Percentage of population by insurance coverage (uninsured, Medicare only, Medicaid only, other insurance only, Medicare+Medicaid, Medicare+Private Insurance);
- (d) Median household income;
- (e) Median home value; and
- (f) Mean population density.

5. Spatial Data. Geographic shapefiles (State- and ZCTA-level) were obtained from the U.S. Census.

Using this data, the main sample construction was as follows:

1. The ZCTA-level shapefiles and zip-to-locality crosswalk were utilized to create locality geographies in ArcGIS. Locality borders were then generated for each combination of within-state adjacent localities (i.e. no cross-state borders were used). Borders that corresponded with large rivers were also omitted, as other characteristics would likely not evolve smoothly across the boundary in these instances.
2. Physician addresses from the Physician and Other Supplier Public Use File were geocoded using Texas A&M GeoServices. Then the shortest straight-line distance was calculated between each border and candidate border point within 10 miles.

3. The Physician and Other Supplier Public Use Variables and Physician Compare variables were combined. Only primary care physicians were retained.³ Next, I dropped all physicians that had outlier geocoordinates (i.e. those outside of the boundaries of their listed zip code) and physicians who practice in nine-digit zip codes that had a different locality than their respective five-digit zip code were also dropped. Finally, I keep only providers who are within one bandwidth of the border.
4. Only borders that have 5 providers on each side of the boundary were maintained.
5. I also created “clusters” of physicians in each border to more finely control for the spatial characteristics of providers (see Table 3.1, Column 4). To do so, I utilized k -means clustering *separately for each border* based on physician longitude and latitude. The algorithm was set to create one cluster for every fifty physicians within the bandwidth’s length of the border (rounded up). After the initial clusters were determined, clusters were divided into “good” clusters—defined as those with 5 physicians on either side of the boundary—and “bad” clusters. The physicians in bad clusters were re-assigned to the nearest good cluster. An illustration of these clusters is displayed in Appendix Figure B.4, which recreates the map from Figure 3.1(b) with the points for each physician and different clusters denoted by a shape-color combination.

This procedure results in the sample of providers that I use for my main analysis. I also repeat the procedure in Steps #3 and #4 for all multiples of bandwidth used in my Section 3.5.2 bandwidth analysis. Finally, I assign controls to each physician based on characteristics of the surrounding Census block groups. Specifically, I calculate an inverse-distance weighted average of characteristics using block groups with centroids within two miles of a given physician’s location that are also in the same payment locality.

³Primary care physicians included those who listed their specialties as: Family Practice, General Practice, Geriatric Medicine, Internal Medicine, Osteopathic Manipulative Medicine, Pediatric Medicine, or Preventive Medicine. Among these primary care physicians, over 95% listed Family Practice or Internal Medicine as their specialty.

C Additional Figures and Tables

A description of the analyses performed in this section are as follows:

- **Figure B.1.** This figure displays other measures of bunching to supplement the analysis in Section 3.4.2. See text of that section and figure notes for more detail.
- **Figure B.2.** In addition to the analysis displayed in Figure 3.4 and discussed in Section 3.4.2, I also show how specific characteristics vary across the cutoff, both on the physician- and block-group-level. Panels A and C display the physician- and block-group-level results for the High-Impact Sample, respectively: of the 22 outcomes examined, only 1 achieves traditional levels of statistically significance, which is roughly the rate expected for incorrectly rejecting the null hypothesis. To emphasize that this single significant result is likely spurious and not a result of intentional sorting, I also present results for a "Low-Impact" Sample, i.e. the borders that have GAF differences of less than 2 percent. As displayed in Panels B and D, there are several characteristics that achieve statistical significance, despite the substantially lower degree of treatment (and thus reduced incentives for physicians to sort around the cutoff).
- **Figure B.3.** This figure recreates Figure 3.5 when no adjustments for border-fixed effects are included and achieves qualitatively similar (though slightly larger) effects.
- **Figure B.4.** This figure recreates Figure 3.1B, with overlaid points representing the location of physicians. Each color-shape combination represents a different physician cluster—see Appendix B, Section B and figure notes for more detail.
- **Figure B.5.** This figure displays results that are analogous to those displayed in Figure 3.6B, but where $Y_{ip} = 1(g < GroupSize \leq g')$, where g and g' represent points the distribution of group sizes (i.e. these points represent the effects on the PDF, rather than CDF).
- **Figure B.6.** This figure displays regression discontinuity plots for other measures of group size discussed in Section 3.5.1. See that section and figure notes for more detail.
- **Figure B.7.** This figure considers two other ways that physician *treatment behavior* may change in response to differential payment. The first row (Panels A and B) represents per-patient utilization (i.e. the intensive-margin treatment decision) and the second row (Panels

C and D) display the number of Medicare patients treated per physician (i.e. the extensive-margin treatment decision). Due to the imprecision of the estimates in the left-hand columns, additional controls were included in the right-hand specifications to reduce noise. (Specifically, a control for the *predicted* level of the outcome, where the prediction is generated from a LASSO design.) While the intensive-margin results do not achieve traditional levels of statistical significance ($p = 0.115$ in my most precise specification), the estimated effect of a 1.7-1.8 percent increase in resource intensity (measured by log relative value units per patient) in response to a 1% change in reimbursement is generally in-line with recent literature on supply-side response to financial incentives (Clemens and Gottlieb, 2014). Likewise, I do not find any statistically significant effects on extensive-margin treatment decisions (measured by log Medicare patients treated), consistent with the literature on the subject.⁴ Nonetheless, the 95% confidence interval for my precise specification rules out responses larger than a 1.1% increase in Medicare beneficiaries treated in response to a 1% change in reimbursement. To frame this in terms of potential *reductions* in Medicare prices, my result suggests that modest decreases in reimbursement would not limit primary care access for Medicare beneficiaries.

- **Figure B.8.** This figure displays regression discontinuity plots for two other measures of practice structure, practice *location* size (Panel A) and log group size (Panel B). See Section 3.5.1 for further discussion.
- **Figure B.9.** This figure exhibits sensitivity of estimates to the effects of individual borders by performing a “leave-one-out analysis” — i.e. re-estimating Equations 3.10 and 3.11 with all observations assigned to the listed border omitted from the sample. As shown in the figure, no individual border seems to exhibit substantial influence on the estimates.
- **Figure B.10.** The purpose of this figure is to perform additional robustness checks of my main analysis:
 - “Cluster level” section: explores alternatives to the tract-level clustering strategy in the main specification.
 - “Kernel choice” section: details alternative kernel choices to the rectangular kernel used.

⁴Clemens and Gottlieb (2014) also are unable to detect any change in the number of patients treated in their setting. In contrast, Alexander and Schnell (2019) argue that an increase in *Medicaid* reimbursement for primary care services meaningfully increased healthcare access for Medicaid beneficiaries. However, because Medicaid reimbursements are substantially less generous than those for Medicare (Zuckerman et al., 2017), estimates of access from this paper are most relevant for fee structures that meet or exceed Medicare’s generosity.

- “Distance function” section: examines a different functional form for distance controls (linear was used in the main specification).
- “Drop medical centers” line: tests robustness to omission of areas of significant provider density. While medical centers are not explicitly indicated in the data, I proxy for them by tabulating the number of providers in grids that are roughly one-third square mile in area (1/100th of a degree of longitude and latitude). I then omit observations from the analysis at the 99th percentile of provider-density among the grids.
- “Drop phys. assigned to >1 borders” line: physicians that were within one bandwidth-distance of multiple borders were assigned to more than one locality-border (roughly 4% of my sample meets this criteria). This line tests robustness to that sample construction choice by dropping these providers.
- “Drop phys. w/ >1 affiliation” line: roughly 10% of doctors in my sample report more than one group affiliation. As part of sample construction, I assigned the smallest listed group size to these physicians; this tests robustness to that choice by dropping providers with multiple affiliations.
- “Drop zip-centroid geocodes” line: certain providers could only be geocoded based on the centroid of their associated zip code (rather than their exact address). Given that these providers are more likely to have a mis-measured location, they are dropped as a robustness check.
- “Specialty group FE” line: in order to more finely control for differences in physician type, I add fixed effects for each major group: Family Practice, Internal Medicine, and Other Primary Care (the Other Primary Care category covers roughly 4% of physicians in my sample).
- “Border-specific spatial controls” line: adds more detailed controls for distance by allowing linear distance controls to vary by each border.

As displayed in the figure, none of these robustness procedures changed the qualitative nature of my estimates and, in most cases, resulted only in extremely small changes to the coefficient and/or standard errors.

- **Figure B.11.** This figure recreates Figure 3.5 when the full sample of borders are included.

See Section 3.5.3 for more detail.

- **Figure B.12.** This figure recreates Figure 3.6 (Panel A) and Appendix Figure B.5 (Panel B) when the full sample of borders are included and the RD-Border Estimation (Equations 3.12 and 3.13) are used. See Sections 3.5.1 and 3.5.3 for more detail.
- **Figure B.13.** The purpose of this figure is display heterogeneity in physician characteristics through a series of separate equations:
 - The first row displays heterogeneity by physician specialty. Given that internal medicine practice physicians generally have a larger share of Medicare beneficiaries (roughly 49% more per the 2010-2015 National Ambulatory Medical Care Surveys), it is expected that they would exhibit a stronger response to differential reimbursement. Panels A and B of the figure confirm this: the RD-IV point estimate is approximately 61% larger for internal medicine physicians than for family practice doctors (although it should be noted that these estimates are not statistically different at traditional levels).
 - The second row divides the sample by whether the physician attended a ranked medical school or not. There do not appear to be meaningful differences in these estimates.
 - The third row divides the sample by whether the physician has over/under twenty years of experience (proxied by the years between 2015 and their medical school graduation). While the RD-IV point estimates are slightly different, they are not statistically distinguishable. (It should also be noted that experience could simply be a proxy for gender or vice versa: the male physicians comprised 69% of those with 20+ years of experience, while female physicians represented 55% of those with fewer than 20 years. This is consistent with the results of the next row.)
 - The fourth and final row divides the sample by physician gender. Again, while the RD-IV point estimates are slightly different, they are not statistically distinguishable.
- **Figure B.14.** The purpose of this figure is display how well group size measures can be used to predict log fixed travel HHI (“FTHHI”). See text of Section 3.6 and figure notes for more detail.
- **Table B.1.** This table reports all payment localities that receive different physician fee schedule payments than their respective states. See Figure 3.1(a) for more information.
- **Table B.2.** This table reports coefficients for other physician practice style outcomes for

different samples and instrumental variables. See discussion previous discussion of Appendix Figure B.7 in this section for more detail regarding outcomes.

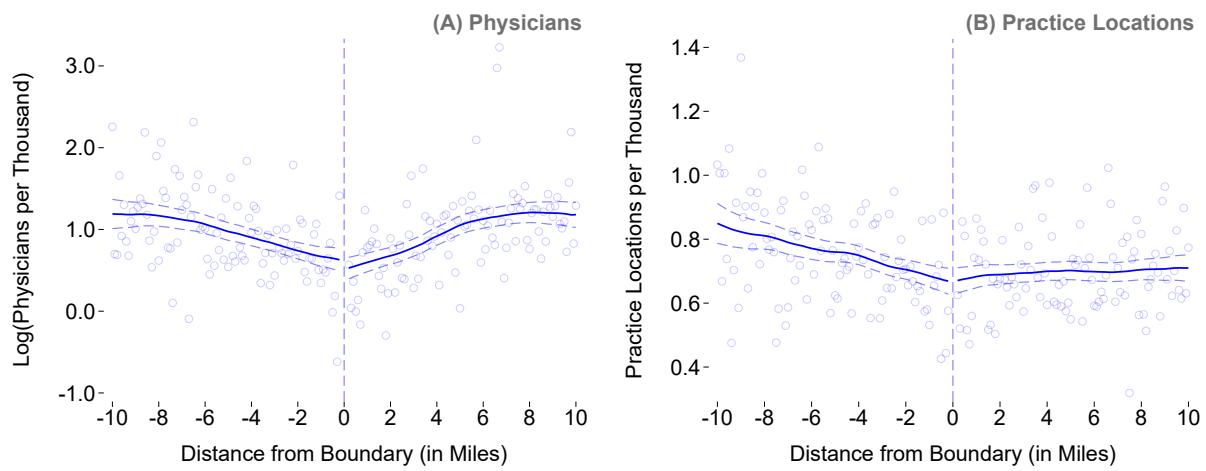


Figure B.1: Validation of Research Design: Alternate Measures of Spatial Provider Density

Notes: The purpose of this figure is to illustrate the absence of bunching around locality boundaries using alternate measures to those presented in Figure 3.3. Panel A is analogous to Figure 3.3B, except with *log* physicians per thousand as an outcome measure. Panel B similarly replaces the outcome so that each point is a practice *location*, rather than an individual physician. This analysis utilizes the High-Impact Sample discussed in Section 3.2.2.

Source: Author calculations using CMS Physician Fee Schedule and CMS Physician and Other Supplier Public Use File.

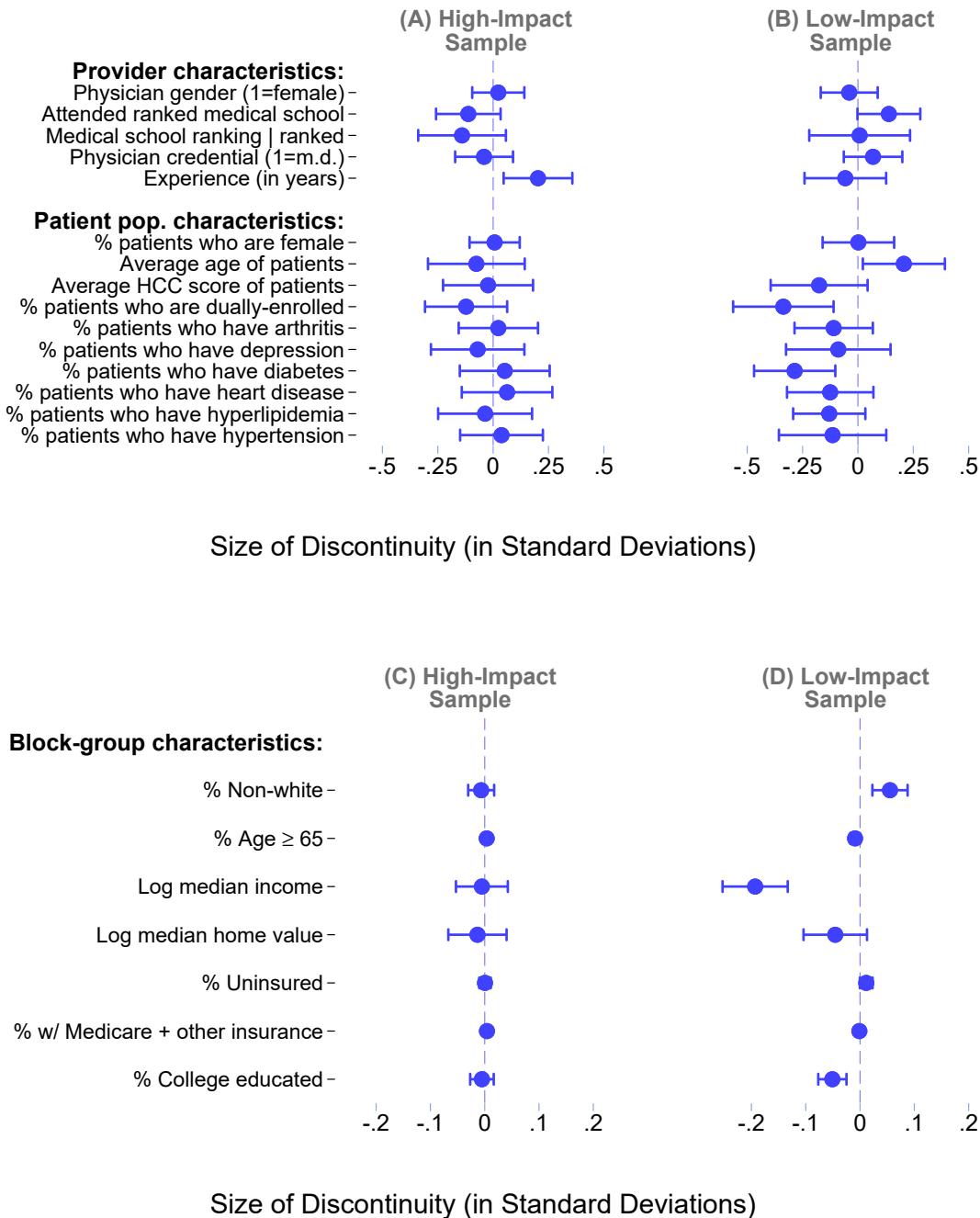


Figure B.2: Validation of Research Design: Covariate Balance Tests

Notes: The purpose of this figure is to illustrate the response of selected covariates to boundary differences in Medicare reimbursement. The top panels (A and B) display provider-level regressions for provider and treated-patient-population-level characteristics, for the High-Impact Sample (borders with differences of at least two percent points) and the Low-Impact Sample (borders with differences less than two percent points), respectively. The bottom panels (C and D) display Census-block-group level regressions for local characteristics for High- and Low-Impact Samples, respectively. All variables have been normalized and used as an outcome measure in Equation 3.8 (i.e., these are discontinuities that have *not* been scaled by the related payment discontinuity). Bars displayed represent 90% confidence intervals. See Appendix Section C for further discussion of this figure.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File, Census Block Group 5-year ACS Estimates (Ruggles et al., 2019).

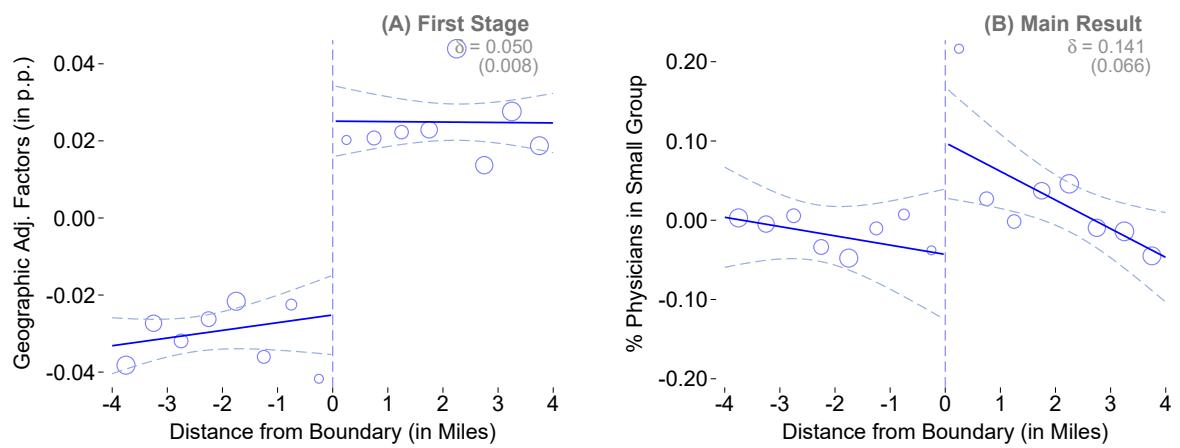


Figure B.3: Main Results: Without Border FE

Notes: The purpose of this figure is to recreate the analysis displayed in Figure 3.5 when no adjustments for border fixed effects are included. See notes to Figure 3.5 for more detail.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

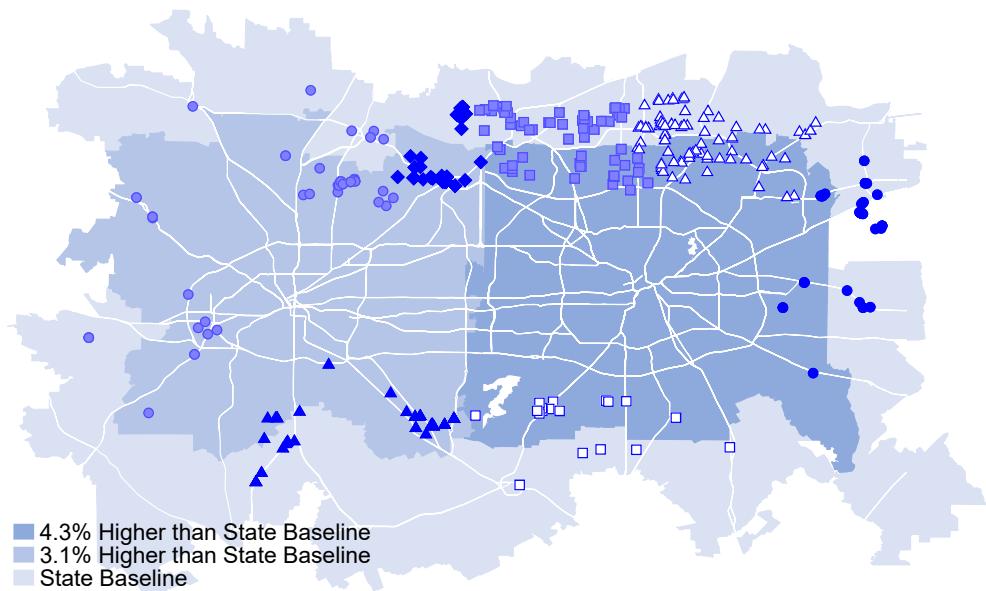


Figure B.4: Localities in the Dallas-Fort Worth Area (Physician Clusters)

Notes: The purpose of this figure is to illustrate the physician clusters used analyses employing cluster-by-border fixed effects (discussed further in Section 3.4). Each shape-by-color combination represents a different set of physicians that are grouped together. Use of border-by-cluster fixed effects effectively forces comparisons within these physician groups. See notes to Figure 3.1(b) for further description of the underlying map characteristics. See notes to Appendix Section B for more detail on how these clusters were constructed.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician and Other Supplier Public Use File.

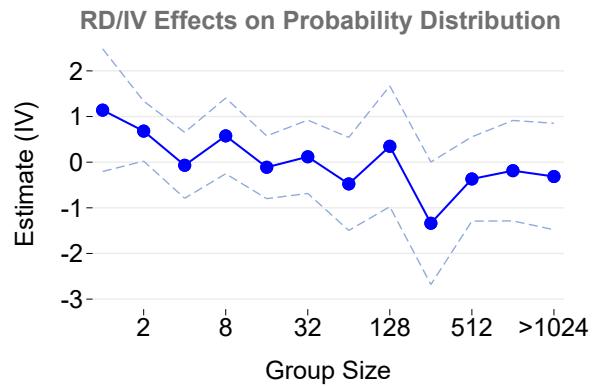


Figure B.5: Distributional Effects on Group Size

Notes: The purpose of this figure is to display results that are analogous to those displayed in Figure 3.6B, but where $Y_{ip} = 1(g < \text{GroupSize} \leq g')$, where g and g' are points in the distribution of group sizes. Coefficient estimated, presented along with 90% confidence intervals, show the impact of a 1 percent difference in Medicare reimbursement is associated with a δ p.p. change in the displayed probability.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, and CMS Physician and Other Supplier Public Use File.

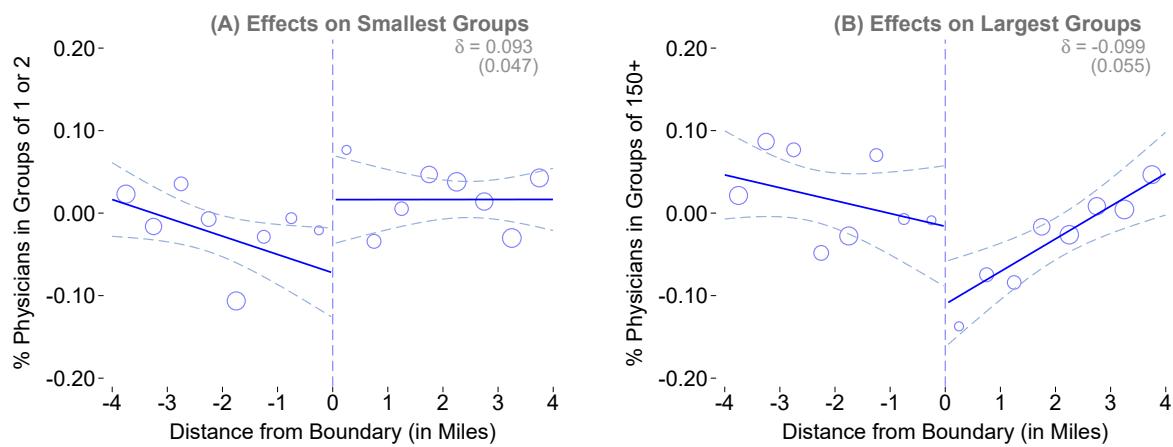


Figure B.6: Distributional Results: Effects on Smallest and Largest Groups

Notes: The purpose of this figure is to present regression discontinuity plots analogous to Figure 3.5 for two other measures of group size. Panel A represents the impact of differential payments on the smallest groups (1-2 providers) while Panel B displays the results for the largest groups (150 providers or more). See Figure 3.5 for more detail on the structure of these plots.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

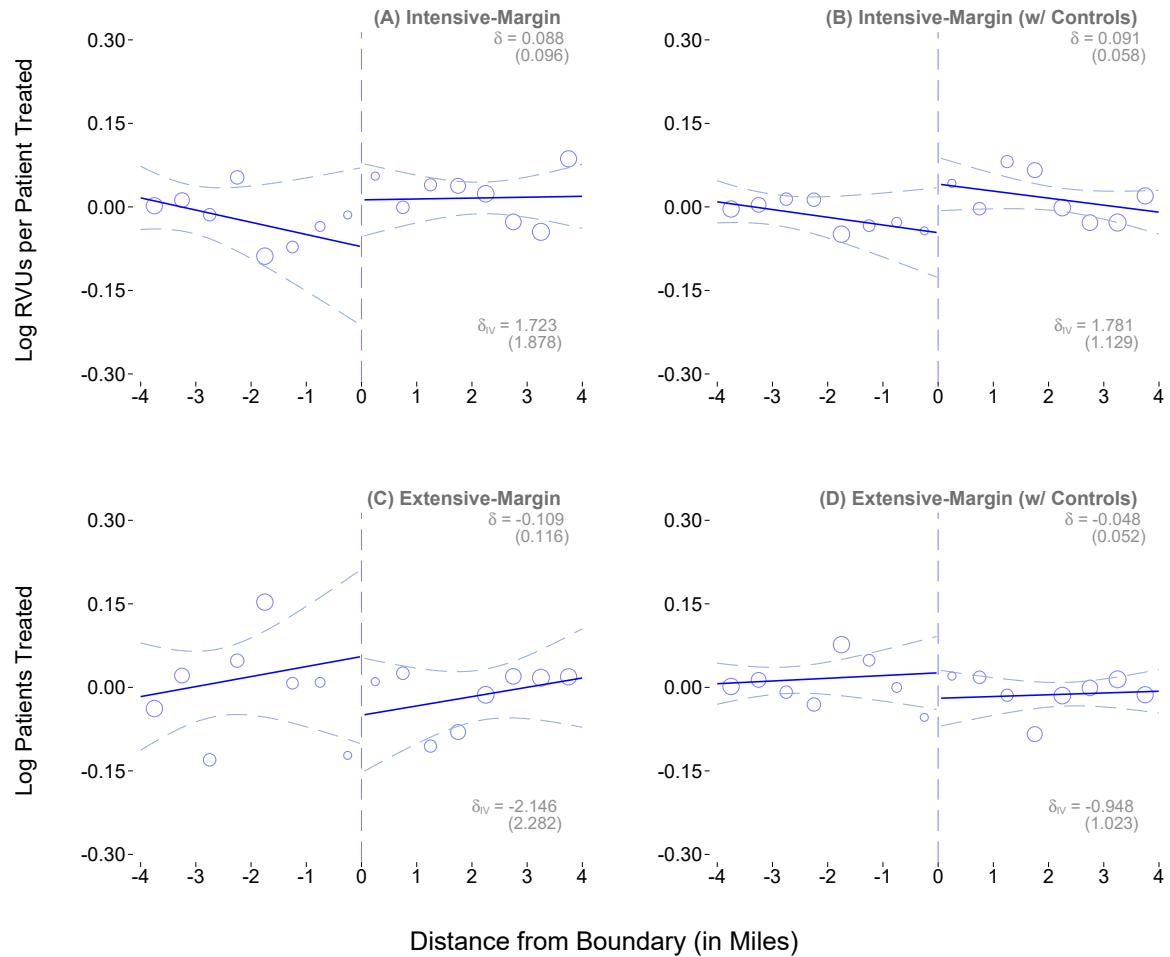


Figure B.7: Other Results: Treatment Outcomes

Notes: The purpose of this figure is to present regression discontinuity plots analogous to Figure 3.5 for measures of physician practice style. See Appendix Section C for further discussion and Figure 3.5 for more detail on the structure of these plots.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

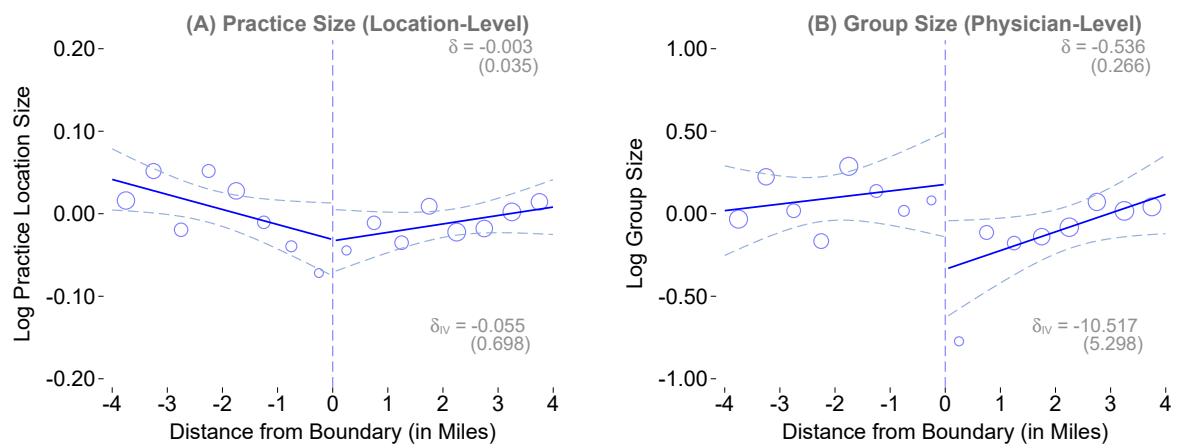


Figure B.8: Other Results: Other Practice Structure Outcomes

Notes: The purpose of this figure is to present regression discontinuity plots analogous to Figure 3.5 for other measures of practice structure. See Section 3.5.1 for further discussion and Figure 3.5 for more detail on the structure of these plots.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

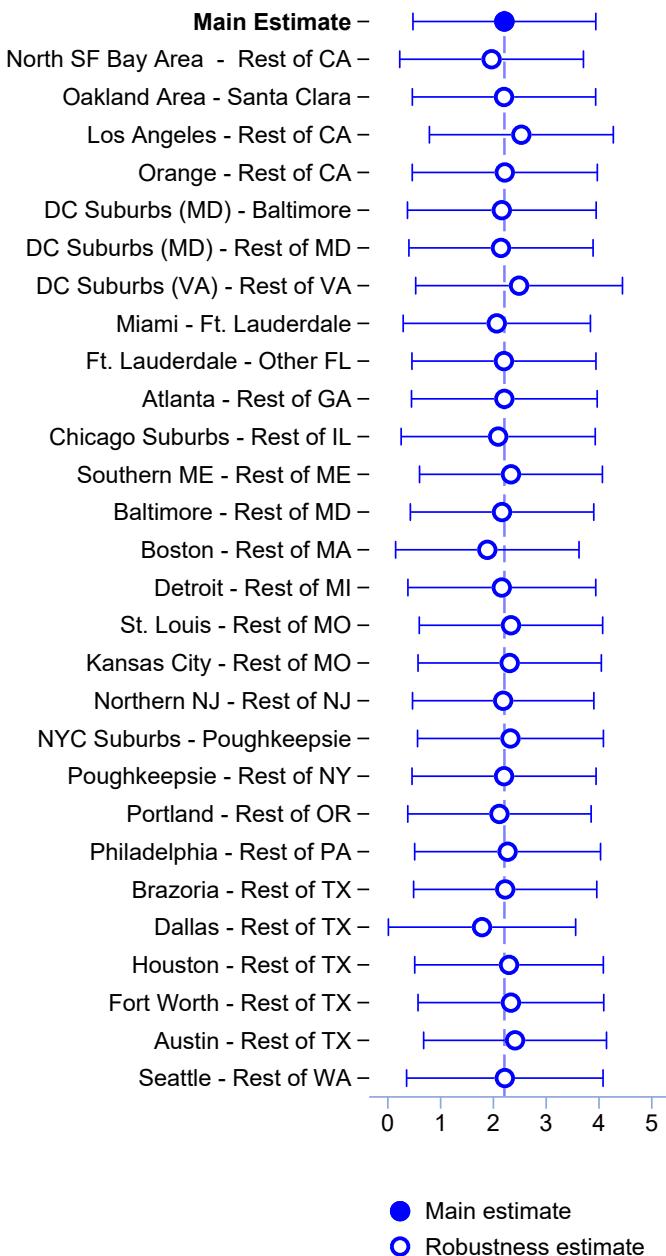


Figure B.9: Border Leave-one-out Analysis

Notes: This figure exhibits sensitivity of estimates to the effects of individual borders. Accordingly, each coefficient presented represents a separate regression of the preferred specification of Equations 3.10 and 3.11 with the listed border name omitted from the sample. Bars indicate 90% confidence intervals.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File, Census Block Group 5-year ACS Estimates (Ruggles et al., 2019).

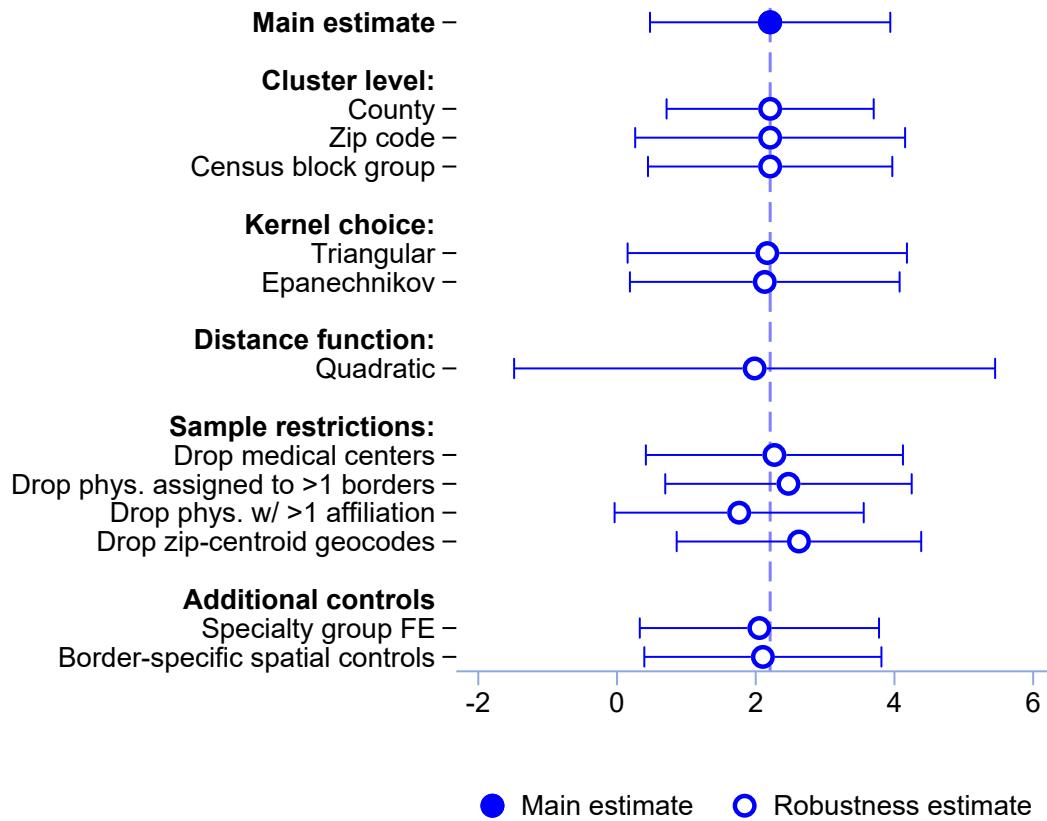


Figure B.10: Other Robustness Checks

Notes: The purpose of this figure is to perform additional robustness checks of my main analysis. Each point represents a separate regression of the preferred specification of Equations 3.10 and 3.11 with the listed modification. Bars indicate 90% confidence intervals. See the text of Appendix Section C for description of each robustness test.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File, Census Block Group 5-year ACS Estimates (Ruggles et al., 2019).

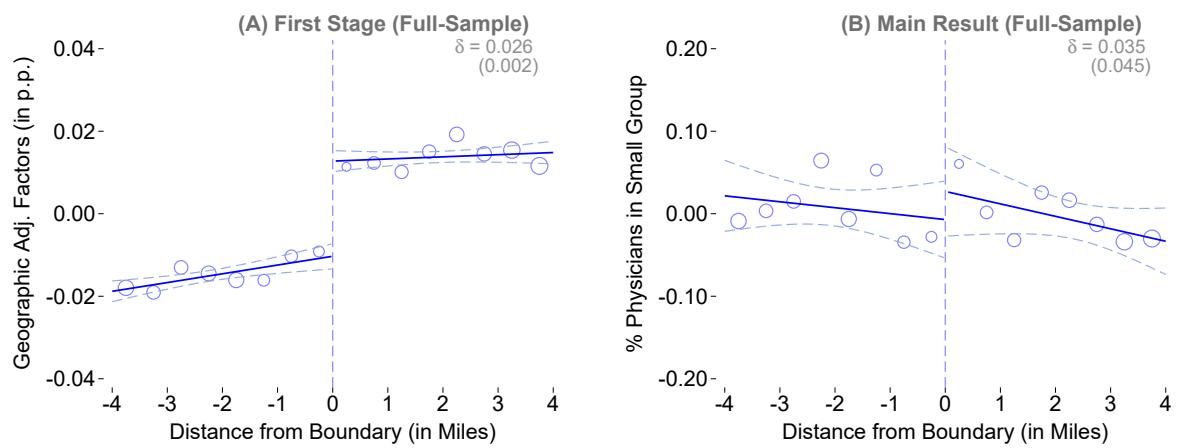


Figure B.11: Main Results: Full Sample

Notes: The purpose of this figure is to recreate the analysis displayed in Figure 3.5 when all borders are included. See discussion in Section 3.5.3 and notes to Figure 3.5 for more detail.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

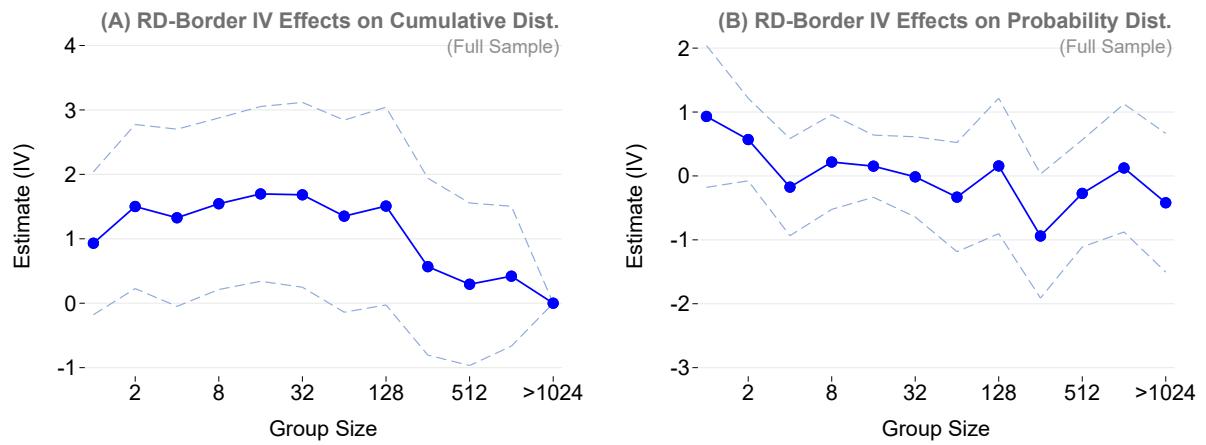


Figure B.12: Distributional Effects on Group Size

Notes: The purpose of this figure is to recreate Figure 3.6 (Panel A) and Appendix Figure B.5 (Panel B) when the full sample of borders are included and the RD-Border Estimation (Equations 3.12 and 3.13) are used. See Sections 3.5.1 and 3.5.3 and notes to referenced figures for more detail.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, and CMS Physician and Other Supplier Public Use File.

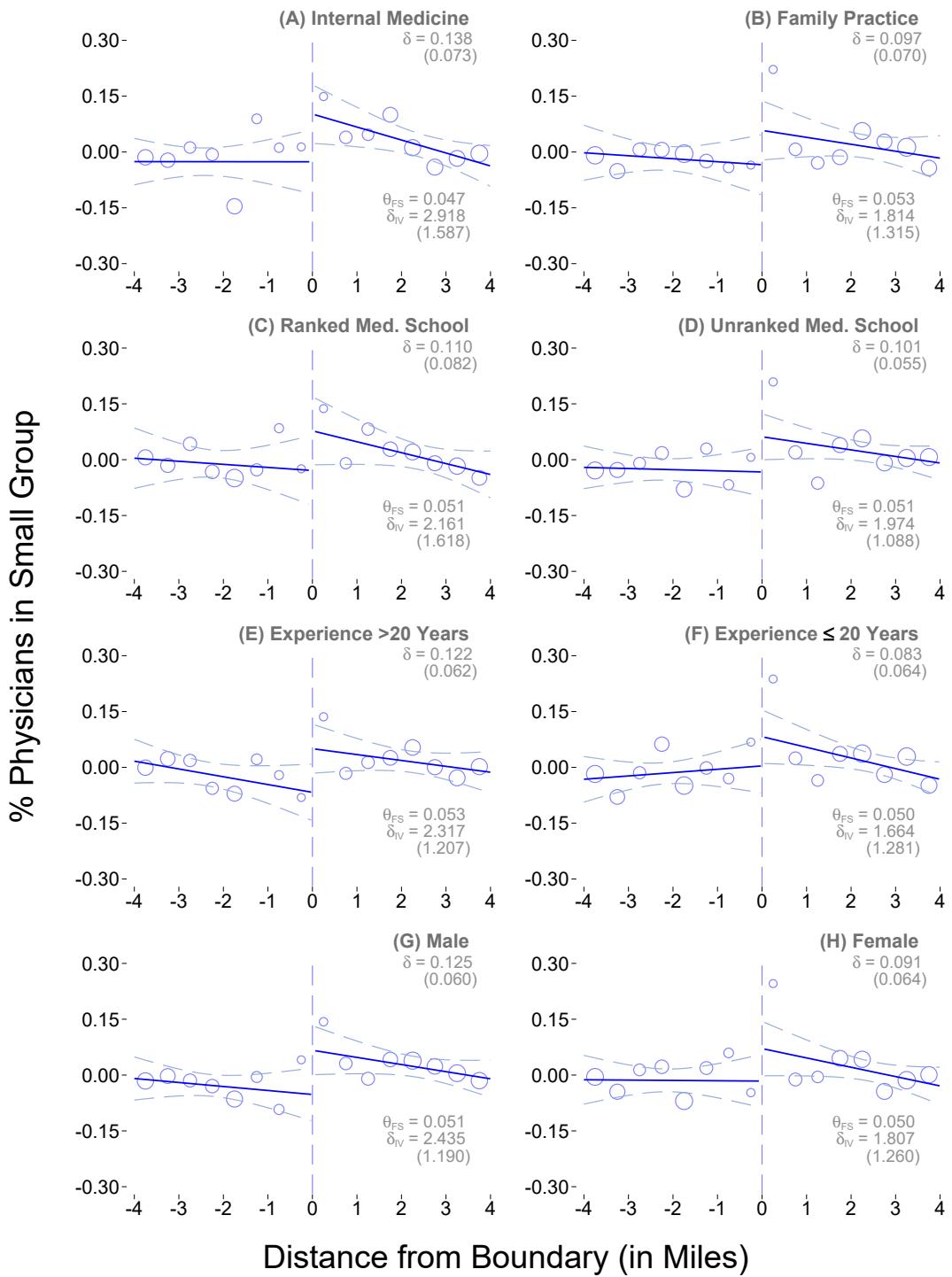


Figure B.13: Heterogeneity by Physician Characteristics

Notes: The purpose of this figure is to display heterogeneity of results by physician characteristics. Within each panel, reduced-form estimates (Equation 3.8) and associated standard errors are in the upper-right corner, while the first-stage (Equation 3.11) and RD-IV estimate (δ_{IV} , from Equation 3.10) and associated standard errors are in the bottom-right corner. See discussion in Appendix Section C for more detail.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

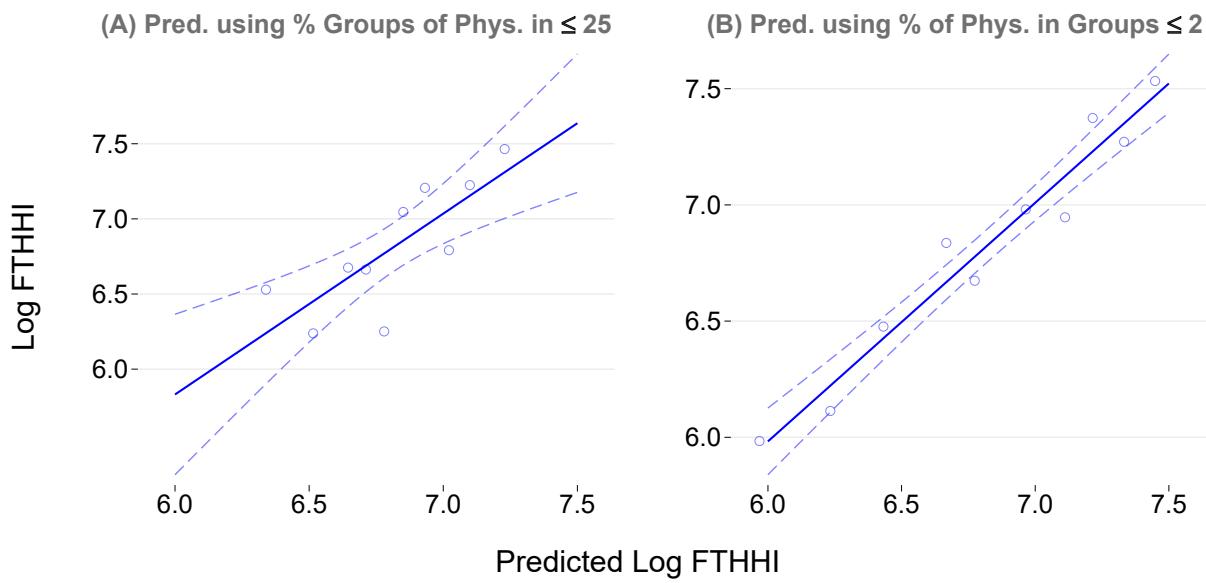


Figure B.14: Prediction of Log FTHHI by Group Size Measure

Notes: The purpose of this figure is display how well group size measures can be used to predict log fixed travel HHI (“FTHHI”). Panel A utilizes the main outcome, % of providers in a small group (defined as 25 providers or fewer), while Panel B utilizes the best predictor among the statistically significant distributional results, the % of providers in groups of 1-2. Each point of the graph represents a binned decile of predicted log FTHHI.

Sources: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File, Census Tract 5-year ACS Estimates (Ruggles et al., 2019).

Table B.1: Medicare-Designated Payment Localities

State	Metro Area or County	State	Metro Area or County	State	Metro Area or County
California	Marin/Napa/Solano	Illinois	Suburban Chicago	New York	Queens
California	San Francisco	Illinois	Chicago	Oregon	Portland
California	San Mateo	Louisiana	New Orleans	Pennsylvania	Metro Philadelphia
California	Oakland/Berkley	Maine	Southern Maine	Texas	Brazoria
California	Santa Clara	Maryland	Baltimore/Surr. Cty's	Texas	Dallas
California	Ventura	Massachusetts	Metro Boston	Texas	Galveston
California	Los Angeles	Michigan	Detroit	Texas	Houston
California	Anaheim/Santa Ana	Missouri	Metro St. Louis	Texas	Beaumont
District Of Columbia	DC + MD/VA Suburbs	Missouri	Metro Kansas City	Texas	Fort Worth
Florida	Fort Lauderdale	New Jersey	Northern NJ	Texas	Austin
Florida	Miami	New York	Manhattan	Washington	Seattle (King County)
Georgia	Atlanta	New York	NYC Suburbs		
Illinois	East St. Louis	New York	North NYC Suburbs		

Notes: This table reports all payment localities that receive different physician fee schedule payments than their respective states. See Figure 3.1(a) for more information. Note that some areas listed here are not included in Figure B.9 as certain borders did not have any points meeting the criteria discussed in Section B.

Sources: CMS Physician Fee Schedule.

Table B.2: Alternate Results: Treatment Outcomes

Outcomes	(1)	(2)	(3)	(4)
Log RVUs per patient	1.723 (1.878)	1.894 (2.250)	1.288 (1.438)	1.242 (1.421)
Log unique patients	-2.146 (2.282)	1.117 (2.836)	-1.062 (1.829)	-0.891 (1.798)
Sample	High Impact	Full	High Impact	Full
Instrumental variable	RD-IV	RD-IV	RD-Border	RD-Border
Observations	5,065	9,685	5,065	9,685
First stage F-statistic	399	148	7,468	5,920

Notes: This table recreates the analysis in Table 3.2 for physician practice style variables. See Section C for further discussion of these variables and Section 3.5.3 for further discussion of alternate samples and instrumental variables. See also Table 3.2 for more information on table structure.

Source: Author calculations using CMS Physician Fee Schedule, CMS Physician Compare Data, CMS Physician and Other Supplier Public Use File.

Bibliography

- Acemoglu, D. and A. Finkelstein (2008). Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector. *Journal of Political Economy* 116(5), 837–880.
- Aguilera, R., T. Corrington, A. Gershunov, and T. Benmarhnia (2021a). Wildfire Smoke Impacts Respiratory Health More Than Fine Particles from Other Sources: Observational Evidence from Southern California. *Nature Communications* 12(1), 1493.
- Aguilera, R., T. Corrington, A. Gershunov, S. Leibel, and T. Benmarhnia (2021b). Fine Particles in Wildfire Smoke and Pediatric Respiratory Health in California. *Pediatrics* 147(4).
- Aizer, A. and J. Currie (2019). Lead and Juvenile Delinquency: New Evidence from Linked Birth, School, and Juvenile Detention Records. *Review of Economics and Statistics* 101(4), 575–587.
- Aizer, A., S. Eli, J. Ferrie, and A. Lleras-Muney (2016). The Long-Run Impact of Cash Transfers to Poor Families. *American Economic Review* 106(4), 935–71.
- Alexander, D. and M. Schnell (2019). The Impacts of Physician Payments on Patient Access, Use, and Health. Working Paper 26095, National Bureau of Economic Research.
- Almond, D., J. Currie, and V. Duque (2018). Childhood Circumstances and Adult Outcomes: Act II. *Journal of Economic Literature* 56(4), 1360–1446.
- Ambrus, A., E. Field, and R. Gonzalez (2020, February). "Loss in the Time of Cholera: Long-run Impact of a Disease Epidemic on the Urban Landscape". *American Economic Review* 110(2), 475–525.
- American Medical Association (2016). Competition in Health Insurance: A Comprehensive Study of U.S. Markets, 15th Edition.

- Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 103(484), 1481–1495.
- Bailey, M., K. Clay, P. Fishback, M. R. Haines, S. Kantor, E. Severnini, and A. Wentz (2018). U.S. County-Level Natality and Mortality Data, 1915-2007. Inter-university Consortium for Political and Social Research [distributor].
- Bailey, M. J., H. W. Hoynes, M. Rossin-Slater, and R. Walker (2020). Is the Social Safety Net a Long-Term Investment? Large-scale Evidence from the Food Stamps Program. Working Paper 26942, National Bureau of Economic Research.
- Bailey, M. J., S. Sun, and B. D. Timpe (2021). Prep School for Poor Kids: The Long-Run Impacts of Head Start on Human Capital and Economic Self-Sufficiency. Working Paper 28268, National Bureau of Economic Research.
- Baker, L. C., M. K. Bundorf, and D. P. Kessler (2014). Vertical Integration: Hospital Ownership Of Physician Practices Is Associated With Higher Prices And Spending. *Health Affairs* (5), 756–763.
- Baker, L. C., M. K. Bundorf, and A. B. Royalty (2019). The Effects of Multispecialty Group Practice on Health Care Spending and Use. Working Paper 25915, National Bureau of Economic Research.
- Barcellos, S. H., L. S. Carvalho, and P. Turley (2019). Distributional Effects of Education on Health. Working Paper 25898, National Bureau of Economic Research.
- Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro (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.
- Barrette, E., G. Gowrisankaran, and R. Town (2020). Countervailing Market Power and Hospital Competition. Working Paper 24354, National Bureau of Economic Research.
- Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006). Adaptive Linear Step-up Procedures That Control the False Discovery Rate. *Biometrika* 93(3), 491–507.

- Benjamini, Y. and D. Yekutieli (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics* 29(4), 1165–1188.
- Billings, S. B. and K. T. Schnepel (2018, July). Life after Lead: Effects of Early Interventions for Children Exposed to Lead. *American Economic Journal: Applied Economics* 10(3), 315–44.
- Bishop, K. C., J. D. Ketcham, and N. V. Kuminoff (2018). Hazed and Confused: The Effect of Air Pollution on Dementia. Working Paper 24970, National Bureau of Economic Research.
- Borgschulte, M., D. Molitor, and E. Zou (2020). Air Pollution and the Labor Market: Evidence from Wildfire Smoke. Working paper. Retrieved from <http://www.davidmolitor.com/research> [Date Accessed: February 22, 2021].
- Boxall, B. (2021). Billions of Dollars Spent on Fighting California Wildfires, but Little on Prevention. *The Los Angeles Times*.
- Buckles, K. S. and D. M. Hungerman (2013). Season of Birth and Later Outcomes: Old Questions, New Answers. *Review of Economics and Statistics* 95(3), 711–724.
- Bureau of Justice Statistics (2016). Annual Determination of Average Cost of Incarceration. *Federal Register* 81 (July 19, 2016): 46957.
- Burke, M., S. Heft-Neal, J. Li, A. Driscoll, P. W. Baylis, M. Stigler, J. Weill, J. Burney, J. Wen, M. Childs, and C. Gould (2021). Exposures and Behavioral Responses to Wildfire Smoke. Working Paper 29380, National Bureau of Economic Research.
- California Air Resources Board (2021). Camp Fire Air Quality Data Analysis. Technical report. <https://ww2.arb.ca.gov/resources/documents/camp-fire-air-quality-data-analysis> (accessed August 17, 2021).
- California Department of Forestry and Fire Protection (2018). Fire Perimeters - California [ds396]. <https://map.dfg.ca.gov/metadata/ds0396.html> [Date Accessed: October 18, 2018].
- Callaway, B., A. Goodman-Bacon, and P. H. C. Sant'Anna (2021). Difference-in-Differences with a Continuous Treatment. Working paper, arXiv.
- Callaway, B. and P. H. Sant'Anna (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics* 225(2), 200–230. Themed Issue: Treatment Effect 1.

- Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82(6), 2295–2326.
- Capps, C., D. Dranove, and C. Ody (2017). Physician Practice Consolidation Driven By Small Acquisitions, So Antitrust Agencies Have Few Tools To Intervene. *Health Affairs* 36(9), 1556–1563.
- Card, D. (1999). The Causal Effect of Education on Earnings. In O. Ashenfelter, D. Card, and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 3A, Chapter 30, pp. 1801–1863. Amsterdam: Elsevier Science & Technology.
- Carleton, T. A., A. Jina, M. T. Delgado, M. Greenstone, T. Houser, S. M. Hsiang, A. Hultgren, R. E. Kopp, K. E. McCusker, I. B. Nath, J. Rising, A. Rode, H. K. Seo, A. Viaene, J. Yuan, and A. T. Zhang (2020). Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits. Working Paper 27599, National Bureau of Economic Research.
- Casalino, L., K. Devers, T. Lake, M. Reed, and J. Stoddard (2003). Benefits of and Barriers to Large Medical Group Practice in the United States. *Archives of Internal Medicine* 163(16), 1958–1964.
- Case, A. and C. Paxson (2010). Causes and Consequences of Early-Life Health. *Demography* 47(1), S65–S85.
- Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The Effect of Minimum Wages on Low-Wage Jobs*. *The Quarterly Journal of Economics* 134(3), 1405–1454.
- Center for Medicare and Medicaid Services (2018). Historical National Health Expenditure Accounts. <https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/Tables.zip>. Accessed 5/3/2018.
- Centers for Medicare and Medicaid Services (2015). Physician Value-Based Payment Modifier and the Physician Feedback Program. <https://www.cms.gov/newsroom/fact-sheets/physician-value-based-payment-modifier-and-physician-feedback-program>. Accessed 8/2/2020.
- Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2013). Evidence On the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy. *Proceedings of the National Academy of Sciences* 110(32), 12936–12941.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA* 315(16), 1750–1766.

Clark, D. and H. Royer (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *American Economic Review* 103(6), 2087–2120.

Clay, K., J. Lewis, and E. Severnini (2016). Canary in a Coal Mine: Infant Mortality Tradeoffs Associated with Mid-20th Century Air Pollution. Working Paper 22155, National Bureau of Economic Research.

Clemens, J. and J. D. Gottlieb (2014). Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health? *The American Economic Review* 104(4), 1320–1349.

Clemens, J. and J. D. Gottlieb (2017). In the Shadow of a Giant: Medicare's Influence on Private Physician Payments. *Journal of Political Economy* 125(1), 1–39.

Colmer, J. and J. Voorheis (2020). The Grandkids Aren't Alright: The Intergenerational Effects of Prenatal Pollution Exposure. Technical report, Working Paper.

Craig, S., M. Grennan, and A. Swanson (2018). Mergers and Marginal Costs: New Evidence on Hospital Buyer Power. Working Paper 24926, National Bureau of Economic Research.

Currie, J. and D. Almond (2011). Human Capital Development Before Age Five. In *Handbook of Labor Economics*, Volume 4, pp. 1315–1486. Elsevier.

Cutler, D. M. and E. Richardson (1999). Your Money and Your Life: The Value of Health and What Affects It. In A. M. Garber (Ed.), *Frontiers in Health Policy Research*, Volume 2, pp. 99–132. MIT.

de Chaisemartin, C. and X. D'Haultfœuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review* 110(9), 2964–96.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica* 78(6), 1863–1903.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif (2019). The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction. *American Economic Review* 109(12), 4178–4219.

- Dranove, D. and C. Ody (2019). Employed for Higher Pay? How Medicare Payment Rules Affect Hospital Employment of Physicians. *American Economic Journal: Economic Policy* 11(4), 249–71.
- Dunn, A. and A. Shapiro (2014). Do Physicians Possess Market Power? *The Journal of Law and Economics* 57(1), 159–193.
- Epstein, A. J., J. D. Ketcham, and S. Nicholson (2010). Specialization and Matching in Professional Services Firms. *The RAND Journal of Economics* 41(4), 811–834.
- Finkelstein, A. (2007). The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare. *The Quarterly Journal of Economics* 122(1), 1–37.
- Finlay, K. and K. R. Genadek (2021). Measuring All-Cause Mortality With the Census Numident File. *American Journal of Public Health* 111(S2), S141–S148. PMID: 34314212.
- Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry (2021). IPUMS, Current Population Survey: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2021. <https://doi.org/10.18128/D030.V9.0>.
- Flynn, P. and M. M. Marcus (2021). A Watershed Moment: The Clean Water Act and Infant Health. Working Paper 29152, National Bureau of Economic Research.
- Gibson, M. (2019). Regulation-Induced Pollution Substitution. *The Review of Economics and Statistics* 101(5), 827–840.
- Gilraine, M. (2020). Air Filters, Pollution and Student Achievement. Working Paper 20-188, Annenberg Institute at Brown University.
- Glied, S. and J. G. Zivin (2002). How Do Doctors Behave When Some (But Not All) of Their Patients Are in Managed Care? *Journal of Health Economics* 21(2), 337 – 353.
- Goldstein, J. R., M. Alexander, C. Breen, A. M. González, F. Menares, M. Osborne, M. Snyder, and U. Yildirim (2021). CenSoc Mortality File: Version 2.0.
- Goodman-Bacon, A. (2021a). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics* 225(2), 254–277. Themed Issue: Treatment Effect 1.

- Goodman-Bacon, A. (2021b). The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes. *American Economic Review* 111(8), 2550–93.
- Gowrisankaran, G., A. Nevo, and R. Town (2015). Mergers When Prices Are Negotiated: Evidence from the Hospital Industry. *American Economic Review* 105(1), 172–203.
- Graff Zivin, J. and M. Neidell (2014). Temperature and the Allocation of Time: Implications for Climate Change. *Journal of Labor Economics* 32(1), 1–26.
- Grainger, C. and T. Ruangmas (Nov 2018). Who Wins from Emissions Trading? Evidence from California. *Environmental and Resource Economics* 71(3), 703–727.
- Grainger, C., A. Schreiber, and W. Chang (2017). Do Regulators Strategically Avoid Pollution Hotspots When Siting Monitors? Evidence from Remote Sensing of Air Pollution. Working paper.
- Greenstone, M. (2002). The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures. *Journal of Political Economy* 110(6), 1175–1219.
- Grönqvist, H., J. P. Nilsson, and P.-O. Robling (2020). Understanding How Low Levels of Early Lead Exposure Affect Children’s Life Trajectories. *Journal of Political Economy* 128(9), 3376–3433.
- Holm, S. M., M. D. Miller, and J. R. Balmes (2021). Health Effects of Wildfire Smoke in Children and Public Health Tools: a Narrative Review. *Journal of Exposure Science & Environmental Epidemiology* 31(1), 1–20.
- Hoynes, H., D. W. Schanzenbach, and D. Almond (2016). Long-Run Impacts of Childhood Access to the Safety Net. *American Economic Review* 106(4), 903–34.
- Hsiang, S., R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. J. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, K. Larsen, and T. Houser (2017). Estimating Economic Damage from Climate Change in the United States. *Science* 356(6345), 1362–1369.
- IPCC (2021). Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Working

- paper. [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press.
- Isen, A., M. Rossin-Slater, and R. Walker (2017a). Relationship Between Season of Birth, Temperature Exposure, and Later Life Wellbeing. *Proceedings of the National Academy of Sciences* 114(51), 13447–13452.
- Isen, A., M. Rossin-Slater, and W. R. Walker (2017b). Every Breath You Take — Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970. *Journal of Political Economy*.
- Jabbarpour, Y., A. Greiner, A. Jetty, M. Coffman, C. Jose, S. Petterson, K. Pivaral, R. Phillips, A. Bazemore, and A. N. Kane (2019). Investing in Primary Care: A State-Level Analysis. Technical report, Patient-Centered Primary Care Collaborative and the Robert Graham Center.
- Jacob, B., L. Lefgren, and E. Moretti (2007). The Dynamics of Criminal Behavior Evidence from Weather Shocks. *Journal of Human resources* 42(3), 489–527.
- Jayachandran, S. (2009). Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires. *Journal of Human Resources* 44(4), 916–954.
- Kane, C. K. (2019). Updated Data on Physician Practice Arrangements: For the First Time, Fewer Physicians are Owners Than Employees. Technical report, American Medical Association.
- Kessler, D. P. and M. B. McClellan (2000). Is Hospital Competition Socially Wasteful? *The Quarterly Journal of Economics* 115(2), 577–615.
- Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica* 75(1), 83–119.
- Landguth, E. L., Z. A. Holden, J. Graham, B. Stark, E. B. Mokhtari, E. Kaleczyc, S. Anderson, S. Urbanski, M. Jolly, E. O. Semmens, D. A. Warren, A. Swanson, E. Stone, and C. Noonan (2020). The Delayed Effect of Wildfire Season Particulate Matter on Subsequent Influenza Season in a Mountain West Region of the USA. *Environment International* 139, 105668.
- Larkin, N. K., S. M. O'Neill, R. Solomon, S. Raffuse, T. Strand, D. C. Sullivan, C. Krull, M. Rorig,

- J. Peterson, and S. A. Ferguson (2009). The BlueSky Smoke Modeling Framework. *International Journal of Wildland Fire* 18, 906–920.
- Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies* 76(3), 1071–1102.
- Liang, Y., D. Sengupta, M. J. Campmier, D. M. Lunderberg, J. S. Apte, and A. H. Goldstein (2021). Wildfire Smoke Impacts on Indoor Air Quality Assessed Using Crowdsourced Data in California. *Proceedings of the National Academy of Sciences* 118(36).
- Lipton, W. (2012). Is a Bigger Medical Practice Always Better? <http://www.physicianspractice.com/overhead/bigger-medical-practice-always-better>. Accessed 5/3/2018.
- Logan, J. and W. Zhang (2020). Developing GIS Maps for U.S. Cities in 1930 and 1940. In I. Gregory, D. DeBats, and D. Lafreniere (Eds.), *The Routledge Companion to Spatial History*. UK: Routledge.
- Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles (2021). IPUMS National Historical Geographic Information System: Version 16.0 [dataset]. Minneapolis, MN. <http://doi.org/10.18128/D050.V16.0>.
- McCrary, J. (2008). Manipulation of the Running variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics* 142(2), 698–714.
- McGuire, T. G. and M. V. Pauly (1991). Physician Response to Fee Changes with Multiple Payers. *Journal of Health Economics* 10(4), 385 – 410.
- Meghir, C., M. Palme, and E. Simeonova (2018). Education and Mortality: Evidence from a Social Experiment. *American Economic Journal: Applied Economics* 10(2), 234–56.
- Miller, N., D. Molitor, and E. Zou (2017). Blowing Smoke: Health Impacts of Wildfire Plume Dynamics. Working paper. Retrieved from <http://www.eric-zou.com/> [Date Accessed: December 2, 2018].
- Moritz, M. A., M.-A. Parisien, E. Batllori, M. A. Krawchuk, J. Van Dorn, D. J. Ganz, and K. Hayhoe (2012). Climate Change and Disruptions to Global Fire Activity. *Ecosphere* 3(6), art49.

- Muhlestein, D. B. and N. J. Smith (2016). Physician Consolidation: Rapid Movement From Small To Large Group Practices, 2013–15. *Health Affairs* 35(9), 1638–1642.
- Nicholson, S. and C. Propper (2012). Medical Workforce. In M. V. Pauly, T. G. McGuire, and P. P. Barros (Eds.), *Handbook of Health Economics*, Volume two, Chapter 14, pp. 873–925. Amsterdam: North Holland.
- Oreopoulos, P. and K. G. Salvanes (2011). Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives* 25(1), 159–84.
- Peeples, L. (2020). How Air Pollution Threatens Brain Health. *Proceedings of the National Academy of Sciences* 117(25), 13856–13860.
- Plantinga, A. J., R. Walsh, and M. Wibbenmeyer (2020). Priorities and Effectiveness in Wildfire Management: Evidence from Fire Spread in the Western United States. Technical Report 20-21, Resources for the Future.
- Radeloff, V. C., D. P. Helmers, H. A. Kramer, M. H. Mockrin, P. M. Alexandre, A. Bar-Massada, V. Butsic, T. J. Hawbaker, S. Martinuzzi, A. D. Syphard, and S. I. Stewart (2018). Rapid Growth of the US Wildland-Urban Interface Raises Wildfire Risk. *Proceedings of the National Academy of Sciences* 115(13), 3314–3319.
- Rangel, M. A. and T. Vogl (2019). Agricultural Fires and Health at Birth. *Review of Economics and Statistics*.
- Reyes, J. W. (2007). Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime. *The B.E. Journal of Economic Analysis & Policy* 7(1).
- Reyes, J. W. (2015). Lead Exposure and Behavior: Effects on Antisocial and Risky Behavior Among Children and Adolescents. *Economic Inquiry* 53(3), 1580–1605.
- Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021). IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. <https://doi.org/10.18128/D010.V11.0>.
- Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2019). IPUMS USA: Version 9.0 [dataset]. Minneapolis, MN. <https://doi.org/10.18128/D010.V9.0>.

- Russ, T. C., M. P. C. Cherrie, C. Dibben, S. Tomlinson, S. Reis, U. Dragosits, M. Vieno, R. Beck, E. Carnell, N. K. Shortt, G. Muniz-Terrera, P. Redmond, A. M. Taylor, T. Clemens, M. van Tongeren, R. M. Agius, J. M. Starr, I. J. Deary, and J. R. Pearce (2021). Life Course Air Pollution Exposure and Cognitive Decline: Modelled Historical Air Pollution Data and the Lothian Birth Cohort 1936. *Journal of Alzheimer's disease : JAD* 79(3), 1063–1074.
- Sanders, N. J. (2012). What Doesn't Kill You Makes You Weaker Prenatal Pollution Exposure and Educational Outcomes. *Journal of Human Resources* 47(3), 826–850.
- Sawyer, W. and P. Wagner (2020). Mass Incarceration: The Whole Pie 2020. Technical report, Prison Policy Initiative.
- Schlenker, W. and M. J. Roberts (2009). Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change. *Proceedings of the National Academy of Sciences* 106(37), 15594–15598.
- Schlenker, W. and W. R. Walker (2015). Airports, Air Pollution, and Contemporaneous Health. *The Review of Economic Studies* 83(2), 768–809.
- Schnell, M. and J. Currie (2018). Addressing the Opioid Epidemic: Is There a Role for Physician Education? *American Journal of Health Economics* 4(3), 383–410.
- Schraufnagel, D. E. (2020). The Health Effects of Ultrafine Particles. *Experimental & Molecular Medicine* 52(3), 311–317.
- Schultz, C. A., S. M. McCaffrey, and H. R. Huber-Stearns (2019). Policy Barriers and Opportunities for Prescribed Fire Application in the Western United States. *International Journal of Wildland Fire* 28(11), 874–884.
- Schwank, H. (2021). Disruptive Effects of Natural Disasters: The 1906 San Francisco Fire. Working paper. <https://sites.google.com/bu.edu/hannaschwank/research> (accessed November 28, 2021).
- Scott, K. W., E. J. Orav, D. M. Cutler, and A. K. Jha (2017). Changes in Hospital–Physician Affiliations in U.S. Hospitals and Their Effect on Quality of Care. *Annals of Internal Medicine* 166(1), 1–8.
- Short, M. N. and V. Ho (2019). Weighing the Effects of Vertical Integration Versus Market Concentration on Hospital Quality. *Medical Care Research and Review*, 1–18. PMID: 30741109.

- Sommer, L. (2021). Why The South Is Decades Ahead Of The West In Wildfire Prevention. <https://www.npr.org/2021/08/31/1029821831/to-stop-extreme-wildfires-california-is-learning-from-florida> (retrieved October 30, 2021).
- Stafford, T. M. (2015). Indoor Air Quality and Academic Performance. *Journal of Environmental Economics and Management* 70, 34–50.
- Stein, A. F., R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan (2015). NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society* 96(12), 2059 – 2077.
- Suades-González, E., M. Gascon, M. Guxens, and J. Sunyer (2015). Air Pollution and Neuropsychological Development: A Review of the Latest Evidence. *Endocrinology* 156(10), 3473–3482.
- Sullivan, D. (2017). The True Cost of Air Pollution: Evidence from the Housing Market. Working paper. Retrieved from <http://www.danielmsullivan.com/> [Date Accessed: October 7, 2021].
- Sun, L. and S. Abraham (2021). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics* 225(2), 175–199. Themed Issue: Treatment Effect 1.
- Tax Policy Center (2021). Historical Average Federal Tax Rates for All Households. <https://www.taxpolicycenter.org/statistics/percent-tax-filers-marginal-tax-rate> [accessed October 22, 2021].
- U.S. Census Bureau (2012). County and City Data Book [United States] Consolidated File: County Data, 1947-1977. Inter-university Consortium for Political and Social Research [distributor]. 10.3886/ICPSR07736.v2.
- Vera Institute of Justice (2012). The Price of Prisons: What Incarceration Costs Taxpayers. Technical report. Technical report, Center on Sentencing and Corrections.
- Vera Institute of Justice (2015). The Price of Jails: Measuring the Taxpayer Cost of Local Incarceration. Technical report. Technical report, Center on Sentencing and Corrections.
- Voorheis, J. (2017). Air Quality, Human Capital Formation and the Long-term Effects of Environmental Inequality at Birth. Technical report, CARRA Working Papers 2017-05, Center for Economic Studies, U.S. Census Bureau.

Vose, J. M., D. L. Peterson, and T. Patel-Weynand (2012). Effects of Climatic Variability and Change on Forest Ecosystems: A Comprehensive Science Synthesis for the U.S. Forest Sector. Gen. tech. rep. pnw-gtr-870. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Walker, W. R. (2013). The Transitional Costs of Sectoral Reallocation: Evidence From the Clean Air Act and the Workforce*. *The Quarterly Journal of Economics* 128(4), 1787–1835.

Wegesser, T. C., L. M. Franz, F. M. Mitloehner, A. Eiguren-Fernandez, and J. A. Last (2010). Lung Antioxidant and Cytokine Responses to Coarse and Fine Particulate Matter from the Great California Wildfires of 2008. *Inhalation Toxicology* 22(7), 561–570.

Willmott, C. J. and K. Matsuura (2001). Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999).

Xierali, I. M. (2018). Physician Multisite Practicing: Impact on Access to Care. *Journal of the American Board of Family Medicine* 31(2), 260–269.

Zhuang, Y., R. Fu, B. D. Santer, R. E. Dickinson, and A. Hall (2021). Quantifying Contributions of Natural Variability and Anthropogenic Forcings on Increased Fire Weather Risk over the Western United States. *Proceedings of the National Academy of Sciences* 118(45).

Zou, E. Y. (2021). Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality. *American Economic Review* 111(7), 2101–26.

Zuckerman, S., L. Skopec, and M. Epstein (2017). Medicaid Physician Fees after the ACA Primary Care Fee Bump. Technical report, Urban Institute.

Contribution Statement

Chapters 1 and 2 of this dissertation are adapted from the manuscript, “Ashes to Ashes: The Lifelong Consequences of Early-Life Wildfire Exposure,” which is jointly authored with Samuel Arenberg. My contribution included: (a) primary conception of the research idea and empirical design, (b) performance of primary statistical analysis, and (c) drafting of the manuscript. Activities conducted by both co-authors included (a) acquisition and cleaning of data, (b) critical revision of the manuscript, and (c) administrative activities.